School of Epidemiology and Public Health

Knowledge Synthesis Group

Breast cancer screening: protocol for an evidence report to inform an update of the Canadian Task Force on Preventive Health Care 2011 Guidelines

March 2017

Ottawa Evidence Review Synthesis Centre (ERSC)

Principal Investigators: David Moher, Julian Little
Centre Staff: Pauline Barbeau, Adrienne Stevens, Andrew Beck, Becky Skidmore (MLS)

Edmonton Evidence Review Synthesis Centre (ERSC)

Principal Investigator: Lisa Hartling
Centre Staff: Jennifer Pillay, Robin Featherstone (MLIS), Ben Vandermeer (Statistician)

Clinical Experts and Methodological Collaborators

Angel Arnaout, Muriel Brackstone, Alexandra Ginty, Amanda Hey, Anat Kornecki, Brian Hutton, Beverley Shea

Suggested citation:

Corresponding authors: Pauline Barbeau and Jennifer Pillay

Author affiliations

Key Question 1 - PROSPERO CRD#: 42017051498
Key Question 2- PROSPERO CRD #: 42017058476

Author contribution
PB, AB, AS and JP drafted the protocol, BS and RF developed the search strategy and provided text for the protocol. JL, DM, and LH contributed to discussions with the CTFPHC and PHAC on the scope for this work and critically reviewed the protocol including all methodological considerations and the Centre's approach for integrating existing systematic reviews.

Acknowledgements
We would like to acknowledge the contribution of the following individuals for their clinical and methodological input as well as their peer review of the protocol:

CTFPHC Working Group
Scott Klarenbach (Chair), Brett Thombs, Harminder Singh, Gabriela Lewin, Marcello Tonelli

PHAC Global Health & Guidelines Division Scientific Research Officers
Alejandra Jaramillo Garcia, Nicki Sims-Jones, Susan Courage, Dana Reid

External Reviewers
Anna M. Chiarelli, PhD, MHSc
Senior Scientist, Prevention and Cancer Control, Cancer Care Ontario
Provincial Lead Scientist, Ontario Breast Screening Program, Cancer Care Ontario
Professor, Dalla Lana School of Public Health, University of Toronto

Jennifer Payne, PhD
Associate Professor, Department of Diagnostic Radiology, Dalhousie University

Declaration of funding
Funding for this protocol and evidence update is provided by the Public Health Agency of Canada. This funding will support the collection of the data, data management, analyses, and writing of the protocol and the technical report.

Role of funder
The funder will provide feedback on the protocol and draft evidence report, but will not be involved in study selection, data extraction or analysis and will not be involved in a decision to seek publication.

Chapter I. Purpose and Background

The purpose of this report is to synthesize evidence about breast cancer screening to inform an update of the 2011 Canadian Task Force on Preventive Health Care (CTFPHC) recommendations (Canadian Task Force on Preventive Health Care, 2011). CTFPHC guidelines are updated approximately every five years or as new evidence becomes available (Canadian Task Force on Preventive Health Care, 2011). This evidence report will 1) seek new evidence syntheses, and assess the quality of evidence to date on the benefits and harms of screening determined by the CTFPHC as critical for decision-making, and 2) systematically review women's willingness to be screened or uptake of screening based on how they value the benefits and harms of screening. The CTFPHC will determine whether these findings change or reaffirm their 2011 recommendations (Canadian Task Force on Preventive Health Care, 2011).

Definition, Prevalence, and Burden

Breast cancer can be differentiated into two types: *in situ* (non-invasive), which can be classified as stage 0, or invasive/infiltrating classified as stages 1-4. *In situ* carcinomas are considered non-metastatic since they are derived from the ductal/lobular epithelium and do not extend past the basement membrane (Richie & Swanson, 2003; National Breast Cancer Foundation, 2015). They include different subtypes (ductal carcinoma in situ (DCIS), lobular carcinoma in situ, or noninvasive Paget disease of the nipple) and are considered highly curable. Some ductal carcinoma *in situ* lesions can progress to invasive types over time (varying by grade but generally between 5-15 years), with reported rates varying between 14% (low-grade DCIS) and 84% (high-grade DCIS) (Mukhtar, Wong, & Esserman, 2015; Feig, 2015). Invasive carcinomas have the potential to spread to the lymph nodes and to metastasis beyond the breast, resulting in greater morbidity and potential mortality (National Breast Cancer Foundation, 2015).

The burden and impact that breast cancer has on the Canadian population and the Canadian healthcare system are substantial. Breast cancer primarily affects women and is the second-leading cause of death within this group (DeSantis, Siegel, Bandi, & Jemal, 2011). One in nine women are expected to develop breast cancer in their lifetime (Canadian Cancer Society's Advisory Committee on Cancer Statistics, 2015). In 2016, the age-standardized incidence rate of breast cancer among women in Canada was 130.1 cases per 100,000 (Canadian Cancer Society's Advisory Committee on Cancer Statistics, 2016). The incidence rate in Canada has been stable since 2004, and variation in incidence between provinces in Canada is low. Death rates from breast cancer in women have been declining, from a peak of 41.7 per 100,000 in 1986 to a projected rate 23.4 per 100,000 in 2016 (Canadian Cancer Society's Advisory Committee on Cancer Statistics, 2016). Across many countries, declines in mortality despite stable incidence rates suggests an improvement in breast cancer diagnostic and care (Kleibl & Kristensen, 2016). In Canada, it is speculated that this relationship may be mediated by both screening with mammography and the use of more effective therapies after surgery (Canadian Cancer Society's Advisory Committee on Cancer Statistics, 2015), although some sources suggest the latter is the primary driver of improved outcomes (Welch et al., 2016).

Substantial economic burden may be experienced by women affected by breast cancer and their families, that may include job termination (16%) or demotion (12%), reduced physical ability to work (45%) among other direct and indirect costs (Canadian Breast Cancer Network, 2010). Out of 22 cancers examined in Ontario, breast cancer was determined to be the fourth costliest in regards to cancer care with an estimated total cost in 2009 of $108 million CAD (de Oliveira et al., 2013).

Risk Factors

The incidence of breast cancer increases with age. The majority of breast cancers occur in females 50–69 years of age (52%), approximately 30% were diagnosed in females aged 70 and over, and 18% occurred in females under age 50 (Canadian Cancer Society's Advisory Committee on Cancer Statistics, 2016). Excess body weight, low levels of physical activity, early age at menarche, late age at full term pregnancy or no pregnancy, late age at menopause, use of combination (estrogen and progestin together) hormone therapy and exposure to diethylstilboestrol are associated with increased risk for breast cancer (Canadian Cancer Society's Advisory Committee on Cancer Statistics, 2016; American Cancer Society, 2013; Cancer Research UK, 2016). Tobacco smoking and alcohol consumption have also been associated with increased risk (Scoccianti et al, 2014; International Agency for Research on Cancer, 2012; Johnson et al., 2010).

In the UK, it has been estimated that 1% of breast cancer is attributable to exposure to ionizing radiation, half of which is due to medical sources (Parkin & Darby, 2011). It has been estimated that 0.1% of breast cancer in women aged under 75 are caused by diagnostic X-rays (de Gonzalez & Darby, 2004), and 0.03-0.06% by receiving a mammogram (de Gonzalez et al., 2010). In an umbrella review of the association between multiple types of cancer and type 2 diabetes, there was a modest positive association between type 2 diabetes and incidence of breast cancer (OR 1.2, 95% prediction interval 1.01-1.43); this association was less susceptible to bias, inferred on the basis of heterogeneity between studies, evidence of small study effects, and excess significance, than for most other types of cancer (Tsilidis, Kasimis, Lopez, Ntzani, & Ioannidis, 2015).

A variety of risk assessment tools are available to calculate a woman's breast cancer risk, including the Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm (BOADICEA), and the Tyrer-Cuzick models (Gail et al., 1989; Claus, Risch, & Thompson, 1994; Berry et al., 2002; Antoniou et al., 2008; Tyrer, Duffy, & Cuzick, 2004). Each of these models is based on different data sets and takes into account various risk factors. Risk factors commonly used to place women in a high-risk category include personal history of cancer or breast abnormalities, previous radiation therapy to the chest, family history, and certain genetic mutations. A personal history of breast (especially at an age under 50) or ovarian cancer significantly increases the risk for diagnosis of contralateral breast cancer (Kleibl & Kristensen, 2016; Garber & Offitt, 2005). Previous diagnoses with benign disease (e.g., breast cysts) also increase the risk for later cancer diagnosis. An increased risk has been observed for women treated as children or young adults with radiation therapy for another form of cancer (life-time risk >20%), and in women who received radiotherapy for treatment of contralateral breast cancer, as compared to surgical treatment solely (Cancer Research UK, 2016; American Cancer Society, 2016).

Hereditary breast cancer may account for 5-10% of all breast cancers (Kleibl & Kristensen, 2016). Studies on family history have found relative risks between 1.5 (second-degree relative) and 3.6 (mother and sister) for developing breast cancer (Kleibl & Kristensen, 2016). For women aged 40-49 years, a recent systematic review and meta-analysis found that these women have an almost two-fold increase in risk if they have a family history of breast cancer (first degree relative) (Nelson et al., 2012). Mutations in high-penetrance genes such as *BRCA1* and *BRCA2* account for 50-70% of hereditary breast cancer, but mutations in other genes (e.g., *TP53, ATM, NBS1, LKB1*) can also increase risk (Mavaddat, Antoniou, Easton, & Garcia-Closas, 2010). Two meta-analyses estimated the cumulative risk of breast cancer by age 70 to be 55% to 65% for women with *BRCA1* mutations and 45% to 47% for women with *BRCA2* mutations (Antoniou et al., 2003). Some carriers elect to undergo prophylactic mastectomy because of this increased risk. For women who do not choose this surgical option, there is debate regarding the optimal surveillience strategy for this high-risk group.

An emerging risk factor not yet used to place women in the high-risk group is the degree of breast density. Breast density, a measure of the extent of radiodense fibroglandular tissue in the breast, may impact screening accuracy and is becoming a well-recognized, and significant risk factor for breast cancer. (Ciatto, Visioli, & Zappa, 2004). Dense breast tissue, reduces the sensitivity of mammographic imaging because of the presence of less fatty tissue (viewed as black) and a greater amount of the fibroglandular tissue which appears similar (white) to breast masses or tumors. Several classification systems can be used to distinguish the amount of density. The four categories of the Breast Imaging Reporting and Data Systems (BI-RADS)- almost entirely fatty, scatted fibroglanduar tissue, heterogeneously dense, and extremely dense- are recommended for reporting by the Canadian Association of Radiologists (Canadian Association of Radiologists, 2012). Besides this classification, there are the 4-grade Wolfe and 5-grade Tabar classifications, and quantitative measures including percentage breast density (percentage of the mammogram with radiodense fibroglandular tissue), area of dense tissue, fractal dimension and skewness (McCormack & dos Santos Silva, 2006). The American College of Radiology notes that there is considerable intra-and inter-observer variation in visually estimating breast density between any two adjacent density categories in the BI-RADS (American College of Radiology, 2013). Women with dense breasts have 3 to 5 times greater lifetime risk of developing breast cancer than women with mostly fatty breasts, regardless of other risk factors (McCormack & dos Santos Silva, 2006; Boyd et al., 2007). The mechanism for this is not well known, but because density reflects the proportion of epithelial and stromal tissue in the breast and breast cancer originate in the epithelial cells, the amount of density may reflect a greater number of cells that are at risk of carcinogenesis and/or an increased rate of epithelial proliferation (Boyd et al., 2005). Between 31% and 43% of women participating in population based mammographic screening have been found to have dense (50-74% breast density) or very dense (≥75%) breasts (McCormack & dos Santos Silva, 2006).

Apart from high-risk groups, some subgroups may merit particular attention as they receive differential benefits and/or harms of screening programs. For example, different baseline incidence rates (e.g., in various age groups) may change the absolute effects for outcomes. There are also certain population groups to consider regarding the need for screening programs to be accessible and equitable throughout a jurisdiction.

Screening Methods

Mammography

Mammography is generally considered to be the primary method for breast cancer screening (Nothacker et al., 2009). During a mammogram, x-rays are used to detect early stages of breast cancer, often before a lump can be physically felt (Gøtzsche & Jorgensen, 2013). Systematic reviews have reported approximately 15-20% reductions in breast-cancer mortality over more than 11 years follow-up from randomized controlled trials (RCTs) of women aged 39-70 years invited to screen with mammography, but they have not found evidence of reduced all-cause mortality (Canadian Task Force on Preventive Health Care, 2011; Gøtzsche & Jorgensen, 2013; Myers et al., 2015, Nelson et al., 2016a; Nelson et al., 2016b; Oeffinger et al., 2015; Siu, 2016). There have been various advancements in mammgoraphy techniques from conventional film, to digital, to the more recent tomosynthesis-3D. Digital mammography (full-field) was developed to overcome some limitations of film mammography (images on sheets of film), such that contrast can be manipulated on a computer to view density in the breast (Pisano et al., 2005). However, limitations of digital mammography include breast structures being superimposed on eachother, leading to a decreased ability to identify abnormal breast tissue

(Friedewald et al., 2014). Tomosynthesis has been introduced as a promising new advancement for breast cancer screening with favourable results; however its implementation in mainstream screening has yet to be endorsed due to its novelty. Tomosynthesis involves the use of computed tomography to create 3-D images without the overlapping of breast structures (Friedewald et al., 2014).

Clinical Breast Examination and Breast Self-Examination

Previous guidelines that have considered clinical breast examination by a health care professional (without concomitant mammography) have concluded that there is insufficient evidence for recommending these procedures when mammography is available (Fitzpatrick-Lewis et al., 2011; Hamashima et al., 2016; Nelson et al., 2009); higher false-positive rates have been observed when comparing clinical breast examination with mammography or adding clinical breast examination to mammography (Elmore et al., 2005; Fitzpatrick-Lewis et al., 2011). In a combined analysis of data from eight controlled trials, the sensitivity and specificity of clinical breast examination were 54% (95%CI, 48%-60%) and 94% (95%CI, 90%-97%), respectively (Barton, Harris, & Fletcher, 1999). Evidence on breast self-examination conducted by women themselves was reviewed in the previous guideline of the CTFPHC. Two trials reported that breast self-examination did not lead to significant differences between breast self-examination and control groups in breast-cancer mortality or all-cause mortality, and increased the rate of benign biopsy results. 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). The sensitivity of breast self-examination has been estimated to be between 20% and 30% (Elmore, Armstrong, Lehman, & Fletcher, 2005).

Magnetic Resonance Imaging (MRI)

MRI uses magnetic fields to create displays of cross-sections of the breast to determine the presence of breast cancer (Elmore, Armstrong, Lehman, & Fletcher, 2005). Most existing systematic reviews have assessed the diagnostic accuracy of MRI as a supplementary test to mammography and in high-risk women only. One systematic review indicated that the sensitivity range of MRI and mammography in high-risk women was 93-100%, much greater than assessing by mammography alone (Lord et al., 2007). Moreover, breast MRI has been consistently found to be more sensitive than mammography or ultrasound in detecting hereditary breast cancer (sensitivity ranges between 89% and 100%), but MRI has reduced specificity (range 35%–64%) (Kriege et al., 2004; Kuhl et al., 2007; Kuhl et al., 2010; MARIBS Study Group, 2005; Sardanelli et al., 2011; Warner et al., 2004; Weinstein et al., 2009). Despite the reported increased sensitivity of MRI compared with mammography, it has been found that some breast cancers identified with mammography are missed with MRI, such as DCIS (Moss-Basha et al., 2010; Lord et al., 2007).

Ultrasound

The false-positive rate has been reported to be higher when using ultrasound alone (2.4% to 12.9%) compared to mammography (0.7%-6%) (Elmore, Armstrong, Lehman, & Fletcher, 2005). Indeed, recommendations to date have generally suggested that ultrasound be used as supplementary to mammography and not as a stand-alone test (Nothacker et al., 2009; Gartlehner et al., 2013). In high-risk women, ultrasound used as part of a sequential test with mammography compared to mammography alone has been found to have higher sensitivity for detection of breast cancer; however, none of the included studies provided sufficient data on specificity so no conclusions could be made on false positive recall rates (Lord et al., 2007).

Harms of Screening

Although screening (via any modality) can benefit women by offering reassurance (if negative result) or by leading to earlier and better treatment options (e.g., breast-conserving surgery [lumpectomy vs. mastectomy], use of less intensive chemotherapy with fewer serious side effects, possibly the option to forego chemotherapy) that can prevent or delay death from breast cancer, screen-positive results may be inaccurate (false-positives) or lead to overdiagnosis (see definition in Table 1), leading to psychological or physical harms including adverse effects from further diagnostic testing or cancer treatments (i.e., overtreatment).

False Positives

A range of rates of false positive screens, which is the number who screen positive, but are not subsequently diagnosed with breast cancer, has been reported with one-time mammography rates of about 10% and cumulative (about 10 years) rates of about 50% (Nelson et al., 2009). Some investigators have found that false-positive screening tests may cause considerable and sustained psychological distress (Gøtzsche and Nielsen, 2009; Salz 2010). This distress may continue even after it is known that there is no cancer present (Brodersen, Thorsen, & Kriener, 2007), and may persist for years after women are declared free from cancer (Brodersen & Siersma, 2013). On the contrary, other investigators have found that the anxiety from a false-positive screen, is relatively mild, short-lived, acceptable to women, and less significant clinically than the anxiety caused by late-stage breast cancer (Feig, 2004). Negative consequences related to biopsies or other invasive procedures, such as surgery, may also result from false positives (Canadian Task Force on Preventive Health Care, 2011).

Overdiagnosis

Overdiagnosis in breast cancer screening refers to screen-detected lesions, histologically diagnosed as breast cancer, that would not have become clinically apparent during a woman's lifetime and would not have caused any health implications (Miller et al., 2014); this term includes cases when the death of a women due to other causes occurs before breast cancer symptoms would have developed (Puliti et al., 2012). Overdiagnosis is usually calculated using data on how frequently screen-detected DCIS becomes invasive in the absence of screening and over what period of time (Morris et al., 2015). The main concern is that women are harmed because of treatment morbidity that occurs without any benefit from any life years gained; because of this the term "overtreatment" is also used, although most studies do not report analyses specific to this phenomenon. Reported estimates of overdiagnosis can vary widely depending on methodology used and assumptions taken (Etzioni & Gulati, 2016; Morris et al., 2015); for example, estimates from RCTs have ranged between 4-7% and 22%, and from epidemiological and observational studies have ranged between 10% or less and as much as 50% (Duffy et al., 2010; Marmot et al., 2013; Miller et al., 2014). A meta-analysis of incidence trends estimated the total overdiagnosis in mammography screening programs at 52% (95% CI, 46%-58%) or 1 in 3 breast cancer (Jorgensen & Gøtzsche, 2009). However, an independent UK panel reviewed RCTs and observational studies and estimated the likelihood of overdiagnosis to be 19% (Marmot et al., 2013).

Rationale for Screening and Currency of Guidelines

The enormous burden of breast cancer warrants efforts to detect invasive breast cancer early, before a patient is symptomatic and at a time when options for treatment and chances of survival from breast cancer are most favorable. Indeed, organized screening programs (definitions vary, however, typically delivered in a more standardized approach with services evaluated) have led to a shift from late-stage to early-stage disease detection (Helvie et al., 2014; Williams, Carter, & Rychetnik, 2014). Nevertheless,

ongoing research struggles to decipher the optimal screening modality, screening intervals, and age to start screening. Moreover, increasing awareness of overdiagnosis and overtreatment and debates on the benefit-harm ratio of screening over more than a decade have added value-laden complexities into recommendations made by guideline panels and, possibly, into decision making by women.

Previous CTFPHC Recommendations and Recommendations from Other Guideline Developers

In 2011, the CTFPHC published a guideline (Canadian Task Force on Preventive Health Care, 2011) based on results from a systematic review (Fitzpatrick-Lewis et al., 2011) on breast cancer screening for women. The following recommendations were provided:

- For women aged 40-49 years, it was recommended to not routinely screen with mammography (Weak recommendation; moderate-quality evidence)
- For women aged 50-69 years, it was recommended to routinely screen with mammography every two to three years" (Weak recommendation; moderate- quality evidence)
- For women aged 70-74 years, it was recommended to routinely screen with mammography every two to three years" (Weak recommendation; low-quality evidence)
- Recommended not routinely screening with MRI scans" (Weak recommendation; no evidence)
- Recommended not routinely performing clinical breast examinations alone or in conjunction with mammography to screen for breast cancer (Weak recommendation; low-quality evidence)
- Recommended not advising women to routinely practice breast self-examination (Weak recommendation; moderate-quality evidence)

There are differences in recommendations between the 2011 CTFPHC guidelines (see **Appendix A-I**) and others published since. Specifically, in regard to mammography, the CTFPHC made a weak recommendation against screening for women aged 40-49 years, suggesting that not undertaking screening would be the best option for most, but not all women; the USPSTF recommended that women under 50 years old use their personal judgments and beliefs on whether to undergo biennial screening; the ACS recommended that women 45 years or older should undergo regular screening and that women aged 45-54 years should do so annually, and the Japanese National Cancer Center recommended that women aged 40 and older should undergo screening. These differences relate in part to 1) differences in the evidence quality, availability, or evidence used by each guideline panel (e.g., differences in sources of breast cancer incidence and harm data as well as data stratifying incidence and outcome data between different age groups), 2) judgments made by each panel in terms of the relative importance of different benefits (e.g., all-cause vs. disease-specific mortality) and harms (e.g., overall recall from false-positive and biopsies from false-positive) and 3) the manner in which the benefit and harm data were weighed and how and whether contextual factors were considered. Because of these unique factors inherent to guideline development, including evidence and contextual considerations within the Canadian screening context, as well as differences in the scope (e.g., desire to examine clinical and breast self-examination) compared with other guidelines, the CTFPHC has decided to update their recommendations rather than consider adopting or adapting recent recommendations of other panels.

Objectives

The objectives are to 1) synthesize up-to-date evidence on the benefits and harms of breast cancer screening by conducting an overview of selected systematic reviews and an updated search for more recent primary studies and to 2) systematically review how women weigh the benefits and harms of screening and how this valuation effects their decision to undergo screening. The findings will be used by the CTFPHC, along with additional considerations of feasibility, acceptability, affordability, and equity, to change or reaffirm their previous recommendations.

Chapter 2. Methods Overview

Updated Analytic Framework

The updated analytic framework (Figure 1) and key questions for this updated report will closely follow the CTFPHC 2011 framework and questions with some modifications discussed below.

The population of interest remains women aged 40 years or older who are not at high-risk for breast cancer: (a) without pre-existing or personal history of breast cancer and (b) not considered to be at risk of breast cancer on the basis of extensive family history of breast or ovarian cancer, abnormal breast pathology, deleterious genetic mutations, or previously received radiation treatment to the chest (such as Hodgkin's) for cancer. Studies of high-risk populations will be excluded because of lack of direct applicability to population screening in Canada. Breast density was considered by the CTFPHC since it increases one's risk for breast cancer and is common (depending on how defined); because this factor is not common to most lists of high-risk groups at this time and is discovered during (rather than known prior to) screening, we will not exclude studies including women found to have dense breasts. However, studies exclusively enrolling women with dense breasts (i.e., ≥75% study population) will not be included because the current guideline targets general population screening. More information on the approach to breast density can be found in Table 1.

The CTFPHC has included mammography (film, digitial, and tomosynthesis), ultrasound, as well as magnetic resonance imaging in combination with clinical breast exam or self-breast exam, as relevant screening modalities for their guideline update.This evidence report will only assess outcomes considered by the CTFPHC as critical for decision making (i.e., rated as 7 or higher out of 9). Since its 2011 guideline, the CTFPHC has re-rated harm outcomes to better reflect their importance in patient decision making.

Women's values and preferences, previously a contextual question in the CTFPHC 2011 review, will now be considered a key question and evaluated with systematic review methodology. The CTFPHC has narrowed the scope to focus on women's valuation of benefits and harms related to screening.

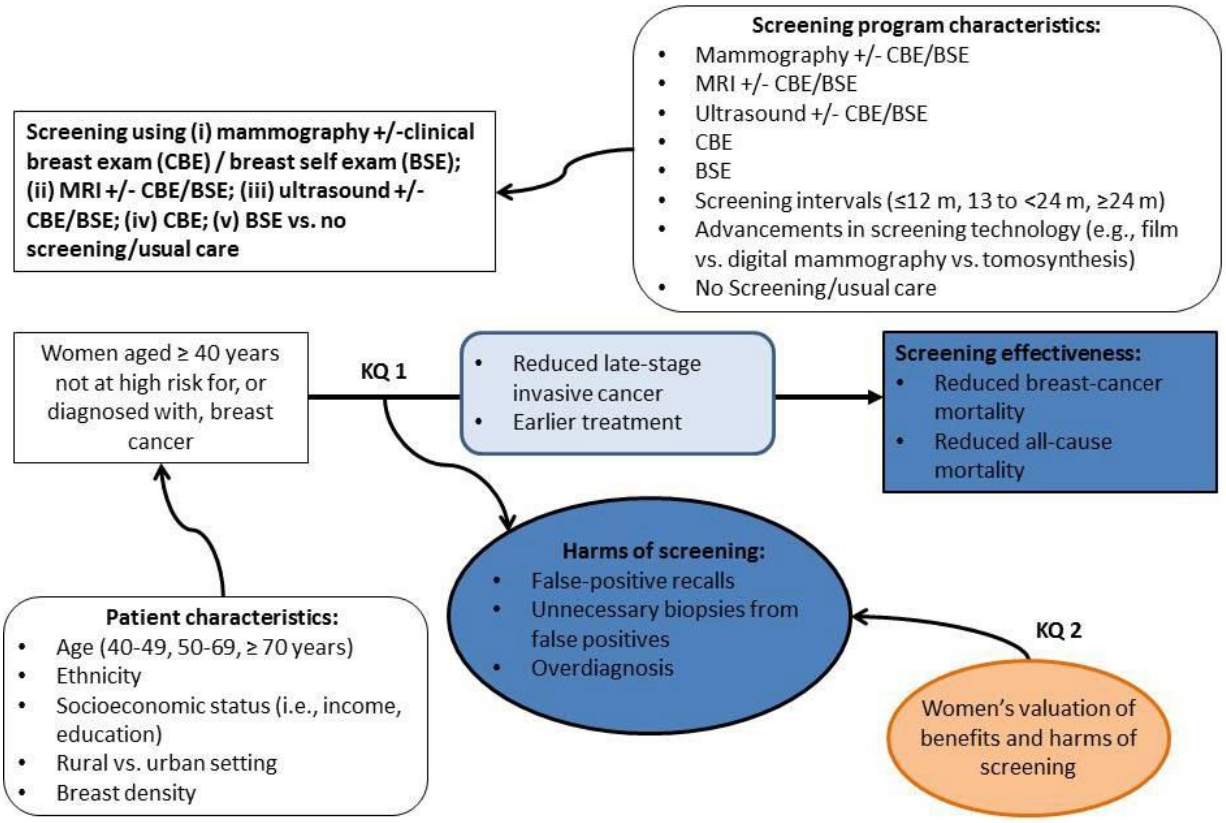

Figure 1: Updated analytic framework for breast cancer screening

Key Questions

KQ1 (Chapter 3): For women ≥40 years of age and not at high-risk, what are the benefits and harms of the following screening modalities: (i) mammography (film, digital, tomosynthesis) with or without clinical breast examination or breast self-examination; (ii) MRI with or without clinical breast examination or breast self-examination; (iii) ultrasound with or without clinical breast examination or breast self-examination; (iv) clinical breast examination; and (v) breast self-examination compared to no screening/usual care?

KQ2 (Chapter 4): How do women weigh the benefits and harms of breast cancer screening, and how do they use this information in their decisions to undergo or not undergo screening?

General Methodology

The methodological approach to each of the two KQs and their respective outcomes has been tailored to ensure the evidence needs of the task force are met. It will identify new evidence on critical benefits and harms of breast cancer screening by examination of pre-identified existing systematic reviews on the topic. It will also identify all evidence on women's values and preferences by a systematic review of

evidence. Results shared in the final summary of evidence report will help the CTPFHC either reaffirm or change its 2011 recommendations on breast cancer screening.

The PRISMA-P guideline was used to develop this protocol (Shamseer et al., 2015) (**Appendix O**). The protocol will be registered in PROSPERO and posted in the Open Science Framework.

For KQ1, our approach largely consists of an overview of selected systematic reviews. We will select the highest quality, up to date, and fit-for-purpose (the best reporting of risk of bias to use for GRADE assessments) systematic review that synthesizes information for a given intervention-outcome pair to report as the evidence base (multiple reviews may be used depending on the available evidence for each intervention-outcome pair). Detailed information in regards to the methodology for KQ1 can be found in Chapter 3. We will update this evidence base by searching for more recent studies made available since the date of the review's last searches. For KQ2, a standard systematic review methodology will be used.

Literature Search

Where applicable to the KQ, draft bibliographic database search strategies (see **Appendix B** and **C**) were developed by information specialists (BS and RF). These searches will be peer-reviewed using the PRESS 2015 checklist (McGowan et al., 2016); results of the PRESS reviews will be provided in an Appendix in the final report. Grey literature will be searched using the Canadian Agency for Drugs and Technologies in Health (CADTH) Grey Matters Checklist, Canadian provincial and territorial websites, by contacting authors and by searching ClinicalTrials.gov and the World Health Organization's International Clinical Trials Registry Platform for ongoing or completed studies. For each KQ, the totality of these supplemental searches will be confined to what can be accomplished within 40 hours of work by one team member. We will document the search details, and potentially relevant studies identified from the supplemental searches will be integrated into the PRISMA flow diagram.

Study Selection

Results from the search strategies will be uploaded into reference management software and duplicates removed. Unique citations will be uploaded into online SR management software, DistillerSR (Evidence Partners, Ottawa, Canada) for the selection phase. During title and abstract screening, the liberal accelerated method will be used: reviewers will review citations presented in random order, and only one reviewer will be needed to deem the record as potentially relevant and pass it through to full-text screening, whereas two reviewers will be required to exclude the record during title and abstract review. For full-text screening, two reviewers will independently assess the articles against all selection criteria. Conflicts will be resolved by consensus or by consulting a third team member.

The reasons for exclusion at full-text screening will be documented in the Distiller file and in a separate document listing excluded studies. Reports that are co-publications or multiple reports of the same study will be included and tabulated, although our evidence report will only cite each unique study using its primary source.

A pilot testing phase among reviewers will be implemented on a sample of records/articles before the commencement of title/abstract (n=50) and full-text (n=20-30 articles, depending on the yield from the title and abstract assessment) screening.

Articles not available electronically will be ordered via interlibrary loan; as a practical consideration, if the article is not received within 30 days it will be excluded and noted as the reason for exclusion.

Studies reported only in abstract form will not be included, although we will try to locate the full report by searching and/or contacting authors. A list of potentially relevant studies available only in abstract form or listed as a completed or on-going study in a study registry (without access to a full report) will be provided in an appendix.

Data Extraction and Management

Details of the data extraction phase for each KQ are provided in subsequent chapters. For both KQs, standardized data extraction forms will be developed a priori and pilot tested on a sample of reviews and studies. Extraction will be completed by one person and verified by a second person; disagreements will be addressed by consensus or third party adjudication. For primary studies identified through searching, authors will be contacted by email twice over two weeks in relation to missing, incomplete, or unclear information.

Risk of Bias Assessment

Included reports (whether reviews or studies) will be assessed for quality (reviews) or for risk of bias (studies). These assessments will be conducted by one reviewer and verified by a second reviewer. Disagreements will be resolved by consensus or third-party adjudication. Additional details are provided in each Chapter below.

Planned Schedule and Timeline

-Draft GRADE tables: May 2017

-Draft Report: June 2017

-Final Report: August 2017

Chapter 3. Effectiveness of Breast Cancer Screening on Clinical and Patient Important Outcomes

Research Questions

KQ1: For women ≥40 years of age and not at high-risk, what are the benefits and harms of the following screening modalities: (i) mammography (film and digital) with or without clinical breast examination or breast self-examination; (ii) magnetic resonance imaging (MRI) with or without clinical breast examination or breast self-examination; (iii) ultrasound with or without clinical breast examination or breast self-examination; (iv) clinical breast examination; and (v) breast self-examination compared to no screening/usual care?

Inclusion and Exclusion Criteria

Studies will be selected according to the inclusion and exclusion criteria outlined in Table 1.

Table 1 – Inclusion and exclusion criteria for Key Question 1

	Inclusion	Exclusion
Population	Women aged ≥40 years of age and not at high-risk for breast cancer ‡. Subgroups: age (40-49 years, 50-69 years, 70 years and older); ethnicity, including whether women are from indigenous populations; socioeconomic status; geographical location (rural vs. urban settings); breast density. ‡If the study population comprises <75% of women with high breast density or the population comprises <20% of women who are high risk, these studies will be included. The definition of 'dense breasts' will be that as defined in each included systematic review or primary study .	Cohorts comprised of 75-100% women with high breast density; men with breast cancer, women with pre-existing or personal history of breast cancer; women considered to be at high-risk for breast cancer on the basis of family history (in a first degree relative) of breast or ovarian cancer or other personal risk factors, such as abnormal breast pathology or BRCA1/BRCA2 genetic mutations, previously received radiation treatment to the chest (such as Hodgkin's) for cancer.
Intervention and comparator	(i) Mammography (film, digital, or tomosynthesis-3D mammography*) with or without clinical breast examination/breast self-examination vs. no screening/usual care (ii) MRI with or without clinical breast examination/breast self-examination vs. no screening/usual care (iii) Ultrasound* with or without clinical breast examination/breast self-examination vs. no screening/usual care (iv) clinical breast examination vs. no screening/usual care (v) breast self-examination vs. no screening/usual care	Combination modalities other than that already indicated.

	Subgroups: screening interval (≤12 months, 13 - <24 months, ≥24 months); advancements in screening technology (comparison of film mammography, digital mammography or tomosynthesis); no screening vs. usual care * Added modality since the CTFPHC 2011 review. Also (i) added +/- CBE/BSE to the following screening modalities: MRI, ultrasound; (ii) explicitly stated 'no screening/usual care' as comparator.	
Outcomes	• Breast cancer related mortality • All-cause mortality • Overdiagnosis* • False-positive results and consequences (e.g., FP recalls, FP recalls requiring unnecessary biopsies) **Overdiagnosis**: screen-detected lesions that would not become clinically apparent during a woman's lifetime in the absence of screening and would not have caused any health implications. These are a subset of true positives where cancer is diagnosed, but would never have become clinically apparent otherwise. Different from false-positive as there is a diagnosis. **False-positive**: positive screening result, but cancer is not present. *Authors' calculations/or reporting of overdiagnosis will be used (various iterations of the same data may be possible due to differing methods of calculation from various authors). We will not undertake calculations for these outcomes. Note: Changes/modifications since CTFPHC 2011 review: -Included 'overdiagnosis' as critical harm outcome.	
Timing	No limit	
Settings	Primary care or other settings generalizable to primary care, including referrals by primary care providers	Any setting where it could not be reasonably generalizable to a Canadian screening context
Databases	Medline, Cochrane Library	
Study designs	**Mortality outcomes*:** Randomized controlled trials , including cluster randomized controlled trials **Harms outcomes:** Overdiagnosis *: RCTs, non-RCTs, cohort studies, ecological studies FP, FP consequences: As per the 2011 CTFPHC guideline, for these outcomes, if feasible, we will use the Canadian Breast Cancer Screening Database Registry data (Canadian Partnership Against Cancer, 2015). The 2011 approach will be repeated with the use of standard tables. Will also consider other Canadian studies captured in our search update and captured in the	Cross-sectional studies , case series, case reports, controlled before-after, case-control studies, diagnostic test accuracy studies, modelling studies

	existing reviews. *Systematic reviews reporting on outcomes using these study designs will be extracted. If the synthesis includes designs other than those of interest, a commentary will be provided.	
Language	English and French	

MRI: magnetic resonance imaging

Literature Search

To conduct the overview of reviews, we have a total of eighteen pre-identified potentially relevant reviews for consideration. These reviews were identified by consulting the CTFPHC. These are listed in **Appendix AI and Appendix AII** and include but are not limited to reviews conducted for guidelines produced by the U.S. Preventive Services Task Force (USPSTF) (Nelson et al., 2016a; Siu, 2016), the American Cancer Society (ACS) (Oeffinger et al., 2015), and the Japanese National Cancer Center (NCC) (Hamashima et al., 2016). The other fourteen potentially relevant systematic reviews were identified through the bibliography of the USPSTF and ACS review since they were the most recent and high quality systematic reviews being considered.

The draft search strategy to update evidence from the existing systematic reviews was developed by an experienced medical information specialist (RS) in consultation with the review team **(Appendix B)**. We will search Ovid MEDLINE® In Process & Other Non-Indexed Citations and the Cochrane Library on Wiley. Strategies will utilize a combination of controlled vocabulary and keywords. The vocabulary and syntax will be adjusted across databases.

Based on knowledge of more recently published systematic reviews we anticipate that we will need to conduct these searches from around October 2014 to the present for all modalities with the exception of self-breast examination, which may need to be searched from as early as October 2010. The time-point of these searches will be confirmed once the overviews work has been completed. For feasibility, we will not conduct diagnostic-style literature searches (i.e., without study design filters employed) to harness information relevant to outcomes such as false-positive results; this will be a noted limitation in the final report.

We plan to use the CADTH Grey Matters checklist to update the grey literature search for unpublished literature and search the following websites: provincial screening programs in Canada, BC Cancer Agency, Cancer Care Ontario, the Canadian Cancer Society, and Canadian Partnership Against Cancer. Authors from identified key harm publications from our search (published in last 1-2 years) will be asked if they are aware of any additional primary studies. Study registry searches will be performed in ClinicalTrials.gov and the World Health Organization's International Clinical Trials Registry Platform. All other supplemental search strategies are outlined in Chapter 2.

Selection Process

Systematic reviews of the evidence will be included if they: i) searched more than one database; ii) reported their selection criteria; iii) conducted quality or risk of bias assessment on included studies; iv) provided a list and synthesis of included studies. Pre-identified potentially relevant reviews from existing guidelines are listed in **Appendix A-I and Appendix A-II**.

We will select the highest quality, up to date, and fit-for-purpose (the best reporting of risk of bias to use for GRADE assessments) systematic review that synthesizes information for a given intervention-outcome pair to report as the evidence base (multiple reviews may be used depending on the available evidence for each intervention-outcome pair). A chosen CTFPHC external content/clinical expert will review the list of included studies form the existing reveiws to help identify any additional primary studies (if missed).

Studies identified in the updating search will be selected according to the eligibility criteria in Table 1. Remaining details of the study selection process can be found previously in Chapter 2, under the Study Selection section. Draft title and abstract and full-text screening forms are shown in **Appendix D**.

Data Extraction and Management

We will extract information at face value for how it was synthesized and/or reported in the relevant reviews. If evidence is synthesized in a manner that does not allow clean extraction according to study designs of interest, the information will be extracted and a commentary provided. Data extractions for any newly identified studies will occur with a separate data extraction form.

We will not conduct any quality control checks to verify the accuracy of the reviews' extractions. However, since we are considering multiple reviews that are likely to overlap in terms of content, this will be used to determine congruency of extraction of information across reviews.

Details for the data extraction and management process can be found in Chapter 2, under the Data Extraction and Management section. Draft items for data extraction for existing systematic reviews and newly identified primary studies can be found in **Appendix E** and **Appendix F.**

For the outcomes of false-positive results and consequences of false-positive results, we will review reports from the Canadian Breast Cancer Screening Database. For the purpose of observing any discrepancies, we will also extract data from other Canadian studies in our search update and similar organizations in the estimates reported from the existing reviews (e.g., USTPSTF).

For multiple events that may occur in one person (i.e., adverse events), we will assume each event represents a unique individual unless otherwise specified. If we were to encounter a study where there is a reason for concern that many events are recorded but likely in a small number of patients, these could be evaluated in sensitivity analyses.

Risk of Bias Assessment

Details for the risk of bias assessment process can be found in Chapter 2, under the Risk of Bias Assessment section. For systematic reviews, AMSTAR will be used to assess the methodological quality of all relevant reviews (**Appendix G**). Quality assessments of the primary studies from relevant systematic reviews will not be completed if not conducted by the review authors or re-done if different tools were used. For newly identified primary studies, the Cochrane Risk of Bias tool (Higgins & Green, 2008) (**Appendix H**) will be used to evaluate the risk of bias in randomized controlled trials and non-randomized controlled trials, and the Newcastle-Ottawa scale (Wells et al., 2007) will be used to

evaluate the risk of bias in cohort studies (**Appendix I**). Quality assessments of the individual studies will inform the GRADE domain of study limitations.

We will not conduct any quality control checks to verify the accuracy of the selected reviews' risk of bias or quality assessments. However, since we are considering multiple reviews that are likely to overlap in terms of content, this will be used to determine congruency of this information from across reviews. Should there be any discrepencies in the risk of bias assessments, we will further examine the methods used to assess risk of bias, and select the systematic review with the most rigorous method used for each intervention-outcome pair. We will narratively provide commentary on the reasons for discrepencies, but will not be revising the original risk of bias assessments.

Analysis Plan

For the overview, we will summarize the characteristics of the included reviews and provide the AMSTAR assessments. We will summarize the syntheses undertaken by authors, including the number and type of studies.

Among the relevant reviews, the decision as to which systematic review will be used as the evidence for a given modality and outcome will be informed by alignment to the eligibility criteria, the currency of the date of the literature search, understanding of the accuracy of reported outcomes data from cross-checking of reviews, and the quality of the conduct of the review via AMSTAR assessment.

For the newly identified primary studies, the characteristics of included studies will be summarized narratively and presented in summary tables.

Where possible, relative and absolute effects with 95% confidence intervals will be presented according to the GRADE summary of findings and evidence profile tables adopted by the CTFPHC. For example, risk ratios and risk differences will be ideally used to report effects for binary data. GRADE guidance will be utilized for presenting continuous data (Guyatt et al., 2013). Where possible, the number needed to treat/harm will be calculated. If not possible, a narrative description of findings will be presented in the GRADE table (e.g., ranges of values for studies not pooled).

Meta-analysis

Where possible, effect estimates (as reported by review authors) will be presented in a forest plot. If a meta-analysis was conducted in an existing review, we will determine whether it will be appropriate and feasible (assessment of heterogeneity according to clinical and methodological characteristics) to update the pooled estimate with data from any newly identified studies. Random effects models will be used in these cases. We will note the limitations of this approach, such as the risk of over- or under-estimating the findings in light of the differing methodologies that may have been employed to collate the body of evidence.

If a meta-analysis was deemed not appropriate by review authors, then we will attempt to present findings (including those of newer studies) visually in a forest plot without a pooled effect estimate; the range of effects will be described in the text.

A decision of de novo meta-analysis with the body of evidence will need to be made in consultation with the CTFPHC, acknowledging the limitations noted above and the limited information may be available from reviews to judge aspects of clinical and methodological heterogeneity in addition to the potential of differing methodologies employed in those reviews and that used here to locate new studies.

The numbers needed to treat/harm for the body of evidence will be calculated where appropriate and possible to do so. Also where appropriate and possible, the hazard ratio for time-to-event data will be pooled using the generic inverse variance method.

For observational studies, use of adjusted estimates of effect (and their corresponding listings of confounders which were modeled), will be an important criterion to evaluate for heterogeneity. Ideally, adjusted estimates will be used in the meta-analysis.

The existence of any unit of analysis errors will be taken into consideration when evaluating the syntheses undertaken in systematic reviews. Unit of analysis errors can occur in studies that employ a cluster design (i.e., a clinical practice) and yet are analyzed at the individual level (i.e., patients), potentially leading to overly precise results and contributing greater weight in a meta-analysis. If empirically-derived intracluster correlation coefficients (ICC) are available for newly identified primary studies, we will adjust the analysis of those data to address these errors (Killip, Mahfoud, & Pearce, 2004).

Sparse binary data and studies with zero events

The following guidance will be employed where feasible and possible to do. When event rates are less than 1%, the Peto odds ratio method will be used. However, when control groups are of unequal sizes, when large magnitude of effect is observed, or when events become more frequent (5%–10%), the Mantel-Haenszel method without correction factor will be used for quantitative synthesis (Fu et al., 2011). When studies report rare events or a meta-analysis is not appropriate a narrative synthesis will be performed.

Statistical heterogeneity

The Cochrane's Q (considered statistically significant at $p<0.10$) and I^2 statistic will be used to assess the statistical heterogeneity of effect estimates among included studies. For the interpretation of I^2, a rough guide of low (0-25%), moderate (25-50%), substantial (50%-75%), and considerable (≥75%) will be used (Higgins, Thompson, Deeks, & Altman, 2003; Sterne et al., 2011). Should considerable statistical heterogeneity exist, we will present all studies in a forest plot, but will not provide the pooled overall estimate.

Sub-group analyses

Subgroup analyses will be conducted if also conducted or presented narratively in existing systematic reviews. These will be updated, where possible, with information from newer studies. Relevant subgroup analyses are listed below:

- Age (40-49 years, 50-69 years, 70 years and older)
- Various ethnic populations, including indigenous populations (will be determined post hoc, depending on populations encountered in studies and/or how conducted in existing reviews)
- Socioeconomic status (e.g., income, level of education; as assessed by study authors)
- Geographical location (rural vs. urban settings)
- Use of clinical breast examination/breast self-examination with mammography, MRI, or ultrasound

- Screening interval (≤12 months, 13 - <24 months, ≥24 months)
- Advancements in screening technology (film vs. digital mammography)

Sensitivity analyses
Sensitivity analyses may be undertaken to restrict analyses to those studies assessed as being of low risk of bias, based on the overall judgment, and may also be performed to address any decisions made regarding the handling of data or to explore heterogeneity.

Small study effects
To assess for small study effects, a combination of graphical aids (i.e., funnel plot) and/or statistical tests (e.g., Egger regression test, Hedges-Olkin) will be performed subject to at least 10 studies in any given analysis.

Meta-analysis Software
If a meta-analysis is undertaken, the Cochrane Review Manager (RevMan) software version 5.3 will be used to calculate effect estimates and conducting meta-analyses. For all analyses not possible in RevMan, we will use Comprehensive Meta-Analysis (CMA) or Stata.

GRADE
For each critical outcome, the Grading of Recommendations Assessment, Development and Evaluation (GRADE) framework (Oxman & GRADE Working Group, 2004; Canadian Task Force on Preventative Health Care, 2014) will be used to assess the quality of the evidence.

Assessment of each of the GRADE domains (study limitations, consistency, precision, directness, reporting bias) will be presented, where possible, based on the information provided in the reviews and incorporated primary studies. A best fit to assess each domain will be made where information provided in reviews is different from guidance recommended by GRADE: for example, where study authors may use risk of bias assessment criteria different than that provided by GRADE. If data from the systematic reviews and newly identified studies are not combined, a narrative description across findings will be presented and assessment on the collective information will be made in the GRADE table.

Chapter 4. Women's Valuation of the Benefits and Harms of Breast Cancer Screening

Research Question

KQ2: How do women (a) weigh the benefits and harms of breast cancer screening, and (b) use this valuation in their decisions to undergo screening?

Inclusion and Exclusion Criteria

Studies will be selected according to the inclusion and exclusion criteria outlined below. The key criterion is that the study must incorporate an assessment (indirectly or directly) by the participants of the relative magnitude/incidence of possible benefits and harms of screening for breast cancer. Weighing of outcomes may be explicit within decision analysis or other studies generating ratings (e.g., preference weights) of various outcomes for consideration during development of recommendations (see Table 2), or implicit such as when women are asked whether they would choose screening based on information on benefits and harms presented within the study, or based on their past experience of screening (e.g., past false positive and after getting accurate information on benefits they are asked choose decide if they will screen again). The ideal study would ask women to decide on an acceptable threshold in terms harms per benefit, such as number of false positive recalls per breast-cancer life-year gained. We will exclude studies comparing uptake/acceptance of screening in groups of women receiving or not receiving decision aids or other decision-based interventions; while the evidence may have some relevance to women's outcome valuation, it is considered to also capture additional aspects (e.g., clinician-patient communication, implementation issues of different forms of information) which would confound the findings for this research question and be difficult to interpret. We will also exclude studies evaluating understanding or preferences for how different risk communication strategies/formats may improve their ability to make informed consent.

If there is a very limited evidence base for this KQ (i.e., in terms of quantity/sample size, methodological quality, inconsistency between studies, or applicability to our population or setting) we will consider including studies examining women's valuation of harms *or* benefits rather than the trade-off between the two; for example, studies examining women's acceptance of screening after experiencing a false positive result could offer some indirect evidence to help the CTFPHC in their deliberations. Likewise, relative ratings/preference weights for different harm outcomes (total recall rate vs. recall leading to biopsy) could be informative.

Table 3– Inclusion and exclusion criteria for Key Question 2

	Inclusion	Exclusion
Population	Women aged ≥40 years of age and not at high-risk for breast cancer Subgroups: age (40-49 years, 50-69 years, 70 years and older), ethnicity, including whether the women are from an indigenous population (will be determined post hoc depending on populations encountered in studies), socioeconomic status, geographical location (rural vs. urban settings)	Women with pre-existing or personal history of breast cancer; women considered to be at high-risk for breast cancer on the basis of extensive family history of breast or ovarian cancer or other personal risk factors, such as abnormal breast pathology or deleterious genetic mutations, having previous radiation treatment to the chest (such as Hodgkin's) for cancer.

		For studies with high-risk and not at high-risk as study groups, we will only use data from the not at high-risk group.
Intervention/Context	Screening for breast cancer using mammography, MRI, ultrasound, clinical breast examination, or breast self-examination. Mammography(film, digital, tomosynthesis-3D mammography), US, or MRI (+/- clinical breast examination) Women must be provided with some information (may not include estimates of effects) on the magnitude and/or types of potential benefits and harms of screening; one alternative is when women who have experienced harms (false positives) are provided with information on benefits to make decisions for future screening. Information can be provided in written form or orally.	
Comparators	Depending on study design, comparator may be no screening or another form of screening, or the study may not have a comparison. When only one arm (e.g. screening) of a comparative study is included in the assessment of patient preferences, this study will be classified as a non-comparative study. Comparator may be based on participant characteristics, such as age or socioeconomic status.	
Outcomes	• Willingness to be screened • Acceptability of screening • Uptake of screening • Willingness to pay for screening • Relative ranking/rating of benefit and harm outcomes (e.g. preference or utility weights) • Factors related to benefit and harm outcome valuation that contribute to choices for screening (e.g. severity of harm, age of women, availability of treatment, perceived risk for breast cancer) • Other outcomes will be considered (e.g. intent to return for another screen).	
Timing	2000-present	
Settings	Primary care or other settings generalizable to primary care, including referrals by primary care providers	Any setting where it could not be reasonably generalizable to a Canadian screening context
Databases	Medline, Cochrane CENTRAL , CINAHL, PsycINFO, PubMed	
Study designs	All experimental study designs, examples including: • Utility-based stated and revealed preference	Commentaries, opinion, editorials, and reviews

	studies (e.g. contingent analysis or valuation studies including discrete choice experiments, willingness to pay) • Studies used to develop health-state utility weights • Surveys • Qualitative studies (These studies may be embedded within RCTs or other controlled study designs)	
Language	English and French	

Literature Search

For the womens' valuation of benefits and harms, we will conduct a de novo systematic review as we are not aware of another group conducting this. The search for literature will start at 2000 since this timeframe captures the period where the magnitude of benefits and harms (e.g., overdiagnosis) of breast cancer screening have come under much scrutiny (Olsen & Gøtzsche, 2001; Gøtzsche & Olsen, 2000). The search strategy will consist of both controlled vocabulary, such as the National Library of Medicine's MeSH (Medical Subject Headings), and keywords, and will be peer-reviewed. Methodological filters will not be applied to limit retrieval by study design. Searches will be restricted by language to include full texts published in English and French, with a publication date restriction of 2000. We will search MEDLINE (1946-) via Ovid; Cochrane Library; CINAHL (1937-present) and PsycINFO via EBSCOhost; and PubMed via NCBI Entrez. The detailed search strategy for MEDLINE is reported in **Appendix C** and will be adapted to accommodate the controlled vocabularies of each database.

Reference lists of all included studies and relevant systematic reviews (see Chapter 2 for selection criteria for systematic reviews) will be scanned for further studies. Authors of all included studies will be contacted and invited to submit papers or reports of other studies of which they are aware. We plan to use the CADTH Grey Matters checklist to update the grey literature search for unpublished literature and search the following websites: provincial cancer screening programs, BC Cancer Agency, Cancer Care Ontario, Canadian Partnership Against Cancers, Canadian Cancer Society, World Conference on Breast Cancer, other relevant stakeholder organization websites. Our clinical experts will be asked to identify 2-3 key conference proceedings to search (2014-2016 only to capture recent but not yet published studies).

Study Selection

Details of the study selection process can be found in Chapter 2, under the Study Selection section. Draft screening forms are shown in **Appendix K.**

Data Extraction and Management

Details for the data extraction and management process can be found in Chapter 2, under the Data Extraction and Management section. Key points for extraction apart from participant characteristics are details of the survey (including attributes included in valuation studies or scenarios used) or interview guide, background information on screening and outcomes provided to participants, methods of analysis including models or theories used when applicable, and any subgroup analyses conducted. Draft items for data extraction can be found in **Appendix L**.

Risk of Bias Assessment

Details of the risk of bias assessment process can be found in Chapter 2, under the Risk of Bias Assessment section. The majority of studies are anticipated to be cross-sectional in nature, even if embedded within a trial or other controlled study. Critical appraisal tools from the Centre for Evidence-Based Management (http://www.cebma.org/resources-and-tools/what-is-critical-appraisal/) will be used for surveys and qualitative studies (**Appendix M**). Some studies may have data from two or more study groups. We will use the Cochrane Risk of Bias tool (Higgins & Green, 2008) (**Appendix H**) to evaluate the risk of bias in randomized controlled trials and non-randomized controlled clinical trials (if data for both groups is relevant to the KQ). We will use the EPOC tool for controlled before-after studies (**Appendix N**), the Newcastle-Ottawa scale (Wells et al., 2007) will be used to evaluate the risk of bias in cohort (**Appendix I**) and case-control (**Appendix J**) studies.

Analysis Plan

Study characteristics and findings will be summarized narratively and presented in summary tables. We will report all relevant quantitative findings, such as percentages (e.g., willingness to screen) and means and standard deviations, and will discuss differences between studies (e.g., design, populations, methods) that may contribute to differing results. We will report qualitative themes in narrative format, if applicable.

Reference List

American Cancer Society (2013). Breast Cancer Facts & Figures 2013-2014. http://www.cancer.org/acs/groups/content/@research/documents/document/acspc-042725.pdf

American Cancer Society (2016). What are the risk factors for breast cancer? http://www.cancer.org/cancer/breastcancer/detailedguide/breast-cancer-risk-factors

American College of Radiology (2013). ACR BI-RADS Atlas- Mammography. http://www.acr.org/~/media/ACR/Documents/PDF/QualitySafety/Resources/BIRADS/01%20Mammography/02%20%20BIRADS%20Mammography%20Reporting.pdf

Antoniou, A. C., Cunningham, A. P., Peto, J., Evans, D. G., Lalloo, F., Narod, S. A. et al. (2008). The BOADICEA model of genetic susceptibility to breast and ovarian cancers: updates and extensions. *British journal of cancer, 98,* 1457-1466.

Antoniou, A., Pharoah, P. D. P., Narod, S., Risch, H. A., Eyfjord, J. E., Hopper, J. L. et al. (2003). Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case series unselected for family history: a combined analysis of 22 studies. *The American Journal of Human Genetics, 72,* 1117-1130.

Armstrong, K., Moye, E., Williams, S. et al. (2007). Screening mammography in women 40 to 49 years of age: a systematic review for the American College of Physicians. *Ann Intern Med, 147*, 516-26.

Barton, M. B., Harris, R., & Fletcher, S. W. (1999). Does This Patient Have Breast Cancer?: The Screening Clinical Breast Examination: Should It Be Done? How? *JAMA, 282,* 1270-1280.

Berry, D. A., Iversen, E. S., Gudbjartsson, D. F., Hiller, E. H., Garber, J. E., Peshkin, B. N. et al. (2002). BRCAPRO validation, sensitivity of genetic testing of BRCA1/BRCA2, and prevalence of other breast cancer susceptibility genes. *Journal of Clinical Oncology, 20,* 2701-2712.

Boyd, N. F., Guo, H., Martin, L. J., Sun, L., Stone, J., Fishell, E. et al. (2007). Mammographic density and the risk and detection of breast cancer. *New England journal of medicine, 356,* 227-236.

Boyd, N. F., Rommens, J. M., Vogt, K., Lee, V., Hooper, J. L., Yaffe, M. J., Peterson, A. D. (2005). Mammographic breast density as an intermediate phenotype of breast cancer. *Lancet Oncol, 6,* 798-808.

Brodersen, J. & Siersma, V. D. (2013). Long-term psychosocial consequences of false-positive screening mammography. *The Annals of Family Medicine, 11,* 106-115.

Brodersen, J., Thorsen, H., & Kreiner, S. (2007). Validation of a Condition–Specific Measure for Women Having an Abnormal Screening Mammography. *Value in health, 10,* 294-304.

Broeders, M., Moss, S., Nystrom L, et al. (2012).The impact of mammographic screening on breast cancer mortality in Europe: a review of observational studies. *J Med Screen, 19* (Suppl 1):14-25.

Canadian Association of Radiologists (2012). CAR Practice Guidelines and Technical Standards for Breast Imaging and Intervention. http://www.car.ca/uploads/standards%20guidelines/20131024_en_breast_imaging_practice_guidelines.pdf

Canadian Breast Cancer Network (2010). Breast Cancer: Economic Impact and Labour Force Re-Entry. http://www.cbcn.ca/index.php?id=2912&lang=en&pageaction=content.page

Canadian Cancer Society's Advisory Committee on Cancer Statistics (2015). Canadian Cancer Statistics. http://www.cancer.ca/~/media/cancer.ca/CW/cancer%20information/cancer%20101/Canadian%20cancer%20statistics/Canadian-Cancer-Statistics-2015-EN.pdf?la=en

Canadian Cancer Society's Advisory Committee on Cancer Statistics (2016). Canadian Cancer Statistics. http://www.cancer.ca/~/media/cancer.ca/CW/cancer%20information/cancer%20101/Canadian%20cancer%20statistics/Canadian-Cancer-Statistics-2016-EN.pdf?la=en

Canadian Partnership Against Cancer (2015). Breast Cancer Screening in Canada: Monitoring and Evalatin of Quality Indicators- Results Report, January 2009-December 2010. Toronto: Canadian Partnerships Against Cancer.

Canadian Task Force on Preventative Health Care (2014). Canadian Task Force on Preventative Health Care. Procedure Manual. http://canadiantaskforce.ca/files/procedural-manual-en.pdf.

Canadian Task Force on Preventive Health Care (2011). Recommendations on screening for breast cancer in average-risk women aged 40–74 years. *Canadian Medical Association Journal, 183,* 1991-2001.

Cancer Research UK (2016). Breast cancer risk factor. http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer/risk-factors

Ciatto, S., Visioli, C., Paci, E., & Zappa, M. (2004). Breast density as a determinant of interval cancer at mammographic screening. *British journal of cancer, 90,* 393-396.

Claus, E. B., Risch, N., & Thompson, W. D. (1994). Autosomal dominant inheritance of early–onset breast cancer. Implications for risk prediction. *Cancer, 73,* 643-651.

de Gonzalez, A. B. & Darby, S. (2004). Risk of cancer from diagnostic X-rays: estimates for the UK and 14 other countries. *The lancet, 363,* 345-351.

de Oliveira, C., Bremner, K. E., Pataky, R., Gunraj, N., Chan, K., Peacock, S. et al. (2013). Understanding the costs of cancer care before and after diagnosis for the 21 most common cancers in Ontario: a population-based descriptive study. *CMAJ open, 1,* E1-E8.

DeSantis, C., Siegel, R., Bandi, P., & Jemal, A. (2011). Breast cancer statistics, 2011. *CA: a cancer journal for clinicians, 61,* 408-418.

Duffy, S. W., Tabar, L., Olsen, A. H., Vitak, B., Allgood, P. C., Chen, T. H. et al. (2010). Absolute numbers of lives saved and overdiagnosis in breast cancer screening, from a randomized trial and from the Breast Screening Programme in England. *Journal of medical screening, 17,* 25-30.

Elmore, J. G., Armstrong, K., Lehman, C. D., & Fletcher, S. W. (2005). Screening for breast cancer. *Jama, 293,* 1245-1256.

Etzioni, R. & Gulati, R. (2016). Recognizing the limitations of cancer overdiagnosis studies: a first step towards overcoming them. *Journal of the National Cancer Institute, 108,* djv345.

Feig, S. A. (2004). Adverse effects of screening mammography. *Radiologic Clinics of North America, 42,* 807-819.

Feig, S. A. (2015). Overdiagnosis of breast cancer at screening is clinically insignificant. *Academic radiology, 22,* 961-966.

Fitzpatrick-Lewis, D., Hodgson, N., Ciliska, D., Peirson, L., Gauld, M., Liu, Y. Y. et al. (2011). Breast cancer screening. *Calgary: Canadian Task Force on Preventive Health Care*.

Friedewald, S.M., Rafferty, E.A., Rose, S.L., et al. (2014). Breast Cancer Screening using tomosynthesis in combination with digital mammography. *JAMA, 24*, 2499-2507.

Fu, R., Gartlehner, G., Grant, M., Shamliyan, T., Sedrakyan, A., Wilt, T. J. et al. (2011). Conducting quantitative synthesis when comparing medical interventions: AHRQ and the Effective Health Care Program. *Journal of clinical epidemiology, 64,* 1187-1197.

Gail, M. H., Brinton, L. A., Byar, D. P., Corle, D. K., Green, S. B., Schairer, C. et al. (1989). Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. *Journal of the National Cancer Institute, 81,* 1879-1886.

Garber, J. E. & Offit, K. (2005). Hereditary cancer predisposition syndromes. *Journal of Clinical Oncology, 23,* 276-292.

Gartlehner, G., Thaler, K., Chapman, A., Kaminski-Hartenthaler, A., Berzaczy, D., Van Noord, M. G. et al. (2013). Mammography in combination with breast ultrasonography versus mammography for breast cancer screening in women at average risk. *The cochrane library*.

Gotzsche, P. C. & Jorgensen, K. J. (2013). Screening for breast cancer with mammography. *The cochrane library*.

Gotzsche, P. C. & Nielsen, M. (2009). Screening for breast cancer with mammography. *The cochrane library*.

Gotzsche, P. C. & Olsen, O. (2000). Is screening for breast cancer with mammography justifiable? *The lancet, 355,* 129-134.

Guyatt, G. H., Thorlund, K., Oxman, A. D., Walter, S. D., Patrick, D., Furukawa, T. A. et al. (2013). GRADE guidelines: 13. Preparing summary of findings tables and evidence profiles–continuous outcomes. *Journal of clinical epidemiology, 66,* 173-183.

Harris, R., Yeatts, J., Kinsinger, L. (2011). Breast cancer screening for women ages 50 to 69 years a systematic review of observational evidence. *Prev Med, 53*, 108-14.

Hamashima, C., Hamashima, C., Hattori, M., Honjo, S., Kasahara, Y., Katayama, T. et al. (2016). The Japanese Guidelines for Breast Cancer Screening. *Japanese journal of clinical oncology,* hyw008.

Helvie, M. A., Chang, J. T., Hendrick, R. E., & Banerjee, M. (2014). Reduction in late–stage breast cancer incidence in the mammography era: Implications for overdiagnosis of invasive cancer. *Cancer, 120,* 2649-2656.

Higgins, J. P. & Green, S. (2008). *Cochrane handbook for systematic reviews of interventions*. (5 ed.) Wiley Online Library.

Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *Bmj, 327,* 557-560.

Hofvind, S., Ponti, A., Patnick, J., et al. (2012) False-positive results in mammographic screening for breast cancer in Europe: a literature review and survey of service screening programmes. *J Med Screen, 19(Supp1)*, 57-66.

Iared, W., Shigueoka, DC., Torloni, MR., et al. (2011) Comparative evaluation of digital mammography and film mammography: systematic review and meta-analysis. *Sao Paulo Med J, 129(4),* 250-60.

Independent U. K. Panel on Breast Cancer Screening. (2012). The benefits and harms of breast cancer screening: an independent review. *Lancet, 380(9855)*, 1778-86.

International Agency for Research on Cancer (2012). IARC MONOGRAPHS ON THE EVALUATION OF CARCINOGENIC RISKS TO HUMANS. http://monographs.iarc.fr/ENG/Monographs/vol100E/

Johnson, K. C., Miller, A. B., Collishaw, N. E., Palmer, J. R., Hammond, S. K., Salmon, A. G. et al. (2010). Active smoking and secondhand smoke increase breast cancer risk: the report of the Canadian Expert Panel on Tobacco Smoke and Breast Cancer Risk (2009). *Tobacco control,* tc-2010.

Jorgensen, K. J. & Gotzsche, P. C. (2009). Overdiagnosis in publicly organised mammography screening programmes: systematic review of incidence trends. *Bmj, 339,* b2587.

Killip, S., Mahfoud, Z., & Pearce, K. (2004). What is an intracluster correlation coefficient? Crucial concepts for primary care researchers. *The Annals of Family Medicine, 2,* 204-208.

Kleibl, Z. & Kristensen, V. N. (2016). Women at high risk of breast cancer: Molecular characteristics, clinical presentation and management. *The Breast, 28,* 136-144.

Kriege, M., Brekelmans, C. T., Boetes, C., Besnard, P. E., Zonderland, H. M., Obdeijn, I. M. et al. (2004). Efficacy of MRI and mammography for breast-cancer screening in women with a familial or genetic predisposition. *New England journal of medicine, 351,* 427-437.

Kuhl, C. (2007). The current status of breast MR imaging part I. Choice of technique, image interpretation, diagnostic accuracy, and transfer to clinical practice 1. *Radiology, 244,* 356-378.

Kuhl, C., Weigel, S., Schrading, S., Arand, B., Bieling, H., K+Ânig, R. et al. (2010). Prospective multicenter cohort study to refine management recommendations for women at elevated familial risk of breast cancer: the EVA trial. *Journal of Clinical Oncology, 28,* 1450-1457.

Lee, C. I., Bassett, L. W., & Lehman, C. D. (2012). Breast density legislation and opportunities for patient-centered outcomes research. *Radiology, 264,* 632-636.

Lord, S. J., Lei, W., Craft, P., Cawson, J. N., Morris, I., Walleser, S. et al. (2007). A systematic review of the effectiveness of magnetic resonance imaging (MRI) as an addition to mammography and ultrasound in screening young women at high risk of breast cancer. *European journal of cancer, 43,* 1905-1917.

Magnus, MC., Ping, M., Shen, MM., et al. (2011). Effectiveness of mammography screening in reducing breast cancer mortality in women aged 39-49 years: a meta-analysis. *J Womens D-39 Health (Larchmt), 20(6)*, 845-52.

MARIBS Study Group (2005). Screening with magnetic resonance imaging and mammography of a UK population at high familial risk of breast cancer: a prospective multicentre cohort study (MARIBS). *The lancet, 365,* 1769-1778.

Marmot, M. G., Altman, D. G., Cameron, D. A., Dewar, J. A., Thompson, S. G., & Wilcox, M. (2013). The benefits and harms of breast cancer screening: an independent review. *BrJ Cancer, 108,* 2205-2240.

Mavaddat, N., Antoniou, A. C., Easton, D. F., & Garcia-Closas, M. (2010). Genetic susceptibility to breast cancer. *Molecular oncology, 4,* 174-191.

McCormack, V. A. & dos Santos Silva, I. (2006). Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. *Cancer Epidemiology Biomarkers & Prevention, 15,* 1159-1169.

McGowan, J., Sampson, M., Salzwedel, D. M., Cogo, E., Foerster, V., & Lefebvre, C. (2016). PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Statement. *Journal of clinical epidemiology*.

Miller, A. B., Wall, C., Baines, C. J., Sun, P., To, T., & Narod, S. A. (2014). Twenty five year follow-up for breast cancer incidence and mortality of the Canadian National Breast Screening Study: randomised screening trial.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of internal medicine, 151,* 264-269.

Morris, E., Feig, S. A., Drexler, M., & Lehman, C. (2015). Implications of overdiagnosis: impact on screening mammography practices. *Population health management, 18,* S-3.

Mossa-Basha, M., Fundaro, G. M., Shah, B. A., Ali, S., & Pantelic, M. V. (2010). Ductal carcinoma in situ of the breast: MR imaging findings with histopathologic correlation 1. *Radiographics, 30,* 1673-1687.

Mukhtar, R. A., Wong, J. M., & Esserman, L. J. (2015). Preventing overdiagnosis and overtreatment: just the next step in the evolution of breast cancer care. *Journal of the National Comprehensive Cancer Network, 13,* 737-743.

Myers, E. R., Moorman, P., Gierisch, J. M., Havrilesky, L. J., Grimm, L. J., Ghate, S. et al. (2015). Benefits and harms of breast cancer screening: a systematic review. *Jama, 314,* 1615-1634.

National Breast Cancer Foundation, Inc. (2015). About Breast Cancer- Stages 0 & 1. http://www.nationalbreastcancer.org/breast-cancer-stage-0-and-stage-1

National Cancer Institute (2015). SEER Stat Fact Sheet: Breast Cancer. http://seer.cancer.gov/statfacts/html/breast.html

Nelson, H. D., Cantor, A., Humphrey, L., Fu, R., Pappas, M., Daeges, M. et al. (2016a). *Screening for Breast Cancer: A Systematic Review to Update the 2009 U.S. Preventive Service Task Force Recommendation*.

Nelson, H. D., Tyne, K., Naik, A., Bougatsos, C., Chan, B., Nygren, P. et al. (2009). *Screening for Breast Cancer: Systematic Evidence Review Update for the U.S. Preventive Services Task Force. Evidence Review Update No. 74*.

Nelson, H. D., Fu, R., Cantor, A., Pappas, M., Daeges, M., & Humphrey, L. (2016b). Effectiveness of breast cancer screening: systematic review and meta-analysis to update the 2009 US Preventive Services Task Force recommendation. *Annals of internal medicine, 164,* 244-255.

Nelson, H. D., Zakher, B., Cantor, A., Fu, R., Griffin, J., O'Meara, E. S. et al. (2012). Risk factors for breast cancer for women aged 40 to 49 years: a systematic review and meta-analysis. *Annals of internal medicine, 156,* 635-648.

Nothacker, M., Duda, V., Hahn, M., Warm, M., Degenhardt, F., Madjar, H. et al. (2009). Early detection of breast cancer: benefits and risks of supplemental breast ultrasound in asymptomatic women with mammographically dense breast tissue. A systematic review. *BMC cancer, 9,* 335.

Oeffinger, K. C., Fontham, E. T., Etzioni, R., Herzig, A., Michaelson, J. S., Shih, Y. C. T. et al. (2015). Breast cancer screening for women at average risk: 2015 guideline update from the American Cancer Society. *Jama, 314,* 1599-1614.

Olsen, O. & Gotzsche, P. C. (2001). Cochrane review on screening for breast cancer with mammography. *The lancet, 358,* 1340-1342.

Otto, S. J., Fracheboud, J., Looman, C. W. N. et al. Intiation of population-based mammography screening in Dutch municipalities and effect on breast-cancer mortality: a systematic review. *Lancet, 361*, 1411-7.

Oxman, A. D. & GRADE Working Group (2004). Grading quality of evidence and strength of recommendations. *Bmj, 328,* 1490-1494.

Parkin, D. M. & Darby, S. C. (2011). 12. Cancers in 2010 attributable to ionising radiation exposure in the UK. *British journal of cancer, 105,* S57-S65.

Paci E. (2012). Summary of the evidence of breast cancer service screening outcomes in Europe and first estimate of the benefit and harm balance sheet. *J Med Screen, 19(Suppl 1),* 5-13.

Pataky, R. (2015). Using linked administrative data to measure the costs of cancer care in BC and Ontario. Canadian Centre for Applied Research in Cancer Control. https://www.cadth.ca/.../Concurrent%20Session%20B1_Cancer%20Care_Reka%20Pat.

Pisano, E.D., Gatsonis, C., Hendrick, E., et al. (2005). Diagnostic performance of digital versus film mammography for breast-cancer screening. *N Engl J Med*, *353*, 1773-1783.

Puliti, D., Duffy, S. W., Miccinesi, G., De Koning, H., Lynge, E., Zappa, M. et al. (2012). Overdiagnosis in mammographic screening for breast cancer in Europe: a literature review. *Journal of medical screening, 19,* 42-56.

Rahbar, H., McDonald, E. S., Lee, J. M., Partridge, S. C., & Lee, C. I. (2016). How Can Advanced Imaging Be Used to Mitigate Potential Breast Cancer Overdiagnosis? *Academic radiology, 23,* 768-773.

Rashidian, A., Barfar, E., Hosseini, H., et al. (2013). Cost effectiveness of breast cancer screening using mammography; a systematic review. *Iran J Public Health, 42(4)*:347-57.

Richie, R. C. & Swanson, J. O. (2003). Breast cancer: a review of the literature. *JOURNAL OF INSURANCE MEDICINE-NEW YORK THEN DENVER--, 35,* 85-101.

Salz, T., Richman, A. R., & Brewer, N. T. (2010). Meta–analyses of the effect of false–positive mammograms on generic and specific psychosocial outcomes. *Psycho-Oncology, 19,* 1026-1034.

Sardanelli, F., Podo, F., Santoro, F., Manoukian, S., Bergonzi, S., Trecate, G. et al. (2011). Multicenter surveillance of women at high genetic breast cancer risk using mammography, ultrasonography, and contrast-enhanced magnetic resonance imaging (the high breast cancer risk italian 1 study): final results. *Investigative radiology, 46,* 94-105.

Scoccianti, C., Lauby-Secretan, B. +., Bello, P. Y., Chajes, V. +., & Romieu, I. (2014). Female breast cancer and alcohol consumption: a review of the literature. *American journal of preventive medicine, 46,* S16-S25.

Shamseer, L., Moher, D., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M. et al. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *Bmj, 349*.

Shea, B., Grimshaw, JM., Wells, GA., Boers, M., Andersson, N., Hamel, C., Porter, AC., Tugwell, P., Moher, D., & Bouter , LM. (2007). Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol*, 15, 10.

Siu, A. L. (2016). Screening for breast cancer: US Preventive Services Task Force recommendation statement. *Annals of internal medicine, 164,* 279-296.

Sterne, J. A., Sutton, A. J., Ioannidis, J. P., Terrin, N., Jones, D. R., Lau, J. et al. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *Bmj, 343,* d4002.

Tsilidis, K. K., Kasimis, J. C., Lopez, D. S., Ntzani, E. E., & Ioannidis, J. P. (2015). Type 2 diabetes and cancer: umbrella review of meta-analyses of observational studies. *Bmj, 350,* g7607.

Tyrer, J., Duffy, S. W., & Cuzick, J. (2004). A breast cancer prediction model incorporating familial and personal risk factors. *Statistics in medicine, 23,* 1111-1130.

Warner, E., Plewes, D. B., Hill, K. A., Causer, P. A., Zubovits, J. T., Jong, R. A. et al. (2004). Surveillance of BRCA1 and BRCA2 mutation carriers with magnetic resonance imaging, ultrasound, mammography, and clinical breast examination. *Jama, 292,* 1317-1325.

Weinstein, S. P., Localio, A. R., Conant, E. F., Rosen, M., Thomas, K. M., & Schnall, M. D. (2009). Multimodality screening of high-risk women: a prospective cohort study. *Journal of Clinical Oncology, 27,* 6124-6128.

Welch, H.G., Prorok, P.C., O'Malley, J., Kramer, B.S. (2016). Breast-cancer tumor size, overdiagnosis, and mammography screening effectiveness. *NEJM, 375,* 1438-1447.

Wells, G. A., Shea, B., O-Connell, D., Peterson, J., Welch, V., Losos, M. et al. (2007). Newcastle-Ottawa Scale. *Hospital Research Institute Ottawa (http://www.ohri.ca/programs/clinical_epidemiology/nosgen.pdf*.

Williams, J. H., Carter, S. M., & Rychetnik, L. (2014). Organised cervical screening 45 years on: How consistent are organised screening practices? *European journal of cancer, 50,* 3029-3038.

Appendices

Appendix A-I. Recommendations from Other Guideline Developers

Organization	Not at high-risk women aged 40 to 49 years	Not at high-risk women aged 50 to 69 years	Not at high-risk women aged 70 to 74 years	Additional recommendations
Canadian Task Force on Preventive Health Care (2011)	For women aged 40-49 years, we recommend not routinely screening with mammography	For women aged 50-69 years, recommend routinely screening with mammography every two to three years	For women aged 70-74 years, we recommend routinely screening with mammography every two to three years	Recommend not routinely screening with MRI scans Recommend not routinely performing clinical breast examinations alone or in conjunction with mammography to screen for breast cancer Recommend not advising women to routinely practice breast self-examination
U.S. Preventive Services Task Force (2016)	For women prior to age 50 years, women who place a higher value on the potential benefit from the potential harms may choose to begin biennial screening mammography	For women aged 50-69 years, recommends biennial screening mammography	No recommendation	
American Cancer Society (2015)	Women should have the opportunity to begin annual screening between the ages of 40-44 years For women at 45 years, recommend undergo regular screening mammography For women aged 45-54 years, recommend screened annually	For women aged 45-54 years, recommend screened annually For women 55 years and older, recommend transition to biennial screening or have the opportunity to continue screening annually	No recommendation	Recommend continuing screening mammography as long as overall health is good and have a life expectancy of 10 years or longer Do not recommend clinical breast examination at any age
Japanese National Cancer Center (2016)	For women aged 40-74 years, mammographic screening without clinical breast examination recommended for population-based and opportunistic screenings For women aged 40–64 years, mammographic screening with clinical breast examination is recommended for population-based and opportunistic screenings	For women aged 40-74 years, mammographic screening without clinical breast examination recommended for population-based and opportunistic screenings For women aged 40–64 years, mammographic screening with clinical breast examination is recommended for population-based and opportunistic screenings	For women aged 40-74 years, mammographic screening without clinical breast examination recommended for population-based and opportunistic screenings	Do not recommend clinical breast examination and ultrasonography for population-based screening

Appendix A-II. Additional potential systematic reviews

Potentially relevant ystematic reviews identified from:

-the USTPSTF 2016 Systematic review:

- Armstrong et al., 2007
- Nothacker et al., 2009
- Harris et al., 2011
- Otto et al., 2003

-the ACS 2014 Systematic review:

- Broeders et al., 2012
- Gotzsche et al., 2013
- Hofvind et al., 2012
- Iared et al., 2011
- Independent U.K. panel on breast cancer screening, 2012
- Magnus et al., 2011
- Marmot et al., 2013
- Paci et al., 2012
- Puliti et al., 2012
- Rashidian et al., 2013

Appendix B. Search Strategies for KQ1

KQ1: Benefits and Harms of Different Screening Modalities

Breast Cancer Screening, Including Breast Self-Examination

Oct 2014-present
2016Oct25

Database: Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) <1946 to Present>
Search Strategy:

---

1 exp Breast Neoplasms/ (252031)
2 ((breast* or mamma or mammar*) adj3 (cancer* or carcinoid* or carcinoma* or carcinogen* or adenocarcinoma* or adeno-carcinoma* or malignan* or neoplasia* or neoplasm* or sarcoma* or tumour* or tumor*)).tw,kw. (288000)
3 exp Carcinoma, Intraductal, Noninfiltrating/ (9164)
4 intraductal carcinoma*.tw,kw. (848)
5 (ductal carcinoma in situ or DCIS).tw,kw. (6471)
6 or/1-5 [BREAST CANCER] (344782)
7 exp Breast Neoplasms/di, pc (42169)
8 exp Mass Screening/ (112536)
9 screen*.tw,kw. (581834)
10 "Early Detection of Cancer"/ (14857)
11 ((early or earlier or earliest) adj3 (detect* or diagnos* or identif* or recogni*)).tw,kw. (202302)
12 exp Self-Examination/ (2294)
13 ((self-exam* or self-detect* or self-screen*) adj5 (breast$1 or mamma or mammary or nipple$1)).tw,kw. (1896)
14 Physical Examination/ (37521)
15 (exam* adj5 (breast? or mamma or mammar* or nipple?)).tw,kw. (13266)
16 exp Breast Neoplasms/ra (13141)
17 exp Mammography/ (26911)
18 (mammograph* or mammogram*).tw,kw. (28228)
19 exp Magnetic Resonance Imaging/ (365078)
20 (fMRI or fMRIs or MRI or MRIs or NMRI or NMRIs or MR imaging or NMR imaging or magnetic resonance imag* or magnetic resonance tomograph* or MR tomograph*).tw,kw. (334228)
21 (chemical shift imaging or proton spin tomograph* or zeugmatograph*).tw,kw. (976)
22 exp Breast Neoplasms/us (3560)
23 (ultrasound* or ultrason* or echograph* or echomammogra* or echo-mammogra* or echotomograph* or echo-tomograph* or sonograph*).tw,kw. (337776)
24 or/7-23 (1612490)
25 6 and 24 [BREAST CANCER SCREENING] (87637)
26 Male/ not (Female/ and Male/) (2523682)
27 25 not 26 [MALE-ONLY REMOVED] (86297)
28 exp Infant/ not (exp Adult/ and exp Infant/) (767911)
29 exp Child/ not (exp Adult/ and exp Child/) (1089813)
30 Adolescent/ not (exp Adult/ and Adolescent/) (537360)
31 or/28-30 (1699968)

32 27 not 31 [CHILD-ONLY REMOVED] (85862)
33 exp Animals/ not (exp Animals/ and Humans/) (4343931)
34 32 not 33 [ANIMAL-ONLY REMOVED] (84601)
35 (comment or editorial or news or newspaper article).pt. (1168329)
36 (letter not (letter and randomized controlled trial)).pt. (948307)
37 34 not (35 or 36) [OPINION PIECES REMOVED] (78356)
38 (201410* or 201411* or 201412* or 2015* or 2016*).dc. (2615459)
39 37 and 38 [UPDATE PERIOD] (9703)
40 (controlled clinical trial or randomized controlled trial or pragmatic clinical trial).pt. (521901)
41 clinical trials as topic.sh. (181352)
42 (randomi#ed or randomly or RCT$1 or placebo*).tw. (774894)
43 ((singl* or doubl* or trebl* or tripl*) adj (mask* or blind* or dumm*)).tw. (149457)
44 trial.ti. (164073)
45 or/40-44 (1125554)
46 39 and 45 [RCTS] (683)
47 controlled clinical trial.pt. (92104)
48 Controlled Clinical Trial/ or Controlled Clinical Trials as Topic/ (97361)
49 (control* adj2 trial*).tw. (197609)
50 Non-Randomized Controlled Trials as Topic/ (90)
51 (nonrandom* or non-random* or quasi-random* or quasi-experiment*).tw. (43541)
52 (nRCT or nRCTs or non-RCT$1).tw. (517)
53 (pre- adj3 post-).tw. (57735)
54 (pretest adj3 posttest).tw. (3911)
55 Historically Controlled Study/ (76)
56 (control* adj2 stud$3).tw. (190073)
57 Control Groups/ (1628)
58 (control$ adj2 group$1).tw. (395697)
59 trial.ti. (164073)
60 or/47-59 (963990)
61 39 and 60 [NON-RCTS] (771)
62 exp Cohort Studies/ (1607449)
63 cohort$1.tw. (394543)
64 Retrospective Studies/ (610873)
65 (longitudinal or prospective or retrospective).tw. (945496)
66 ((followup or follow-up) adj (study or studies)).tw. (43921)
67 Observational study.pt. (27345)
68 (observation$2 adj (study or studies)).tw. (68552)
69 ((population or population-based) adj (study or studies or analys#s)).tw. (13809)
70 ((multidimensional or multi-dimensional) adj (study or studies)).tw. (90)
71 Comparative Study.pt. (1778832)
72 ((comparative or comparison) adj (study or studies)).tw. (93034)
73 exp Case-Control Studies/ (822290)
74 ((case-control* or case-based or case-comparison) adj (study or studies)).tw. (84965)
75 or/62-74 (3798948)
76 39 and 75 [OBSERVATIONAL STUDIES] (3025)
77 46 or 61 or 76 [ALL STUDY DESIGNS] (3563)

****

Cochrane Library

Search Name: CTFPHC - Breast Cancer Screening - All Modalities
Date Run: 25/10/16 15:47:59.102
Description: 2016 Oct 21 (OHRI) - Oct 2014-present – Post-PRESS

ID Search Hits
#1 [mh "Breast Neoplasms"] 9857
#2 ((breast* or mamma or mammar*) near/3 (cancer* or carcinoid* or carcinoma* or carcinogen* or adenocarcinoma* or adeno-carcinoma* or malignan* or neoplasia* or neoplasm* or sarcoma* or tumour* or tumor*)):ti,ab,kw 22276
#3 [mh "Carcinoma, Intraductal, Noninfiltrating"] 116
#4 (intraductal next carcinoma*):ti,ab,kw 161
#5 ("ductal carcinoma in situ" or DCIS):ti,ab,kw 292
#6 {or #1-#5} 22323
#7 [mh "Breast Neoplasms"/DI,PC] 1446
#8 [mh "Mass Screening"] 5513
#9 screen*:ti,ab,kw 28604
#10 [mh "Early Detection of Cancer"] 872
#11 ((early or earlier or earliest) near/3 (detect* or diagnos* or identif* or recogni*)):ti,ab,kw 5437
#12 [mh Self-Examination] 200
#13 ((self next (exam* or detect* or screen*)) near/5 (breast* or mamma or mammary or nipple*)):ti,ab,kw 206
#14 [mh ^"Physical Examination"] 907
#15 (exam* near/5 (breast* or mamma or mammar* or nipple*)) .tw,kw. 2
#16 [mh "Breast Neoplasms"/ra] 380
#17 [mh Mammography] 1020
#18 (mammograph* or mammogram*):ti,ab,kw 1828
#19 [mh "Magnetic Resonance Imaging"] 6998
#20 (fMRI or fMRIs or MRI or MRIs or NMRI or NMRIs or "MR imaging" or "NMR imaging" or ("magnetic resonance" next imaging) or ("magnetic resonance" next tomograph*) or (MR next tomograph*)):ti,ab,kw 14187
#21 ("chemical shift imaging" or ("proton spin" next tomograph*) or zeugmatograph*):ti,ab,kw 19
#22 [mh "Breast Neoplasms"/US] 85
#23 (ultrasound* or ultrason* or echograph* or echomammogra* or echo-mammogra* or echotomograph* or echo-tomograph* or sonograph*):ti,ab,kw 20853
#24 {or #7-#23} 67190
#25 #6 and #24 Publication Year from 2014 to 2016 689

DSR – 12 [Reviews]
DARE – 26 [Reviews]
CENTRAL – 615 [RCTs]
HTA – 22 [Reviews]
NHS EED – 14 [*Economic studies - do not download*]

Breast Self-Examination (Oct 2010 – Sep 2014)
2016Oct25
OVID MEDLINE
Database: Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) <1946 to Present>

Search Strategy:
---
1 exp Breast Neoplasms/ (252031)

2 ((breast* or mamma or mammar*) adj3 (cancer* or carcinoid* or carcinoma* or carcinogen* or adenocarcinoma* or adeno-carcinoma* or malignan* or neoplasia* or neoplasm* or sarcoma* or tumour* or tumor*)).tw,kw. (288000)
3 exp Carcinoma, Intraductal, Noninfiltrating/ (9164)
4 intraductal carcinoma*.tw,kw. (848)
5 (ductal carcinoma in situ or DCIS).tw,kw. (6471)
6 or/1-5 [BREAST CANCER] (344782)
7 exp Self-Examination/ (2294)
8 ((self-exam* or self-detect* or self-screen*) adj5 (breast$1 or mamma or mammary or nipple$1)).tw,kw. (1896)
9 or/7-8 (3288)
10 6 and 9 [BREAST SELF-EXAMINATION] (2111)
11 Male/ not (Female/ and Male/) (2523682)
12 10 not 11 (2104)
13 exp Infant/ not (exp Adult/ and exp Infant/) (767911)
14 exp Child/ not (exp Adult/ and exp Child/) (1089813)
15 Adolescent/ not (exp Adult/ and Adolescent/) (537360)
16 or/13-15 (1699968)
17 12 not 16 (2083)
18 exp Animals/ not (exp Animals/ and Humans/) (4343931)
19 17 not 18 [ANIMAL-ONLY REMOVED] (2083)
20 (comment or editorial or news or newspaper article).pt. (1168329)
21 (letter not (letter and randomized controlled trial)).pt. (948307)
22 19 not (20 or 21) [OPINION PIECES REMOVED] (1947)
23 (201010* or 201011* or 201012* or 2011* or 2012* or 2013* or 201401* or 201402* or 201403* or 201404* or 201405* or 201406* or 201407* or 201408* or 201409*).dc. (4117599)
24 22 and 23 [UPDATE PERIOD] (279)
25 (controlled clinical trial or randomized controlled trial or pragmatic clinical trial).pt. (521901)
26 clinical trials as topic.sh. (181352)
27 (randomi#ed or randomly or RCT$1 or placebo*).tw. (774894)
28 ((singl* or doubl* or trebl* or tripl*) adj (mask* or blind* or dumm*)).tw. (149457)
29 trial.ti. (164073)
30 or/25-29 (1125554)
31 24 and 30 [RCTS] (24)
32 controlled clinical trial.pt. (92104)
33 Controlled Clinical Trial/ or Controlled Clinical Trials as Topic/ (97361)
34 (control* adj2 trial*).tw. (197609)
35 Non-Randomized Controlled Trials as Topic/ (90)
36 (nonrandom* or non-random* or quasi-random* or quasi-experiment*).tw. (43541)
37 (nRCT or nRCTs or non-RCT$1).tw. (517)
38 (pre- adj3 post-).tw. (57735)
39 (pretest adj3 posttest).tw. (3911)
40 Historically Controlled Study/ (76)
41 (control* adj2 stud$3).tw. (190073)
42 Control Groups/ (1628)
43 (control$ adj2 group$1).tw. (395697)
44 trial.ti. (164073)
45 or/32-44 (963990)

46 24 and 45 [NON-RCTS] (16)
47 exp Cohort Studies/ (1607449)
48 cohort$1.tw. (394543)
49 Retrospective Studies/ (610873)
50 (longitudinal or prospective or retrospective).tw. (945496)
51 ((followup or follow-up) adj (study or studies)).tw. (43921)
52 Observational study.pt. (27345)
53 (observation$2 adj (study or studies)).tw. (68552)
54 ((population or population-based) adj (study or studies or analys#s)).tw. (13809)
55 ((multidimensional or multi-dimensional) adj (study or studies)).tw. (90)
56 Comparative Study.pt. (1778832)
57 ((comparative or comparison) adj (study or studies)).tw. (93034)
58 exp Case-Control Studies/ (822290)
59 ((case-control* or case-based or case-comparison) adj (study or studies)).tw. (84965)
60 or/47-59 [OBSERVATIONAL STUDIES] (3798948)
61 24 and 60 [OBSERVATIONAL STUDIES] (62)
62 31 or 46 or 61 [ALL STUDY DESIGNS] (88)

****
Cochrane Library

Search Name: CTFPHC - Breast Cancer Screening - Self-Examination
Date Run: 25/10/16 15:45:31.291
Description: 2016 Oct 21 - 2010-2014 - Post-PRESS

ID Search Hits
#1 [mh "Breast Neoplasms"] 9857
#2 ((breast* or mamma or mammar*) near/3 (cancer* or carcinoid* or carcinoma* or carcinogen* or adenocarcinoma* or adeno-carcinoma* or malignan* or neoplasia* or neoplasm* or sarcoma* or tumour* or tumor*)):ti,ab,kw 22276
#3 [mh "Carcinoma, Intraductal, Noninfiltrating"] 116
#4 (intraductal next carcinoma*):ti,ab,kw 161
#5 ("ductal carcinoma in situ" or DCIS):ti,ab,kw 292
#6 {or #1-#5} 22323
#7 [mh Self-Examination] 200
#8 ((self next (exam* or detect* or screen*)) near/5 (breast* or mamma or mammary or nipple*)):ti,ab,kw 206
#9 #7 or #8 301
#10 #6 and #9 Publication Year from 2010 to 2014 23

CENTRAL – 23 [RCTs]

Breast Cancer Screening – Harms

Draft Strategy
Oct 2014-present
2016 Oct25
Database: Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) <1946 to Present>

Database: Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) <1946 to Present>
Search Strategy:

---

1 exp Breast Neoplasms/ (252031)
2 ((breast* or mamma or mammar*) adj3 (cancer* or carcinoid* or carcinoma* or carcinogen* or adenocarcinoma* or adeno-carcinoma* or malignan* or neoplasia* or neoplasm* or sarcoma* or tumour* or tumor*)).tw,kw. (288000)
3 exp Carcinoma, Intraductal, Noninfiltrating/ (9164)
4 intraductal carcinoma*.tw,kw. (848)
5 (ductal carcinoma in situ or DCIS).tw,kw. (6471)
6 or/1-5 [BREAST CANCER] (344782)
7 exp Breast Neoplasms/di, pc (42169)
8 exp Mass Screening/ (112536)
9 screen*.tw,kw. (581834)
10 "Early Detection of Cancer"/ (14857)
11 ((early or earlier or earliest) adj3 (detect* or diagnos* or identif* or recogni*)).tw,kw. (202302)
12 exp Self-Examination/ (2294)
13 ((self-exam* or self-detect* or self-screen*) adj5 (breast$1 or mamma or mammary or nipple$1)).tw,kw. (1896)
14 Physical Examination/ (37521)
15 (exam* adj5 (breast? or mamma or mammar* or nipple?)).tw,kw. (13266)
16 exp Breast Neoplasms/ra (13141)
17 exp Mammography/ (26911)
18 (mammograph* or mammogram*).tw,kw. (28228)
19 exp Magnetic Resonance Imaging/ (365078)
20 (fMRI or fMRIs or MRI or MRIs or NMRI or NMRIs or MR imaging or NMR imaging or magnetic resonance imag* or magnetic resonance tomograph* or MR tomograph*).tw,kw. (334228)
21 (chemical shift imaging or proton spin tomograph* or zeugmatograph*).tw,kw. (976)
22 exp Breast Neoplasms/us (3560)
23 (ultrasound* or ultrason* or echograph* or echomammogra* or echo-mammogra* or echotomograph* or echo-tomograph* or sonograph*).tw,kw. (337776)
24 or/7-23 (1612490)
25 6 and 24 [BREAST CANCER SCREENING] (87637)
26 Male/ not (Female/ and Male/) (2523682)
27 25 not 26 [MALE-ONLY REMOVED] (86297)
28 exp Infant/ not (exp Adult/ and exp Infant/) (767911)
29 exp Child/ not (exp Adult/ and exp Child/) (1089813)
30 Adolescent/ not (exp Adult/ and Adolescent/) (537360)
31 or/28-30 (1699968)
32 27 not 31 [CHILD-ONLY REMOVED] (85862)
33 exp Animals/ not (exp Animals/ and Humans/) (4343931)

34 32 not 33 [ANIMAL-ONLY REMOVED] (84601)
35 (comment or editorial or news or newspaper article).pt. (1168329)
36 (letter not (letter and randomized controlled trial)).pt. (948307)
37 34 not (35 or 36) [OPINION PIECES REMOVED] (78356)
38 (201410* or 201411* or 201412* or 2015* or 2016*).dc. (2615459)
39 37 and 38 [UPDATE PERIOD] (9703)
40 (controlled clinical trial or randomized controlled trial or pragmatic clinical trial).pt. (521901)
41 clinical trials as topic.sh. (181352)
42 (randomi#ed or randomly or RCT$1 or placebo*).tw. (774894)
43 ((singl* or doubl* or trebl* or tripl*) adj (mask* or blind* or dumm*)).tw. (149457)
44 trial.ti. (164073)
45 or/40-44 (1125554)
46 39 and 45 [RCTS] (683)
47 controlled clinical trial.pt. (92104)
48 Controlled Clinical Trial/ or Controlled Clinical Trials as Topic/ (97361)
49 (control* adj2 trial*).tw. (197609)
50 Non-Randomized Controlled Trials as Topic/ (90)
51 (nonrandom* or non-random* or quasi-random* or quasi-experiment*).tw. (43541)
52 (nRCT or nRCTs or non-RCT$1).tw. (517)
53 (pre- adj3 post-).tw. (57735)
54 (pretest adj3 posttest).tw. (3911)
55 Historically Controlled Study/ (76)
56 (control* adj2 stud$3).tw. (190073)
57 Control Groups/ (1628)
58 (control$ adj2 group$1).tw. (395697)
59 trial.ti. (164073)
60 or/47-59 (963990)
61 39 and 60 [NON-RCTS] (771)
62 exp Cohort Studies/ (1607449)
63 cohort$1.tw. (394543)
64 Retrospective Studies/ (610873)
65 (longitudinal or prospective or retrospective).tw. (945496)
66 ((followup or follow-up) adj (study or studies)).tw. (43921)
67 Observational study.pt. (27345)
68 (observation$2 adj (study or studies)).tw. (68552)
69 ((population or population-based) adj (study or studies or analys#s)).tw. (13809)
70 ((multidimensional or multi-dimensional) adj (study or studies)).tw. (90)
71 Comparative Study.pt. (1778832)
72 ((comparative or comparison) adj (study or studies)).tw. (93034)
73 exp Case-Control Studies/ (822290)
74 ((case-control* or case-based or case-comparison) adj (study or studies)).tw. (84965)
75 (ecolog* adj (study or studies)).tw. (4307)
76 or/62-75 (3802298)
77 39 and 76 [OBSERVATIONAL STUDIES] (3032)
78 exp Mass Screening/ae [Adverse Effects] (729)
79 "Early Detection of Cancer"/ae [Adverse Effects] (200)
80 exp Self-Examination/ae [Adverse Effects] (2)
81 exp Mammography/ae [Adverse Effects] (666)

82 exp Diagnostic Errors/ (105457)
83 misdiagnos*.tw,kw. (25253)
84 (miss$2 adj3 diagnos*).tw,kw. (4326)
85 (overdiagnos* or over diagnos*).tw,kw. (4030)
86 (false adj (negative* or positive*)).tw,kw. (65013)
87 ((error* or false$2 or wrong$2) adj3 (alarm* or detect* or diagnos*)).tw,kw. (19779)
88 exp Medical Overuse/ (4832)
89 overtreat*.tw,kw. (3466)
90 ((inappropriate* or unnecessar*) adj3 (followup or follow-up or procedur* or therap* or treatment*)).tw,kw. (10266)
91 (inappropriate* or unnecessar* or safe or adverse or adversely or undesirabl* or unintend* or unintent* or unwanted or harm* or injurious* or risk or risks or reaction* or complication*).ti. (761629)
92 ((adverse* or undesirabl* or unintend* or unintent* or unwanted or harm* or toxic or injurious* or serious* or fatal) adj5 (affect or affected or affecting or affects or consequence* or effect* or react or reacts or reacted or reacting or reaction* or event* or outcome* or incident*)).tw,kw. (485571)
93 ((adverse* or inappropriat* or unnecessar* or undesirabl* or unintend* or unintent* or unwanted or injurious* or serious*) adj5 (alarm* or anxiet* or anxious* or distress* or emotion* or feeling* or psycholog* or uncertaint*)).tw,kw. (6694)
94 iatrogen*.tw,kw. (26879)
95 or/78-94 (1407212)
96 39 and 95 [HARMS OF BREAST CANCER SCREENING] (1645)
97 96 and 46 [HARMS OF BREAST CANCER SCREENING - RCTS] (183)
98 96 and 61 [HARMS OF BREAST CANCER SCREENING - NON-RCTS] (228)
99 96 and 77 [HARMS OF BREAST CANCER SCREENING - OBSERVATIONAL STUDIES] (714)
100 or/97-99 [HARMS OF BREAST CANCER SCREENING - ALL STUDY DESIGNS] (829)

****

Cochrane Library

Search Name: CTFPHC - Breast Cancer Screening - All Modalities - Harms
Date Run: 25/10/16 15:55:31.636
Description: 2016 Oct 21 (OHRI) - Oct 2014-present - Post-PRESS

ID Search Hits
#1 [mh "Breast Neoplasms"] 9857
#2 ((breast* or mamma or mammar*) near/3 (cancer* or carcinoid* or carcinoma* or carcinogen* or adenocarcinoma* or adeno-carcinoma* or malignan* or neoplasia* or neoplasm* or sarcoma* or tumour* or tumor*)):ti,ab,kw 22276
#3 [mh "Carcinoma, Intraductal, Noninfiltrating"] 116
#4 (intraductal next carcinoma*):ti,ab,kw 161
#5 ("ductal carcinoma in situ" or DCIS):ti,ab,kw 292
#6 {or #1-#5} 22323
#7 [mh "Breast Neoplasms"/DI,PC] 1446
#8 [mh "Mass Screening"] 5513
#9 screen*:ti,ab,kw 28604
#10 [mh "Early Detection of Cancer"] 872
#11 ((early or earlier or earliest) near/3 (detect* or diagnos* or identif* or recogni*)):ti,ab,kw 5437

#12 [mh Self-Examination] 200
#13 ((self next (exam* or detect* or screen*)) near/5 (breast* or mamma or mammary or nipple*)):ti,ab,kw 206
#14 [mh ^"Physical Examination"] 907
#15 (exam* near/5 (breast* or mamma or mammar* or nipple*)) .tw,kw. 2
#16 [mh "Breast Neoplasms"/RA] 380
#17 [mh Mammography] 1020
#18 (mammograph* or mammogram*):ti,ab,kw 1828
#19 [mh "Magnetic Resonance Imaging"] 6998
#20 (fMRI or fMRIs or MRI or MRIs or NMRI or NMRIs or "MR imaging" or "NMR imaging" or ("magnetic resonance" next imaging) or ("magnetic resonance" next tomograph*) or (MR next tomograph*)):ti,ab,kw 14187
#21 ("chemical shift imaging" or ("proton spin" next tomograph*) or zeugmatograph*):ti,ab,kw 19
#22 [mh "Breast Neoplasms"/US] 85
#23 (ultrasound* or ultrason* or echograph* or echomammogra* or echo-mammogra* or echotomograph* or echo-tomograph* or sonograph*):ti,ab,kw 20853
#24 {or #7-#23} 67190
#25 #6 and #24 3707
#26 [mh "Mass Screening"/AE] 45
#27 [mh "Early Detection of Cancer"/AE] 13
#28 [mh Self-Examination/AE] 0
#29 [mh Mammography/AE] 27
#30 [mh "Diagnostic Errors"] 2893
#31 misdiagnos*:ti,ab,kw 198
#32 (miss* near/3 diagnos*):ti,ab,kw 89
#33 (overdiagnos* or (over next diagnos*)):ti,ab,kw 181
#34 (false next (negative* or positive*)):ti,ab,kw 2484
#35 ((error* or false* or wrong*) near/3 (alarm* or detect* or diagnos*)):ti,ab,kw 1143
#36 [mh "Medical Overuse"] 136
#37 overtreat*:ti,ab,kw 184
#38 ((inappropriate* or unnecessar*) near/3 (followup or "follow-up" or procedur* or therap* or treatment*)):ti,ab,kw 540
#39 (inappropriate* or unnecessar* or safe or adverse or adversely or undesirabl* or unintend* or unintent* or unwanted or harm* or injurious* or risk or risks or reaction* or complication*):ti 37781
#40 ((adverse* or undesirabl* or unintend* or unintent* or unwanted or harm* or toxic or injurious* or serious* or fatal) near/5 (affect or affected or affecting or affects or consequence* or effect* or react or reacts or reacted or reacting or reaction* or event* or outcome* or incident*)):ti,ab,kw 118857
#41 ((adverse* or inappropriat* or unnecessar* or undesirabl* or unintend* or unintent* or unwanted or injurious* or serious*) near/5 (alarm* or anxiet* or anxious* or distress* or emotion* or feeling* or psycholog* or uncertaint*)):ti,ab,kw 1157
#42 iatrogen*:ti,ab,kw 670
#43 {or #26-#42} 156154
#44 #25 and #43 Publication Year from 2014 to 2016201

DSR – 6 [Reviews]
DARE – 2 [Reviews]

CENTRAL – 190 [RCTs]
NHS EED – 3 [*Economic studies - do not download*]

Appendix C. Search Strategies for KQ2

KQ2: Womens' valuations and harms in the decision to undergo screening

Draft MEDLINE search

1. exp Breast Neoplasms/
2. exp Carcinoma, Intraductal, Noninfiltrating/
3. *Neoplasms/pc
4. *Precancerous Conditions/pc
5. ((adenocarcinoma* or adenoma* or cancer* or carcinogen* or carcinoid* or carcinoma* or malignan* or metasta*or neoplas* or sarcoma* or tumour* or tumor*) adj3 (breast? or mamma or mammar*)).tw,kf.
6. (DCIS or (ductal carcinoma adj1 in situ)).tw,kf.
7. intraductal carcinoma*.tw,kf.
8. or/1-7 [Combined MeSH & textwords for breast cancer]
9. exp *Breast Neoplasms/di, pc, ra, us
10. "Early Detection of Cancer"/
11. exp Magnetic Resonance Imaging/
12. exp Mammography/
13. exp Mass Screening/
14. Physical Examination/
15. exp Self-Examination/
16. Ultrasonography, Mammary/
17. ((breast? or mamma or mammar* or nipple?) adj5 (exam* or selfexam*)).tw,kf.
18. ((earlier or earliest or early or rapid) adj3 (detect* or diagnos* or identif* or recogni*)).tw,kf.
19. (echograph* or echo-mammogra* or echo-tomograph* or echomammogra* or echotomograph* or sonograph* or ultra-son* or ultra-sound* or ultrason* or ultrasound*).tw,kf.
20. (magnetic resonance imag* or magnetic resonance tomograph* or MR tomograph* or MRI or MRIs or NMRI or NMRIs).tw,kf.
21. (mammogram* or mammograph*).tw,kf.
22. screen*.tw,kf.
23. or/9-22 [Combined MeSH & textwords for screening]
24. and/8,23 [Combined concepts for breast cancer and screening]
25. Choice Behavior/
26. *Consumer Behavior/
27. exp Consumer Participation/
28. Cooperative Behavior/
29. exp Decision Making/
30. Focus Groups/
31. Health Care Surveys/
32. exp Informed Consent/
33. Interviews as Topic/
34. Patient Acceptance of Health Care/
35. exp Patient Education as Topic/
36. Patient Participation/
37. Patient Preference/
38. Social Values/
39. "Surveys and Questionnaires"/
40. Treatment Refusal/
41. (15D* and (HRQoL or QoL or "quality of life")).mp.
42. ((accept* or consider* or choice? or choos* or chose? or decid* or decis* or input* or involv* or opinion* or participat* or perceiv* or percepti* or perspective? or prefer* or refus* or respons* or valuation or value? or valuing or view*) adj3 (citizen? or client? or consumer? or female? or male? or men or patient? or public or stake?holder* or user? or wom#n)).tw,kf.
43. ((analys#s or valuation? or value? or valuing) adj3 (conjoint or contingent)).tw,kf.
44. (choice? adj2 (behavio?r* or discrete or experiment*)).tw,kf.

45. ((choice? or choos* or consent* or decision*) adj1 informed).tw,kf.
46. ((choice? or choos* or decision*) adj2 (made or make or makes or making or shar* or support*)).tw,kf.
47. (EQ 5D or EQ5D or EuroQoL 5D or EuroQoL5D).mp.
48. (focus group? or interview* or questionnaire? or survey*).tw,kf.
49. gambl*.tw,kf.
50. health utilit*.tw,kf.
51. HUI.tw,kf.
52. (multi?attribute or multi?criteria).tw,kf.
53. (preference? adj1 (elicit* or scor* or state*)).tw,kf.
54. prospect theor*.tw,kf.
55. (SF 12 or SF 36 or SF 6D or SF12 or SF36 or SF6D).mp.
56. (trade off? or tradeoff?).tw,kf.
57. (willing* adj2 pay*).tw,kf.
58. or/25-57 [Combined MeSH & text words for patient preferences & values]
59. and/24,58 [Combined concepts for breast cancer screening and patient preferences/values]
60. Male/ not (Female/ and Male/)
61. 59 not 60 [Male only records excluded]
62. Adolescent/ not (exp Adult/ and Adolescent/)
63. exp Child/ not (exp Adult/ and exp Child/)
64. exp Infant/ not (exp Adult/ and exp Infant/)
65. 61 not (62 or 63 or 64) [Adolescent/Infant/Child only records excluded]
66. exp Animals/ not Humans/
67. 65 not 66 [Animal only records excluded]
68. (comment or editorial or news or newspaper article).pt.
69. (letter not (letter and randomized controlled trial)).pt.
70. 67 not (68 or 69) [Opinion pieces excluded]
71. case reports.pt.
72. (case report* or case stud*).ti.
73. 70 not (71 or 72) [Case studies excluded]
74. limit 73 to (english or french)
75. limit 74 to yr="2000-Current"

Appendix D. Draft Screening Forms for KQ1

Level 1 and Level 2 example for KQ1. Effectiveness of screening

Level 1 – Title and abstract screening

1. Does this record discuss breast cancer screening?
   - ❍ **Yes/possibly**
   - ❍ No
   - ❍ **Unclear/no abstract**

*Reasons for selecting 'no':
1) Does not focus on breast cancer screening in a population screening context (**If >20% of the population are high risk- then exclude. For now, include all studies which assess dense breasts populations**). *High Risk*: women with pre-existing or personal history of breast cancer, family history (in a first degree relative) of breast or ovarian cancer or other personal risk factors, such as abnormal breast pathology or BRCA1/BRCA2 genetic mutations, previously receivied radiation treatment to the chest (such as Hodgkin's) for cancer. 2) Animal/in vivo studies 2) It focuses on breast cancer screening but it is clearly obvious that it is one of the following: **CPG, SRs, Narrative literature review, commentary (without primary data), editorials (without primary data), protocol**

*Those answered yes/unclear will be passed through to full-text screening.

Level 2 – Full-text screening

1. Is the full-text available?
   - ☐ **Yes**
   - ☐ No
   - ☐ abstract only

2. Is the article published in English or French?
   - ☐ **Yes**
   - ☐ No

3. Is the article any of the following study designs?

   RCTs (including cluster), or novel/extended analysis of RCT data.
   Non-RCTs
   Comparative cohort studies (including adminstrative database studies/registries)
   Ecological studies
   Example of studies to exclude:
   case-control,
   cross-sectional studies,
   case-series,
   controlled before-after,
   diagnostic test accuracy studies

modelling studies.

Also exclude narrative reviews, systematic reviews/meta-analysiss, commentaries & Editorals (without primary data), protocols, papers on study design

- **Yes**
- No
- Diagnostic Type Accuracy Study

4. Is the article focused on breast cancer screening?
   - **Yes**
   - No

5. Is it the population of interest?
   - No- women <40 years (exclusively)
   - No- women ≥ 40 years who are high –risk (based on family history and other personal risk factors- genetic mutations, abnormal pathology, previous history of cancer, etc).
   - **Yes- women ≥ 40 years who are 'not at high risk'- i.e., average risk**
   - **Yes- women ≥ 40 years who have dense breasts**
   - **Unclear- mixed aged population who are 'not at high risk' or who have dense breasts**
   - No- mixed aged population who are at 'high risk'

6. Does this article include a relevant intervention?
   Mammography (film, digital, tomosynthesis) with or without CBE/BSE
   MRI with or without CBE/BSE
   Ultrasound with or without CBE/BSE
   CBE
   BSE
   - **Yes**
   - No

7. Is the comparator: "no screening", "usual care", or if mammography (film vs. digitial vs tomosynthesis)?
   - **Yes**
   - No

Typically, these questions are nested. If an answer allows us to proceed in the inclusion criteria, the next question will appear. Those bolded would be those that would pass through to the following question. If question 7 is 'Yes', this article would be passed through to a post-hoc evaluation, ensuring it has outcomes of interest.

Appendix E. Draft Items for Data Extraction for KQ1 (existing systematic reviews)

Publication details: year of publication, language, publication status

Search details: Databases searched and years searched

Selection criteria: PICOTs of the review

Study Characteristics: Number of included studies, type of study design, population, sample sizes, quality of included studies (must align with the CTFPHC PICOTs)

Results of the systematic review: Summarize qualitatively body of evidence

Results of the meta-analysis: Pooled estimate, heterogeneity tests

Strengths and limitations of the review

AMSTAR quality

Appendix F. Draft Items for Data Extraction for KQ1 (primary studies)

Publication details: year of publication, language, publication status
Characteristics of study: study design, methods, country, setting, sample size, number of centres [if applicable], duration of follow-up, source of funding

Characteristics of population: age, gender, ethnicity, other risk factors, information regarding respondent bias/representativeness of the included population

Details about the exposure/intervention: type of screening test performed, frequency/interval of screening

Details about comparator: type of screening test performed (or no screening), frequency/interval of screening (if applicable)

Outcomes of interest: definitions, measurement methods, data, adjusted and unadjusted effect estimates

Confounding factors that were taken into consideration

Risk of bias items

Appendix G. The AMSTAR Tool

1. Was an 'a priori' design provided?
The research question and inclusion criteria should be established before the conduct of the review.

Note: Need to refer to a protocol, ethics approval, or pre-determined/a priori published research objectives to score a "yes."

☐ Yes
☐ No
☐ Can't answer
☐ Not applicable

2. Was there duplicate study selection and data extraction?
There should be at least two independent data extractors and a consensus procedure for disagreements should be in place.

Note: 2 people do study selection, 2 people do data extraction, consensus process or one person checks the other's work.

☐ Yes
☐ No
☐ Can't answer
☐ Not applicable

3. Was a comprehensive literature search performed?
At least two electronic sources should be searched. The report must include years and databases used (e.g., Central, EMBASE, and MEDLINE). Key words and/or MESH terms must be stated and where feasible the search strategy should be provided. All searches should be supplemented by consulting current contents, reviews, textbooks, specialized registers, or experts in the particular field of study, and by reviewing the references in the studies found.

Note: If at least 2 sources + one supplementary strategy used, select "yes" (Cochrane register/Central counts as 2 sources; a grey literature search counts as supplementary).

☐ Yes
☐ No
☐ Can't answer
☐ Not applicable

4. Was the status of publication (i.e. grey literature) used as an inclusion criterion?
The authors should state that they searched for reports regardless of their publication type. The authors should state whether or not they excluded any reports (from the systematic review), based on

☐ Yes
☐ No
☐ Can't answer
☐ Not

their publication status, language etc. *Note: If review indicates that there was a search for "grey literature" or "unpublished literature," indicate "yes." SINGLE database, dissertations, conference proceedings, and trial registries are all considered grey for this purpose. If searching a source that contains both grey and non-grey, must specify that they were searching for grey/unpublished lit.*	applicable
5. Was a list of studies (included and excluded) provided? A list of included and excluded studies should be provided. *Note: Acceptable if the excluded studies are referenced. If there is an electronic link to the list but the link is dead, select "no."*	☐ Yes No Can't answer ☐ Not applicable
6. Were the characteristics of the included studies provided? In an aggregated form such as a table, data from the original studies should be provided on the participants, interventions and outcomes. The ranges of characteristics in all the studies analyzed e.g., age, race, sex, relevant socioeconomic data, disease status, duration, severity, or other diseases should be reported. *Note: Acceptable if not in table format as long as they are described as above.*	☐ Yes No Can't answer ☐ Not applicable
7. Was the scientific quality of the included studies assessed and documented? 'A priori' methods of assessment should be provided (e.g., for effectiveness studies if the author(s) chose to include only randomized, double-blind, placebo controlled studies, or allocation concealment as inclusion criteria); for other types of studies alternative items will be relevant. *Note: Can include use of a quality scoring tool or checklist, e.g., Jadad scale, risk of bias, sensitivity analysis, etc., or a description of quality items, with*	☐ Yes No Can't answer ☐ Not applicable

some kind of result for EACH study ("low" or "high" is fine, as long as it is clear which studies scored "low" and which scored "high"; a summary score/range for all studies is not acceptable).

---

8. Was the scientific quality of the included studies used appropriately in formulating conclusions?

The results of the methodological rigor and scientific quality should be considered in the analysis and the conclusions of the review, and explicitly stated in formulating recommendations.

Note: Might say something such as "the results should be interpreted with caution due to poor quality of included studies." Cannot score "yes" for this question if scored "no" for question 7.

☐ Yes
No
Can't answer
☐ Not applicable

---

9. Were the methods used to combine the findings of studies appropriate?

For the pooled results, a test should be done to ensure the studies were combinable, to assess their homogeneity (i.e., Chi-squared test for homogeneity, I2). If heterogeneity exists a random effects model should be used and/or the clinical appropriateness of combining should be taken into consideration (i.e., is it sensible to combine?).

Note: Indicate "yes" if they mention or describe heterogeneity, i.e., if they explain that they cannot pool because of heterogeneity/variability between interventions.

☐ Yes
No
Can't answer
☐ Not applicable

---

10. Was the likelihood of publication bias assessed?

An assessment of publication bias should include a combination of graphical aids (e.g., funnel plot, other available tests) and/or statistical tests (e.g., Egger regression test, Hedges-Olken).

Note: If no test values or funnel plot included, score "no". Score "yes" if mentions that publication bias could not be assessed because there were fewer than 10 included studies.

☐ Yes
No
Can't answer
☐ Not applicable

---

11. Was the conflict of interest included?
Potential sources of support should be clearly acknowledged in both the systematic review and the included studies.

Note: To get a "yes," must indicate source of funding or support for the systematic review AND for each of the included studies.

☐ Yes
☐ No
☐ Can't answer
☐ Not applicable

Shea *et al. BMC Medical Research Methodology* 2007 7:10 doi:10.1186/1471-2288-7-10

Appendix H. Draft Cochrane Risk of Bias Piloting Form

1. **Selection bias domain**: Random sequence generation
    - ❍ Low risk
    - ❍ Unclear risk
    - ❍ High risk

Support for judgement:

2. **Selection bias domain**: Allocation concealment
    - ❍ Low risk
    - ❍ Unclear risk
    - ❍ High risk

   Support for judgement:

3. **Performance bias domain**: Blinding of participants and personnel (for each outcome)
    - ❍ Low risk
    - ❍ Unclear risk
    - ❍ High risk

   Support for judgement:

4. **Detection bias domain**: Blinding of outcome assessment (for each outcome)
    - ❍ Low risk
    - ❍ Unclear risk
    - ❍ High risk

   Support for judgement:

5. **Attrition bias domain**: Incomplete outcome data (for each outcome)
    - ❍ Low risk
    - ❍ Unclear risk
    - ❍ High risk

   Support for judgement:

6. **Reporting bias domain**: Selective reporting
    - ❍ Low risk
    - ❍ Unclear risk
    - ❍ High risk

   Support for judgement:

7. **Other sources of bias**
    - ❍ Low risk
    - ❍ Unclear risk
    - ❍ High risk

   Support for judgement:

Appendix I. Risk of Bias for Observational Studies: Newcastle Ottawa Scale (NOS)-Cohort Studies

Note: A study can be awarded a maximum of one star for each numbered item within the Selection and Outcome categories. A maximum of two stars can be given for Comparability

Selection

1) Representativeness of the exposed cohort

a) truly representative of the average_____(describe) in the community ✸
b) somewhat representative of the average_____in the community ✸
c) selected group of users eg nurses, volunteers
d) no description of the derivation of the cohort

2) Selection of the non-exposed cohort
a) drawn from the same community as the exposed cohort ✸
b) drawn from a different source
c) no description of the derivation of the non-exposed cohort

3) Ascertainment of exposure
a) secure record (eg surgical records) ✸
b) structured interview ✸
c) written self-report
d) no description

4) Demonstration that outcome of interest was not present at start of study
a) yes ✸
b) no

Comparability

1) Comparability of cohorts on the basis of the design or analysis
a) study controls for_____(select the most important factor) ✸
b) study controls for any additional factor ✸ (This criteria could be modified to indicate specific control for a second important factor.)

Outcome

1) Assessment of outcome
a) independent blind assessment ✸
b) record linkage ✸
c) self report
d) no description

2) Was follow-up long enough for outcomes to occur
a) yes (select an adequate follow up period for outcome of interest) ✸
b) no

3) Adequacy of follow up of cohorts
a) complete follow up - all subjects accounted for ✸
b) subjects lost to follow up unlikely to introduce bias - small number lost - > _____ % (select an adequate %) follow up, or description provided of those lost) ✸
c) follow up rate <_____% (select an adequate %) and no description of those lost
d) no statement

Appendix J. Risk of Bias for Observational Studies: Newcastle Ottawa Scale (NOS)-Case-Control Studies

Note: A study can be awarded a maximum of one star for each numbered item within the Selection and Exposure categories. A maximum of two stars can be given for Comparability.

Selection

1) Is the case definition adequate?
   a) yes, with independent validation ✸
   b) yes, eg record linkage or based on self-reports
   c) no description

2) Representativeness of the cases
   a) consecutive or obviously representative series of cases ✸
   b) potential for selection biases or not stated

3) Selection of Controls
   a) community controls ✸
   b) hospital controls
   c) no description

4) Definition of Controls
   a) no history of disease (endpoint) ✸
   b) no description of source

Comparability

1) Comparability of cases and controls on the basis of the design or analysis
   a) study controls for_____(Select the most important factor.) ✸
   b) study controls for any additional factor • (This criteria could be modified to indicate specific control for a second important factor.)

Exposure

1) Ascertainment of exposure
   a) secure record (eg surgical records) ✸
   b) structured interview where blind to case/control status ✸
   c) interview not blinded to case/control status
   d) written self-report or medical record only
   e) no description

2) Same method of ascertainment for cases and controls
   a) yes ✸
   b) no

3) Non-Response rate
   a) same rate for both groups ✸
   b) non respondents described
   c) rate different and no designation

Appendix K. Draft Screening Forms for KQ2

1. Is the article potentially published in English or French?
   - **Yes**
   - No

2. Is this article a report of a research study?
   - **Yes**
   - No
   - Unclear

3. Is the population specific to (or containing data from) women ≥ 40 years old?
   - **Yes**
   - No
   - Unclear

4. Does the study assess women who are at not at high-risk for breast cancer, or at least not only women with pre-existing cancer or at high-risk for cancer? High-risk: on the basis of extensive family history of breast or ovarian cancer or other personal risk factors, such as abnormal pathology or deleterious genetic mutations.
   - **Yes**
   - No (i.e. specific to high-risk)
   - Unclear

5. Does the study relate to one of our screening modalities of interest?
   - **Yes (Mammography (film or digital), MRI, ultrasound, CBE, CBE & mammography, BSE)**
   - No
   - Unclear

6. Does this article include an assessment/appraisal by women of at least one benefit and one harm from breast cancer screening? Benefits and harms do not need to be specific to our critical outcomes, e.g. may be related to harms such as anxiety, radiation exposure etc.
   - **Yes**
   - No
   - Unclear

For title/abstract screening if one or more answers are no the article will be excluded; in all other cases it will pass to full-text screening. For full text selection, any no answers will exclude citation, and unclears will lead to consensus or author contact for clarification. Typically, these questions are nested. If an answer allows us to proceed in the inclusion criteria, the next question will appear.

Appendix L. Draft Items for Data Extraction for KQ2

Publication details: year of publication, language, publication status

Characteristics of study: study design, data collection and analysis methods, country, setting, sample size, response rate (if applicable), number of centres [if applicable], source of funding

Characteristics of population: age, gender, ethnicity, education, income, other risk factors, information regarding respondent bias/representativeness of the included population

Details about the study methods: attributes of screening/outcomes, background information/definitions provided, interview questions, analytical approach

Outcomes of interest: definitions, measurement methods, data/findings including key quotes if applicable

Confounding factors that were taken into consideration or within analyses and findings

Appendix M. CEBM Tools for Surveys and Qualitative Studies

https://www.cebma.org/resources-and-tools/what-is-critical-appraisal/

Surveys

1. Did the study address a clearly focused question / issue?
2. Is the research method (study design) appropriate for answering the research question?
3. Is the method of selection of the subjects clearly described?
4. Could the way the sample was obtained introduce (selection) bias?
5. Was the sample of subjects representative with regard to the population to which the findings will be referred?
6. Was the sample size based on pre-study considerations of statistical power?
7. Was a satisfactory response rate achieved?
8. Are the measurements (questionnaires) likely to be valid and reliable?
9. Was the statistical significance assessed?
10. Are confidence intervals given for the main results?
11. Could there be confounding factors that haven't been accounted for?
12. Can the results be applied to your organization?

Qualitative studies

1. Did the study address a clearly focused question / issue?
2. Is the research method (study design) appropriate for answering the research question?
3. Was the context clearly described?
4. How was the fieldwork undertaken? Was it described in detail? Are the methods for collecting data clearly described?
5. Could the evidence (fieldwork notes, interview transcripts, recordings, documentary analysis, etc.) be inspected independently by others?
6. Are the procedures for data analysis reliable and theoretically justified? Are quality control measures used?
7. Was the analysis repeated by more than one researcher to ensure reliability?
8. Are the results credible, and if so, are they relevant for practice?
9. Are the conclusions drawn justified by the results?
10. Are the findings of the study transferable to other settings?

Adapted from Crombie, *The Pocket Guide to Critical Appraisal*; the critical appraisal approach used by the Oxford Centre for Evidence Medicine, checklists of the Dutch Cochrane Centre, BMJ editor's checklists and the checklists of the EPPI Centre.

Appendix N. EPOC Risk of Bias Tool for Controlled Before-After Studies

Was the allocation sequence adequately generated?

Score “Low risk” if a random component in the sequence generation process is described (eg Referring to a random number table). Score “High risk” when a nonrandom method is used (eg performed by date of admission). NRCTs and CBA studies should be scored “High risk”. Score “Unclear risk” if not specified in the paper.

Was the allocation adequately concealed?

Score “Low risk” if the unit of allocation was by institution, team or professional and allocation was performed on all units at the start of the study; or if the unit of allocation was by patient or episode of care and there was some form of centralised randomisation scheme, an on-site computer system or sealed opaque envelopes were used. CBA studies should be scored “High risk”. Score “Unclear risk” if not specified in the paper.

Were baseline outcome measurements similar?

Score “Low risk” if performance or patient outcomes were measured prior to the intervention, and no important differences were present across study groups. In RCTs, score “Low risk” if imbalanced but appropriate adjusted analysis was performed (e.g. Analysis of covariance). Score “High risk” if important differences were present and not adjusted for in analysis. If RCTs have no baseline measure of outcome, score “Unclear risk”.

Were baseline characteristics similar?

Score “Low risk” if baseline characteristics of the study and control providers are reported and similar. Score “Unclear risk” if it is not clear in the paper (e.g. characteristics are mentioned in text but no data were presented). Score “High risk” if there is no report of characteristics in text or tables or if there are differences between control and intervention providers. Note that in some cases imbalance in patient characteristics may be due to recruitment bias whereby the provider was responsible for recruiting patients into the trial.

Were incomplete outcome data adequately addressed?

Score “Low risk” if missing outcome measures were unlikely to bias the results (e.g. the proportion of missing data was similar in the intervention and control groups or the proportion of missing data was less than the effect size i.e. unlikely to overturn the study result). Score “High risk” if missing outcome data was likely to bias the results. Score “Unclear risk” if not specified in the paper (Do not assume 100% follow up unless stated explicitly).

Was knowledge of the allocated interventions adequately prevented during the study?

Score “Low risk” if the authors state explicitly that the primary outcome variables were assessed blindly, or the outcomes are objective, e.g. length of hospital stay. Primary outcomes are those variables that correspond to the primary hypothesis or question as defined by the authors. Score “High risk” if the outcomes were not assessed blindly. Score “Unclear risk” if not specified in the paper.

Was the study adequately protected against contamination?

Score “Low risk” if allocation was by community, institution or practice and it is unlikely that the control group received the intervention. Score “High risk” if it is likely that the control group received the intervention (e.g. if patients rather than professionals were randomised). Score “Unclear risk” if professionals were allocated within a clinic or practice and it is possible that communication between

intervention and control professionals could have occurred (e.g. physicians within practices were allocated to intervention or control)

Was the study free from selective outcome reporting?
Score “Low risk” if there is no evidence that outcomes were selectively reported (e.g. all relevant outcomes in the methods section are reported in the results section). Score “High risk” if some important outcomes are subsequently omitted from the results. Score “Unclear risk” if not specified in the paper.

Was the study free from other risks of bias?
Score “Low risk” if there is no evidence of other risk of biases.

Appendix O. PRISMA-P

(Preferred Reporting Items for Systematic review and Meta-Analysis Protocols) 2015 checklist: recommended items to address in a systematic review protocol*

Section and topic	Item No	Checklist item	Reported on page #
ADMINISTRATIVE INFORMATION			
Title:			
Identification	1a	Identify the report as a protocol of a systematic review	21
Update	1b	If the protocol is for an update of a previous systematic review, identify as such	3
Registration	2	If registered, provide the name of the registry (such as PROSPERO) and registration number	Register after CTF approval
Authors:			
Contact	3a	Provide name, institutional affiliation, e-mail address of all protocol authors; provide physical mailing address of corresponding author	1
Contributions	3b	Describe contributions of protocol authors and identify the guarantor of the review	2
Amendments	4	If the protocol represents an amendment of a previously completed or published protocol, identify as such and list changes; otherwise, state plan for documenting important protocol amendments	-
Support:			
Sources	5a	Indicate sources of financial or other support for the review	2
Sponsor	5b	Provide name for the review funder and/or sponsor	2
Role of sponsor or funder	5c	Describe roles of funder(s), sponsor(s), and/or institution(s), if any, in developing the protocol	2
INTRODUCTION			
Rationale	6	Describe the rationale for the review in the context of what is already known	3
Objectives	7	Provide an explicit statement of the question(s) the review will address with reference to participants, interventions, comparators, and outcomes (PICO)	12-13, 19-20
METHODS			
Eligibility criteria	8	Specify the study characteristics (such as PICO, study design, setting, time frame) and report characteristics (such as years considered, language, publication status) to be used as criteria for eligibility for the review	12-13, 19-20

Information sources	9	Describe all intended information sources (such as electronic databases, contact with study authors, trial registers or other grey literature sources) with planned dates of coverage	10, 21
Search strategy	10	Present draft of search strategy to be used for at least one electronic database, including planned limits, such that it could be repeated	33-46
Study records:			
Data management	11a	Describe the mechanism(s) that will be used to manage records and data throughout the review	10
Selection process	11b	State the process that will be used for selecting studies (such as two independent reviewers) through each phase of the review (that is, screening, eligibility and inclusion in meta-analysis)	10-11
Data collection process	11c	Describe planned method of extracting data from reports (such as piloting forms, done independently, in duplicate), any processes for obtaining and confirming data from investigators	11, 14-15, 21
Data items	12	List and define all variables for which data will be sought (such as PICO items, funding sources), any pre-planned data assumptions and simplifications	14-15, 18-20
Outcomes and prioritization	13	List and define all outcomes for which data will be sought, including prioritization of main and additional outcomes, with rationale	13, 20
Risk of bias in individual studies	14	Describe anticipated methods for assessing risk of bias of individual studies, including whether this will be done at the outcome or study level, or both; state how this information will be used in data synthesis	11, 15, 21
Data synthesis	15a	Describe criteria under which study data will be quantitatively synthesised	15-18
	15b	If data are appropriate for quantitative synthesis, describe planned summary measures, methods of handling data and methods of combining data from studies, including any planned exploration of consistency (such as I^2, Kendall's τ)	15-18
	15c	Describe any proposed additional analyses (such as sensitivity or subgroup analyses, meta-regression)	15-18
	15d	If quantitative synthesis is not appropriate, describe the type of summary planned	15-18
Meta-bias(es)	16	Specify any planned assessment of meta-bias(es) (such as publication bias across studies, selective reporting within studies)	15-18
Confidence in cumulative evidence	17	Describe how the strength of the body of evidence will be assessed (such as GRADE)	15-18

*** It is strongly recommended that this checklist be read in conjunction with the PRISMA-P Explanation and Elaboration (cite when available) for important clarification on the items. Amendments to a review protocol should be tracked and dated. The copyright for PRISMA-P (including checklist) is held by the PRISMA-P Group and is distributed under a Creative Commons Attribution Licence 4.0.**

From: Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart L, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ. 2015 Jan 2;349(jan02 1):g7647.