Mammography screening for breast cancer: A replication of a systematic review of observational studies to inform an update of the Canadian Task Force on Preventive Health Care guidelines

Authors: Institute of Health Economics, Edmonton, AB, Canada.

Acknowledgement: The Institute of Health Economics would like to acknowledge the Ottawa Evidence Review Synthesis Centre (ERSC) for developing the protocol on which this systematic review was based. We are also grateful to the Ottawa ERSC for developing the literature searches and to the University of Manitoba for support in conducting the meta-analyses.

Supplementary Information

- Supplementary File 1: Search Strategy for Searches of Bibliographic Databases
- Supplementary File 2: PRISMA Diagram
- Supplementary File 3: Data Tables
- Supplementary File 4: Summary GRADE Tables

Abstract

Background: In 2018 the Canadian Task Force on Preventive Health Care recommended screening mammography every two to three years for women aged 50–74 years. More recent international guidelines recommend that women should commence breast cancer (BC) screening at an even younger age. Evidence from randomized controlled trials informed the previous Task Force guidelines, but no new trials have since been conducted.

Objective: To conduct a systematic review of observational studies comparing the effectiveness of screening mammography with no screening in women aged 40 years or older who are at an average or moderately increased risk of BC.

Methods: MEDLINE, EMBASE, and the Cochrane Central Register of Controlled Trials were searched in July 2023 to identify relevant non-randomized comparative studies published from 2014 onwards. Separate meta-analyses were conducted for each outcome by study design, accrual type, population, and age subgroup (40–44, 45–49, 50–59, 60–69, 70–74, ≥75 years), where possible. Grading of Recommendations Assessment, Development and Evaluation (GRADE) was used to assess the certainty of evidence.

Results

Cohort and population-based studies (very low certainty evidence) indicated that screening reduced BC mortality in women older than 49 years of age (risk ratio range 0.10 to 0.73). Results were more equivocal for older and younger women, although there was some indication of benefit from screening for women aged 45–49 years. All-cause mortality was lower in screened women, but this effect varied inversely with increasing age. Women who undergo screening generally had less advanced cancer at diagnosis and tended to receive less aggressive treatment. Although the degree of effect differed between studies of general screening populations and BC only populations, the direction of effect was generally similar. None of the studies reported harms associated with screening, and the length of follow up in the few prospective studies was not adequate to calculate overdiagnosis.

Conclusion

Evidence supports the use of screening mammography in women aged 50–69 years who have an average or moderately increased risk for BC, and it may provide benefit for women aged 45-49 years. However, any significant reductions observed in BC mortality in the populations studied may be attenuated in Canadian populations due to differences in baseline risk. Questions about the effectiveness of mammography with respect to screening interval, modality, risk factors, and specific age subgroups, and in the Canadian population in general, could not be answered by the currently available non-randomized comparative evidence.

Introduction

Breast cancer (BC) is the most commonly diagnosed non-skin cancer and the second leading cause of death from cancer among Canadian women.[1] It was estimated that 29,400 women will be diagnosed with BC in 2023, accounting for 26% of all cancer diagnoses in women, and that 5,400 will die of the disease (13% of cancer deaths among women).[1] BC outcomes have been steadily improving over recent decades, which has been attributed to the introduction of organized mammography screening programs and better treatment options.[2] Ideally, screening can detect lesions at an earlier stage before they become clinically apparent. While regular breast screening with mammography can reduce BC mortality,[3] population-based screening is not without risk and places additional demands on the healthcare system.[2,4] False positive results require further testing and workup, lead to unnecessary biopsies and potentially harmful invasive treatments, and can cause psychological distress, while overdiagnosis can detect lesions that may never progress or cause harm.

The burden of BC and BC screening programs underscores the importance of developing and updating evidence-based clinical practice guidelines to inform public health screening initiatives. Worldwide, current guidelines for screening women at average risk of BC vary with respect to the ages to start and stop screening, the screening interval, and the emphasis placed on shared decision-making.[5] The 2018 guidelines of the Canadian Task Force on Preventive Health Care (hereafter, "Task Force") recommend mammography screening in women aged 50–74 years every two to three years; screening is not recommended in women aged 40–49 years.[6] These recommendations are conditional on the value a woman places on the relative potential benefits and risks of screening.[6] Based on guidelines developed by the Alberta Breast Cancer Screening Clinical Practice Guidelines Committee in 2022, Alberta recently became the first province in Canada to lower the age of screening to include average-risk women aged 45–49 years.[7] The European Commission Initiative on Breast Cancer similarly recommends biennial mammography screening of asymptomatic women at average risk aged 45–74 years.[8] The U.S. Preventive Services Task Force released draft recommendations in May 2023 that recommend biennial BC screening for all women aged 40–74 years.

This replication of a systematic review was one component of a broader evidence review conducted to inform an update of the 2018 Task Force clinical practice guideline for BC screening, along with a parallel systematic review conducted using the same protocol. As such, this review was one source of evidence that the Task Force considered when developing their 2024 screening recommendations. Since the evidence report that informed the previous Task Force guidelines focused on randomized controlled trials (RCTs)[9] and no new RCTs have since been conducted, this review focused only on evidence from observational studies.

Methods

This review followed the methods outlined in a protocol developed by the Ottawa Evidence Review Synthesis Centre (ERSC), which is available at https://osf.io/he9zg. This original protocol followed guidance of the Cochrane Handbook,[10] the Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group,[11] and the Task Force Methods Manual.[12] The Task Force helped develop and approved the key questions and study selection criteria but was not involved in the conduct of this systematic review.

Search strategy

An experienced information specialist, together with the Ottawa ERSC, developed search strategies for three bibliographic databases (Ovid MEDLINE, EMBASE, and the Cochrane Central Register of Controlled Trials) using a combination of controlled vocabulary terms and keywords. A second information specialist

peer-reviewed the MEDLINE search strategy using the Peer Review of Electronic Search Strategies (PRESS) checklist.[13] The searches were conducted in September 2023 and were limited to English and French language articles published from 2014 onwards. The search strategies for the three databases are available in Supplementary File 1. Grey literature searches and scanning of the reference lists of relevant studies were not conducted.

Study selection

Eligible studies compared the effectiveness of a mammography-based screening strategy for BC with no screening in women aged 40 years or older who were at an average or moderately increased risk of BC. The study selection criteria for this systematic review were developed by the Ottawa ERSC and are available in the protocol cited above. The following criteria were modified to clarify the selection decision, in consultation with the Task Force and the Public Health Agency of Canada (as the organization who established and provides support to the Task Force):

- Eligible date ranges were defined as a median screening initiation date of 2000—that is, 50% of the date range for screening was in 2000 or later to increase applicability to current technologies and cancer treatments.
- Studies that do not report the breast cancer risk level of participants or reported a combination of risk levels were included. Studies in which less than 15% of participants were under age 40 and those that provided separate data for participants 40 years of age or older were included.
- Studies that used a combination of film and digital mammography or that had only one round of screening (or did not report the number of screening rounds) were included; however, studies using film mammography and those with less than two rounds were downgraded in GRADE.

Studies that examined a general screening population and those that enrolled only participants who had BC at study initiation (hereafter, "BC only" population) were considered eligible for this systematic review. The latter compared women with BC who had a history of screening versus those with no history of screening. Although studies using a general screening population are more relevant to the Canadian screening context, BC only studies may be at least partially representative of real-world clinical settings and were included for comprehensiveness.

We pilot tested the study selection process using a random sample of 50 records and 25 full-text articles until at least 95% agreement between reviewers was reached. Two researchers independently screened the titles and abstracts and identified relevant full-text articles using DistillerSR literature review software.[14] For the title and abstract screening, the DistillerSR re-rank feature was used to prioritize relevant records, and reviewers manually screened abstracts until the highest remaining re-rank score was 0.1. The remaining unreviewed records with a relevance score under 0.1 were screened by DistillerSR Artificial Intelligence (DistillerSR Inc.). Full-text articles of potentially relevant records were retrieved and reviewed by two reviewers using a standardized form. Disagreements were resolved by consensus.

Data extraction and quality assessment

One researcher extracted data into a pilot-tested form in Excel. The data were verified by a second researcher. Descriptive data (including year of publication, country, screening and follow-up dates, case accrual, age and number of screening participants, risk level, and mammography information [type, screening interval, and number of rounds]) and relevant outcome data pertaining to the pre-specified age groups (40–44 years, 45–49 years, 50–59 years, 60–69 years, 70–74 years, ≥75 years) for the longest available follow-up duration using any measure reported were extracted for each study. Disagreements

that could not be resolved by consensus were referred to a third researcher. The characteristics of included studies and outcome data were presented narratively and in tables.

One researcher assessed the quality of the included studies using the Newcastle-Ottawa quality assessment scales for cohort and case-control studies[15] and the NICE Quality Appraisal Checklist for quantitative studies reporting correlations and associations,[16] which was applied to the population-based studies. A second researcher verified the assessment results. Disagreements were resolved by consensus.

Data analysis and grading the evidence

An experienced statistician used R (The R Foundation) to pool relevant studies for each outcome using a random effects model. Within an outcome, separate meta-analyses were conducted for different studies designs (cohort, case control, population-based), accrual types (short and long case), and study populations (general screening and BC only). Forest plots were created for all studies and subgroups by age group and screening interval, as specified in the protocol, whenever data were available. Results of the analyses were used to calculate the absolute effect per 1,000 individuals. All relevant studies were included in "all age" comparisons, while only age groups of interest (as specified above) were reported as subgroups. Weighted risk ratios (RR) were used to pool data for dichotomous outcomes, and absolute effects were calculated for all outcomes. Data were pooled regardless of degree of statistical heterogeneity, but studies with high heterogeneity were later downgraded for consistency in GRADE. Subgroup analyses by age groups of interest and screening interval were conducted to explore possible causes of heterogeneity, where possible.

The GRADE framework was used to assess the quality of the evidence across the five GRADE domains (risk of bias, inconsistency, indirectness, imprecision, and other considerations) for each outcome.[11] The evidence was graded by one researcher and verified by a second researcher using a predefined list of considerations for each GRADE category that was piloted and refined by two researchers. Disagreements in GRADE assessments that could not be resolved by consensus were referred to a third researcher. Summary tables with the GRADE assessments and explanations for all judgements were created for every comparison within each outcome.

Results

Study selection and characteristics

Of the 9,623 abstracts identified by electronic searches, 40 studies met the inclusion criteria for this review: 31 cohort (28 retrospective), two case-control, and seven population-based studies (Supplementary File 2). A list of excluded studies with reasons for exclusion is available on request. The characteristics of the included studies and a list of outcomes by study design are available in the supplement (Supplementary File 3, Tables S-1 and S-2). The median year of publication was 2018 (interquartile range 2017 to 2020), and most studies used a long-case accrual (n=27). Seven studies assessed a general screening population,[17-23] while the remainder compared people with breast cancer who had a history of screening with those who had not been screened. Among the 29 studies that reported the screening interval, the typical length was two years (25 studies). In four other studies the interval was either annual[20] or varied from 1 to 3 years.[24-26]

The most frequently reported outcome was BC stage at diagnosis (70% of the studies), followed by BC specific mortality (65% of studies). Less than one-third of studies reported other outcomes, and only one study reported on quality of life. None of the included studies reported on BC morbidity, overdiagnosis, false-positive rates, or life years gained. Some studies reported incidence rates or effect measures rather than proportions, which meant that not all studies could be included in the meta-analysis. Mean study

follow-up length ranged from 3.1 years to 22 years; 12 studies did not report this parameter. The sample sizes varied widely across the studies. Cohort studies examining a general screening population had the highest median sample size (median 263,312; range 54,635 to over 13 million), followed by population-based studies (median 55,216; range 4,903 to 149,833), case-control studies (median 10,783; range 681 to 20,885), and BC only cohort studies (median 1,648; range 230 to 549,091). Control groups encompassed contemporaneous cohorts not invited to screening, regions where screening was not yet available, and historical cohorts prior to the widespread advent of screening. In one study,[27] the comparison group was a different age cohort that was not routinely invited to screening.

The included studies varied in the age ranges of participants included in the study populations. Overall, 11 studies consider only women between the ages of 50 to 69 years. Other studies included women aged 40–49 years (n=18) and/or women aged 70 years and older or had no maximum age limit (n=18).

Evidence quality

Across all outcomes and comparisons, the certainty of the evidence was rated as "very low" due to the observational nature of the studies and other risks to internal validity, including retrospective data collection and follow-up that was either inadequate (<80% of included participants) or insufficient in length (minimum of 5-year follow up for mortality outcomes) (Supplementary File 4). In many cases, the effect estimates had wide confidence intervals (CI) that failed to exclude important benefit or harm, or there was high variability in point estimates or direction of effect across the studies. Most of the studies also had poor external validity because they did not report the risk level of participants (90% of studies); selectively included participants with BC (83%); or did not report the screening modality (73% of studies), number of screening rounds (80%), or screening interval (28%). All the case-control studies, most of the cohort studies (61%), and some of the population-based studies (29%) controlled for the main prognostic of age and other factors such as tumor size and nodal status. A further 10% of cohort studies and none of the population-based studies controlled for age only or another prognostic factor(s) (16% of cohort studies and 43% of population-based studies). Many of the subgroup analyses often included only a single study because outcome data were reported variously across studies, with some not providing raw data or only providing results for all women without age stratification.

Study findings

Where possible, results are reported for the following age groups: all ages, 50 to 74 years (the recommended screening ages in Canada), and the age stratifications listed above.

BC mortality

BC only population

Cohort studies

The pooled RR for three cohort studies with short case accrual was 0.37 (95% CI 0.24, 0.57) for women across all age groups (107 to 60 fewer deaths per 1,000 women) (Figure 1).[25, 28, 29] Single cohort studies reporting various age subgroups (50–65[29], 40–49[25], 50–59[25], 60–69[25] and ≥70 years[28]) had a similar result (Table 2). For the six long case accrual cohort studies, the pooled RR was 0.28 (95% CI 0.18, 0.44; 12 to 8 fewer deaths per 1,000 women) for all ages (Figure 2)[30-35] and 0.26 (95% CI 0.20, 0.34) for women aged 50–69 years (104 to 85 fewer deaths per 1,000 women)[30, 32, 34] (Table 2). The single study that did not provide raw data reported a hazard ratio of 2.81 (95% CI 1.57, 5.04) in favour of screening.[36] One small study found no discernible benefit for screening in the 40–49-year age subgroup,[35] whereas another study[33] reported a lower risk of BC mortality among women older than 74 years who had undergone screening (Table 2). Across the short and long case cohort studies, the median sample size was 1,834 (range 230 to 549,091) and the longest follow up ranged from 5 to 25 years for this outcome.

Other study designs

One of two short case accrual population-based studies found a lower risk of BC specific mortality among women aged 40-49 years once screening was introduced (incidence-based mortality rate ratio 0.92; 95% CI 0.85, 0.99).[37] A second study of the same age group found that the benefit was confined to the 45–49-year age group rather than to women in their early 40s (Supplementary File 3, Table S-3).[38] In one study of women aged 50-69 years, a lower risk of BC specific mortality was only observed in the 50-64-year age grouping; there was no discernible difference in risk among women aged 65–69 years.[38] Among the various incidence measures reported in the long case accrual studies, the rates of BC specific mortality were lower among screened participants compared with those who did not undergo screening for all age groupings analyzed (all ages, 50–69 years, and ≥70 years) (Supplementary File 3, Table S-3).[39, 40] The sample size for the three studies that reported it ranged from 4,903 to 71,990, while the longest follow up ranged from 8 to 11.2 years for this outcome.

Two case-control studies reported on breast-specific mortality in a total of 21,566 women aged 50-69 and 50-84 years (one short case[41] and one long case accrual[42]). Neither study found any discernible difference between the screening and non-screening groups after 15 and 19.9 years of follow up, respectively.

General screening population (cohort studies)

The risk of BC specific mortality was lower for screened women aged 50–65 years than for those who were not screened in one study (RR 0.73, 95% CI 0.65, 0.81; 2 to 1 fewer deaths per 1,000 women)[19] (Supplementary File 3, Table S-4). However, there was no discernible difference between the two groups in women aged 70–74 years, 75–84 years, or older.[23] Combined results from two long case accrual cohort studies showed a reduced risk of mortality (RR 0.40, 95% CI 0.33 to 0.47; 0 fewer deaths per 1,000 women) in women aged 40–79 years.[17, 22] An additional study (long case accrual) found that the risk of mortality was reduced in women aged 70–74 years who had participated in screening (annual screening), compared with no screening, but not in women older than 74 years of age.[20] For this outcome, the sample sizes ranged from 54,635 to over 13 million; the longest follow up ranged from 8 to 16 years across the studies.

All-cause mortality

BC only population

The pooled RR for the two cohort studies with short case accrual was 0.51 (95% CI 0.40, 0.64; I^2=80%) for women across all age groups (143 to 86 fewer deaths per 1,000 women).[25, 28] A similar result was found when data from the four long case accrual cohort studies was pooled (RR 0.33, 95% CI 0.21, 0.51; 53 to 33 fewer deaths per 1,000 women) (Figure 3) regardless of age[30, 32, 43, 44] and for the 50–69-year age group[30, 32] (RR 0.26, 95% CI 0.18, 0.39; I^2=0%) (61 to 45 fewer deaths per 1,000 women). For this outcome, the sample sizes ranged from 904 to 24,387 (median 1,975) and the longest follow up ranged from median 5.1 to 10.5 years across the studies.

General screening population

In the two short case accrual cohort studies that reported on all-cause mortality, the risk of death was lower among screened women (RR 0.57, 95% CI 0.35 to 0.92; 30 to 4 fewer deaths per 1,000 women).[18, 23] The reduction in mortality risk was consistent across all age subgroups reported (50-68 years, 70-74 years, ≥75 years, and ≥85 years), but the degree of effect declined as age increased (Supplementary File 3, Table S-5).[23] The longest follow up for the 1,302,554 study participants included in the two studies ranged from mean 3.9 years to median 13.7 years.

Treatment received after diagnosis

BC only population

Cohort studies using short case accrual found that screening made no difference with respect to all types of radiotherapy (and adjuvant therapy alone among women of all ages (RR 1.00, 95% CI 0.93, 1.08) or the subgroup of women aged 50-69 years (RR 1.04, 95% CI 0.97, 1.13; 59 fewer to 68 more per 1,000 women)[45-47] (Supplementary File 3, Table S-6). This was also the case for pooled data from the long case accrual cohort studies for women of all ages and for the 40–49-year age subgroup. However, slightly more women in the screening group underwent radiotherapy in the 50–69-year age group compared with the unscreened group (RR 1.08, 95% CI 1.01, 1.16; 7 to 117 more per 1,000 women).

In both short and long case accrual studies, women of all ages (RR range 0.69–0.82; 174 to 100 fewer per 1,000 women) and those aged 40- 49 (long case accrual; RR 0.67, 95% CI 0.52, 0.85; 314 to 98 fewer per 1,000 women) and 50–69 years (short and long case accrual; RR range 0.61–0.80; 216 to 106 fewer per 1,000 women) in screened populations were less likely to receive chemotherapy than those who had not been screened (Supplementary File 3, Table S-7). There was no difference between the two groups with respect to rates of adjuvant (all ages) or neoadjuvant (50-69 years) chemotherapy in the long case accrual studies [32, 48, 49].

None of the short or long case accrual studies demonstrated any discernible effect of screening on rates of breast surgery (all ages and 50–69-year age groups) or breast conserving surgery (all ages) (Supplementary File 3, Table S-8). However, significantly fewer mastectomies were performed in younger women (40–49 years) who had been screened compared with those unscreened in one small study using long case accrual (RR 0.63, 95%CI 0.45, 0.87; 265 to 63 fewer per 1,000 women) [35]. In contrast, women older than 74 years were four times more likely to undergo mastectomy (RR 4.07, 95% CI 3.02, 5.49; 357 to 793 more women per 1,000) and half as likely to receive breast conserving surgery (RR 0.48, 95% CI 0.30, 0.77; 359 to 118 fewer women per 1,000), compared with women in the non-screening group in another small study (long case accrual) [33] (Supplementary File 3, Table S-8). Pooled analysis of women of all ages indicated that those who participated in screening were more likely to receive sentinel lymph node biopsy than those who did not undergo screening (RR 1.44, 95%CI 1.35, 1.54; 184 to 284 more women per 1,000) (Supplementary File 2, Table S-9).[32, 43, 49] There was also a non-statistically significant trend toward a higher risk of axillary surgery in the non-screened population, compared with the screened population.

General screening population

A single cohort study (long case accrual) reported on various treatment-related outcomes among women aged 70–84 years [20]. Women in the screened group were more likely to receive radiotherapy (41–51% versus 32–40%) but less likely to receive chemotherapy (9–15% versus 12–21%) compared with the unscreened group. Breast conserving surgery was also conducted more frequently among screened women (49–53% versus 33–37%). However, rates of mastectomy were similar between the two groups in this age strata (range 10–18%).

Cancer staging at diagnosis

BC only population

Pooled analysis of four cohort studies (short case accrual) indicated that women across all ages were less likely to be diagnosed with stage II or higher cancer if they participated in screening[25, 29, 46, 50] (Supplementary File 3, Table S-10). This was also a consistent finding across the cohort studies with long case accrual, including the 40–49-year age subgroup[35] and those older than 74 years.[33] The results among the long case accrual population-based studies were more equivocal. Generally, lower rates of

advanced cancer at diagnosis were observed in the screened populations across various age groupings compared with the unscreened population (Supplementary File 3, Table S-11). However, one study from Denmark found that BC screening was not associated with a reduction in the incidence of advanced cancer among women aged 50–69 years (incidence rate ratio 0.96; 95% CI 0.90, 1.02).[51] In a study from Australia, the age-standardized incidence of stage IV cancer per 100,000 women was consistently higher in the screened group, compared with unscreened women: 7.2 versus 4.1 for all ages; 9.3 versus 5.8 for 40–49 years; 16.9 versus 9.4 for 50–69 years; 22.9 versus 12.5 for women older than 69 years [52].

General screening population

In one cohort study (short case accrual) of women older than 69 years, the adjusted cumulative incidence per 100 women of localized invasive cancer was higher among the screened group compared with the unscreened group (3.2–3.8 versus 1.5–2.6), whereas the differences were less distinct for more advanced cancers (0.8–1.0 versus 0.7–0.9)[23]. Another cohort study (long case accrual) of women aged 40-79 years found the risk of having stage III or higher cancer at diagnosis was lower in the screened group (RR 0.84, 95% CI 0.83 to 0.86; 61 to 50 fewer women per 1,000).[17] An additional study (long case accrual) that characterized cancer stage only in terms of invasive or non-invasive cancer reported no difference between screened and unscreened women aged 49–79 years (difference in cumulative incidence rate -0.2%; 95% CI: -9.1, 8.8).[21]

Quality of life

BC only population

A single cohort study (n=647; short case accrual) from Germany found no difference between screening and no screening with respect to health-related quality of life for women with BC aged 50–69 years, and for the 50–59- and 60–69-year age subgroups.[45]

General screening population

None of the studies reported quality of life outcomes in a general screening population.

Interval cancer among screening participants

BC only population

Two short case accrual cohort studies found a rate of 22% (range 5 to 48%) for interval cancer among 8,159 women who participated in screening (Supplementary File 3, Table S-12).[29, 45] The rates of interval cancers were highly variable across the long case accrual cohort studies, ranging from 6% to 23% (mean 13%) in a sample of 191,043 women (Supplementary File 3, Figure S-1).[30, 32, 34-36, 49, 53] The two studies that measured cancer stage at diagnosis among individuals with interval cancer found that 45% were at least stage II, 20% were at least stage III, and 6% were stage IV [34, 53]. One small study reported a rate of 17% among women aged 40-49 years, most of which were invasive cancers (96%).[35]

General screening population

None of the studies reported rates of interval cancer in a general screening population.

Other harms and subgroup analyses

None of the studies reported on BC morbidity (such as adverse effects of treatment) or false positive rates of screening, and the length of follow up in the few prospective studies was not adequate to calculate overdiagnosis. Analysis of outcomes by risk factors, demographic characteristics, treatment type, and screening Intervals and modalities was not possible due to the limited reporting of these aspects and the lack of common outcome measures and data stratified by age among the studies.

Discussion

Main findings

This review is unique in comparing results from general screening populations to populations that had BC at study initiation. This was done not only to ensure comprehensive data collection, but also to assess consistencies and divergences in trends for BC screening between two populations, where more favorable estimates would be expected from studies comparing screening history status among people with BC than from general screening study populations. Studies in BC only populations provide an indication of the upper limit of the effectiveness of BC screening, while those conducted in general screening populations are more likely to provide results closer to the lower limit of what can be achieved. The true effects of BC screening likely lie somewhere along the spectrum of results for BC only populations and RCTs among the general screening population (which were not addressed in this review). More confidence could, therefore, be placed in results where the direction of effect is the same in both populations and aligned with those of RCTs.

Cohort and population-based studies of women with BC (very low certainty evidence) indicated that screening reduced BC mortality in those older than 49 years of age, with the effect being generally larger in studies using long case accrual. Although there was no discernible benefit for women aged 40–49 years in the cohort studies, results from the population-based studies suggested that the main benefit accrues to the 45–49-year portion of the 40–49-year age group. Therefore, the lack of benefit in the 40–49-year age bracket may be due to inclusion of data from women in their early 40s, leading to a dilution of effect that was also noted in a recent review of incident-based mortality studies.[54] Cohort studies of general screening populations also reported reductions, albeit smaller, in the risk of BC mortality among screened women, but found no benefit for those older than 69 years. All-cause mortality was also lowered by screening in general screening populations and those with a screening history in the BC only populations (ages 39 to 90 years), although the degree of reduction appeared to be inversely proportional to age in the latter group.

In cohort studies of BC only populations (very low certainty evidence), participants with a history of screening were less likely to receive chemotherapy and more likely to undergo sentinel lymph node biopsy than axillary dissection than those with no history of screening, regardless of age. However, pooled data across all ages suggested that screening had no discernible effect on rates of radiotherapy or breast conserving surgery, or on health-related quality of life in the 50–69-year age group. Women in the 40–49-year age group who underwent screening had lower rates of mastectomy, whereas those older than 74 years had significantly higher rates of invasive surgery. However, it was unclear from the studies what influence patient age had on these treatment decisions. Limited data from a single study of a general screening population aged 70–84 years indicated that screening participants were more likely to receive radiotherapy and breast conserving surgery, but less likely to receive chemotherapy, than those who were not screened. This potentially reflects current clinical recommendations cautioning the use of aggressive treatments in older populations.[20]

Pooled analyses of cohort studies (very low certainty evidence) of BC only populations indicated a statistically significant reduction in advanced BC for women undergoing screening, regardless of age. This was also generally the case in the population-based studies. However, two studies found either no difference between the groups or an increased incidence of stage IV disease among screened women.[51, 52] The authors of these two studies also noted evidence of substantial overdiagnosis. Results from three studies of general screening populations were contradictory with respect to stage at diagnosis.[17, 21, 23] The rates and degree of severity of interval cancers were highly variable across the studies.

Comparison with other reviews

Support for mammography screening among women aged 50–69 years is based on RCTs conducted over 20 years ago. However, changes in population health, imaging technologies, and available treatments potentially limit the applicability of these results to screening mammography in the current era.[55, 56] For example, a 2013 Cochrane review of RCTs found that screening had no discernible effect on BC or all-cause mortality after 10 and 13 years, respectively.[56] Over 50 systematic reviews on routine mammography for BC screening in asymptomatic women have been published since 2000, reflecting the controversies that still surround BC screening and the continuing uncertainty of when and how often to conduct screening.[56, 57]

Since no new RCTs have been published since 2017, data from non-randomized comparative studies have been increasingly used to gain a better understanding of the effects of BC screening on health outcomes. Two recent reviews that included RCT and observational data from a general screening population reported RRs for BC mortality (short case accrual) of 0.87–0.88 for women younger than 50 years (48 to 29 fewer deaths per 100,000 women); 0.77–0.86 for those aged 50-69 years (138 to 45 fewer deaths per 100,000 women); and 0.77–0.90 for those aged 70–74 years (207 to 122 fewer deaths per 100,000 women); results for the long case accrual studies were similar.[58, 59] The observational studies in one of these reviews yielded a 25–31% risk reduction for women aged 50–69 years.[58] Screening had no effect on all-cause mortality.[58, 59] In one review, advanced BC was reduced for women older than 49 years (RR 0.62), but not those aged 39–49 years (RR 0.98).[58] In the other review, the likelihood of being diagnosed with advanced BC was reduced by screening across all age groups.[59]

Two recent reviews of cohort and incidence-based mortality studies found that screening reduced BC mortality for women aged 50–69 years (up to 22%) and younger than 50 years (up to 20%).[54, 60] No benefit was found for women older than 69 years by either review.

The results from this review are consistent with previous publications, indicating equivocal results for mortality reductions beyond the usual screening age of 50–69 years. Since previous reviews included general screening populations, and in some cases RCTs, the risk reductions were lower than for the dataset on women with BC at study outset in the current review. However, the RRs derived from the limited data on mortality rates from a general screening population were similar to previous reviews that included only non-randomized studies. The degree of effect varied according to study design, accrual type, and comparator group (contemporaneous, historical, or geographic), which is in alignment with previous reviews of observational studies.[54, 60]

Canadian context

Absolute measures are very dependent on the baseline incidence in the population being studied, so it would be useful to apply the effect estimates from this review for specific and all-cause mortality to a Canadian general population. A pan-Canadian study by Coldman et al.[61] reported the risk of BC specific mortality in a general population of Canadian women for the age groups 40–49, 50–59, 60–69, and 70–79 years. However, only three of the included cohort studies reported BC specific mortality rates for these age groupings (Table 3).[25, 33, 35] Applying the results from these studies, which were derived from BC only populations, to a Canadian baseline risk substantially diminishes the effect of screening on this outcome, particularly for individuals younger than 70 years of age. The studies reporting rates of all-cause mortality did not provide data in age subgroupings that were amenable to comparison with available Canadian data.

Limitations

Conclusions from this paper are limited for several reasons. Although the literature search was exhaustive, only articles published in the English or French language were included, which could have led to a language bias. Also, only including studies from countries with a very high Human Development Index may have omitted relevant studies. However, this was done to ensure inclusion of studies with high applicability to Canadian current practice.

Although study selection was not limited by design type, indiscriminate pooling was avoided by grouping data, where possible, according to study design, type of accrual and population, and relevant subgroups identified a priori. However, doing so meant that data were sparse or non-existent for some questions on effectiveness related to risk factors, screening intervals and modalities, and some age group strata. It has been suggested that studies assessing screening require follow-up of at least 20 years.[62] None of the studies reported on harms of screening, such as overdiagnosis and false positive rates, or on BC morbidity, which is a serious limitation since it precludes assessment of the downstream harms associated with receiving a positive test result.

Follow-up in most of the included studies was either not reported or shorter than 10 years, which may either underestimate or overestimate the effects of screening. In contrast, trials with lower levels of internal validity tend to overestimate the effects of screening.[56, 60] To what degree these factors offset each other in the current dataset is unclear. In addition, the shorter follow-up periods in specific study groupings and lack of reporting of participation rates precluded the calculation of overdiagnosis or adjustment for self-selection bias. The latter bias can be particularly problematic in non-randomized studies reporting mortality outcomes because participants of mammography screening tend to be much healthier than non-participants—screening participants are 24% to 56% less likely to die from cancers other than breast cancer than are non-participants.[18, 63]

Conclusion

Although the included studies exhibited heterogeneity in various aspects such as outcome measures, age of participants, and screening practices, they nonetheless assessed service delivery models and populations may be at least partially representative of real-world clinical settings. BC screening appears to reduce the risk of BC mortality in women aged 50–69 years with average or moderately increased risk for BC in both general screening and BC only populations. Although the results were more equivocal for older and younger women, there is some indication that women aged 45–49 years may also accrue some benefit from screening. However, the lack of data on morbidity associated with screening, such as the negative effects of false positive results and overtreatment, meant that it was not possible to gauge how the benefits of screening are balanced by the potential harms. In addition, any significant effects observed in the included studies may be attenuated in Canadian populations. All-cause mortality was lower in screened women, although this effect varied inversely with increasing age in women older than 49 years. Women who undergo screening generally have less advanced cancer at diagnosis and, consequently, tend to receive less aggressive treatment. Although the degree of effect differed between studies of general screening populations and BC only populations, the direction of effect was generally similar. Questions about the effectiveness of screening with respect to screening interval, modality, risk factors, and specific age subgroups, and in the Canadian population in general, could not be answered by the currently available non-randomized comparative evidence.

References

1. Canadian Cancer Society. *Breast cancer statistics* [Internet]. Canadian Cancer Society 2023 [cited 2024 Feb 26]. Available from: https://cancer.ca/en/cancer-information/cancer-types/breast/statistics.
2. Brenner DR, Poirier A, Woods RR, Ellison LF, Billette JM, Demers AA, et al. Projected estimates of cancer in Canada in 2022. *CMAJ* 2022;194(17):E601-E607.
3. Diwanji D, Joe B. An evidence-based and inclusive breast cancer screening strategy: Summary of current guidelines from the American College of Radiology and Society of Breast Imaging. *Radiol Imaging Cancer* 2022;4(2):e229003.
4. Smith RA, Duffy SW, Gabe R, Tabar L, Yen AM, Chen TH. The randomized trials of breast cancer screening: What have we learned? *Radiol Clin North Am* 2004;42(5):793-806.
5. Ren W, Chen M, Qiao Y, Zhao F. Global guidelines for breast cancer screening: A systematic review. *Breast* 2022;64:85-99.
6. Klarenbach S, Sims-Jones N, Lewin G, Singh H, Thériault G, Tonelli M, et al; Canadian Task Force on Preventive Health Care. Recommendations on screening for breast cancer in women aged 40–74 years who are not at increased risk for breast cancer. *CMAJ* 2018;190(49):E1441-E1451.
7. Alberta Health Services (AHS). *AHS lowers recommended age for breast cancer screening* [Internet] 2022 [cited 2024 Feb 27]. Available from: https://www.albertahealthservices.ca/news/Page16912.aspx.
8. European Commission Initiative on Breast Cancer Screening. *European guidelines on breast cancer screening and diagnosis* [Internet]. European Commission; 2016 [cited 2024 Feb 26]. Available from: https://cancer-screening-and-care.jrc.ec.europa.eu/en/ecibc/european-breast-cancer-guidelines.
9. Moher D, Little J. *Breast cancer screening: Part A. An evidence report to inform an update to the Canadian Task Force on Preventive Health Care 2011 guidelines*. Public Health Agency of Canada [Internet] 2017 [cited 2024 Feb 28]. Available from: https://canadiantaskforce.ca/wp-content/uploads/2018/11/Evidence-Report-Breast-Cancer-Screening_FINAL.pdf.
10. Higgins JPT, Green S. *Cochrane handbook for systematic reviews of interventions version 6.4 (updated August 2023)* [Internet] 2023 [cited 2024 Feb 29]. Available from: www.training.cochrane.org/handbook.
11. GRADE Working Group. *Handbook for grading the quality of evidence and the strength of recommendations using the GRADE approach* [Internet] 2013 [cited 2024 Feb 29]. Available from: https://gdt.gradepro.org/app/handbook/handbook.html.
12. Canadian Task Force on Preventive Health Care. Canadian Task Force on Preventive Health Care procedure manual [Internet] 2014 [cited 2024 Feb 29]. Available from: https://canadiantaskforce.ca/wp-content/uploads/2016/12/procedural-manual-en_2014_Archived.pdf.
13. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS peer review of electronic search strategies: 2015 guideline statement. *J Clin Epidemiol* 2016;75:40-46.
14. DistillerSR. Version 2.35. DistillerSR Inc. 2023. Available from: https://www.distillersr.com/.

15. Wells G, Shea B, O'Connell D, Peterson J, Welch V, Losos M, et al. *The Newcastle–Ottawa Scale (NOS) for assessing the quality of non-randomized studies in meta-analysis* [Internet] 2000 [cited 2024 Feb 29]. Available from:https://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.

16. National Institute for Health and Care Excellence (NICE). Methods for the development of NICE public health guidance (third edition). *Appendix G: Quality appraisal checklist - quantitative studies reporting correlations and associations*. [Internet] 2012 NICE; [cited 2024 Feb 29]. Available from: https://www.nice.org.uk/process/pmg4/chapter/appendix-g-quality-appraisal-checklist-quantitative-studies-reporting-correlations-and#checklist.

17. Choi E, Jun JK, Suh M, Jung KW, Park B, Lee K, et al. Effectiveness of the Korean National Cancer Screening Program in reducing breast cancer mortality. *NPJ Breast Cancer* 2021;7(1):83.

18. Czwikla J, Giersiepen K, Langner I, Enders D, Heinze F, Rothgang H, et al. A cohort study of mammography screening finds that comorbidity measures are insufficient for controlling selection bias. *J Clin Epidemiol* 2018;104:1-7.

19. Dunn N, Youl P, Moore J, Harden H, Walpole E, Evans E, et al. Breast-cancer mortality in screened versus unscreened women: Long-term results from a population-based study in Queensland, Australia. *J Med Screen* 2021;28(2):193-199.

20. Garcia-Albeniz X, Hernan MA, Logan RW, Price M, Armstrong K, Hsu J. Continuation of annual screening mammography and breast cancer mortality in women older than 70 years. *Ann Intern Med* 2020;172(6):381-389.

21. Lund E, Nakamura A, Thalabard JC. No overdiagnosis in the Norwegian breast cancer screening program estimated by combining record linkage and questionnaire information in the Norwegian women and cancer study. *Eur J Cancer* 2018;89:102-112.

22. Morrell S, Taylor R, Roder D, Robson B, Gregory M, Craig K. Mammography service screening and breast cancer mortality in New Zealand: A national cohort study 1999-2011. *Br J Cancer* 2017;116(6):828-839.

23. Richman IB, Long JB, Soulos PR, Wang SY, Gross CP. Estimating breast cancer overdiagnosis after screening mammography among older women in the United States. *Ann Intern Med* 2023;176(9):1172-1180.

24. Duffy SW, Tabar L, Yen AMF, Dean PB, Smith RA, Jonsson H, et al. Mammography screening reduces rates of advanced and fatal breast cancers: Results in 549,091 women. *Cancer* 2020;126(13):2971-2979.

25. Luu XQ, Lee K, Jun JK, Suh M, Jung KW, Choi KS. Effect of mammography screening on the long-term survival of breast cancer patients: Results from the National Cancer Screening Program in Korea. *Epidemiol Health* 2022;44:e2022094.

26. Oberaigner W, Geiger-Gritsch S, Edlinger M, Daniaux M, Knapp R, Hubalek M, et al. Reduction in advanced breast cancer after introduction of a mammography screening program in Tyrol/Austria. *Breast* 2017;33:178-182.

27. de Glas NA, de Craen AJM, Bastiaannet E, Op 't Land EG, Kiderlen M, van de Water W, et al. Effect of implementation of the mass breast cancer screening programme in older women in the Netherlands: Population based study. *BMJ* 2014;349:g5410.

28. Ip EC, Cohen-Hallaleh RB, Ng AK. Extending screening in "elderly" patients: Should we consider a selective approach? *Clin Breast Cancer* 2020;20(5):377-381.

29. Woods LM, Rachet B, O'Connell DL, Lawrence G, Coleman MP. Are international differences in breast cancer survival between Australia and the UK present amongst both screen-detected women and non-screen-detected women? Survival estimates for women diagnosed in West Midlands and New South Wales 1997-2006. *Int J Cancer* 2016;138(10):2404-2414.

30. Barco I, Chabrera C, Garcia Font M, Gimenez N, Fraile M, Lain JM, et al. Comparison of screened and nonscreened breast cancer patients in relation to age: A 2-institution study. *Clin Breast Cancer* 2015;15(6):482-489.

31. Duffy SW, Tabar L, Yen AMF, Dean PB, Smith RA, Jonsson H, et al. Beneficial effect of consecutive screening mammography examinations on mortality from breast cancer: A prospective study. *Radiology* 2021;299(3):541-547.

32. Garcia Fernandez A, Chabrera C, Garcia Font M, Fraile M, Lain JM, Gonzalez S, et al. Mortality and recurrence patterns of breast cancer patients diagnosed under a screening programme versus comparable non-screened breast cancer patients from the same population: Analytical survey from 2002 to 2012. *Tumour Biol* 2014;35(3):1945-1953.

33. Ilenko A, Sergent F, Mercuzot A, Zitoun M, Chauffert B, Foulon A, et al. Could patients older than 75 years benefit from a systematic breast cancer screening program? *Anticancer Res* 2017;37(2):903-907.

34. Niraula S, Biswanger N, Hu P, Lambert P, Decker K. Incidence, characteristics, and outcomes of interval breast cancers compared with screening-detected breast cancers. *JAMA Netw Open* 2020;3(9):e2018179.

35. Plecha D, Salem N, Kremer M, Pham R, Downs-Holmes C, Sattar A, et al. Neglecting to screen women between 40 and 49 years old with mammography: What is the impact on treatment morbidity and potential risk reduction? *AJR Am J Roentgenol* 2014;202(2):282-288.

36. Seneviratne S, Campbell I, Scott N, Shirley R, Lawrenson R. Impact of mammographic screening on ethnic and socioeconomic inequities in breast cancer stage at diagnosis and survival in New Zealand: A cohort study. *BMC Public Health* 2015;15:46.

37. Wilkinson AN, Ellison LF, Billette JM, Seely JM. Impact of breast cancer screening on 10-year net survival in Canadian women age 40-49 years. *J Clin Oncol* 2023;41(29):4669-4677.

38. Taylor R, Gregory M, Sexton K, Wharton J, Sharma N, Amoyal G, et al. Breast cancer mortality and screening mammography in New Zealand: Incidence-based and aggregate analyses. *J Med Screen* 2019;26(1):35-43.

39. Hubner J, Katalinic A, Waldmann A, Kraywinkel K. Long-term incidence and mortality trends for breast cancer in Germany. *Geburtshilfe Frauenheilkd* 2020;80(6):611-618.

40. Moller MH, Lousdal ML, Kristiansen IS, Stovring H. Effect of organized mammography screening on breast cancer mortality: A population-based cohort study in Norway. *Int J Cancer* 2019;144(4):697-706.

41. Bastos J, Rodrigues V, Paap E, Broeders M, Pina M, Cruz D, et al. Breast cancer screening effectiveness in Portugal central Region. *Eur J Cancer Prev* 2017;26 Joining forces for better cancer registration in Europe: S204-S207.

42. Heinavaara S, Sarkeala T, Anttila A. Impact of organised mammography screening on breast cancer mortality in a case-control and cohort study. *Br J Cancer* 2016;114(9):1038-1044.

43. Kobayashi N, Hikichi M, Ushimado K, Sugioka A, Kiriyama Y, Kuroda M, et al. Differences in subtype distribution between screen-detected and symptomatic invasive breast cancer and their impact on survival. *Clin Transl Oncol* 2017;19(10):1232-1240.

44. Lim ZL, Ho PJ, Khng AJ, Yeoh YS, Ong ATW, Tan BKT, et al. Mammography screening is associated with more favourable breast cancer tumour characteristics and better overall survival: Case-only analysis of 3739 Asian breast cancer patients. *BMC Med* 2022;20(1):239.

45. Braun B, Kurosinski MA, Khil L, Tio J, Krause-Bergmann B, Hense HW. The mode of detection is not associated with quality of life in women with breast cancer. *Breast Care (Basel)* 2020;15(5):498-505.

46. Ujhelyi M, Pukancsik D, Kelemen P, Kovacs E, Kenessey I, Udvarhelyi N, et al. Does breast screening offer a survival benefit? A retrospective comparative study of oncological outcomes of screen-detected and symptomatic early stage breast cancer cases. *Eur J Surg Oncol* 2016;42(12):1814-1820.

47. Varga Z, Balog K, Sebo E, Arkosy P, Toth D. Beyond a decade: A comparative study of 15-year survival rates in screen-detected vs. symptomatic breast cancer patients in Hungary. *Ir J Med Sci* 2024;193(1):57-63.

48. Bayard S, Fasano G, Chen Y, Davis M, Drotman M, Bensenhaver J, et al. Screening mammography mitigates breast cancer disparities through early detection of triple negative breast cancer. *Clin Imaging* 2021;80:430-437.

49. Elder K, Nickson C, Pattanasri M, Cooke S, Machalek D, Rose A, et al. Treatment intensity differences after early-stage breast cancer (ESBC) diagnosis depending on participation in a screening program. *Ann Surg Oncol* 2018;25(9):2563-2572.

50. Tomsic S, Zagar T, Mihor A, Mlakar M, Lokar K, Jarm K, et al. Prognostic factors and outcomes in women with breast cancer in Slovenia in relation to step-wise implementation of organized screening. *PLoS One* 2022;17(11):e0278384.

51. Jorgensen KJ, Gotzsche PC, Kalager M, Zahl PH. Breast cancer screening in Denmark: A cohort study of tumor size and overdiagnosis. *Ann Intern Med* 2017;166(5):313-323.

52. Jacklyn G, McGeechan K, Irwig L, Houssami N, Morrell S, Bell K, et al. Trends in stage-specific breast cancer incidence in New South Wales, Australia: Insights into the effects of 25 years of screening mammography. *Breast Cancer Res Treat* 2017;166(3):843-854.

53. de Munck L, Siesling S, Fracheboud J, den Heeten GJ, Broeders MJM, de Bock GH. Impact of mammographic screening and advanced cancer definition on the percentage of advanced-stage cancers in a steady-state breast screening programme in the Netherlands. *Br J Cancer* 2020;123(7):1191-1197.

54. Dibden A, Offman J, Duffy SW, Gabe R. Worldwide review and meta-analysis of cohort studies measuring the effect of mammography screening programmes on incidence-based breast cancer mortality. *Cancers (Basel)* 2020;12(4):976.

55. Henderson JT, Webber EM, Weyrich M, Miller M, Melnikow J. *Screening for breast cancer: A comparative effectiveness review for the US. Preventive Services Task Force*. Evidence Synthesis No. 231. Rockville, MD: Agency for Healthcare Research and Quality (AHRQ), 2024. Available from: https://www.ncbi.nlm.nih.gov/books/NBK603789/.

56. Gotzsche PC, Jorgensen KJ. Screening for breast cancer with mammography. *Cochrane Database Syst Rev* 2013;(6):CD001877.

57. Raichand S, Dunn AG, Ong MS, Bourgeois FT, Coiera E, Mandl KD. Conclusions in systematic reviews of mammography for breast cancer screening and associations with review design and author characteristics. *Syst Rev* 2017;6(1):105.

58. Nelson HD, Fu R, Cantor A, Pappas M, Daeges M, Humphrey L. Effectiveness of breast cancer screening: Systematic review and meta-analysis to update the 2009 US Preventive Services Task Force recommendation. *Ann Intern Med* 2016;164(4):244-255.

59. Canelo-Aybar C, Ferreira DS, Ballesteros M, Posso M, Montero N, Sola I, et al. Benefits and harms of breast cancer mammography screening for women at average risk of breast cancer: A systematic review for the European Commission initiative on breast cancer. *J Med Screen* 2021;28(4):389-404.

60. Irvin VL, Kaplan RM. Screening mammography & breast cancer mortality: Meta-analysis of quasi-experimental studies. *PLoS One* 2014;9(6):e98105.

61. Coldman A, Phillips N, Wilson C, Decker K, Chiarelli AM, Brisson J, et al. Pan-Canadian study of mammography screening and mortality from breast cancer. *J Natl Cancer Inst* 2014;106(11):dju261.

62. Tabar L, Vitak B, Chen TH, Yen AM, Cohen A, Tot T, et al. Swedish two-county trial: impact of mammographic screening on breast cancer mortality during 3 decades. *Radiology* 2011;260(3):658-663.

63. Phillips N, Coldman A. Comparison of nonbreast cancer incidence, survival and mortality between breast screening program participants and nonparticipants. *Int J Cancer* 2008;122(1):197-201.