Pan-Canadian Study of Mammography Screening and Mortality from Breast Cancer

Andrew Coldman, Norm Phillips, Christine Wilson, Kathleen Decker, Anna M. Chiarelli, Jacques Brisson, Bin Zhang, Jennifer Payne, Gregory Doyle, Rukshanda Ahmad

Manuscript received November 15, 2013; revised February 4, 2014; accepted July 17, 2014.

Correspondence to: Andrew Coldman, PhD, British Columbia Cancer Agency, #800 – 686 W Broadway, Vancouver, BC V5Z 1G1, Canada (e-mail: acoldman@bccancer.bc.ca).

Background
Screening with mammography has been shown by randomized controlled trials to reduce breast cancer mortality in women aged 40 to 74 years. Estimates from observational studies following screening implementation in different countries have produced varied findings. We report findings for seven Canadian breast screening programs.

Methods
Canadian breast screening programs were invited to participate in a study aimed at comparing breast cancer mortality in participants and nonparticipants. Seven of 12 programs, representing 85% of the Canadian population, participated in the study. Data were obtained from the screening programs and corresponding cancer registries on screening mammograms and breast cancer diagnoses and deaths for the period between 1990 and 2009. Standardized mortality ratios were calculated comparing observed mortality in participants to that expected based on nonparticipant rates. A substudy using data from British Columbia women aged 35 to 44 years was conducted to assess the potential effect of self-selection participation bias. All statistical tests were two-sided.

Results
Data were obtained on 2796472 screening participants. The average breast cancer mortality among participants was 40% (95% confidence interval [CI] = 33% to 48%), lower than expected with a range across provinces of 27% to 59%. Age at entry into screening did not greatly affect the magnitude of the average reduction in mortality, which varied between 35% and 44% overall. The substudy found no evidence that self-selection biased the reported mortality results, although the confidence intervals of this assessment were wide.

Conclusion
Participation in mammography screening programs in Canada was associated with substantially reduced breast cancer mortality.


Background
Randomized controlled trials demonstrating a reduced mortality from breast cancer among those invited to be screened with mammography were first reported almost 50 years ago (1). Trials have continued to be performed to demonstrate the reproducibility of early findings and address specific questions about efficacy at different ages and different frequencies (2). Recent structured reviews conducted by independent task forces (3–5) have concluded that breast cancer mortality rates are reduced among women offered screening between the ages of 40 and 74 years.

While demonstration of mortality reductions in clinical trials provides an assessment of efficacy, there is a need to demonstrate impact when screening is implemented in general populations. Demonstration of mortality reductions in population implementations is of current relevance, as some authors have argued that screening is less effective in more recent years because of improved treatment and heightened public and professional awareness of the early symptoms of breast cancer (6–8). Furthermore, the lack of mortality benefit found in the two Canadian trials (9,10) makes an evaluation of the benefits of program screening in Canada of particular relevance. Evaluations of screening programs have been conducted in a number of countries, and published results have been summarized by various authors (11–13). The summaries of these reviews found widespread statistically significant reductions in breast cancer mortality, although a few individual studies found no benefit (14–17). Study attributes likely to affect results are: inclusion of women not eligible for screening in mortality counts, short follow-up after initiation of screening, and comparisons based on eligibility for screening rather than participation in screening. Models of US trends in breast cancer mortality have attributed a portion of the reductions to the effect of mammography screening (18).

We report here the results of a study of breast cancer mortality among screening participants in Canadian organized breast
screening programs and also report a substudy of the effect of self-selection among participants on breast cancer mortality.

### Methods

The study was proposed under the auspices of the Canadian Breast Cancer Screening Initiative (CBCSI), which was supported by the Public Health Agency of Canada (PHAC). The CBCSI includes representation from professional organizations, regulatory authorities, and all provincial and territorial screening programs (19).

Each of the 12 screening programs was provided a study outline and invited to participate in the study. Three programs indicated they were unable to provide the data required, and two declined. The following seven provincial screening programs participated in the study: British Columbia (BC), Manitoba (MB), Ontario (ON), Québec (QC), New Brunswick (NB), Nova Scotia (NS), and Newfoundland and Labrador (NL). All the programs used a two-view, mostly analog mammography provided at designated centers with a single radiologist interpretation. Entry to all of the programs was based upon self-referral with province-wide geographic access provided using a mixture of clinics and mobile services. Depending on time period and province, women may have received personalized invitation letters to participate in screening prior to enrollment. After enrollment, women received periodic reminder letters for screening following the provincial schedule for screening. In all provinces, breast care for women, including those participating in screening, was managed through family physicians who received results of diagnostic tests and who then managed referral to tertiary services. Each participating program had the study approved by its local research ethics review process.

The study was designed to compare the mortality experience of women who participated in an organized screening program with an estimate of what would have occurred had they not enrolled, based upon the experience of women in the same jurisdiction who did not. Separate cohorts were assembled for each screening program, consisting of women who had at least one program screen between ages 40 and 79 years in the period between January 1, 1990 and December 31, 2009. Women were considered to enter the participant cohort at their first screen in that province and remained in the cohort until death or December 31, 2009. Women were considered to enter the participant cohort at their first screen in that province and remained in the cohort until death or December 31, 2009. An incidence-based mortality approach was used based on the formula described by Saseini (20), where expected breast cancer mortality for a cancer-free individual is calculated as the probability of being diagnosed and dying from cancer within a time interval using incidence and survival rates from a referent group. Referent rates were derived from nonparticipants in each province defined to be those not in, or before entry to, the participant cohort (21). Identification and follow-up of individual nonparticipants is not required, because only the age-specific years-at-risk and survival experience following a breast cancer diagnosis is required. Nonparticipant values were obtained by subtracting the values for the participant cohort from the overall totals obtained from population demographic and provincial cancer registry data, respectively.

Each province arranged to link its screening program database to its provincial cancer registry and provincial mortality database. Linkage was performed using provincial health numbers and full name and date of birth where required. The linkage was used to identify cases and deaths from breast cancer that occurred within each cohort prior to the study end date. Each participating program was required to assemble raw data in a format described in the study protocol. The formatted data were then transformed to create analytic tabular data using two computer programs supplied by the study center. The first program extracted breast cancer cases, deaths, and computed time-at-risk in arrays with dimensions of age (five-year groups) and time (by year) since entry for the screening cohort. The program also tabulated breast cancer deaths and person-years by age and time since diagnosis for the screening program cancer cases. The second program computed the same quantities for the total population and for all the breast cancer cases in each province. The resulting tabular nonidentifiable data sets were securely transmitted to the study center, where statistical analysis was conducted. Screening program person-years were adjusted for provincial out-migration using age-, sex-, and province-specific migration rates supplied by Statistics Canada. Provincial age-specific breast cancer incidence rates are years-at-risk weighted averages of the participant and nonparticipant rates. This relationship was used to calculate the rates in nonparticipants given the known participant and provincial rates and the known weights. An analogous relationship was used to calculate the nonparticipant survival rates from the participant and provincial survival rates, where the appropriate weights are the numbers of cancer cases in the respective groups.

### Statistical Analysis

Standardized mortality ratios (SMRs) were calculated as the ratio of observed mortality in each participant group to the province-specific expected mortality, calculated using the nonparticipant incidence and survival rates. Confidence intervals (CIs) for the SMR incorporated stochastic variability in both the observed and expected values using the delta method. Expected mortality variability was derived from the variability of nonparticipant rates, assuming a Poisson distribution. Forest plots were used to display results, and tests of homogeneity of rates were made using Q (22) and F (23,24) and were two-sided, with analysis performed using the rmeta routine in the R statistical package (25). Q provides a statistical test for the presence of inhomogeneity, and F measures the degree of inhomogeneity.

Estimates of absolute benefit of screening were expressed as the number needed to participate, NNP(10), to prevent a single breast cancer death within 10 years of entry using the following formula:

$$\text{NNP}(10) = \frac{\text{SMR}}{(1-\text{SMR}) \times 10 \text{ year breast cancer mortality rate in participants}}$$

For use in the above formula, the ages at first participation in screening-specific 10 year mortality rates were calculated for all provinces using a life-table approach, where out-of-province migration or loss to follow-up reduced the years-at-risk while deaths from other causes did not. Confidence intervals for the age-specific NNP(10)s were derived from the confidence limits for the corresponding SMRs.

Comparing screening participants to nonparticipants is subject to self-selection bias. In the context of this study, bias occurs in the calculation of expected mortality if the incidence and survival rates of nonparticipants are not equal to the rates that would have been observed in participants had they not chosen to participate...
in screening (counterfactual rates). To investigate the degree of bias due to self-selection we undertook a substudy using data on British Columbian women aged 35 to 44 years who were selected to represent a narrow age range, including screen-eligible (aged 40 to 44 years after 1987) and ineligible (aged 35 to 39 years) women. Screening was uncommon in British Columbia women prior to 1985 (26). The ratio of incidence rates was calculated for the age group 35 to 39 years from 2000 through 2009 compared with 1975 to 1984, providing an estimate of trends in breast cancer incidence unrelated to screening. The ratio was applied to the incidence rate for the age group 40 to 44 years in the prescreening period 1975 to 1984 to give predicted rates for 2000 to 2009 in the absence of screening. These predicted rates for 2000 to 2009 are weighted (proportional to years at risk) averages of observed incidence rates in those who did not participate in screening and unknown counterfactual rates for those who did. This relationship permitted estimation of the counterfactual incidence rates at ages 40 to 44 years for screening participants. A similar approach was taken to the estimation of counterfactual survival rates. The resulting estimates of counterfactual incidence and survival rates were used to calculate the expected 10-year mortality in British Columbia women aged 40 to 44 years entering screening from 2000 to 2009 and compared with the same calculation using nonparticipant rates. If the nonparticipant- and counterfactual-based estimates of mortality reduction are equal, then bias is not present. Further details of the method used are described in the Supplementary Methods (available online).

**Results**

The screening participant group included observations on 2,796,472 women with a total of 20.2 million person-years of observation. Province-specific characteristics of the screening programs and screening cohorts are summarized in Table 1. Maximum years post entry of each provincial cohort ranged from 12 to 20. More women report bilateral mammography than are reported as participating in these programs, as not all screening is performed in program-affiliated centers and because of potential misclassification of diagnostic mammography by respondents.

Province-specific breast cancer incidence rates were computed by five-year age groups and calendar periods 1990 to 1994, 1995 to 1999, 2000 to 2004, and 2005 and 2009. Similarly, survival hazards were calculated by single year following diagnosis by age, calendar period, and province. Age-specific breast cancer incidence rates and cancer-specific survival rates were anticipated to be elevated among screening participants because of the effects of lead time and overdiagnosis. The anticipated relationships are illustrated by the cumulative incidence rates between ages 50 and 69 years and five-year survival rates for the same ages calculated for participants and nonparticipants in each province (Table 2). Participant survival rates are biased both by overdiagnosis and lead-time effects.

The expected numbers of deaths and SMRs with confidence intervals were calculated for each provincial screening cohort for all ages combined (Figure 1). The average breast cancer mortality among participants was 40% (95% CI = 33% to 48%), lower than expected, with a range across provinces of 27% to 59%. Results are presented separately by age at screening cohort entry in decades...
Per 100 000 among nonparticipants of the same age at that time. Among women aged 35 to 39 years, the five-year survival rate from 1975 through 1984 increased from 78.6% to 87.7% in 2000 to 2009. In 1975 to 1984, the five-year survival rates among women aged 40 to 44 years was 83.0%. The change in survival between 1975 and 1984 and 2000 and 2009 for women aged 35 to 39 years resulted in a predicted counterfactual five-year survival rate of 90.6% for women aged 40 to 44 years in 2000 to 2009, as compared with 89.7% for nonparticipant women of the same age in the same period. Among women aged 40 to 44 years beginning screening between 2000 and 2009 and followed until 2010, there were 27 deaths from breast cancer observed, with 62.5 (95% CI = 60 to 65) deaths expected based upon nonparticipant rates and 68.7 (95% CI = 32 to 120) expected based on estimated counterfactual rates. The mortality reduction associated with screening was estimated to be four percentage points greater when counterfactual rates were used (61%, 95% CI = 9 to 81) than when nonparticipant rates were used (57%, 95% CI = 39 to 72) in this cohort.

**Discussion**

We found that participation in a provincial breast screening program was associated with 40% lower breast cancer death rates than expected in all provinces included and did not vary greatly with age. The number needed to participate in screening to prevent a single breast cancer death within 10 years decreased with age from 1247 for women first screened at age 40 to 49 to 498 for those first screened at age 70 to 79. These results are not influenced by trends in the efficacy of treatment or population awareness of breast cancer symptoms, as comparisons are made between contemporaneous cases in the same populations. Results were not homogeneous among the provinces. Programs are not identical, having different age eligibility, some providing annual screening for higher risk women (MB, ON and NL), some screening average risk women annually during part of the study period (BC), and some including clinical breast examination (MB, ON, NS, NL) in addition to mammography for part of the study period. The utilization of screening mammography outside of organized programs also varied across the provinces (19) and Table 1. No evidence of overestimation of the mortality reduction associated with screening because of self-selection was found in the BC substudy, although the confidence intervals of this assessment were wide.

---

Table 2. Cumulative incidence rates of breast cancer and five-year survival rates for screening program participants and nonparticipants aged 50 to 69 years by province*

<table>
<thead>
<tr>
<th>Province</th>
<th>Cumulative incidence rate of invasive breast cancer between ages 50 and 69 years (%)</th>
<th>Breast cancer-specific five-year survival rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Participant</td>
<td>Nonparticipant</td>
</tr>
<tr>
<td>BC</td>
<td>72</td>
<td>5.2</td>
</tr>
<tr>
<td>MB</td>
<td>79</td>
<td>5.5</td>
</tr>
<tr>
<td>ON</td>
<td>6.9</td>
<td>5.4</td>
</tr>
<tr>
<td>QC</td>
<td>8.1</td>
<td>5.7</td>
</tr>
<tr>
<td>NB</td>
<td>70</td>
<td>5.3</td>
</tr>
<tr>
<td>NS</td>
<td>78</td>
<td>5.1</td>
</tr>
<tr>
<td>NL</td>
<td>79</td>
<td>5.1</td>
</tr>
</tbody>
</table>

* BC = British Columbia; MB = Manitoba; NB = New Brunswick; NL = Newfoundland and Labrador; NS = Nova Scotia; ON = Ontario; QC = Québec.

---

Figure 1. Forest plot of standardized mortality ratios (SMRs) by province, all ages combined. Summary estimate is based upon a random effects model. All statistical tests were two-sided. CI = confidence interval; SMR = standardized mortality ratio.

---

Figure 2. Age at entry into screening did not greatly affect the magnitude of the average reduction in mortality, which varied between 35% and 44%. There was evidence of nonhomogeneity between provincial programs in the core screening age ranges of 50 to 59 years (I² = 88%, P < .001) and 60 to 69 years (I² = 75%, P < .001), so that a random effects model was used to summarize effects. Other ages did not demonstrate significant nonhomogeneity; ages 40 to 49 years (I² = 50%, P = .14), 70 to 79 years (I² = 0, P = .83), but there were few contributing provinces and reduced deaths. Summary estimates of SMRs across provinces were 0.56 (95% CI = 0.45 to 0.67), 0.60 (95% CI = 0.49 to 0.70), 0.58 (95% CI = 0.50 to 0.67), 0.65 (95% CI = 0.56 to 0.74), and 0.60 (95% CI = 0.52 to 0.67) for women aged 40 to 49, 50 to 59, 60 to 69, 70 to 79 (Figure 2) and 40 to 79 years (Figure 1), respectively.

The 10-year breast cancer mortality rates were 1.02, 1.98, 2.50, and 3.97 per 1000 participants first screened in the period between 1995 and 1999 at ages 40 to 49, 50 to 59, 60 to 69 and 70 to 79 years, respectively. The numbers needed to prevent a death at 10 years, NNP(10), were 1247 (95% CI = 802 to 1990), 757 (95% CI = 485 to 1178), 552 (95% CI = 400 to 812), and 498 (95% CI = 321 to 717) for those first participating at ages 40 to 49, 50 to 59, 60 to 69, and 70 to 79 years, respectively.

The substudy of British Columbia (BC) data examining the potential influence of self-selection produced results as follows. Among women aged 35 to 39 years, the ratio of incidence rates between 2000 and 2009 compared with 1975 to 1984 was 1.014. Applying this ratio to rates for women aged 40 to 44 years from 1975 to 1984 led to an estimated counterfactual incidence rate of 138 per 100 000 from 2000 to 2009 compared with a rate of 115
Observational studies of mortality reductions associated with breast screening have used differing methodologies, which have been reviewed by several authors (11,12,13). A methodology that is close to what's used in clinical trials is to be preferred. Screened and referent populations should be comparable, observations should be contemporaneous, and only breast cancer deaths that could be affected by screening should be counted. A recent review of observational studies conducted in Europe found that studies with better control of these factors had mortality reductions that were generally greater than for the clinical trials (11). In the review of European studies, mortality reduction estimates ranged from 38% to 48% for screened women. Possible reasons why greater reductions are reported in observational studies compared with randomized trials include bias, improved treatment of early-stage breast cancer, improved screening technology, and actual participation vs invitation. Analysis of Canadian data showed increasing

Figure 2. Forest plot of standardized mortality ratios (SMRs) by province for ages at entry: 40 to 49 years (A), 50 to 59 years (B), 60 to 69 years (C), and 70 to 79 years (D). Summary estimates are based upon random effects models. All statistical tests were two-sided. CI = confidence interval; SMR = standardized mortality ratio.
survival for cancers diagnosed in the period between 1990 and 2000 among screening participants, while no change was seen in the same period in nonparticipant cancer survival rates (27).

There have not been any systematic reviews of self-selection bias in programmatic breast screening. Moss et al (28) compared nonparticipants in breast screening in one region of the United Kingdom to all subjects in a comparable region where screening was not offered. They found that the SMR of breast cancer death among nonparticipants was 1.13 compared with uninvited. A case-control study in the Netherlands (29,30) compared nonparticipants to comparable women not invited in five regions. For three out of five regions, there was little difference between nonparticipants and uninvited, but for two regions there was a considerable difference, with nonparticipants having reduced breast cancer mortality compared with uninvited. From these findings, it would appear that generalizations regarding the effects of self-selection are unclear, and estimates should not be transferred from one setting to another, as has been done before (21). Lower socioeconomic status is associated with reduced risk of breast cancer (31) and reduced participation in screening (32), so that women presenting for screening may be at elevated risk, as was found in the substudy presented here. However, survival of screening nonparticipants has been observed to be inferior to that of women not invited to screening (33,34) and is therefore inferior to counterfactual participant survival.

This study used death from cancer as the outcome. It has been proposed that death from all causes be used as a supporting outcome in trials to avoid potential problems with the accuracy of death certification or unrecognized indirect effects of increased breast cancer detection on other causes of death (35). Unfortunately, the methodology used here did not permit this analysis, and other authors have pointed out that the value of doing this is very limited (36,37). This study did not measure the extent of participation, because women were only required to have a single-program screen to become a participant; therefore, the effect of regular screening may be underestimated. Also, women who did not participate in screening programs may still have been screened, as mammography was accessible outside of the programs, especially in the earlier part of the study period, which would tend to diminish the measured effect of screening. Self-selection to participate may cause bias in the results, although there was no evidence that this bias favored screening in the substudy. The substudy did not indicate that screening effects were overestimated because of self-selection bias, but the methodology used did not permit examination of all age-groups and was only performed using data from one region. In addition, the substudy was limited to mostly premenopausal women, assumed that incidence and survival trends in adjacent age groups were the same over the study period, and resulted in wide confidence intervals on the estimated effect of self-selection bias. The BC substudy is likely generalizable to other provinces and other ages, as each provincial program provided similar access to screening, and age at entry was primarily determined by age when screening became available. Since breast cancer deaths occurring in screening participants who have migrated out of province are not included in the province-specific mortality counts, it is necessary to adjust for migration in the calculation of the expected rates. This was done by assuming participants to have the same age-specific migration rates as the provincial average. The effect of this adjustment was small, but it is possible that if migration rates of screened women greatly exceeded that of other women, then the expected number of breast cancer deaths may be overestimated.

The value of breast cancer screening has attracted a large number of polarized comments (5). The present study found statistically significantly lower breast mortality rates among screening participants of all ages in multiple regions. Monitoring the performance of screening programs using proximal measures, such as cancer detection rates and stage distribution at diagnosis, is well developed (19). While such monitoring is valuable, it is limited in its ability to demonstrate the risks and benefits of screening. There is a need to undertake future studies using multiple methodologies in order to provide assurance that programs are delivering their anticipated mortality benefits.

References


**Funding**

Funding was provided for data extraction by the Public Health Agency of Canada. Data collection was performed by each province, analysis was performed at the British Columbia center, and writing was performed by the provincial representatives. The authors declare no direct financial interests in the performance of this research. All authors (except RA) are employees of organizations involved in the delivery of screening in their province.

**Notes**

The authors would like to acknowledge the assistance of the following individuals:
Suzanne Leonfeller, S. Edswar Kumar, Department of Health, New Brunswick, Fredericton; Vicky Majpruz, Prevention and Control Cancer Care Ontario, Toronto; Mohamed Abdolell, Dept of Diagnostic Radiology, Dalhousie University, and Ron Dewar, Cancer Care, Nova Scotia; Eric Pelletier, Jean-Marc Daigle and André Langlois, Institut National de Santé Publique du Québec; Natalie Biswanger and Pascal Lambert, Cancer Care Manitoba, Beth Halyard, Newfoundland and Labrador Centre for Health Information, Newfoundland, St. John’s.

**Affiliations of authors:** Cancer Surveillance and Outcomes (AC, NP) and Screening Mammography Program of BC (CW), BC Cancer Agency, Vancouver, British Columbia; Screening Programs, Cancer Care Manitoba, Winnipeg, Manitoba (KD); Ontario Breast Screening Program, Cancer Care Ontario, Toronto, Ontario (AMC); Québec Breast Cancer Screening Program, Institut National de Santé Publique du Québec, Quebec City, Quebec (UB); New Brunswick Breast Cancer Screening Services Program, Fredericton, New Brunswick (BZ); Nova Scotia Breast Screening Program, Halifax, Nova Scotia (JP); Newfoundland and Labrador Breast Screening Program, Eastern Health, St. John’s, Newfoundland (GD); Public Health Agency of Canada, Ottawa, Ontario (RA).