The Cost-Effectiveness of Screening Mammography in Canada

By

Ngoc-Thy Dinh

A thesis submitted to the Faculty of Graduate and Postdoctoral Studies

in conformity with the requirements

for the degree of Doctor of Philosophy (Population Health)

University of Ottawa

Ottawa, Ontario, Canada

March 2015

Authorizations

No authorization for the use of figures and tables that have been published in scientific journals was required. The manuscripts included in this dissertation have not yet been submitted for publication.

Abstract

This work includes a series of studies that examines the health and economic impacts of screening mammography from international and Canadian perspectives. This work is a compendium of several researched chapters that include an introduction, four body chapters, and a discussion. The body chapters include a systematic review of the health economic literature on screening mammography, a review of quantitative models used to examine the consequences of breast cancer screening, and cost-effectiveness analyses of screening mammography programs in Canada for the general female population and for subgroups of the population at high-risk for breast cancer. There are three analytic chapters that will be submitted as manuscripts for peer-reviewed publication. The main results of this research show that current screening mammography practices in Canada may extend life at an acceptable cost to the health care system. Due to the outlined methodological limitations of this research the results should be interpreted with caution.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

CEA	Cost-effectiveness analysis
CUA	Cost-utility analysis
CBA	Cost-benefit analysis
LYG	Life-years gained
CER	Cost-effectiveness ratio
ICER	Incremental cost-effectiveness ratio
PHAC	Public Health Agency of Canada
QALY	Quality-adjusted life year
QOL	Quality of life
CAD	Canadian dollar

Acknowledgements

The student wishes to acknowledge the following individuals and organizations, and especially the thesis supervisors for their significant contributions to this work.

Thank you to the thesis supervisors for their excellent guidance and mentorship, and for their patience:

- Dr. Kevin Brand, Telfer School of Management, University of Ottawa
- Dr. Douglas Coyle, Department of Epidemiology and Community Medicine, University of Ottawa

Thank you to the thesis internal committee members for their time and useful advice:

- Dr. Catherine Deri Armstrong, Department of Economics, University of Ottawa
- Dr. Howard Morrison, Department of Epidemiology and Community Medicine, University of Ottawa, and formerly (retired) Senior Science Advisor in the Centre for Chronic Disease Prevention, The Public Health Agency of Canada.
- Dr. Kevin Pottie, Department of Family Medicine, University of Ottawa

Thank you to the thesis external examiner for his constructive feedback and taking the time to participate in the thesis defense in person:

- Dr. Andrew Coldman, School of Population and Public Health, University of British Columbia

This work was financially supported by the Canadian Cancer Society Research Studentship from 2008 to 2010. In addition, this work could not be completed without the financial and in-kind contributions of the Public Health Agency of Canada, Statistics Canada, and the following employees of these institutions:

- Mr. William Flanagan, Health Analysis Division, Statistics Canada
- Mr. Jay Onysko former Manager, Screening and Early Detection Section, Chronic Disease Surveillance Division, Centre for Chronic Disease Prevention, The Public Health Agency of Canada.

The student also wishes to thank Roseline Savage and teaching and administrative faculty within the PhD in Population Health Program at the University of Ottawa for their support. She also thanks her colleagues at the Conference Board of Canada, friends, and family members for their tireless support.

Chapter 1: Introduction

Preamble

This dissertation is composed of six chapters.

The first chapter is an introduction to the work as a whole and includes the main overall research questions, objectives and rationale. The introduction also provides a description of the subject within the context of a population health framework.

The second chapter of the dissertation provides an overview of quantitative models that have been used to estimate the effectiveness and/or cost-effectiveness of screening mammography.

The third chapter and first manuscript is intended to be submitted to the journal Chronic Diseases and Injuries in Canada (CDIC) for peer-reviewed publication. It is a systematic review on the impact of age, screening frequency, and disease risk on the cost-effectiveness of screening mammography and provides a basis for methodological insight into the design of a cost-effectiveness study for the Canadian context.

The fourth chapter and second manuscript is a cost-effectiveness analysis study of the impact of screening age and frequency on cost-effectiveness of screening mammography in Canada and is intended to be submitted for peer-reviewed publication in the Canadian Medical Association Journal (CMAJ).

The fifth chapter and third manuscript is a cost-effectiveness analysis study of the impact of screening age and frequency on the cost-effectiveness of screening mammography for high-risk sub-groups of the Canadian female population – specifically women with family history of disease and extremely dense breasts. This article is intended to be submitted to the Journal of Health Services and Policy Research (JHSPR) for peer-reviewed publication.

The sixth chapter is a discussion of the overall work and provides a summary of findings, discusses the general strengths and limitations research, and describes the implications for research and policy.

The Research Questions

The following research questions were explored in each chapter:

1. What is known about screening mammography and the issues pertaining to screening in Canada? (Chapter 1)
2. What are the types of quantitative models used in the evaluation of the effectiveness and cost-effectiveness of screening mammography internationally, which could be used for a cost-effectiveness study for Canada? (Chapter 2)
3. What guidance do the relevant previous cost-effectiveness analysis studies provide regarding the appropriate choice of age of eligibility, screening frequency, and other target population characteristics such as the baseline incidence of breast cancer in the population, for the assessment of the optimal screening in Canada? (Chapter 3-manuscript 1)
4. What is the impact of age eligibility and screening frequency on the cost-effectiveness of screening mammography in Canada? (Chapter 4-manuscript 2)
5. What is the impact of targeted screening for high-risk subgroups of the Canadian population on the cost-effectiveness of screening mammography? (Chapter 5-manuscript 3)
6. What are the main research findings and recommendations for action in terms of research and policy in breast cancer screening? (Chapter 6)

Objectives

The specific objectives of this research in correspondence to each of the dissertation questions were to:

1. Review the literature and briefly summarize the knowledge on breast cancer and screening mammography for the secondary prevention of breast cancer in order to better inform the health economic evaluation of screening mammography in Canada;
2. Review the literature on quantitative models previously used to evaluate the health and economic impact of screening mammography in order to inform the health economic evaluation of screening mammography in Canada;

3. Appraise the published literature concerning the health economic impact of age eligibility, screening frequency, and other target population characteristics on the cost-effectiveness of screening mammography in order to inform the health economic evaluation of screening mammography in Canada;
4. Determine the optimal screening mammography policy option for Canada based on varying age eligibility criteria and screening frequency in the general risk population using a cost-effectiveness analysis paradigm;
5. Determine the optimal screening mammography policy option for Canada based on varying age eligibility criteria and screening frequency in high-risk subgroups using a cost-effectiveness analysis paradigm; and
6. Summarize the research findings, discuss other issues that were not addressed in the research, and provide insight into future directions in terms of research and policy recommendations.

Rationale

Breast cancer is the second leading incident cancer (second to non-melanoma skin cancer) and the second-leading cause of cancer mortality (second to lung cancer) among women in Canada (Canadian Cancer Society, 2014). Population-based breast cancer screening programs aim to reduce breast-cancer related deaths through earlier detection and treatment. Currently there are organized breast cancer screening programs in all provinces and territories of Canada, except in Nunavut, with varying program practices (Public Health Agency of Canada, 2008).

Differences in the effectiveness and cost-effectiveness of breast cancer screening can be attributable to differences in context in terms of incidence and prevalence of breast cancer (possibly traceable to different risk factor exposures), differences in quantitative modeling methodology and health care system characteristics. Some screening program design characteristics of interest in studies to date include changing age eligibility for screening within the target population and screening frequency, in addition to more targeted interventions for select high-risk populations such as women with familial or genetic risk.

The merits of population-based screening mammography have been and continue to be heavily debated. In many developed countries, routine screening mammography has been recommended and sometimes fully government-funded for women of moderate to high risk of breast cancer, characterized by older age generally of 50 to 70 years, on the basis that several randomized controlled trials and quasi-experimental studies having shown a survival benefit. For example, the study by Kalager et al. (2010) reported a 10 per cent reduction in breast cancer mortality attributable to screening mammography. Reducing screening intervals (increasing screening frequency)[1] and expanding age of eligibility (screening under the age of 50 and over the age of 70 years)[2], there would be an increase in the number of cancers detected and a presumed decrease in breast cancer mortality due to earlier detection and treatment. However, the costs and any associated harms due to more frequent screening over a women's lifetime will also increase. The tradeoffs between the anticipated benefits (expedited detection, reduction in breast cancer mortality, and improved life expectancy) and anticipated harms (false-positive and false-negative screen tests, iatrogenic radiation exposure, anxiety, and costs) still need to be examined within the Canadian context. Other questions include the benefits of more tailored screening programs for specific subgroups of the population, such as populations deemed at higher risk for breast cancer such as women with a genetic predisposition and/or family history of disease, and whether more targeted screening should be in addition to or a replacement for routine screening for women of moderate risk characterized by age only.

To date there have been no satisfactory economic evaluations on the cost-effectiveness of population-based breast cancer screening program designs in Canada. From a preliminary review of the literature, only one Canadian study was found assessing potential effectiveness and total costs of specifically screening Ontario women aged 40-49 (Hunter et al., 2004). This study is limited however as there is no comparison of program alternatives and the outcomes of interest are considered intermediate (costs per cancer detected as opposed to

[1] Screening interval is the duration of time between screens. Depending on age and disease risk, screening intervals can vary from 1-year for women at high risk to 2 or 3 years for women at moderately high risk. If recommended screening intervals are reduced a woman will experience higher frequency of screening over her lifetime provided age eligibility does not change.

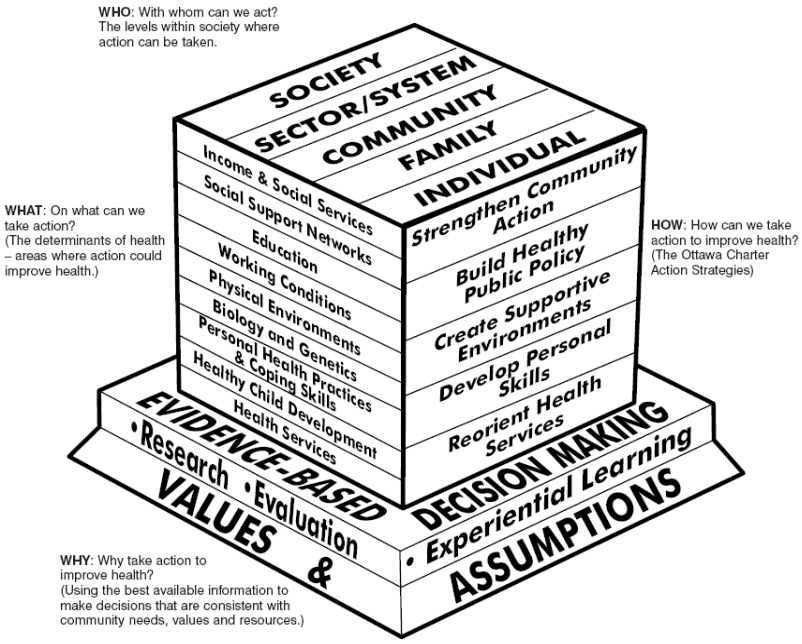
[2] Age eligibility is the recommended age at which a women would be invited for screening in a population-based screening mammography program. Age eligibility is characterized by an age range that reflects an elevated risk or probability of breast cancer. The status quo in Canada is screening for women ages 50 to 70 years of age.

cost per breast cancer death prevented or life year saved). It is uncertain how well the current literature published from other countries and for other populations is generalizable (in terms of methodology, assumptions, and results) to the Canadian setting. In addition, there is no published work that specifically examines the effects of population heterogeneity on the effectiveness and cost-effectiveness of screening within the Canadian context. This proposed research not only adds to the existing literature by drawing on previous work and available data, but has important public health policy implications for Canada.

Prior to the research conducted for this dissertation, one could only speculate that previous studies abroad would be comparable to the Canadian context which may not hold true. Despite the large amount of Canadian data on breast cancer and breast cancer screening, as well as capable tools such as the Population Health Model – Breast Cancer Screening Module (Statistics Canada) that could be leveraged to assess the cost-effectiveness of screening mammography, very few analyses have been conducted. Further, the general consensus in the health economic literature is that systematic reviews of previous health economic studies are not a substitute for conducting a study particular to the jurisdiction of interest. At most such reviews can inform the design of a study, such as is the case for the review conducted for this dissertation, but can never be seen as a replacement.

A Population Health Framework and Conceptual Model

Figure 1. Hamilton and Bhatti Integrated Framework for Population Health Promotion



Note: This model is otherwise known as "the cube" and highlights the various factors or determinants that affect population health (Hamilton and Bhatti, 1995)

Acknowledging that there are a multitude of health determinants, this research focuses on the impact of health services and the potential to reorient health services in order to improve population health. Other important health determinants include: population and individual income, availability and quality of social services, social support networks, education, working conditions, physical environments, individual biology and genetics, personal health practices and coping skills, and healthy child development. Environmental characteristics (not conceptualized or measurable as characteristics of individuals but as characteristics of places and aggregates of people) can affect the health of individual people. These determinants of health are believed to impact physiological processes in individuals' bodies that manifest into healthy or unhealthy outcomes. A *population health approach* focuses on the improvement and maintenance of health through action directed toward the health of an entire population, or sub-population, rather than individuals. Population-based breast cancer screening is a public health intervention used in the secondary prevention of breast cancer.

Secondary prevention is "a set of measures available to individuals and communities for the early detection and prompt intervention to control disease and minimize disability", such as screening programs (Last, 2007).

Within the Canadian health care system, it is intended that all Canadians have equal and equitable access to a range of health care service. This has come to include population-based screening mammography. However in reality, just like many other health services, access to screening mammography is not uniform. Inequities exist for specific subgroups of the at-risk (women age 40 and over) population. Health inequalities refer to differences, variations, and disparities in the health achievements of individuals and groups (Kawachi et al., 2002). Health inequities, on the other hand, refer to those inequalities in health that are deemed to be unfair or stemming from some form of injustice.

Economic approaches can be leveraged to support decisions regarding the use of scarce resources (monetary and non-monetary) that could have alternative uses. Specifically, it can be used to analyze and assess the costs and benefits of improving patterns of allocation of resources. In the context of health as a commodity or service, health economics can be considered a study of scarcity and choice with the objective of providing analysis to inform the best combination of resources to deliver optimal care, i.e. optimal allocation of a given quantity of resources between alternative options for improving health. Economics applies to all activities where scarcity and choice exist, which are quite evident in health. Although our research is focussed on the reorientation of health services, there is an understanding that health services are only one of many health determinants that interact to exert effects on population health. The following research is but a small piece of a larger picture. Nonetheless, as public health interventions can have a substantial impact on population health so do decisions regarding health resource allocation.

To describe the conceptual model that supports this research the proposed study operates on the following assumptions and hypotheses:

1) Breast cancer screening reduces breast cancer mortality via early detection and by shifting the distribution of stage at detection (stage shifting) with more cancers detected that are more amenable to successful treatment and improved prognosis compared to no screening;

2) Screening more women (i.e. extending age eligibility) will increase the number of cancers detected but will have downsides including greater anxiety from false-positives and other potential harms as well as increased costs;

3) Increasing screening frequency (decreasing the amount of time between screening episodes) will increase the number of cancers detected but will incur greater screening costs. However decreasing screening frequency (increasing the amount of time between screening episodes) may decrease the number of cancers detected at earlier-stage resulting in increased breast cancer mortality;

4) Breast cancer incidence and survival is affected by multiple factors including genetic profile/susceptibility, environmental factors. lifestyle factors, and health care system factors including access to health services including screening, diagnosis, and treatment, as well as the effectiveness of these services in detecting and treating disease;

5) Resource scarcity is a reality that affects healthy public policy. Cost-effectiveness analyses can be used to evaluate the trade-offs between benefits and costs of interventions as well as determine the optimization of available resources;

6) An intervention that is effective may not necessarily be cost-effective. Health economic evaluations provide "value for money" insight to supplement evidence of intervention efficacy. To be cost-effective is to provide a health gain at an "affordable" price and is contingent upon a set threshold cost or willingness to pay.

Health and Economic Burden of Breast Cancer in Canada

Cancer is described as a malignant neoplasm that occurs when disruption of cellular growth causes cells of an organ or tissue to develop and reproduce abnormally. These cells typically invade and destroy tissue and metastasize to distant sites in the body if left untreated (Last, 2007). In the case of breast cancer, neoplasm starts in the cells of the breast (Canadian Cancer Society, 2014).

For 2014 it was estimated that there would be about 24,400 new cases of female breast cancer and 5,000 breast cancer deaths in Canada, representing 14 per cent of all deaths in

female cancer deaths in that year (Canadian Cancer Society, 2014). It is the second most incident female cancer in Canada, with an age-standardized incidence rate of 99 cases per 100,000 female population (Canadian Cancer Society, 2014). The lifetime probability of developing breast cancer for Canadian women is 1 in 9, and of women diagnosed with breast cancer 1 in 28 will eventually die of it (Canadian Cancer Society, 2014). The highest age-standardized incidence rates (ASIR) of breast cancer in Canada are in Quebec and Nova Scotia (109 and 105 cases per 100,000, respectively) and the lowest incidence rates in British Columbia and Saskatchewan (90 and 95 cases per 100,000, respectively). Between 1979 and 1999, the incidence of breast cancer in Canada steadily increased but since then it has been declining by a rate of about 1.7 per cent per year. The age-standardized breast cancer mortality rate has decreased by more than 27 per cent, dropping from 32 deaths per 100,000 in 1986 to 23.1 deaths per 100,000 population in 2010.

Secular trends in risk factors during this time period, such as later age at first birth, increase rates of nulliparity, as well as earlier age of menarche ought to have increased breast cancer incidence. It has been suggested however that the recent decline in incidence (dropping back closer to pre-screening levels) may be attributable to screening uptake which may have eventually exhausted the pool of prevalent cancer in the screened population. Also, changes in risk and protective factors, such as the decline in use of hormone replacement therapy and changes in lifestyle such as reduction in excessive alcohol consumption, a risk factor for breast cancer, may have played a role in this decline. Age standardized cancer mortality rates have been decreasing since mid-1980 with advancements in treatment and earlier detection through screening mammography. According to Brown et al. (1993), although breast cancer has relatively high survivability, it remains high on the lists of total person-years and average life-years lost because the median age of diagnosis for this cancer is usually relatively young.

The Public Health Agency of Canada estimated that the direct costs and indirect costs in 2008 from breast cancer totalled almost $500 million of which 40 per cent were on drugs, 3 per cent on physician care, 25 per cent on hospital costs, and 3 per cent on productivity losses (mortality) (Public Health Agency of Canada, 2008). Lost wages for among woman with breast cancer have often been used as a measure of indirect cost (Lidgren et al., 2007;

Lauzier et al., 2008; Gordon et al., 2007; Arozullah et al., 2004). Lauzier et al. (2008) found that on average women with breast cancer in their study lost 27 per cent of their projected usual annual wages due to their disease. They also found that a higher percentage of lost wages was statistically associated with a lower level of education, living at a large distance from health care services, lower social support, having invasive disease, receipt of chemotherapy, self-employment, shorter tenure in the job, and part-time work.

Breast Cancer Etiology

A risk factor is an component of personal behaviour or lifestyle, an environment exposure, or an inborn or inherited characteristic, that, on the basis of epidemiologic evidence, is known to be associated with health-related conditions (Last, 2001). Sex and age are the strongest risk factors for breast cancer where the disease occurs mostly in women and increases with age. Other strong risk factors for breast cancer range from family history of disease and genetics to lifestyle factors to in utero exposures (Hankinson et al., 2004). The most well-documented risk factors for breast cancer include: increasing age; having a family history of breast cancer and having increased number of affected relatives; genetic predisposition, early age of menarche; late age at menopause; null parity (having never given birth) and late age at first birth (Nelson et al., 2012). In terms of dietary risk factors, increased consumption of saturated fat has a modest relationship with breast cancer (McPherson et al., 2000; Nelson et al., 2012). Increased circulating estrogen levels in postmenopausal women and the increased use in postmenopausal estrogens (particularly when combined with a progestin) such as with the use of hormone replacement therapy, have also been found to be positively associated with breast cancer risk (Collaborative Group on Hormonal Factors in Breast Cancer, 2002). Other risk factors include obesity in post-menopausal women as well as radiation exposure.

Cancer is often defined as a genetic disease with an accumulation of genetic and epigenetic deviations that result in a malignant phenotype (Domcheck and Weber, 2002). The disease is a component of several chromosomal dominant cancer syndromes, of which the most common are BRCA1 or BRCA2 mutation syndromes. More recently, causes of breast cancer have been further explained using epigenetics, defined as "any heritable influence (in the

progeny of cells or of individuals) on gene activity, unaccompanied by a change in DNA sequence" (Nature.com glossary, 2007). Epigenetics, especially in the field of breast cancer, is gaining popularity as an area of study into the causes of breast cancer due to the fact that only a small proportion of cancers are known to be attributable to the inheritance of susceptibility genes, BRCA1 and BRCA2 (Phipps et al., 2005). Gene-environment interactions have received more emphasis as plausible explanations of breast cancer etiology whereby risk of disease is affected by interacting genetic and environmental determinants. The effects of some genetic variants may be magnified and may only become noticeable in the presence of certain environmental exposures. As an example, some studies have also shown an association between ionizing radiation exposure and genetic susceptibility in the etiology of breast cancer (Chakaborty et al., 1998; Hall et al., 2000).

Generally, data shows that higher incidence of breast cancer is reported in white women and women of higher socioeconomic status (SES) compared to women of other races and lower SES. It has been suggested that the magnitude of the disparity in breast cancer incidence between races decreases with increasing SES (Vainstein, 2008). Much of the research from the United States has shown that although breast cancer incidence is lower in African American women, case-fatality rates are higher compared to white women, in part explained by poorer prognostic characteristics of breast cancers in African American women and differences in access to health care services and differences in the quality of care that is delivered once access is attained (Cheblowski et al., 2005; Richardson et al., 2005). Based on the Canadian Community Health Survey, visible minorities access cancer screening services significantly less than Caucasian Canadians (Quan et al., 2006). Although incidence may be lower in these groups compared to Caucasian Canadians, their survival could be poorer due to later stage diagnosis potentially due barriers to access to health care services. According to research findings from the First Nations Cancer Research and Surveillance Priorities for Canada Workshop Report, cancer has historically occurred at a lower rate in Aboriginal populations; however rates are increasing and dramatically in some regions of Ontario for some cancers, including breast cancer (Cancer Care Ontario, 2004). The survival rate for breast cancer is significantly worse in First Nations women compared to the Ontario population. One plausible explanation for this higher case-fatality rate is later stage at diagnosis among First Nations women.

Breast Cancer Detection and Diagnosis in Canada

Mammography is a technique that uses an x-ray to detect breast cancer at an early stage. Organized, population-based mammography in Canada is offered through government sponsored programs at dedicated screening centres at no (out-of-pocket) cost and without physician referral for women within the target age group. Distinct from "organized screening", opportunistic screening refers to mammograms conducted upon referral. Unlike organized screening, under opportunistic screening, women must be referred by a health care provider (usually a general practitioner physician) to a radiologist for testing and who is then reimbursed on a fee-for-service basis by the provincial health care system.

After a lump has been detected by a screening mammogram more definitive tests for diagnosis of breast cancer are usually arranged (Public Health Agency of Canada, 2008). Tests include imaging studies, biopsy, and laboratory tests. Imaging studies allow tissues, organs and bones to be viewed in more detail regarding the size and spread (metastasis), if any, of the possible tumour. Imaging studies can include x-rays, ultrasounds, CT scans or bone scans, and diagnostic mammogram (in addition to the screening mammogram). Biopsy is usually needed to make a definite diagnosis of cancer. This process involves removing cells from the body and examining them under a microscope. Fine needle aspiration, core needle biopsy, and surgical biopsy, are methods used to perform breast biopsies. Laboratory tests are performed on the breast tissue after cancer cells are found in the biopsy sample. These tests allow the physician to know more about the cancer and to plan the best treatment for the patient. Laboratory tests include: hormone receptor status tests (estrogen and progesterone), HER2 test, and sometimes blood tests (shows how well the organs are working, whether there is cancer, and if it has metastasized). All of these tests may also be used to stage and grade the cancer (Public Health Agency of Canada, 2008).

Once cancer is diagnosed, the tumour is assigned a stage and grade, which helps determine the type of treatment required. The cancer stage describes tumour size and indicates whether the tumour has spread from its original location (Canadian Cancer Society, 2014). In the earliest stage of breast cancer, cancer cells are found only in the milk ducts or lobules (in situ cancer). If there is a correct diagnosis of in situ cancer before the cells spread to the surrounding tissue, then the risk from that tumor should be eliminated once it is surgically

removed. When breast cancer spreads out of the duct or lobule, it is then considered invasive cancer, and can at this point still be treated effectively if diagnosed early. The currently accepted staging scheme for breast cancer is the TNM (Tumour Nodes Metastases) classification, which includes several stages (Canadian Cancer Society, 2014). These stages are described below.

Stage 0: includes two types: ductal carcinoma in situ (DCIS) where abnormal cells lining of a milk duct and have not spread outside the duct and lobular carcinoma in situ (LCIS) where abnormal cells are in the lining of a lobule.

Stage I: tumour is 2 cm or smaller and the cancer has not spread outside the breast.

Stage IIA: T0, N1, M0 – no tumour in breast, cancer found in 1-3 axillary lymph nodes and no cancer spread to distant sites. T1, N1, M0 – tumour is 2 cm or less in diameter, cancer spread to 1-3 axillary lymph nodes and/or internal mammary lymph nodes, and cancer not spread to distant sites. T2, N0, M0 – tumour larger than 2 cm and less than or equal to 5 cm in diameter, cancer not spread to distant sites.

Stage IIB: T2, N1, M0 – tumour larger than 2 cm and less than or equal to 5 cm in diameter, cancer spread to 1-3 axillary lymph nodes and/or internal mammary lymph nodes, cancer spread to distant sites. T3, N0, M0 – tumor more than 5 cm in diameter, cancer not spread to lymph nodes, and cancer not spread to distant sites.

Stage IIIA: T0, N2, M0 – no tumour in breast, cancer in 4-9 axillary lymph nodes or in internal mammary lymph nodes, and cancer not spread to distant sites. T1, n2, M0 – tumour 2 cm or less in diameter, cancer spread to 4-9 axillary lymph nodes or to internal mammary lymph nodes, and cancer not spread to distant sites. T2, N2, M0 – tumour larger than 2 cm and less than or equal to 5 cm in diameter, cancer spread to 4-9 axillary lymph nodes or to internal mammary lymph nodes, and cancer not spread to distant sites. T3, N1/2, M0 – tumour tumor is more than 5 cm in diameter, cancer spread to 1-9 axillary lymph nodes or to internal mammary lymph nodes, and cancer not spread to distant sites.

Stage IIIB: T4, N0/1/2, M0 – tumour spread to chest wall or skin, cancer not spread to any lymph nodes or cancer spread to 1-9 axillary lymph nodes or cancer may/may not have spread to internal mammary lymph nodes, and cancer not spread to distant sites.

Stage IIIC: T1/2/3/4, N3, M0 – tumour any size, cancer spread to 10 or more axillary lymph nodes or cancer spread to 1 or more infraclavicular or supraclavicular lymph nodes or cancer spread to more than 3 axillary lymph nodes and to internal mammary lymph nodes, and cancer not spread to distant sites. T4d – inflammatory breast cancer cliassifed as stage III unless has spread to distant sites or lymph nodes far from breast, in which case it is stage IV.

Stage IV: T1/2/3/4, N1/2/3, M1 – tumour is any size, any degree of lymph node involvement, and cancer spread to distant sites e.g. bone, liver, lung, brain or lymph nodes far from breast.

It should be noted that for stage 0 tumours, Canada only retains surveillance information for cases diagnosed as DCIS, as LCIS is considered a benign disease (i.e. not an outcome of interest). Tumour grade is determined by examining a biopsy sample under a microscope and comparing the look and behaviour of tumour cells compared to normal cells (Canadian Cancer Society, 2014). Grade 1 tumours are low-grade and have cells that are slow-growing and less likely to spread. Grade 2 tumours are moderate grade. Grade 3 tumours are high grade and tend to grow quickly and are more likely to spread.

General Principles of Screening

Tumours detected by screening tend to be at an earlier stage of their development than those detected otherwise – this is referred to as stage shift (Shen et al., 2005). Survival benefit from early detection is reflected through improved stage distributions. However, stage shift does not account for all improved survival since improved survival can be obtained for cancers that are diagnosed at the same stage (within stage shift). For example, a cancer that is detected at early stage II node negative may have better prognosis than if detected at later stage II node negative. Early detection can occur when screening takes place during the sojourn time (time spent in the preclinical phase -detectable but no symptoms present- of the natural history of disease) (Boer et al., 2004). Screening is defined as testing for a condition when the person has no overt signs or symptoms of that condition. However, the purpose of screening is not to detect a disease for detection's sake, but rather to improve prognosis. The detection of earlier disease on its own is not enough to justify a screening program as it must

also demonstrate an additional benefit of extended life and quality of life due to earlier detection (Harris et al., 2006).

There are a number of factors that determine the appropriateness of adopting a population-based screening program. These include: 1) burden of suffering from the target population – incidence and prevalence of disease, mortality, morbidity, costs of care; 2) accuracy of the screening test in detecting early-stage disease i.e. sensitivity and specificity of the test; 3) costs and suffering caused by the test; 3) effectiveness of early detection in improving outcomes and the trade-off between benefits and potential harms; 4) resource constraints; 5) ability of clinicians, patients, and the healthcare system to implement a screening program i.e. feasibility, acceptability, uptake, patient education, adherence, and follow-up; 6) philosophical and moral objections; and 7) opportunity costs i.e. does screening displace resources needed for other more effective health care services?

Characteristics of an effective organized screening program include: the presence of a valid and acceptable screening test; the earlier and efficient diagnosis of the disease; minimal diagnosis of non-progressive disease (low number of false-positives); the availability of effective therapy for the detected disease and; favourable compliance of the at-risk population with screening.

Organized and Opportunistic Screening Mammography in Canada

For almost three decades in Canada, screening mammography for the early detection of breast cancer has been considered an important technology in reducing mortality from the disease. In Canada, women can participate in screening as part of an organized screening program or are screened opportunistically.

Organized population-based breast cancer screening programs have been implemented in most provinces and territories across Canada since 1988. As before mentioned, the only jurisdiction to not have an organized breast cancer screening program to date is Nunavut. Table 2 provides an overview of usual practices of breast cancer screening programs in Canada by province/territory for the 2007 and 2008 screen years. Typically, there are three stages of a breast cancer screening program: 1) identification and invitation of the target

population; 2) provision of the screening examination and; 3) if an abnormality is detected, further investigation. Women within the target age for the program are recruited through a letter of invitation, a physician referral or self-referral. The target or status quo age range for screening is 50-69 years as the benefits of regular screening mammography in reducing breast cancer mortality have been most strongly demonstrated for this age group (Kerlikowske et al., 1995; de Koning et al.,1995). The targeted screening population is characterised as asymptomatic women between the specified ages according to the screening policy, with no prior diagnosis of breast cancer; however all programs across Canada screen some women, although not actively recruited, outside the target age group. Screening facilities may be either a mobile unit or a fixed site. Women who do not have any breast cancer symptoms receive two-view mammography of each breast. In addition, some programs also offer clinical breast examination (CBE) performed by a trained health professional (nurse or technologist). The pathway by which a woman traverses through a screening program is diagrammed in Figure 3.

Opportunistic screening is screening that is offered outside a screening program. Across all Canadian provinces and territories, women can access screening outside the organized program by being referred by a general practitioner or family physician. The proportion of all breast cancer screening in any given province or territory that is opportunistic or organized screening varies across all jurisdictions. The usual method of funding screens outside of programs is via fee for service billing with specific billing codes for each mammogram performed for the purpose of screening. Billing code is the only way in which opportunistic screens can be tracked. A drawback of this type of administrative data collection is the lack of consistent tracking over time and the collection of demographic and other detailed information on each patient who receives screening. In addition, women who access screening opportunistically may less likely to adhere to recommended screening guidelines that is screening at appropriate intervals, due to the lack of reminder systems at the primary care level. In figure 2, the process by which the moderate and higher risk population are screened for breast cancer in Canada.

Table 1 Breast cancer screening programs in Canadaa - usual practices, 2007 and 2008 screen years

Province/Territory	Program start date	Clinical breast examination on site	Program practices for women outside the 50-69 age group		
			Age group	Accept	Screen Frequency
Northwest Territories	2003	No	30-39 40-49 70+	No Yes Yes	N/A Annual Biennial
Yukon Territory	1990	No	30-39 40-49 70+	No Yes Yes	N/A None Biennial
British Columbia	1988	No	30-39 40-49 70-79 80+	Accept with physician referral Yes Yes Accept with physician referral	None Annual Biennial None
Alberta	1990	No	30-39 40-49 70-74 75+	No Yes Yes Yes	N/A Annual Biennial None
Saskatchewan	1990	No	30-39 40-49 70-74 75+	No No[b] Yes Yes	N/A N/A Biennial None
Manitoba	1995	No[c]	30-39 40-49 70+	No Accept to mobile unit with physician referral Accept to mobile unit with physician referral	N/A Biennial None
Ontario	1990	Yes[d]	30-49 70-74 75+	Accept high risk women with physician referral who meet the eligibility criteria Yes Yes	Annual Biennial None
Quebec	1998	No	30-34 35-49 70+	No Accept with physician referral[e] Accept with physician referral[e]	N/A None None
New Brunswick	1995	No	30-39 40-49 70+	Accept high risk women with physician referral who meet the eligibility criteria Accept with physician referral Accept with physician referral	N/A None None
Nova Scotia	1991	Yes[f]	30-39 40-49 70+	No Yes Yes	N/A Annual None
Prince Edward Island	1998	No	30-39 40-49	Accept high risk women with physician referral who meet the eligibility criteria Yes	Annual Annual

			70-74	Yes	Biennial
Newfoundland and Labrador	1996	Yes[g]	30-49 70+	No Accept if previously enrolled in program	N/A None

Source: Canadian Partnership Against Cancer, "Organized Breast Cancer Screening Programs in Canada, 2007-08"

a Nunavut has not developed an organized breast cancer screening program.

b Accept age 49 on the mobile if they would be 50 in that calendar year.

c Nurse or technologist provided CBE service until October 2005.

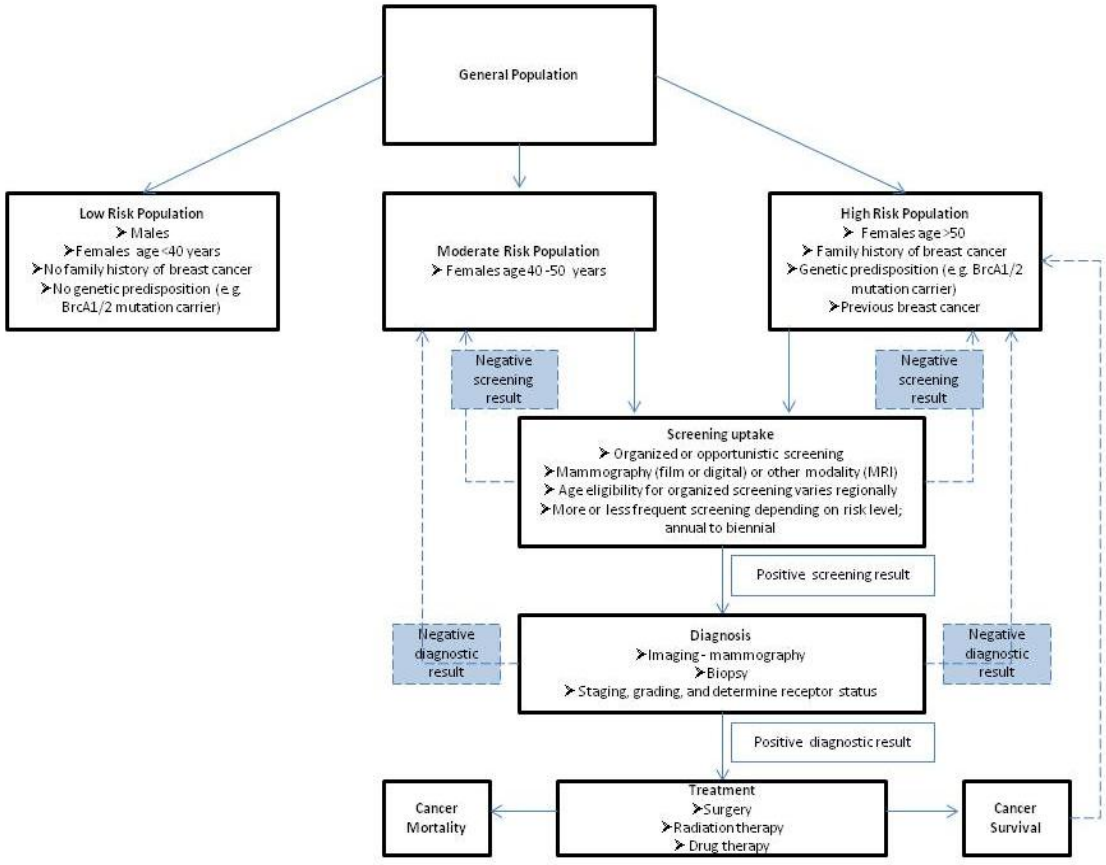
d Nurse provides clinical breast examination at 52 per cent of sites.

e Accept with physician referral if done at a program screening centre, but is not officially considered within the program.

f Modified examination only, performed by technologist at time of mammography.

g Nurse.

Figure 2 Conceptual Model of Breast Cancer Screening in Canada



Screening Mammography Utilization or Participation

Targets for screening participation have generally been reported at 70 per cent for a first screen and 90 per cent for subsequent screens among the target population age 50-69 years (Public Health Agency of Canada, 2007). For the year 2005-06, actual screening mammography utilization (opportunistic and organized program screens combined) did not reach this level, rather the rates varied slightly across provinces from 60.1 per cent in British Columbia to 64.6 per cent in Quebec. (Public Health Agency of Canada, 2011). The variation in the proportion of utilization attributable to organized screening was greater, from 14.5 per cent in Alberta to 85.2 per cent in British Columbia (Public Health Agency of Canada, 2011).

Benefits and Harms of Screening Mammography

The main benefit and principle for breast cancer screening is that it can prevent breast cancer mortality. Mammography screening has been found to reduce breast cancer mortality as much as by 10 per cent, according to a recent study (Kalager, 2010). As mentioned before, breast cancer screening is assumed to reduce mortality through stage shift. We assume that having adequate screening compared to a situation where there was little or no screening, will incur more early stage breast cancers being detected than late stage thus improving prognosis of the disease, as well as reduced treatment and time spent in recovery. There are also perceived benefits of screening which relate to beliefs about the positive outcomes associated with behaviour in response to a real or perceived threat (Champion, 2008) The perceived benefit construct is most often applied to health behaviours and is specific to an individual's perception of the benefits that will accrue by engaging in a specific health action. Champion writes that perceived benefits of screening mammography includes a woman's beliefs about the benefits of obtaining a mammogram. The perception of benefits is theoretically related to the woman's beliefs about her own individual health outcomes and not those that might be experienced by others. However, it could also be true that a woman may feel that mammography would help find breast cancer early for others but not necessarily believe it would do so for herself, and in this case, may be less likely to participate in screening mammography. This perceived benefit construct is included in

several health behaviour models including the Health Belief Model, the Transtheoretical Model, Protection Motivation Theory, and the Theory of Planned Behaviour (Hochbaum, 1958; Velicer, 1985; Maddux, 1985; Shifter, 1985).

The potential harms of screening include false negative test result, false positive test results, over-diagnosis and attendant over treatment, and iatrogenic radiation exposure. A false negative result is a test result that indicates that a person does not have a specific disease or condition when the person actually does have the disease or condition (National Cancer Institute, 2014). False-negative results are more frequent among younger women (under age 50) than older women because younger women tend to have denser breasts. Mammograms are less sensitive for women with dense breasts and as a woman ages, her breasts usually become fattier and false-negative results become less likely. The most important issue with false-negative results is that they can lead to a false sense of security, delays in treatment, and potentially premature death and reduced quality of life. Given-Wilson et al. (1997) found this to be true when they studied the causes and consequences of false negative mammograms among 29 women in the UK. They found that negative mammogram reports were significantly associated with being under the age of 50 as well as premenopausal status, and breast density. A negative mammogram resulted in a significant delay to definitive treatment where the median time to treatment for negative versus positive mammography was 7 weeks versus 3 weeks. In addition, there was significant clinical progression in over half of the women whose treatment was delayed.

A false-positive is a test result that suggests a person has a specific disease or condition when the person actually does not have the disease or condition. This usually occurs at the level of reading where a radiologist will call a screen as abnormal, but at diagnosis there is no cancer. False-positive results are also more common for younger women, women who have had previous breast biopsies, women with a family history of breast cancer, and women who are taking estrogen, such as in hormone replacement therapy. According to Miller Croswell et al. (2009), the cumulative risk of a false-positive test in women after 14 screens (which might happen if one is screened on an annual basis starting at age is 40 to 54, for example) is 48.8 per cent with just under half (45 per cent) of those false positive cases resulting in an invasive diagnostic (Miller et al., 2009). A systematic review by Armstrong et

al. (2007) examined cumulative risk of a false-positive among women screened between the ages 40 to 49; finding a cumulative risk of between 20 per cent and 56 per cent after 10 mammograms.

False-positive mammograms can be detrimental as they can lead to anxiety and other forms of psychological distress in affected women. In addition, the workup and extra tests involved in confirming diagnosis can be costly, time consuming, and physically uncomfortable. A recent meta-analysis by Salz et al. (2010) on the impact of false positive mammography found that having a false-positive mammogram result is associated with greater anxiety and distress about breast cancer as well as more frequent breast self-exams and higher perceived effectiveness of screening mammography (Salz et al., 2010). False positives were associated with the generic outcome of generalized anxiety; however this effect size was small. Research suggests that people are willing to receive false-positive results from cancer tests and even expect it (Schwartz et al., 2004).

There has been much debate over the extent of over-diagnosis of breast cancer, particularly ductal carcinoma in situ (DCIS), as a result of screening. (Zackrisson et al., 2006) It has been estimated that one in eight women would not have had their breast cancer diagnosed if they had not gone for screening (Advisory Committee on Breast Cancer Screening, 2006). There are some researchers that believe over-diagnosis is likely to cause more harm than benefit, while others take the position that high detection of DCIS represents a substantial proportion of invasive cancers avoided (Mitra et al., 2000; Cady et al., 1998). As a consequence, the proportion of DCIS among screen-detected cancers is often used as an indicator representing benefit or harm when comparing screening programs, depending on the position taken (Yen et al., 2003). It is still unclear however what the recommended prevalence of DCIS should be and what proportion of detected cases of DCIS would actually progress to invasive disease. It is generally agreed however that a proportion of DCIS will not have progressed to invasive carcinoma of breast in the absence of screening. In Canada, DCIS is diagnosed as a stage in cancer and treated when found. As such, there are some who argue that the rate of over-diagnosis in Canada is high because DCIS is screened for and treated.

Factors Related to Screening Mammography Uptake

According to a review by Schueler et al. (2008), who studied factors associated with mammography utilization, the influential factors include physician access barriers, past screening behaviour, economic, ethnicity, and women's screening knowledge. In this review literature on factors associated with receipt of mammography, including data sources in English and published between 1988 and 2007 was synthesized. Included were 221 studies that described results for almost 5 million women internationally. A calculation of odds ratios (OR) associated with mammography use and a random effects model was used to assess trends in mammography utilization and to calculate summary multivariate point estimates. It was found that physician access barriers such as not having a physician-recommend mammography and having no primary care provider were strongly related to not obtaining a mammogram. Past screening behaviour such as clinical breast examination and Pap test were correlated strongly with receipt of mammography. The only socioeconomic factor that was found to be correlated with access to mammography was having no insurance. There was an ethnic variation in regards to concerns regarding cost, mammography safety, and pain, which were more important to African American and Latina women, and having no insurance was more important to white and Chinese women. Cost concerns and the presence of a family history of breast cancer were less important to older women, whereas screening knowledge had a stronger impact on mammography use in women aged 65 years and older. The authors also found that mammography uptake increased over time.

Breast Cancer Screening Modalities

Simplified methods of screening for breast cancer include breast self-examination (BSE) and clinical breast examination (CBE). Kosters and Gotzsche found no improvement in breast cancer mortality rates in those screened using BSE and CBE compared with a no screening control group. They reported that the screening group resulted in more biopsies (indeed twice the number under no screening). Their data suggests that the doubling of biopsies may have had no salutary impact on breast cancer mortality.

Mammography is the most common technology used to screen for breast cancer in population. It is an imaging technique that uses x-rays to provide an image of the internal structure of the breast and can show abnormal growths or changes in breast tissue before they become symptomatic or clinical. Mammograms can be done for diagnostic or screening purposes. Based on a meta-analysis, mammography has a true-positive rate of 83 per cent to 95 per cent and a false positive rate of 0.9 per cent to 6.5 per cent (Mushlin et al., 1998). If a suspicious lesion is identified on a mammogram, other techniques may be used for further investigation, including ultrasound, biopsy magnetic resonance imaging (MRI) and laser scanning (Health Canada, 2011). The sensitivity (the proportion of diseased subjects who test positive with the screening test) and specificity (the proportion of healthy subjects who test negative with the screening test) of mammography are affected by breast density, which in turn may partly be affected by age, hormone replacement therapy (HRT) use, parity, body mass index, and family history or genetic predisposition (Carney et al., 2003). The sensitivity of mammography is much lower in women with radiographically dense breasts, with values ranging from 62.9 per cent in extremely dense-breasted women to 87 per cent in women with breasts of higher fat composition, whereas specificity values ranged from 89.1 per cent to 96.9 per cent, respectively. Despite mammography being the gold-standard for breast cancer screening, there are certain limitations with the technology. For example, mammograms require many resources, including a dedicated machine, radiologic film and developing chemicals (in the case of film mammography), a trained x-ray technologist, and a skilled radiologist to assess the image (Nover et al., 2009). More recently, full-field digital mammography (FFDM) has been used for screening in several organized screening centers. Rather than recording an image on film, FFDM records an image in an electronic file. FFDM has a higher sensitivity and specificity compared to film mammography. Although the technology is 10 to 40 times more expensive to purchase compared to film screen mammography units, there is a potential for cost-savings in regards to increases in efficiency, effectiveness, and reduction in resource requirements. According to study by Pisano et al. (2005), FFDM is more accurate in women age less than 50 years with radiographically dense breasts, and premenopausal or peri-menopausal women. The technology itself may cause patient discomfort and, as before mentioned, the imperfect accuracy of the technology and/or the reading may lead to unnecessary biopsy. Other

drawbacks of mammography include exposure of the breast to radiation which may cause radiation carcinogenesis. Despite these limitations, mammography is still recommended for use in breast cancer screening, depending on the age group, frequency of screening, and other target population characteristics, as the benefits have been considered to outweigh the harms.

The American Cancer Society recommends annual magnetic resonance imaging (MRI) screening for individuals with BRCA1/2 mutations, those having a first-degree relative with a BRCA1/2 mutation, or those with a lifetime risk for breast cancer of 20 per cent to 25 per cent. The sensitivity of MRI in visualizing invasive cancer is nearly 100 per cent, yet specificity varies (Orel and Schnall, 2001).

The technologies that have been reviewed in this section are only but a few of the possible technologies that have been developed over time and are being currently developed. The review by Nover et al. (2009) provides a more exhaustive review of the technologies that are currently available, not all in Canada however, for breast cancer screening.

Breast Cancer Treatment

Over the past couple of decades, there have been significant gains made in systematic treatment of breast cancer, especially in advanced breast cancer, including the introduction of new chemotherapeutic agents (O'Shaughnessy et al., 2002; Thomas et al., 2007). The treatment regimen for breast cancer is dependent upon the stage at which the cancer is diagnosed as well as specific characteristics and wishes of the affected person. The common treatments for breast cancer include surgery, radiation therapy, chemotherapy, hormone therapy, and biological therapy (Canadian Cancer Society, 2011).

Surgery and surgery with radiation therapy is often considered primary treatment of breast cancer. Chemotherapy, hormone therapy, and biological therapies are often considered adjuvant treatments for the disease. In terms of surgery, there are two different types: 1) lumpectomy (breast-conserving surgery), which is the removal of a lump and some tissue, but not the whole breast, and 2) mastectomy, which is the removal of the whole breast. The choice between the two is based on personal choice. Lumpectomy is less invasive than mastectomy; however several factors may influence an affected woman's choice. For

example, there is a lower chance of recurrence of cancer with a mastectomy compared a lumpectomy; however a lumpectomy has a better cosmetic result compared to a mastectomy (Canadian Cancer Society, 2011).

Radiation therapy is the use of radiation (beams) to kill tumour cells. During this process, normal cells may also be affected. Chemotherapy is the use of specific drugs that interfere with the ability of cancer cells to grow and spread, however this treatment may also affect healthy body cells. Some chemotherapy drugs have been known to interfere with the ability for a woman to become pregnant, therefore some women may opt for a treatment plan that may not involve chemotherapy, such as mastectomy without chemotherapy. Hormone therapy is a treatment that removes hormones from the body or blocks their action and stops cancer cells from growing. This treatment is often used on women who have been found to have a tumour that is hormone receptor positive(Canadian Cancer Society, 2011).

More recent research has shown that pharmacogenetics may play a significant role in breast cancer therapy. Pharmacogenomics involves the study and use of information regarding how a person's genetic profile can affect their response to treatments. It combines pharmacology and genomics to develop to allow for planning and administration of specific treatment regimens that are tailored to the individual. Although it is an emerging field of study, it is not yet widely applied in practice. In their recent study on pharmacogenomics of breast cancer therapy, Westbrook and Stearns (2013), concluded that ongoing prospective studies and increasing understanding of pharmacogenetics will assist in better predicting the risks of toxicity or probabilities of response to specific treatments as well as to provide more personalized therapy for women diagnosed with breast cancer.

Health Economic Evaluation to Inform Breast Cancer Screening Policy

Health economic evaluations are important because resources, including people, time, facilities, equipment, and knowledge, are scarce. Choices have to be made on how they are allocated and used. The cost-effectiveness of any intervention (treatment, technology, program, policy, etc.) can be assessed using an economic evaluation. Economic analyses compares both the costs and consequences of interventions. Types of economic evaluations include cost-effectiveness analyses, cost-benefit analyses, and cost-utility analyses.

In cost-effectiveness analyses (CEA) the health effects are measured in natural units related to the objective of the intervention, such as cases of disease averted, lives saved, or life-years gained. CEA is most widely used in situations requiring a choice under the constraint of a budget. In the context of breast cancer screening, a program director may be interested in the cost per case detected, however, this type of health outcome is often considered only intermediate and not directly linkable to the overall objective of breast cancer screening which is to reduce premature mortality. Cost-effectiveness analyses often use the unit of analysis of incremental cost per life-year gained or an incremental cost-effectiveness ratio (ICER), which is the ratio of the difference in the cost of the two health intervention options being compared to the difference in the health consequences of the two health intervention options being compared.

In cost-utility analyses (CUA), the evaluation focuses on the quality of the health outcome produced or forgone by a health intervention. In CUA, the incremental cost of an intervention is compared to the incremental health improvement attributable to the intervention, where the health improvement is measured in quality-adjusted life years (QALYs) gained, or some variant such as disability-adjusted life-years (DALYs) gained. Cost-benefit analyses (CBA) require program consequences to be valued in monetary units. This type of analysis is not as frequently performed as CEAs or CUAs due to the challenge of putting a dollar value on a human (statistical) life. The benefit of CBA is that it enables the direct comparison of the intervention's incremental cost with its incremental consequences in the same units of measurement. Table 3 provides a general comparison of CEA, CUA, and CBA, adapted from the work by Kaplan et al. (2002).

Table 2 Comparison of the three main health economic evaluation approaches: Cost-effectiveness analysis, cost-utility analysis, and cost-benefit analysis

Type of analysis	Compares	With	Strengths and limitations
Cost-effectiveness	Monetary value of resources used	Health effects: clinically based (death rate, blood pressure, test performance, lives saved, life-years gained)	Relatively easier to measure and interpret. Not comprehensive; not based on consumer preference for health states.
Cost-utility	Monetary value of resources used	Health effects: preference based (health-related quality of life)	Comprehensive; based on consumer preference for health states. Difficult to elicit health utilities.
Cost-benefit	Monetary value of resources used	Monetary value: resources saved or created, including health	Places monetary value on lives and the quality of life. Can shift focus away from health improvement. Can be difficult to interpret and measure.

Equity

In terms of breast cancer mortality, although certain groups may not have as high breast cancer incidence as others, they may actually experience differential mortality from the disease as a result of some type of inequity, usually in the form of access to preventive and/or curative care.

Economic evaluations of health interventions often ignore heterogeneity and equity by using average or representative values. Ignoring that the population is heterogeneous and made up various sub-groups with respect to the risks of having disease can lead to optimistic or pessimistic estimates of cost-effectiveness ratios (depending on the intervention), and sometimes the parameter values. (Zaric et al., 2003). Zaric et al. (2003) recommends that all information that is known, including information about the distribution of heterogeneous characteristics should be incorporated into economic evaluations in order to avoid biased gains in life expectancy and impacts on resulting cost-effectiveness ratios. However, this approach does not allow identification of the optimal treatment strategies within subgroups. Another alternative is to perform stratified analysis by subgroups of the population, recognizing that cost-effectiveness is dependent on people, place and to some extent time (Coyle et al., 2003). This method allows for the consideration of a direct trade-off between concerns for equity and for efficiency.

Health equity has long been an objective of public health policy worldwide; however economic evaluations in public health continue to focus more on efficiency and improving total population health, while typically ignoring explicit consideration of health inequalities. It has been proposed that "distributional weights" might be applied to monetary costs and benefits to reflect decreasing marginal individual value of income (that a gain or loss of a certain amount of money matters more to a poor individual than a rich individual). In focussing on an individual's income however and converting that into individual value, there is little consideration of the social value of improving health of different individuals, such as the poor who generally experience lower life expectancy compared to that of the wealthier. Analysis by vulnerable or disadvantaged subgroups is one method of examining the inequalities across sub-populations. There are approaches to incorporating equity considerations which include: reviewing background information on equity; health inequality impact assessment; opportunity cost analysis of equity, and; equity weighting of health outcomes (Cookson et al, 2007).

Another method of exploring equity in economic evaluation of health interventions is evaluating the trade off between equity and efficiency using the "net benefit framework for cost-effectiveness". “Limited use criteria” (LUC) is a policy where decision makers restrict public funding for healthcare to a subgroup of the population for whom it can be used with the objective of improving value for money. The work by Coyle et al. (2003) demonstrates how the "net benefit framework for cost-effectiveness" can be used to estimate the efficiency gains from stratification, inherent in LUC, based on heterogeneity between patients in terms of costs, outcomes or both. “Net health benefit” is the net benefit (measured in units of health) of investing resources in an intervention rather than investing those resources in a marginally cost-effective program (Stinnett, 1998). This procedure enables decision makers to explicitly examine the trade off between equity and efficiency by assessing the opportunity cost of an equity position. Coyle et al. (2003) note that the more subgroup stratifications possible, the greater the opportunity for efficiency gains. Therefore if stratification is rejected by decision makers based on equity reasons, there will be an associated opportunity cost that can be expressed as a “net benefit loss” or “net benefit reduction”. For instance if there was one possibility for stratification, such as by age, the loss in total net benefit is the difference between the total net benefit of stratification by age and

the total net of no stratification. If a decision maker chooses not to stratify, the loss in total net benefit represents the minimum willingness to pay for equal access to the intervention regardless of the variable of stratification, in this example by age. Although equity was not directly evaluated within the cost-effectiveness analyses in the subsequent dissertation chapters, it will be addressed as a discussion piece within the conclusion chapter. As previously stated, although risk of breast cancer does not appear to differ between socially and economically disadvantaged groups there is some evidence to show variation in mortality. For example, research from the US found women of lower socio-economic status have higher breast cancer mortality rates compared to those in higher socio-economic status (Sprague et al., 2010). It is uncertain whether this inequity exists in Canada, which has a more social health care system compared to the more private health care system in the US.

Current Debate and Next Steps

Since its inception, screening mammography for masses has received little opposition from the general public. However, discussions surrounding the true benefits and harms of screening mammography have emerged over time. The advancement of treatment, the reduction in other risk factors such as hormone replacement therapy use, and more women taking control over their individual health, has resulted in improved breast cancer survival. As we previously mentioned, in 2010 Kalager et al. reported a 10 per cent reduction in breast cancer mortality attributable to screening mammography. This was a disappointing result according to the authors who expected a reduction of 20 per cent or more. Other researchers have publicly questioned the effectiveness or value for money conferred by population-based screening mammography based on certain claims of harms outweighing benefits, including excessive use of lumpectomies, mastectomies, and radiotherapy, high rate of false positive tests, and over-diagnosis.

In the fall of 2009, the U.S. Preventive Task Force updated their screening mammography guidelines by advising screening on a biennial basis for women aged 50-74 only (U.S. Preventive Services Task Force, 2009). This garnered much displeasure among women's groups who have argued that women aged 40-49 should also be screened, despite a lack of evidence for success or cost-effectiveness to support screening for this age group

(Rosenquist et al., 1994). Others argue that mammography is effective in reducing breast cancer mortality in countries such as Canada that have relatively high incidence. They point out that a 10 per cent reduction in disease-related mortality is a considerable benefit. The question that still remains, however, is whether this magnitude of effect is worth the associated costs. Trade-offs between the benefits, harms, and costs associated with various screening guidelines should be considered when making recommendations for routine screening. There is also an issue of resource capacity; a recommendation in which more women are to be screened on a more frequent basis will increase backlog and result in longer wait-times for all women, including those who are at increased risk. The mainstream media has to some extent picked up on this debate regarding a concern over harms versus benefits of screening (Salahi, 2010).

Over time there have been a number of important shifts in the way women are screened within organized programs in Canada, and these policies vary regionally. For instance, the program in British Columbia actively screens women on self-referral who are aged 40-49 annually, and women aged 50-79 biennially (Public Health Agency of Canada, 2011). This province also accepts women under 40, provided that they have a referral from a physician. In contrast, Ontario only actively screens women aged 50-74 on a biennial basis. In addition, some provinces are phasing out the use of analog or film mammography for digital mammography, which has been found to be more sensitive in picking up true cancers as opposed to false positives (suspected cancers after screen that are negative at diagnosis) (Pisano et al., 2005). These varying policies have significant impacts on a number of outcomes, including the ability for a program to obtain adequate coverage of the at-risk population, wait-times, and costs related to screening, diagnosis, and treatment (Gunes et al., 2004).

The most recent Canadian Task Force on Preventive Health (2011) breast cancer screening guidelines was almost identical to those published by the U.S. Preventive Task Force. Based on a systematic review of the literature the Canadian Task Force recommended the following for Canadian women of average risk:

- Age 40–49 years: no routinely screening with mammography. (Weak recommendation; moderate-quality evidence)

- Age 50–69 years: routine screening with mammography every two to three years. (Weak recommendation; moderate-quality evidence)
- Age 70–74 years: routine screening with mammography every two to three years. (Weak recommendation; low-quality evidence)

The pertinent concern that needs to be addressed is why there is so little consensus around population-based screening mammography. This is most likely due to the lack of strong evidence available to support the current practices in terms of effectiveness and efficiency. Other considerations include the assessment of the potential impact of longer screening intervals for women of moderate risk, such as screening every 3 years, or the impact of tailored screening for women at high-risk. The high-risk category would be comprised of women according to age, as well as other high risk factors including family history and/or genetic predisposition. We must also consider the impact of screening vulnerable sub-groups of the population, including women with mental and physical disabilities who face challenges with not only accessing preventive care, but also accessing the health care system in general. Within the context of a publically funded health care system, decisions regarding which services should or can be funded, and by how much, are particularly difficult to make.

To date, there have been very few studies that assess the efficiency or cost-effectiveness of population-based screening mammography in Canada. Decision-makers require sound evidence to support these difficult choices and therefore it is essential that we do not accept the current state of affairs and justify activities based on what has been done in the past. Rather, time should be invested to periodically evaluate these programs to ensure that the benefits outweigh the harms, and that the related costs are reasonable or within society's willingness to pay. The subsequent chapters of this dissertation will explore the important considerations in determining the impact of age, screening frequency, and disease risk on the cost-effectiveness of screening mammography for the secondary prevention of breast cancer in Canada through: 1) a review of the literature (cost-effectiveness models and health economic evaluation studies), 2) a cost-effectiveness analysis to predict the impact of age and screening frequency on the cost-effectiveness of mammography in Canada, and 3) a sub-group analysis of the impact of disease risk on the cost-effectiveness of screening mammography in Canada.

Chapter 2: Review of Quantitative Models for the Measure of Health and Economic Impact of Screening Mammography

Preamble

This chapter describes different types of models used internationally to evaluate the impact of age, screening frequency, and disease risk on the effectiveness and cost-effectiveness of screening mammography in the secondary prevention of breast cancer. Simulation modeling is a general term for the creation of a hypothetical representation of a real system or real world and these population health models represent entire or general populations of people, including people who are healthy and sick (Simulation Technology for Applied Research Glossary). This chapter provides an inventory of different models used to evaluate the effectiveness of screening mammography with a focus on microsimulation models. The purpose of this exercise was to identify a model that could be used to study the cost-effectiveness of screening mammography in Canada.

Previous Health Economic Evaluation Studies on Screening Mammography in Canada

At the time of this research, only one published study evaluating the health and economic impact of screening mammography in Canada was identified. This study by Hunter et al. (2004) only reported on intermediate outcomes i.e., cost per breast cancer detected as opposed to cost per life-year gained or quality-adjusted life-year gained, to assess the costs and benefits of screening women ages 40-49 years in Ontario, Canada. The study cannot be considered a full economic evaluation, but rather a partial economic evaluation, since there was no direct comparison of alternatives in this study. The investigators therefore could not conclude based on study results whether screening women between the ages of 40-49 is cost-effective compared to some alternative, such as usual care or compared to no screening.

Generally, many published economic evaluations of screening mammography are either cost-effectiveness or cost-utility analyses from the United States (US), United Kingdom (UK), Netherlands, France, and other European and Asian countries (see chapter 3 for the systematic review conducted in this dissertation). The comparability across studies remains a

challenge as most if not all studies do not provide adequate information on their methods used and model assumptions. When information is provided studies differ considerably in either design, population characteristics, or methods used, meaning that efforts to combine findings, such as in meta-analysis form, or to generalize results from one context to another is very difficult. To date systematic reviews of economic evaluations have taken the form of narrative reviews as opposed to meta-analyses due to these challenges in combining unlike data.

Many of the published health economic evaluation of screening mammography studies concluded that, in general, population-based screening mammography would be cost-effective compared with no screening. However, none of those studies evaluated the gap in screening performance between recommended guidelines for screening mammography and screening as it has been implemented in Canada historically because none included the observed screening dissemination as a basis for comparison.

Since no appropriate health economic studies were found for Canada, a systematic review was conducted to assess the impact of age, screening frequency, and targeted screening for subgroups of the population, on the cost-effectiveness of screening mammography. Chapter 3 of this dissertation is a systematic review of published studies that examine the impact of age of screening and screening frequency on the cost-effectiveness of mammography in the secondary prevention of breast cancer.

An Inventory of Quantitative Models Used for the Health and Economic Evaluation of Screening Mammography

A variety of quantitative or analytic models have been used to assess the impact of population-based breast cancer screening policies. This is a review of the models that have been previously employed in the evaluation of the impact of age, screening frequency, and disease risk on the cost-effectiveness of screening mammography.

General types of models include decision analysis, life tables, macrosimulation, and microsimulation models. Decision analysis models estimate the utility of an intervention from the viewpoint of the decision maker. The models follow the impact of a decision

through the use of decision branches that represent alternative pathways that can be taken by choice or by chance. The outcomes that have been assessed using this approach to modeling include marginal improvements in life expectancy, costs, or cost per life year or quality adjusted life years gained. Due to their individualistic approach to estimating impact of taking a decision or decisions, decision analyses are not easily amenable to population-level policy analyses. Life table or life expectancy-based models, calculate the number of life years lost or gained at each age taking into account mortality rate, average life expectancy, and morality rate reduction attributable to the intervention. The simplicity of the life table approach limits its capacity to account for complexities in screening, examine multiple types of outcomes, and difficulties in applying discounting to future benefits. Macrosimulation models use aggregated data as opposed to individual-level data to estimate impacts of interventions on population health outcomes over a specified period of time. Due to their high-level nature, macrosimulation modeling may have limited capacity to answer very specific health policy questions. Microsimulation models are computer models that operate at the level of the individual behavioural entity, such as a person, family, or firm, as opposed to aggregate models whose explanatory variables already represent collective properties. These models simulate large representative populations of these low-level entities in order to draw conclusions that apply to higher levels of aggregation such as an entire country (Simulation Technology for Applied Research Glossary).

Microsimulation models are advantageous as it is comparatively less complex to including interactions than with macrosimulation. Model extensions and adjustments can be added such as varying in programs, policies, and screening patterns (van Oortmarssen et al., 1995). This flexibility of microsimulation is particularly useful in obtaining detailed estimates of the health effects and costs of different screening programs or policies in the setting of a real population. In the subsequent sections, we provide a brief review of several types of microsimulation breast cancer models. Most of the models are part of the Cancer Intervention and Surveillance Modeling Network (CISNET), a consortium of National Cancer Institute-sponsored investigators that includes modeling to improve our understanding of the impact of cancer control interventions, including prevention, screening, and treatment, on population trends in incidence and mortality. This consortium, although based in the US, also includes other international models. CISNET publishes very

detailed descriptions of each of their models in documents posted on their website (CISNET, 2013). These documents are provided in two formats: a "model profiles document" and a "reader's guide". These documents are updated periodically with the most recent update published in the summer of 2013. We base the majority of the descriptive content for the CISNET models using these documents. The only non-CISNET model included in this review is the Population Health Model (POHEM), a Canadian model developed at Statistics Canada. The descriptive information for POHEM is not as transparent or consistently updated as it is for the CISNET models, therefore the information in this chapter in regards to POHEM was extracted from published articles as well as unpublished documents provided to us from researchers at Statistics Canada.

Statistics Canada's Population Health Model (POHEM)

The Population Health Model (POHEM) is a microsimulation model developed at Statistics Canada that has been previously used to assess the impacts of colorectal cancer screening, breast cancer treatment, lung cancer, obesity, osteoarthritis, and other population health issues. For the POHEM breast cancer screening model (POHEM-BCS), "what if" scenarios can be constructed to compare the impact of screening policy options on breast cancer incidence and mortality. The scenarios can be constructed in the following ways: the absence of screening mammography, screening with base parameter assumptions, and screening with sensitivity analyses around uncertain parameters. The model integrates health and population data. Information incorporated into the model includes: etiology (risk factors), disease onset and progression, and resource utilization. POHEM-BCS uses Monte Carlo simulation[3] to generate a sample of synthetic individuals to whom demographic characteristics, risk profiles, and health histories of Canadians are assigned. This synthetic population, generated from birth, ages over time, with number of person years (life years), breast cancer cases detected, number of undetected breast cancer cases, and other population outcomes accumulated until the entire population is exhausted i.e. until death (Flanagan et al., 2006).

[3] Monte Carlo (MC) simulations use computational algorithms based on repeated random sampling; sample probability distributions for each variable or parameter to produce many possible outcomes. It is used to help understand the impact of uncertainty by providing results that can be analyzed to obtain probabilities of the occurrence of different outcomes.

The total number of women in the model may be set at any number. The output estimates are reported in aggregate.

In POHEM-BCS, TNM classification[4] is used to split cancers into stage I, II node negative (IIn-), II node positive (IIn+) and stages III and IV combined. This staging was chosen because survival probabilities vary between stages and within certain stages, namely stage II. The number of different stages allowed was also limited by the model's original architecture (allowance of up to four disease states) which existed for use in modeling colorectal cancer screening. The staging allows for the accounting of both stage shifting and within-stage shifting on the impact on breast cancer survival.

Almost all of the data used in the modeling was pre-existing in POHEM for previous work on the construction of scenarios to compare what might have occurred in the absence of screening by mammography, screening with base parameter assumptions, and screening with sensitivity analyses around parameters that were difficult to estimate. Flanagan et al. (2006) compared the base scenario estimates to observed incidence and mortality to benchmark the model. They estimated the impact of screening on incidence and mortality over the period 1986 to 2002.

The model input data include: pre-screening incidence rates from National Cancer Incidence Reporting System database, age-specific stage distribution from the Saskatchewan Cancer Registry and the Canadian Breast Cancer Screening Database (CBCSD), pre-screening survival from the Saskatchewan Cancer Registry, and participation in screening mammography rates which were set at 70 per cent for initial screen (first screen) and 90 per cent for subsequent screening. Sensitivity and specificity of mammography were estimated from the CBCSD data and treatment effects on survival (by stage and treatment regimen) were pre-existing in the model, using estimates published elsewhere (Will et al., 1999). The costing data in this model include: costs associated with diagnosis and initial treatment.

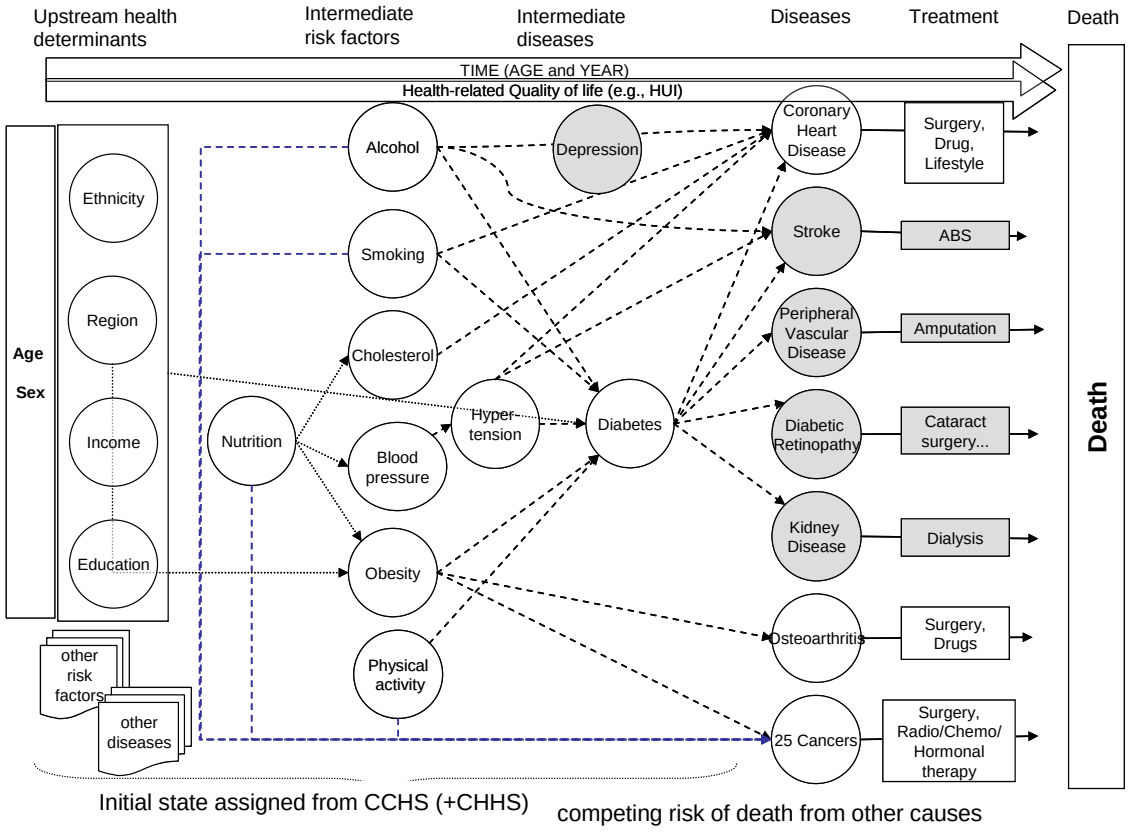
[4] The TNM classification system describes the extent of many types of solid tumour cancers. It gives a common language to describe a cancer. T stands for “tumour” and indicates the size of the primary tumour and the degree of spread to nearby tissues; N stands for lymph nodes and indicates whether or not cancer has spread to nearby lymph nodes, the size of the nodes that contain cancer, and how many lymph nodes contain cancer, and M stands for metastasis and indicates whether or not cancer has spread to distant organis..
Source:: http://www.cancer.ca/en/cancer-information/diagnosis-and-treatment/staging-and-grading/staging/tnm-staging/?region=on#ixzz30V86tE9e

These cost estimates were pre-existing in the model. Further details on these are found in the work by Will et al. (2000).

POHEM-BCS allows for multiple screening tests over each individual's lifetime, depending on the set participation rates, screening interval (number of years between screens), year of start and end of screening, and the age at which screening is to be offered.

Specific details of the model are further provided in chapter 4 of this dissertation.

Figure 3 High-level overview of POHEM



Source: Statistics Canada, Presentation at the International Microsimulation Association Conference, June 9, 2009. Note that greyed boxes are not currently implemented.

The Erasmus MC MISCAN Model

The MIcrosimulation SCreening ANalysis (MISCAN) model was developed in the Netherlands and was first introduced in 1985 by Habbema et al. (1985). It may be considered

the oldest and most frequently cited breast cancer microsimulation model internationally. MISCAN has since been used as a screening model for a number of diseases in addition to breast cancer including cancers of the colon, cervix, and prostate (CISNET, 2013). CISNET researchers that contributed to this model include J DIk F Habbema, Rob Boer, Harry J de Koning, Gerrit J Oortmarssen and Sita Tan, all of whome are from the Deaprtment of Public Health, Erasmus MC, University Medical Center Rotterdam, while Rob Boer is also with the RAND Corporation.

The model is now known as the MISCAN-Fadia model as it has been adapted to integrate the influence of tumour growth (a biologic parameter) and the concept of "fatal diameter", instead of discrete tumour stage as was used in the standard (previous) model (CISNET, 2013). The integration of a biologic parameter to predict survival is what distinguishes MISCAN-Fadia from other models. The MISCAN standard model and MISCAN-Fadia has been used to predict and compare the effectiveness and cost-effectiveness of different screening policy options. MISCAN-Fadia is a microsimulation model which uses inputs and tracks life histories that include cancer history and the effectiveness of treatment and screening. The most marked difference between this model and others, including POHEM, is the detailed natural history component based on a continuous tumour growth model as opposed to discrete stages. The majority of this section describes the MISCAN-Fadia model.

To contrast the standard MISCAN with the MISCAN-Fadia model, several differences should be noted. Firstly, in the standard MISCAN, the screening test result is dependent upon the stage-specific sensitivity of the test (mammography for example), whereas in MISCAN-Fadia, if the diameter of a tumour at the time of screening is larger than the assigned threshold diameter (differs between tumours), the test will result in a positive test. In the standard MISCAN, the benefits of screening relate to the individual's disease history in the absence of screening in that a proportion of screen-detected cancers, proportional to stage, will be cured. In MISCAN-Fadia, on the other hand, models survival of both clinically diagnosed and screen-detected cancers using the "fatal diameter" concept, and the diameter at which the cancer becomes fatal is related to the treatment. In contrast to the standard MISCAN, MISCAN-Fadia integrates differential survival for alternative adjuvant

treatments[5], as well as differences in their use over time. The standard MISCAN allowed for multiple disease histories within the same person whereas in MISCAN-Fadia only one history can be generated for each person in the model. Finally, MISCAN-Fadia models the dissemination of screening (participation) outside or externally to the model using a different program.

There are several referenced limitations in regards to the MISCAN-Fadia model assumptions (CISNET, 2013). First, the model allows only one tumour per woman therefore it is not possible for any woman to have more than one cancer. Although not reflective of true life, this is a consistent limitation across many microsimulation models, including POHEM. Secondly, the model only allows the dissemination of one screening test for each individual which is different from POHEM which allows multiple screening tests throughout an individual's modeled lifetime. Thirdly, the result of the screening test is completely determined by the size of the tumour as well as the tumour-diameter threshold for the screening test. The model does not take into consideration random variation in test sensitivity or impact of human error that may be introduced in the reading of the test. Fourthly, estrogen receptor (ER) status is not modeled. ER status is a predictor of how well a woman with breast cancer will respond to hormone therapy, how a tumour may behave, and what other treatments might be effective (Canadian Cancer Society: "Hormone receptor status testing", 2014). ER status is therefore a predictor of breast cancer survival.

The MISCAN-Fadia model's major components include a natural history component, population component, screening component, and treatment component (CISNET, 2013). The natural history component simulates the natural history of a breast cancer tumour and includes sub-components: cancer incidence and survival/mortality. The population component simulates the demography of the simulated cohort. The screening component simulates screening mammography participation and its impact on that simulated cohort. The treatment component simulates the use of adjuvant treatment and its effects on the simulated cohort.

[5] An adjuvant treatment or therapy is any "additional treatment given after the primary treatment to lower the risk that the cancer will come back. Adjuvant therapy may include chemotherapy, radiation therapy, hormone therapy, targeted therapy, or biological therapy." Source: National Cancer Institute Dictionary of Cancer Terms, http://www.cancer.gov/dictionary?cdrid=45587

To estimate the natural history component, parameters were estimated based on data from the Two County trial study in the "Cohort Model" (see chart 2). The natural history component simulates invasive tumours as well as ductal carcinoma in-situ (DCIS), which is considered non-invasive but can develop into invasive cancer. Sub-components or states of DCIS include: regressive, clinically diagnosed, and progress to invasive tumour. Tumour growth rate, although constant, and tumour size differs between tumours. The tumours are characterized by different sizes at which treatment will not result in a cure – it is called the "fatal diameter", which is a threshold that determines the individual's cancer survival/mortality. Whether a tumour is detected clinically depends on signs or symptoms resulting from the primary tumour or by symptoms related to distant metastases. These probabilities are assumed to depend on the primary tumour's diameter (a set diameter at which point symptoms will occur to result in a clinical diagnosis of the disease). Probability of distant metastases symptoms is dependent upon time from when the disease is characterized as being fatal (once the fatal diameter is reached). This time period is applied to cases where breast cancer is diagnosed clinically as well as to when cancer is screen-detected (CISNET, 2013). For example, a cancer may be screen-detected but at a point where the fatal diameter has already been reached (signifying a very aggressive tumour). It is expected that these types of screen-detected cases would not occur as often as they would in clinically diagnosed cases.

The life-course of a tumour in the model follows the parameters: tumour growth rate, tumour fatal diameter, survival time after reaching the fatal diameter, tumour diameter threshold for screen-detected cancers (diameter at which time a tumour is detectable by a screening test), tumour diameter at clinical diagnosis, and time when distant metastases becomes clinically diagnosed (modeled as a constant proportion of survival time after reaching fatal diameter). If diagnosis occurs before the fatal diameter is reached, in a screen-detected or clinically-detected case, it is assumed treatment will be initiated and the cancer cured. The impact of screening is captured for a detected tumour based on the tumour diameter and the threshold diameter for the screening test for the particular tumour.

Cancer incidence is modeled as a probability distribution for the onset of pre-clinical disease (before there are symptoms or signs of disease and which can only be detected by a

screening test) by age. This stage is the first possible disease state in the model and is labelled as "pre-clinical DCIS". The model assumes that many individuals will not have a detectable DCIS before having invasive cancer and are modeled has having no dwelling time in this stage. In the model, one cancer per woman can occur over a lifetime. The incidence of the onset of preclinical breast cancer data are from birth cohorts from 1895 to 1971, representing cumulative probability estimates from age 0 to 85. The age distribution of the incidence of the onset of pre-clinical breast cancer including DCIS is given by 5 year age groups from age 20 to 85. The survival and mortality benefits of screening occur when the tumour is detected by screening before it becomes fatal (reaches the fatal diameter), assuming that the tumour would have been otherwise detected clinically after it had become fatal. When a cancer is screen-detected before it has reached fatal diameter, it is assumed to be treated and then cured.

The population component simulates the demographic characteristics of the simulated population in the MISCAN-Fadia. The model researchers aimed to simulate from birth to death, the US population made up of 5-year birth cohorts from 1895-99 up to 1965-1969 and 1970, which was a 1-year cohort used to simulate the year 2000. Life tables with 1-year age steps were generated for each cohort for deaths from other causes using the model's base case data (1973–1975 SEER mortality data) at the mid-year of each cohort. The relative size of each cohort represented the size of the population in 1975, correcting for the probability of dying before that year, and then translated into a proportion of the simulated population for each. The demographic parameters in the population component include number of birth cohorts, distribution of the population among the birth cohorts, distribution of dates of birth within each birth cohort, life table parameters for each birth cohort, and lift time breast cancer risk for each birth cohort (CISNET, 2013).

Screening participation by age was modeled outside of MISCAN-Fadia and the output of that screening model was applied to MISCAN-Fadia to predict the impact of screening on various outcomes (CISNET, 2013). As part of CISNET, screening participation data is common amongst all CISNET models (MISCAN-Fadia, Dana Farber Model, Georgetown-Einstein Model, MD Anderson Model, Stanford Model, Wisconsin-Harvard Model, and the University of Rochester Models). Within the MISCAN-Fadia model, two screening

participation scenarios are observed: 1) simulation of a regular invitation-based screening program based on a specific screening period, screening ages and participation rates; 2) CISNET age-specific screening program simulating actual mammography participation in the US from 1975 to 2000 (CISNET, 2013). In terms of assumptions, screening was modeled to pick-up a cancer if the tumour size reaches a specific threshold. If screening occurs before that threshold is reached, then a tumour will not be detected and if the threshold is reached the screening test will always pick-up the cancer. Threshold size is dependent upon year of diagnosis and the individual's age. DCIS (pre-clinical, non-invasive cancer) is detected by screening depending on its sensitivity to DCIS.

The treatment component simulates the use and impact of adjuvant treatment and is integrated into MISCAN-Fadia as a probability of being treated with a specified type of treatment: chemotherapy or tamoxifen or both for two years, tamoxifen for 5 years, chemotherapy and tamoxifen for 5 years, or no adjuvant therapy (CISNET, 2013). Chemotherapy is a treatment for cancer that uses drugs to destroy cancer cells and tamoxifen is a selective estrogen receptor modulator (SERM) that binds to estrogen receptors to prevent estrogen from binding and is used to prevent the recurrence of breast cancer. The data used to model the benefits of adjuvant treatment on all-cause mortality were taken from a meta-analysis. The impact of chemotherapy and tamoxifen were assumed to be independent of each other. Neoadjuvant and primary treatment use and impact do not appear to be integrated into the model, whereas they are integrated into POHEM. Neoadjuvant therapy is treatment that is given prior to primary treatment in order to shrink the tumour so that it is easier to treat with primary treatment. Chemotherapy and radiation may be used as neoadjuvant treatments. A primary treatment is the main type of therapy used for breast cancer which varies by patient. As mentioned previously, primary treatment for breast cancer is surgery.

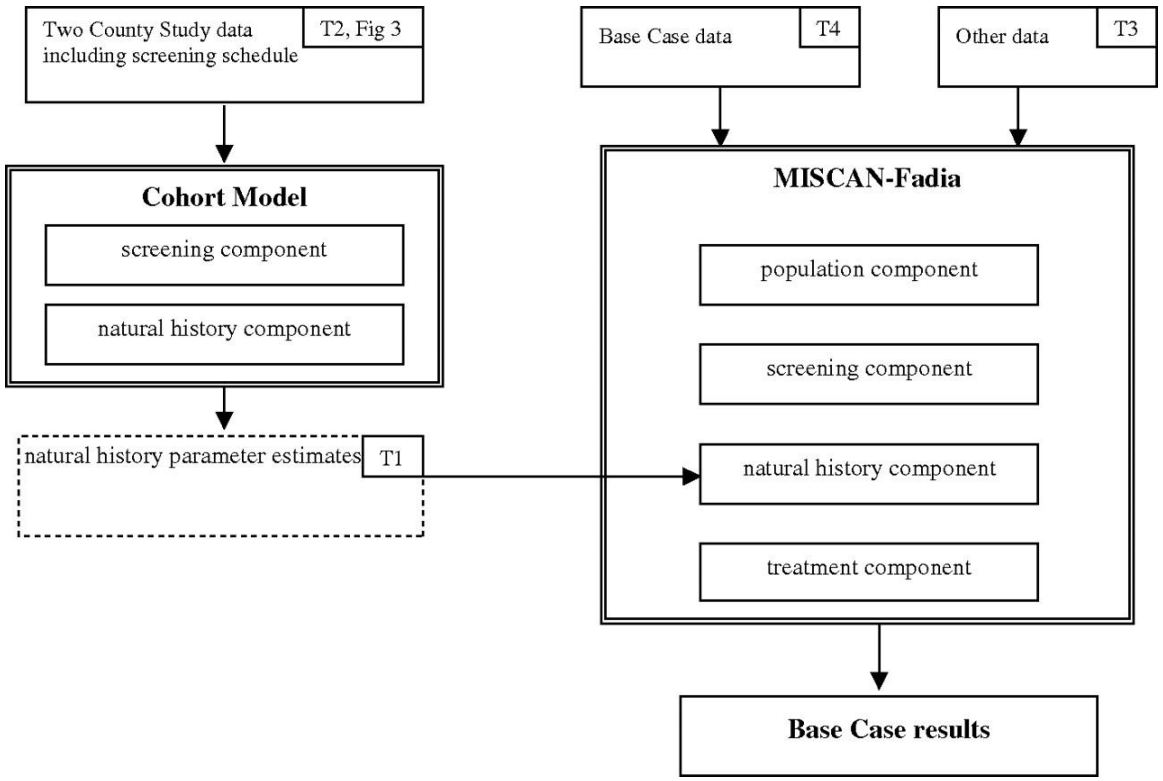
The MISCAN-Fadia outputs include:

- Breast cancer incidence by calendar year from 1975-2000 by stage and 5-year age groups from age 30 to 84;
- Number of deaths by calendar year from 1975-1999;
- Population on July 1st of each calendar year from 1975 to 1999 by 5-year age groups from age 30-84;

- Mean lead time by age (30–84, 30–39, 40–49, 50–59, 60–69, 70–84) –the time from screen detection to the time a tumour would have been clinically detected in the absence of screening;
- Percent over-diagnosis by age (30–84, 30–39, 40–49, 50–59, 60–69, 70–84) – the ratio of the number of screen-detected cancers (women) that never would have been clinically detected to the number of cancers (women) who were screen-detected;
- Number of over-diagnosis cases by 5-year age groups and by calendar year;
- Detection rate at first (initial screen) by age (30–84, 30–39, 40–49, 50–59, 60–69, 70–84) – ratio of the number of cancers detected to the number of women screened;
- Detection rate at subsequent screen by age (30–84, 30–39, 40–49, 50–59. 60–69,
- 70–84);
- Screen sensitivity by age at screening using one year interval (ages 30–84, 30–39,40–49, 50–59, 60–69, 70–84) - (number of screen-detected cancers)/(number of screen-detected cancers + number of interval cancers);
- Screen sensitivity by age at screening using two year interval or biennial screening frequency (ages 30–84, 30–39, 40–49, 50–59, 60–69, 70–84).

The cohort model is used to estimate the parameters of the natural history of breast cancer, using the data from the "Two County" trial for breast cancer screening, by simulating the screening schedule of this trial. These natural history estimates are used in the population model, in combination with the base case data and other data, to run the simulations that produce the base case results for the U.S. breast cancer incidence and mortality in the period 1975–2000. T1 through T4 are tables that give an overview of the data used by the two models, and Fig. 3 refers to the survival data in Fig. 3. See the CISNET Erasmus Breast Cancer Model Profile for details (CISNET, 2013).

Figure 4 The MISCAN-Fadia Continuous Tumor Growth Model for Breast Cancer.



Dana Farber Model

The Dana Farber model is a stochastic model[6] of the natural history of breast cancer that was developed to predict the impact of early detection of breast cancer in women due to screening mammography, physical breast examination, or a combination of both, on mortality (CISNET, 2013). For the purposes of CISNET, the population under study were U.S. women in the chronological time period between 1975 and 1999. The earliest publication using this theoretical model was published by Lee et al. (2003). Although the CISNET profile for this model was dated as an update in 2013, the reference information is quite old which leave us to believe that this model has not been used for some time or at least the information in the profile has not really been updated to reflect current changes to the model or uses of the model. The basic assumptions of the model, similar to other models including POHEM-BCS are that breast cancer is a progressive disease and that any mortality benefit from screening is due to stage shifting (an individual with breast cancer is detected at an earlier stage, resulting in earlier treatment and therefore better prognosis). Other applications of the model include the study of the impact of different screening policy options, similar to how POHEM-BCS is used to answer "what if" questions.

The Dana Farber model is described as having two components: 1) a natural history component and 2) a survival and mortality component (CISNET, 2013). Both components are described in the CISNET breast cancer model profiles as theoretical as of 2013. The natural history of disease component is described as being based on the assumption that breast cancer is a progressive disease with four or five health states:

- Individual is breast cancer-free or has breast cancer but is asymptomatic and would not be detectable through any means;
- Individual has breast cancer but is asymptomatic and could be diagnosed by some means;
- Individual is diagnosed with invasive cancer that has been clinically detected (sought care because of symptoms);
- Death from breast cancer;

[6] Use of random probability distributions for model parameters.

- Death from other cause.

The model inputs include:

- age-dependent breast cancer incidence rates pre-screening (or at a time when screening was not commonly used);
- age-dependent transition rates for the previously listed health states;
- stage distribution of cases when there is no screening and stage distribution of screen-detected cancers, stage distribution of interval cancers, survival distribution by stage, year (to account for treatment effects);
- screening participation rates, screening age, screening frequency;
- mammography sensitivity and physical examination sensitivity by age;
- birth cohort year(s) from which mortality probabilities are assigned.

In the description of this model, data to be used for these parameters would come from randomized trials and from databases such as SEER. The age-dependent transition rates would be estimated from age incidence rates using methods described in the article by Lee and Zelen (2003).

The survival and mortality component of the model integrates the differential modeling of individuals without screening and those who participate in screening. The model allows for the modeling of birth cohorts to predict age-specific breast cancer mortality. Mortality rates are applied to the models that are characterized as age-specific probabilities of death. Different mortality rates are assigned to screen-detected and interval cancers (individual has a history of screening but cancer was not detected at screen but rather picked up between screens). The model is described to take into account lead time bias and length bias, whereby the effect of screening assumes that diagnosis of screen-detected cancers would change the stage distribution beyond what would be expected due to length bias. This refers to the previous identified assumption that screening results in a favourable stage shift and have a more cancers with better prognosis than in the absence of screening. Based on a set reference point (such as from birth), cumulative mortality from that reference point is applied and is dependent upon the age of the individual at that reference point in time.

The model outputs include total breast cancer mortality and reduction in mortality relative to some base (scenario). The researchers note that the most accurate use of the model is to determine relative mortality reductions as this approach would be able to control for other factors not taken into account in the model. There are several other stated limitations reported with this model including the uncertainty around whether earlier detection really results in reduced mortality and that survival is dependent upon the mode of detection and stage (CISNET, 2013). The model has been applied to predict mortality reductions in eight randomized controlled trials and predicted mortality reductions that were starkly different from the reported mortality reductions in those studies (see table 1). Results of the model have also been generated for model validation and sensitivity analysis work (CISNET, 2013).

Table 3 Summary of reported and predicted mortality reductions (MR) using the Dana Farber Model for Eight Randomized Controlled Trials of Breast Cancer Screening.

Trial	FU(yrs)	Reported	MR(CI)	Prediction MR
Malmo–1	19	19%	(0,34%)	17%
Stockholm	15	10%	(–28%,37%)	21%
Gothenburg	13.5	22%	(–7,43%)	20%
Ostergotland	17	11%	(–9,28%)	22%
Edinburgh	14	17%	(–18,42%)	11%
HIP	10	30%	(?)	3%
Canada–1 (40–49)	7	–36%	(–121,16%)	1%
Canada–2 (50–59)	7	3%	(–52,38%)	1%

Source: Dana-Farer Cancer Institute Predicted Mortality Reductions (CISNET, 2013).

Georgetown University-Einstein Model

The Georgetown University Model is a CISNET consortium model that also simulates the incidence of and mortality from breast cancer between 1975 and 2000 in the U.S. female population, and specifically estimates the effectiveness of screening and treatment during this time period (CISNET, 2013). This model, originally developed in order to evaluate the cost-effectiveness of breast cancer screening programs, was developed by Jeanne Mandelblatt, Aimee M. Near, Clyde B. Schechter, and Michael A. Stoto. The model is described as being event-driven and using continuous-time state transitions. Similar to some of the previous and subsequent models in this review, including POHEM-BCS, individuals from different birth cohorts are simulated one by one. Times when events (or health states) occur are based on sampling from pre-specified time-interval distributions.

There are four components of the model: Population or demographic component, natural history of disease component, screening component, and treatment component. The population component simulates women to represent the age distribution of US women in 1975, their breast cancer incidence by age, and their overall mortality experience. The natural history component is part of the model that simulates the performance of screening and whether this results in the early detection of cancer and at what stage of cancer. Performance of screening is governed by screen test sensitivity and specificity. Part of the model also estimates stage shift for the tumour which is dependent upon screen test lead time. The screening component simulates screening participation for each woman – when they get screened is based on a model of observed utilization in the population from 1975 and 1999. The treatment component is applied whenever disease is either screen-detected or clinically detected and assigns a treatment for each woman with breast cancer and assigns a corresponding breast-cancer survival time. These probabilities are based on SEER data by age, stage, ER status, and survival linked to treatment.

The number of women simulated in this model was 55 million.US census data were used to extract information on women born in or after 1890 in order to simulate the population alive in 1975. Several health states are modeled including: development of breast cancer that is screen-detected, development of clinically-detected breast cancer, or death of other causes before breast cancer is diagnosed. When a cancer presents, it is assigned a stage depending

on whether it was screen-detected or clinically-detected. For screen-detected cancers, stage is estimated from the stage the cancer would have been if clinically-detected and the lead time gained from screening based on a formula derived from Bayes' theorem (CISNET, 2013). The cancers are further characterised as being estrogen-receptor (ER) positive or negative. Survival is dependent upon age and stage at diagnosis, ER status, and treatment. This is in contrast to POHEM-BCS, for which survival is only dependent upon age and stage at diagnosis.

The inputs of the model include:

- age distribution of US women;
- projections of breast cancer incidence in the absence of screening by age and year;
- annual mortality among US women from other causes by year;
- age–specific stage-distribution of clinically-diagnosed cancers (SEER data in 1975);
- age–specific stage-distributions of screen-detected cancers (SEER data in the 1990's);
- age-specific mammography sensitivity;
- mean tumor sojourn time by age (sojourn time is the duration of a disease before clinical symptoms become apparent but during a time when it could be detectable with some screening tool);
- mean tumor dwell time in each stage (DCIS, local, regional, distant);
- mammography participation by age and year (from the National Cancer Institute – NCI);
- ER specific distributions of treatment choices in different calendar years by age and stage (from the National Cancer Institute –NCI);
- ER specific breast cancer survival curves by age and stage;
- odds ratio estimates of survival associated with adjuvant therapy (tamoxifen and chemotherapy).

Similar to POHEM-BCS, a life history is generated for each individual in the model which identifies, at some point in time, whether the individual is diagnosed with breast cancer, at what stage of cancer, what treatment they were given, whether they die and when, whether the death is from breast cancer or other cause. In addition, the total number of screens is

provided as well as the number of positive screens. This information then contributes to the estimate of breast cancer incidence and mortality outcomes that are provided in aggregate by year and by decade of age (POHEM-BCS generates outcomes by 5-year age groups).

The reported key limitations of the model are that it assumes the only benefit from early detection is stage shifting and that all breast cancers, including non-invasive DCIS is progressive (CISNET, 2013). The latter assumption is the most concerning because empirical evidence does not necessarily support this assumption. For example, the study by Welch and Black (1997) reported that a proportion of studied women who were not known to have breast cancer throughout their life had DCIS upon autopsy (almost 9 per cent mean prevalence). In contrast, POHEM-BCS also assume stage shifting as the benefit of screening but does not include DCIS as stage in the modeling therefore the issue regarding over-diagnosis due to DCIS would not be considered a limitation.

MD Anderson Model

The MD Anderson Model, as part of the CISNET consortium, aimed to provide estimates and their associated uncertainties of the relative impact of screening mammography, tamoxifen, and improvements in chemotherapy on breast cancer mortality in the U.S. female population since the year 1990. This model also enables the prediction of impact of mortality from different screening policy options and changes in the use of tamoxifen and improvements in chemotherapy. Therefore, similar to POHEM, it can be used to predict real impact of screening on mortality and projected impact of "what if" screening policy questions on mortality. The principal investigator for this model is Donald A. Berry from the University of Texas M.D. Anderson Cancer Center.

Six model components are described in the MD Anderson Model CISNET profile: population component, screening component, cancer incidence component, treatment component, survival and mortality component, and a results component. The population component uses 1975 as the start year for simulation, at which time there would have been prevalent cases of breast cancer. The researchers noted that they had to identify these prevalent cases and omit them from the model. They then simulated 2 million women representing the age distribution of women in 1975 and followed them to the year 2000.

During this modeled period of time, new births, deaths, and migration were allowed. At each year, the number of women with diagnosed breast cancer is tracked. The screening component assigns each woman a screening profile what she would follow throughout her lifetime. The screening profile is characterized by start and end age of screening and frequency of screening. Each year from 1975 to 2000, it was determined whether each women participated in screening mammography based on her screening profile. The cancer incidence component determines whether women receive a breast cancer diagnosis. The probability of being diagnosed with breast cancer depends on her screening profile where a probability of breast cancer is conditional on how long it has been since her last screen (interval). Interval cases are integrated into the model – these are cases of disease that occur between screens. The characteristics of the tumours depend on the method of breast cancer detection – by screen or clinically. The treatment component includes assigned treatment depending on the characteristics of the woman and her cancer (tumour), and the calendar year to reflect changes in treatment over time. In contrast, POHEM-BCS does not adjust treatment by year nor does it model survival based on treatment. The survival and mortality component of involves each breast cancer case being assigned a lifetime with cause of death from breast cancer, as well each woman (with or without breast cancer) has a "natural" lifetime assigned to her as she enters the cohort (her life profile is set at the time she enters the cohort). In this component, survival is defined as the shorter of these two lifetimes. The results component is essentially the accounting process of the model. It aggregates or tallies the simulated breast cancer mortality attributable to screening and treatment.

In terms of the general structure of the model, the MD Anderson Model is described as using Bayesian (conditional probability) updating to estimate the impact of mammography, chemotherapy, and tamoxifen on breast cancer mortality (CISNET, 2013). The simulated cohort in this model is considered dynamic by allowing births, immigration and deaths and emigration in the population each year. Breast cancer events are tracked over time and depend on the age, screening participation, and treatment for breast cancer, which all change over time. In an example of applied conditional probability, a woman's participation in screening in any given year is dependent upon her screening history (a probability is applied based on previous screening). This is similar to POHEM-BCS, which assigns a probability of first screen and probability of subsequent screen. Having a breast cancer diagnosis or not

is condition on her age, the mode of detection (screen or clinical), time since her last screen, and the calendar. If breast cancer is detected, then a stage is assigned to the cancer, as well as nodal status, and ER status with frequencies according to age, mode of detection, and time since last screening. The treatment assigned is determined by the current treatment standards and by the characteristics of the woman and the cancer. The impact of treatment is estimated based on observed effectiveness in previous research. The probability of dying for each woman is based on actuarial survival data and the mortality from breast cancer observed in the simulated cohort is meant to represent that observed from 1975 to 2000. As part of the researchers overall objectives for this model, the estimated results were compared to observed results.

The reported intermediate model outputs included:

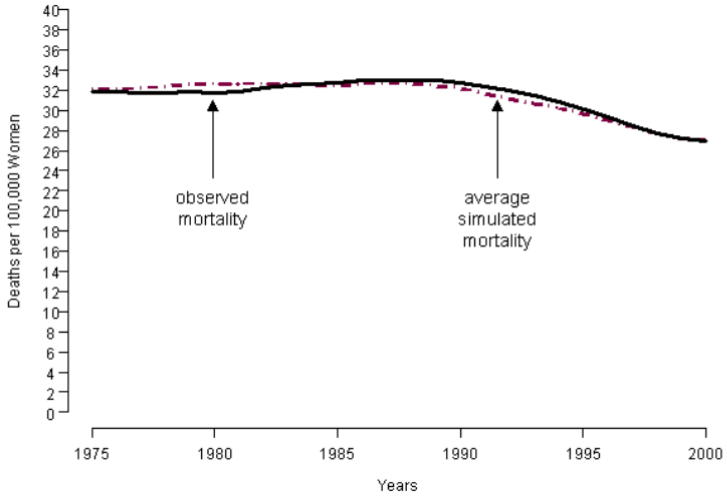
- age distribution of US women each year from 1975–2000;
- prevalence of breast cancer in 1975;
- tumour characteristics for each breast cancer detected each year from 1975–2000;
- survival distribution of breast cancer;
- survival distribution without breast cancer;
- screening mammography participation (age and frequency);
- proportion of women who have ever participated in screening mammography;
- stage-, age-, year- , and mode of detection (screen or clinical)- specific breast cancer incidence; and
- breast cancer mortality by year of diagnosis in 1975 (prevalence), or in 1975 or later (incidence).

The primary model outputs included:

- age–adjusted breast cancer mortality for each year from 1975–2000
- age–adjusted all cause mortality for each year from 1975–2000
- posterior distributions for parameters drawn from prior distributions such as the benefits of adjuvant tamoxifen and adjuvant chemotherapy

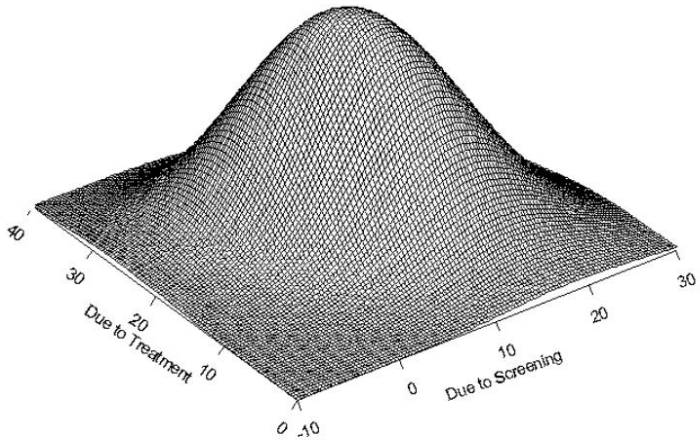
The model results showed that the estimated and observed mortality were reasonably similar (see chart 3). It was also found that there is a negative correlation (-0.40) between the percent reduction in breast cancer mortality due to screening versus treatment.

Figure 5 MD Anderson Model - Simulated mortality (average) and observed mortality



Source: CISNET, 2013

Figure 6 MD Anderson Model - Per cent reduction in breast cancer morality due to treatment and screening (joint contribution)



Source: CISNET, 2013

Stanford Model

The Stanford Model was also included in the CISNET consortium. The stochastic model simulates breast-cancer specific events and tracks the health outcomes using a natural history of disease approach. The stated primary goal of this simulation model was to explain the effect of breast cancer screening and treatment on breast cancer incidence and mortality (SEER) in the US female population between 1975 and 2000 (CISNET, 2013). The developers of the model note that it can be used to predict breast cancer incidence and mortality under various "what if" policy scenarios, such as alternative screening programs characterized by age and screening frequencies or intervals, which is similar to the use of POHEM-BCS. However, the Stanford model can also be leveraged to examine the impact of changing the groups targeted for adjuvant treatment on mortality.

The CISNET profile for the Stanford Model outlines the following algorithm (or health states) applied to birth cohorts form 1887 to 1970 and for each woman in the birth cohort:

- Generate natural history of breast cancer;
- Calculate life history without screening and adjuvant treatment;
- Calculate life history with screening but without adjuvant treatment;
- Calculate life history without screening but with adjuvant treatment; and
- Calculate life history with screening and adjuvant treatment.

Similar to POHEM-BCS, the Stanford Model estimates population-level breast cancer mortality trends through the simulation of individual life histories and then aggregating the generated outcomes to estimate a population-level estimate of disease and death (CISNET, 2013) The model simulates breast-cancer specific events and tracks the health outcomes using a natural history of disease approach. The model uses Monte Carlo simulation to generate the following characteristics for each breast cancer case: 1) date of birth, 2) age of death from other cause, 3) ages when screening occurred, 4) age when invasive breast cancer would have been detected in the absence of screening, 5) age when invasive breast cancer would have been detected in the presence of screening, 6) primary tumour size, extent of nodal and distant involvement, and ER status at the time of detection in the presence and absence of screening, 8) breast cancer survival time given disease stage, tumour size, age at

detection, and mode of detection (screen or clinical), and 9) cause of death (breast cancer, other cause) (CISNET, 2013). The results of the model can never be generated for any one individual – results are always reported by the model at an aggregate level (CISNET, 2013).

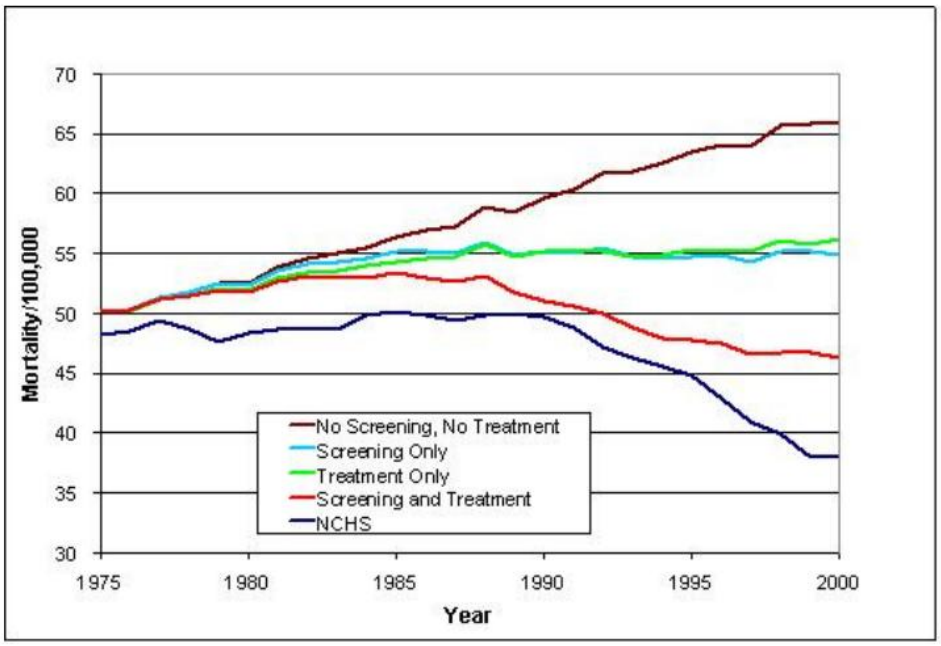
The model assumptions are based on tumour size, similar to the MISCAN-Fadia model. A patient in the model has a screen-detected breast cancer only if the size of the tumour is at or above the tumour size detection threshold of mammography at the time of screening (CISNET, 2013). Once breast cancer is detected, the case is assigned a specific survival time based on age, tumour size, SEER historic stage mode of detection and use of adjuvant treatment. In this way, survival is modeled using several parameters in contrast to POHEM-BCS which relies solely on age and stage of disease.

The model's CISNET profile reports five components: population component, natural history component, screening component, cancer incidence component, treatment component, and survival and mortality component. The population component specifies the birth cohorts for the simulation that are representative of US women born between 1887 and 1970. The developers used a sample size of two million women in each cohort which was found to be a sufficiently large enough number to model in order to reduce variability associated with the use of Monte Carlo simulation. The size of each cohort is kept constant in the simulation, despite the real effects of immigration and emigration

The outputs of the model include breast cancer incidence and mortality by 5-year age groups and by year and screening program characteristics based on cancers generated in years 1975 to 2000 (lead time, over-diagnosis, detection rates for first and subsequent screens, sensitivity – all by 5-year age groups) (CISNET, 2013). The outputs are produced for four different modeled scenarios: background risk only (no screening, no treatment), treatment only, screening only, screening and treatment. The results of the model show that compared to the predicted mortality rate in the absence of screening and adjuvant therapy in the year 2000, in the presence of both screening and adjuvant therapy mortality rate is reduced 29.9 with 16.9 per cent attributable to screening, 6.9 per cent attributable to chemotherapy and attributable due to adjuvant therapy. The estimated relative contributions of screening and adjuvant therapy to the mortality reduction were similar: 53 per cent attributable to screening compared 47 per cent attributable to adjuvant therapy (see chart 5). The observed mortality

rates appear to be lower than all four scenarios therefore the model results may be more meaningful in telling us the relative difference in impact across scenarios than predicting the real absolute impact in the population. A similar analysis could not be conducted using POHEM-BCS as treatment is not integrated in the prediction of survival.

Figure 7 Stanford Model – simulated age-adjusted mortality from breast ancer under scenarios. Observed rates are plotted for comparison (NCHS)



Source: CISNET, 2013

Wisconsin-Harvard Model

Another CISNET microsimulation model, the Wisconsin-Harvard Model is descdribed as generating cancer registry-like datasets over time by manipulating parametric input assumptions for natural history of disease, screening, and treatment in order to answer screening policy questions (CISNET, 2013). The Wisconson-Harvard Model evolved from a model constructed 10 years prior by a PhD student for their dissertation in order to observe breast cancer incidence and mortality in the state of Wisconsin from 1982 to 1992 (Chang, 1982). This base, deterministic model, included a population demography component, natural history of disease component (biologic onset and disease progression), a screening component, and a treatment component. Based on his original model, Change concluded that

a significant proportion of all breast cancers are pre–destined to grow only to a limited size of about 1 cm diameter and would not pose a lethal threat to an affected woman if left untreated. Chang also concluded that this indolent tumour would be indistinguishable from potentially lethal tumors of similar size and termed them "limited malignant potential (LMP)" tumors. We believe that DCIS non-invasive cancers would fit under this category of tumour.

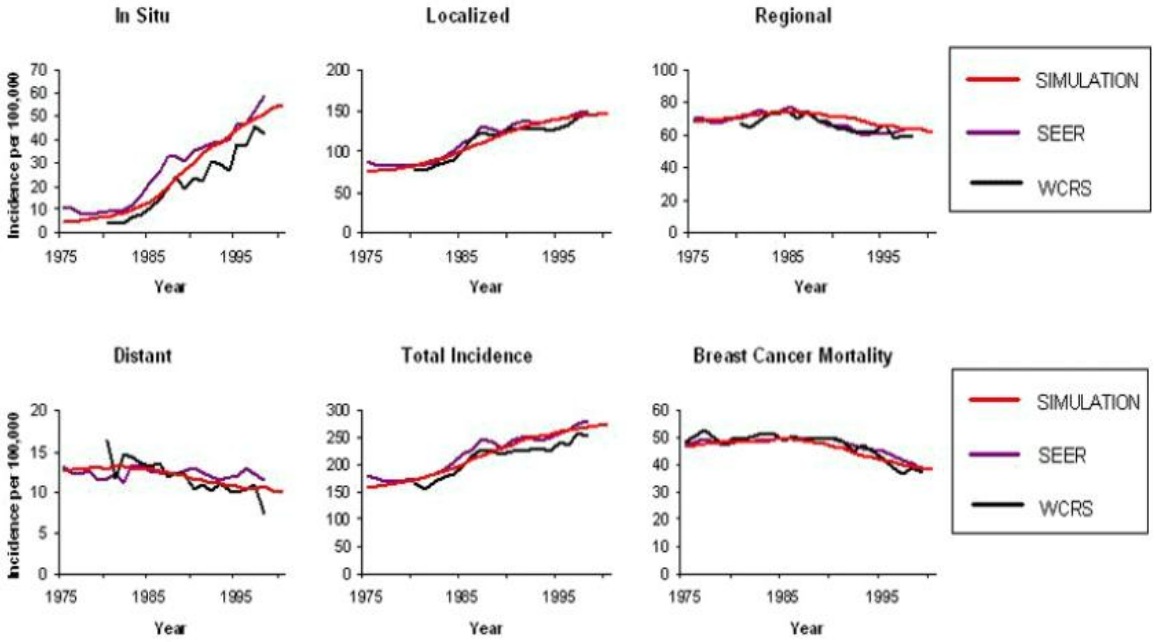
The current model is a redesign of Chang's model in order to generate a realistic virtual Wisconsin cancer registry of incident breast cancers for resident women from 1975 to 2000 and to simultaneously replicate age-specific breast cancer mortality in this population during the same time period using a microsimulation model, and to produce a model that can be used to explore the impact of alternative screening programs and treatment for breast cancer – to answer "what if" policy questions (CISNET, 2013). Briefly, the model is a discrete event simulation with fixed cycle time of 6 months with a start calendar year 1950. It uses a population size of 2.95 million women divided into birth cohorts making up the female population age 20 to 100 years old and representative of women in Wisconsin between 1950 and 2000. Similar to previous models discussed in this chapter, each individual woman are simulated in this model from calendar year 1950 or in the year in which they were age 20 until she dies a simulated death or achieves age 100, or the simulated year reaches 2000 – whichever comes first. The basic processes in the model include natural history of breast cancer from diagnosis to death from breast cancer, screen-detected breast cancer or clinically-detected breast cancer cases, improvements in treatment and utilization of treatments over time, and death from other causes (CISNET, 2013).

The components of the model are briefly described in the CISNET profile as follows. In terms of the natural history of disease component, the onset of breast cancer may occur at some point in a woman's life or not at all. A cancer will grow over time in a progressive manner, depending on the tumour's characteristics, but not all tumours are lethal and some may actually regress. This is different from previous models that assume all diagnosed breast cancers to be progressive. Death from breast cancer occurs as an endpoint of a process whereby tumour growth is uncontrolled and therefore spreads. In terms of the screening component, a breast cancer may be detected by screening mammography whereby

participation in screening may be stochastic or systematic over time. Clinically-detected breast cancers are not differentiated in the model. In terms of the treatment component, breast cancer that could be lethal could be stopped or regressed with treatment and interventions may include surgery with or without radiation therapy. This is considered primary therapy for breast cancer. Adjuvant therapy with tamoxifen and/or chemotherapy is introduced over time within the model. The model uses an all or nothing cure model for treatment for each simulated individual woman – the component approximates observed population-level treatment effectiveness. In terms of deaths from other causes, women can die from other non-breast cancer causes in the model. Demographic data used in the model were from census while mortality was derived from the Berkeley tables (CISNET, 2013). Cancer data came from the Wisconsin Cancer Reporting System and from SEER, while mammography participation data from Wisconsin and from National Cancer Institute (NCI) CISNET base case analysis were used. Treatment use was provided by the NCI and treatment effectiveness from meta-analyses. Characteristics of screening mammography were based on findings in the published literature with expert judgement as supplementary information.

In terms of results, the CISNET profile show the Wisconsin-Harvard model to fit relatively well with observed data from the SEER and the Wisconsin Cancer Reporting System by cancer stage (in situ, localized, regional, and distant), and for breast cancer incidence and breast cancer mortality (WCRS) (see chart 6).

Figure 8 Wisconsin-Harvard Model – Fit of final model against SEER and WCRS.



Source: CISNET, 2013

Conclusion

Modeling can be useful in the assessment of the benefits and harms of different screening policy options when conducting randomized controlled trials or other epidemiologic studies are not feasible. The CISNET breast cancer consortium research groups have worked closely with the U.S. Preventive Services Task Force (USPSTF) to provide modeling input to supplement USPSTF's usual evidence review for updating breast cancer screening recommendations in the US. The evidence that the CISNET modeling has been able to contribute include the added benefits and harms of screening at earlier and later ages and varying the time interval between successive screening exams (or screening frequency). In the US, modeling appears to be valued as a significant contribution to the greater understanding of the harms and benefits of screening policy options in order to inform screening guidelines in that country. In 2011, the Canadian Preventive Task Force updated recommendations for breast cancer screening in Canada based on a systematic review of the literature but did not consult the use of microsimulation modeling using Canadian data. As POHEM-BCS is a Canadian model that is comparable to the models in the CISNET

consortium, this model could be leveraged to inform screening guidelines for Canada. POHEM-BCS could therefore be used to answer "what if" policy questions in regards to identifying optimal screening based on age, screening frequency or intervals, and risk-groups for the Canadian population. We aim to explore this further in the subsequent chapters of this dissertation: a systematic review of impact of breast cancer screening age eligibility, screening frequency, and screening population risk-profile in the cost-effectiveness literature (chapter 3); a cost-effectiveness analysis of screening mammography policy options characterized by screening age and frequency (chapter 4); and the impact of age and screening frequency on the cost-effectiveness of targeted screening for women with high-risk of breast cancer compared with the base case (screening of the general female population) (chapter 5).

Chapter 3 (Manuscript 1): A Systematic Review of the Cost-Effectiveness of Screening Mammography

Authors:

Ngoc-Thy Dinh[1], Kevin Brand[2], Douglas Coyle[3], Howard Morrison[3], Catherine Deri-Armstron[4], and Azim Kasmani[5]

Affiliations:

1. Health Economics, Public Policy Division, Conference Board of Canada
2. Telfer School of Management, University of Ottawa
3. Department of Epidemiology and Community Medicine, University of Ottawa
4. Department of Economics, University of Ottawa
5. Michael G. DeGroote School of Medicine, McMaster University

Abstract and Keywords

The objective of this systematic review was to assess the cost-effectiveness of screening mammography. Studies published in English from 2000 to 2010 were identified from a search of relevant databases including Medline, EMBASE, HealthSTAR, NHSEED, CEA Registry, and Econlit. The review was conducted in June 2010 and focussed on studies evaluating the health and economic of population-based screening programs using film or digital mammography, deemed of high methodological quality. Two blinded reviewers were used to identify relevant studies through a four-stage review process. The search strategy identified a total of 437 articles, of which 11 were included in the analysis. To compare the cost-effectiveness of screening program options across studies, costs were converted to 2013 Canadian dollars. Sequential analysis of incremental cost-effectiveness ratios and a willingness-to-pay threshold of $50,000 per life year gained or quality adjusted life year gained were used to identify optimal screening characterized by screening age eligibility and frequency. In terms of study characteristics, all included studies used quantitative or analytic modeling such as microsimulation or cohort simulation, which included a natural history of disease component. The included studies also employed long time horizons (20 years or lifetime) and an analytic perspective of either the health care system or statutory health insurer or societal viewpoint, where stated. The results of 4 included studies could not be commented on in the context of the other included studies. Another 4 studies found screening to be not be optimal given the set willingness-to-pay threshold and incremental cost-effectiveness. The 3 remaining studies found the following screening programs cost-effective: 1) annual screening for women starting at age 25 years if she has a genetic predisposition to breast cancer; 2) screening every 2 or 3 years for women ages 40 to 49 years; 3) screening every 3 years for women ages 50 to 80. In countries where breast cancer incidence is relatively high, annual screening mammography appears to incur much higher costs for an incremental benefit compared to less frequent screening. Breast cancer screening was found not to be cost-effective in countries where breast cancer prevalence and incidence are low and for chronically-ill women for whom any benefits of cancer screening may be eradicated by excess mortality and morbidity associated with their illness.

MeSH: screening; mammography; model; breast; cancer; breast neoplasms; cost-effectiveness; systematic review

Introduction

Economic evaluations of health interventions are a subset of health economics studies in which comparative data are collected or reported on both costs and effects, thus permitting an incremental comparative analysis between a set of decisions or options (Drummond et al., 2005). Thus, health economic evaluations are particularly useful in providing supportive evidence for health policy and health care decision making. Anderson et al. (2010) suggests that systematic reviews of economic evaluations are meant to inform the development of a new health economic evaluation through the identification of the most relevant studies to inform a particular decision in a given jurisdiction, or to identify the key tradeoffs implicit in a given intervention choice. Currently, there appears to be no satisfactory health economic evaluations of screening mammography for breast cancer program designs within the Canadian context.

The objective of this study is to systematically review the published health economic literature on screening mammography cost-effectiveness. The review aims to specifically assess the need for, and consolidate information, to inform an analysis of the tradeoffs between costs and benefits of screening mammography programs for Canadian women. Of particular interest, the review aims to describe and summarize the characteristics of relevant studies and their findings in regards to the health and economic impact of varying screening age eligibility and frequency or intervals between screens, as well as the health and economic impacts of screening of specific population subgroups such as women.

Methods

Research Questions

The research questions set at the outset of this study were as follows:

1. Are there any Canadian cost-effectiveness studies that examine the impact of age and screening frequency on benefits versus costs of screening mammography?

2. What do the studies tell us about the benefits versus costs of screening based on screening age?

3. What do the studies tell us about the benefits versus costs of screening based on screening frequency?

4. What do the studies tell us about the benefits versus costs of screening in special populations/population subgroups?

5. What lessons can be learned to inform a cost-effectiveness analysis study to identify optimal screening based on age eligibility and screening frequency in the Canadian context?

Design

This systematic review describes and summarizes the characteristics and results of published cost-effectiveness of screening mammography studies in a narrative. At the time of the review, there were no agreed-upon methods for pooling combined estimates of cost-effectiveness in a meta-analytic approach; therefore a narrative summary approach was selected for this review. A narrative summary allows for the examination of the extent to which results and conclusions are homogeneous across studies. This review was conducted in the fall of 2010.

Eligibility Criteria

In this review, studies were included if they: 1) evaluated the health and economic impact of screening mammography for the secondary prevention of breast cancer; 2) were full health economic evaluations such as cost-effectiveness analyses, cost-utility analyses, or cost-benefit analyses (CBA); 3) compared two or more screening programs of varying design, such as different screening age eligibility criteria screening frequency; 4) reported on both costs and health benefits of breast cancer screening; 5) reported on comparative units of analysis including cost per life year gained, cost per quality adjusted life year gained, and cost per disability adjusted life year gained or net benefits in monetary terms, and 6) were published between the year 2000 and 2010 in the English language.

Search Strategy

Studies were identified from relevant databases of peer-reviewed scientific journal articles including EMBASE, MEDLINE, HealthSTAR, the National Health Service Economic Evaluation Database (NHSEED), the Cost-Effectiveness Analysis Registry (CEA Registry) and Econlit. As there is a lack of agreed search terms for economic analyses, we chose to use the search terms: "breast" AND "cancer or neoplasm" AND ("screen$" OR "mammogra$) AND "cost$ or cost analysis or economic$", with restriction to studies published in the English language. Prior to this review, the U.S. Preventive Task Force conducted a review of breast cancer screening studies published between 1994 and 2001. That review found only one cost-effectiveness study which focussed on the cost-effectiveness of mammography for older women age 70 years and older which concluded that screening women beyond 70 was not cost-effective (Mandelblatt et al, 2003). The literature search for our study was restricted to studies published from January 2000 to December 2010, the year in which this review was conducted.

Study Selection

The review process was conducted by two blinded reviewers (ND and AK) in a four-stage process: screening (two phases), quality assessment, and data abstraction. In the first screening phase, using the eligibility criteria, two reviewers independently screened the titles and abstracts of all identified studies. For the second screening phase, all studies that were independently identified as relevant after screening phase 1 were included in the eligibility assessment process where full-text articles were retrieved. In screening phase 2, two reviewers independently reviewed the full-text of the studies for eligibility. Of the retrieved articles, each reviewer independently selected articles for inclusion based on the inclusion and exclusion criteria. Studies that were deemed eligible after the second screening phase were included the quality assessment phase. If there was any discordance in the selection of eligible studies for inclusion, a consensus between the two reviewers was used to reach an agreement. If no agreement could be reached by the two reviewers, a mediating third person was used to obtain consensus (DC). Percentage (observer) agreement was used to measure the level of inter-reviewer agreement (Cicchetti and Conn, 1976.)

Quality Assessment

Those studies that were identified as relevant after the two screening phases were then evaluated for methodological and reporting quality. The quality of included studies was evaluated using a checklist (for the assessment of the quality of modeling in health economic evaluations) developed by Phillips et al. (2006), consisting of 57 questions relating to a health economic evaluation study's characteristics in relation to structure, data, and consistency. Studies that were deemed to be of moderate to high quality and transparent in their methodological reporting were included in the data abstraction and evidence synthesis phases of the review. Only studies that provided information on at least 40 per cent of the checklist items were retained for data abstraction.

Data Abstraction

The data abstraction was undertaken by two reviewers. This stage involves the identification of pre-specified data elements from individual studies and entering the data into a database. A standard data abstraction form was used to record more detailed information on each of the quality assessment criteria from the Philips checklist including study identification information, study design, analytic perspective, time horizon, screening options and comparators, the country of origin and currency reported, main outcomes, the sources of data, primary results including costs, benefits, cost-effectiveness ratios (CER), incremental cost effectiveness ratio (ICER), and secondary results including uncertainty and sensitivity analyses (Phillips et al, 2006). Other additional information relating to the design of and reporting in each study, including the study's strengths and limitations, were also recorded.

Analysis

A narrative summary format was used to describe and compare the included studies' characteristics and findings. The included studies' findings were synthesized and organized in tables according to commonalities and distinctions, grouping similar studies together. The summary tables include detailed information from the data abstraction. The study characteristics recorded include primary author and year of publication, study design, analytic perspective, time horizon, population characteristics (country, age, and any other defining population characteristics), modeling approach, policy or program comparators or screening options evaluated, general methods, and main base-case results in terms of cost-

effectiveness ratios (CER) and incremental cost-effectiveness ratios (ICER), which are presented in 2013 Canadian dollars (CAD).

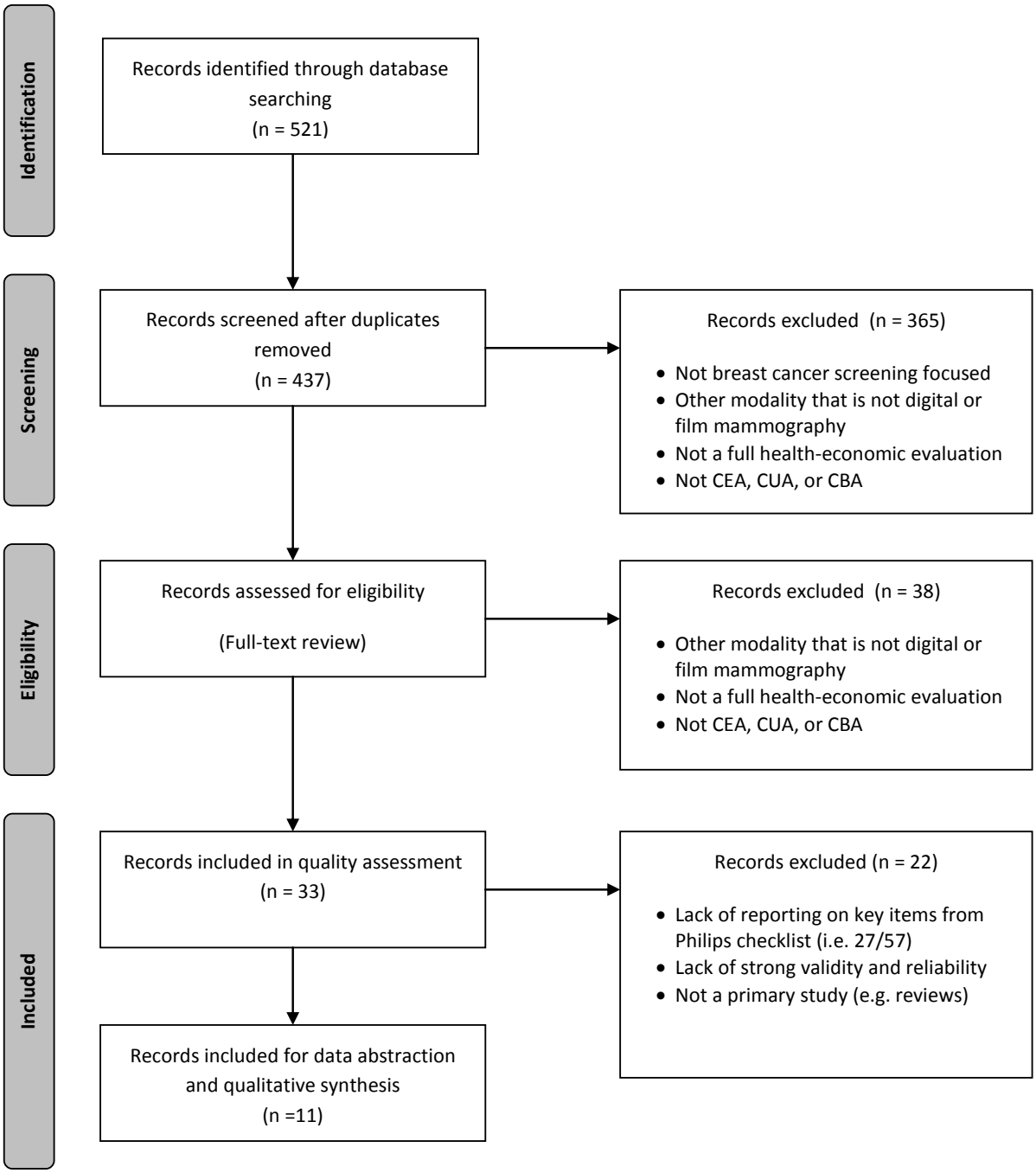
Costs in each included study were converted to 2013 CAD (Canadian) dollars to facilitate comparability of results across studies. Each study's costs were first converted from the original country and year to CAD dollars in the same year using the OANDA Corporation currency converter, and then adjusted for inflation using the 2011 Canadian Consumer Price Index (CPI). Sequential analysis was used to compare cost-effectiveness of various screening programs across comparable studies in relation to study design type (cost-effectiveness analysis, cost-utility analysis, cost-benefit analysis, etc.) and screened population (general female population, diabetic). Sequential analysis is an approach to estimating the incremental cost-effectiveness of a screening program versus the next least expensive program in sequence (from most costly to least costly). A screening program is dominated if another program or combination of other programs is more effective at the same or lower costs.

Results

Literature Search and Selection

A total of 437 unique articles were identified. After a first screen of title and abstract for relevance using the inclusion and exclusion criteria, 72 articles were further retained for full-text review for relevance. A total of 33 articles were retained for quality assessment, of which 11 articles were retained for data abstraction and synthesis. Figure 9 is a diagram summarizing the flow of information through the different phases of the systematic review. Little disagreement between reviewers was observed: inter-reviewer agreement was 99% in the quality assessment.

Figure 9 Flow of information throughout the difference stages of the systematic review



Study Characteristics

Tables 4 and 5 summarize the characteristics of the 11 included studies. Of the 11 included studies, five were cost-effectiveness analyses only (Arveux et al., 2003; Mandelblatt et al., 2005; Neeser et al., 2007; Okonkwo et al., 2008; Wong et al., 2008), four were cost-utility analyses only (Lee et al., 2010; Rojnik et al., 2008; Stout et al., 2006; Tosteson et al., 2008), and two were both cost-effectiveness and cost-utility analyses (De Gelder et al., 2009; Wong et al., 2007). There were no cost-benefit analyses identified in this review.

The study questions addressed across included studies varied substantially. Arveux et al. (2003) set out to evaluate the current screening practices in a region of France, whereas the objectives in the studies by Mandelblatt et al. (2005), Wong I et al. (2007), Rojnik et al. (2008), and Stout et al. (2006) were to examine the cost-effectiveness of various screening program designs by altering age eligibility and screening frequency. The studies by Lee et al. (2010) and Wong et al.(2008) aimed to assess the cost-effectiveness of screening within specific subgroups of the population, specifically women with genetic predisposition and women on dialysis treatment, respectively. Studies by de Gelder et al.(2009) and Neeser et al. (2007) were comparative analyses of the cost-effectiveness of organized screening mammography program (MSP) to opportunistic screening (OS).

Table 4 Summary of included studies - general female population

Reference	Design, Perspective, Time Horizon, Country, Population	Model type, Comparators/ Program designs	Brief methodology	Mammography screening scenario comparison	Cost CAD 2013	Effectiveness	ICER ($/LYG or /QALY)	Findings
Arveux et al., 2003	CEA Perspective not specified 20 years France Women age 50–65 years	Markov-based decision model Mammography screening (screening interval not specified) vs no screening	Model incorporates regional screening program data, morbidity and mortality data, and demographic data. Includes direct costs pertaining to screening, diagnosis, initial treatment, and breast cancer surveillance. Analysis of cost per life-years saved. Costs discounted at 5%.	No screening, 5% cost discounting	100.5 million	(not provided)	-	Screening women age 50-65 the most cost-effective program at a WTP threshold $50,000/LYG. Screening interval not specified. Results did not allow for assessment of the relative contribution of age or screening frequency on screening cost-effectiveness.
				Sc1b: Current screening program women age 50-65 versus no screening, 5% cost discounting	159.1 million	no screening + 1,522 LY	38,536 /LYG	
De Gelder et al., 2009	CEA & CUA Perspective not specified 20 years Switzerland Women age 50-69	Microsimulation model (MISCAN)	Screening ages, interval and attendance and the type of screening (opportunistic or organised), as well as the sensitivity and specificity of mammography are defined in the model.	No screening	2.05 billion	21.3 million LY 21.2 million QALY	-	Screening not cost-effective in terms of LYG. Is cost-effective when using QALY. Results did not allow for assessment of the relative contribution of age or screening frequency on screening cost-effectiveness.
				Sc3: biennial mammogram screening program (MSP) vs no screening, 3% discounting	2.70 billion	21.3 million LY 21.3 million QALY	650/QALY	
Mandelblatt et al., 2005	CEA Societal perspective Lifetime United States Women age 50 and older	Event-driven continuous time Monte Carlo simulation model Biennial screening from 50 with no age limit, and screening starting at age 50	Natural history of disease model using proxies of age-dependent biology. Data includes age-specific incidence rates, stage distributions, probability of disease progression between stages, dwell times, screening, diagnosis and	Biennial screening age 50-69	4,164 per woman	19.453 LY per woman	-	Screening for older women (past 70) is not cost-effective.
				Sc1: Biennial screening age 50-79 vs biennial screening age 50-69 (contribution of age 70-79)	4,520 per woman	19.455 LY per woman	178,235 /LYG	

Reference	Design, Perspective, Time Horizon, Country, Population	Model type, Comparators/ Program designs	Brief methodology	Mammography screening scenario comparison	Cost CAD 2013	Effectiveness	ICER ($/LYG or /QALY)	Findings
		and stopping at age 70, or 79.	treatment, life expectancy, and costs.	Sc2: Biennial screening over lifetime starting at age 50 and older vs biennial screening age 50-79 (contribution of screening age 80+)	4,774 per woman	19.456 LY per woman	203,567 /LYG	
Neeser et al., 2007	CEA Statutory health care insurance perspective Lifetime Switzerland Women age 40 and older	Markov-based decision model Biennial screening mammography starting at different ages: 40, 50, 60, and 70.	Model based on health states linked via transition probabilities derived from various data sources, including cancer registry, clinical data, and published literature. Model simulates annual occurrence of malignant and benign breast tumours, and detection rate. Mortality, screening, diagnostic and treatment costs accumulated over time in model. Discounts effects at 1.5% and costs at 3%.	Start age 40	6,696 per woman	30.674 LY per woman	-	Screening for women age 40 to 49 incurs greater health benefits at nominal cost per woman. Adding younger women to the population that is screened biennially is cost-effective.
				Start age 40 vs 50 (contribution of age 40-49)	6,363 per woman	25.322 LY per woman	62/LYG	
Okonkwo et al., 2008	CEA Perspective not specified Lifetime India Women age 40 and older	Microsimulation model (MISCAN) Biennial screening mammography starting at age 40-60 vs no screening	Natural history of disease model. Changes to MISCAN to account for Indian context, i.e. substitution of a lower cumulative incidence of breast cancer and delayed diagnosis. Simulation of one million women. Unless specified otherwise, assumed that screening programs were of 25 years duration, that the attendance rate was 100%, and that there were 100 years of follow-up. Extrapolated Dutch screening costs. 3 % discount rate of effects and costs.	-	-	-	-	The screening policy options in this study were not comparable in order to assess the relative contributions of age eligibility or screening frequency on the cost-effectiveness of screening mammography.

Reference	Design, Perspective, Time Horizon, Country, Population	Model type, Comparators/ Program designs	Brief methodology	Mammography screening scenario comparison	Cost CAD 2013	Effectiveness	ICER ($/LYG or /QALY)	Findings
Rojnik et al., 2008	CUA Health care sector perspective Lifetime Slovenia Women age 40 years and older	Time-dependent Markov model 36 screening scenarios starting ages 40, 45 and 55 years, ending ages 65, 70, 75 and 80 years, and screening intervals of 1, 2 and 3 years vs no screening	Model characterizes natural history of the disease as having four preclinical stages when breast cancer can be detected by screening but shows no clinical symptoms: localized, regional, and distant invasive stages. Breast cancer incidence, mammography sensitivity, mortality, and breast cancer relative survival modeled as time-dependent transition probabilities. Clinical data obtained from cancer registry. Costs obtained from the Institute of Oncology Ljubljana. Quality of life weights derived from literature. Discounting at 3% for effects and costs.	No screening, discounted at 3%	432 per woman	23.1 LY per woman 23.0 QALY per woman	-	Screening at earlier ages (40-45, 45-50) every three years was considered cost-effective. Screening age 50 to 65 every three years was considered cost-effective. Screening older ages (65-70, 70-75, 75-80) every three years was considered cost-effective. Screening every 2 years was not cost-effective in comparison to screening every 3 years (age 40 to 80).
				Sc1: Mammography screening age 50 to 65 every 3 years (policy 33) vs no screening	756 per woman	23.1 403LY per woman 23.0359 QALY per woman	7,521 /LYG 9,009 /QALY	
				Sc2: Mammography screening 45 to 65 every 3 years (policy 29) vs screening age 50 to 65 every 3 years (policy 33). (Contribution of age 45-50)	864 per woman	23.1518 LY per woman 23.0465 QALY per woman	9,470 /LYG 10,207 /QALY	
				Sc3: Mammography screening age 45 to 70 every 3 years (policy 30) vs screening 45 to 65 every 3 years (policy 29). (Contribution of age 65-70)	934 per woman	23.1583 LY per woman 23.0521 QALY per woman	10,828 /LYG 12,568 /QALY	
				Sc4: Mammography screening age 40 to 70 every 3 years (policy 26) vs screening age 45 to 70 every 3 years (policy 30). (Contribution of age 40-45)	1,103 per woman	23.1701 LY per woman 23.0626 QALY per woman	14,324 /LYG 16,097 /QALY	

Reference	Design, Perspective, Time Horizon, Country, Population	Model type, Comparators/ Program designs	Brief methodology	Mammography screening scenario comparison	Cost CAD 2013	Effectiveness	ICER ($/LYG or /QALY)	Findings
				Sc5: Mammography screening age 40 to 75 every 3 years (policy 27) vs screening age 40 to 70 every 3 years (policy 26). (Contribution of age 70 to 75)	1,130 per woman	23.1718 LY per woman 23.0640 QALY per woman	15,635 /LYG 18,986 /QALY	
				Sc6: Mammography screening age 40 to 80 every 3 years (policy 28) vs screening age 40 to 75 every 3 years (policy 27). (Contribution of age 75-80)	1,170 per woman	23.1737 LY per woman 23.07654 QALY per woman	21,279 /LYG 28,879 /QALY	
				Sc7: Mammography screening age 40 to 80 every 2 years (policy 16) vs screening age 40 to 80 every 3 years (policy 28). (Contribution of screening every 2 years vs 3 years)	1,527 per woman	23.1797 LY per woman 23.0697 QALY per woman	59,490 /LYG 84,986 / QALY	

Stout et al., 2006	CUA Perspective not specified Lifetime United States Women age 40 years and older	Discrete-event simulation model No screening vs sixty-four additional scenarios, each with a particular fixed screening schedule varied by the age at the first screen (40, 45, 50, or 55 years) and at the last screen (65, 70, 75, or 80 years) and by the screening interval (1, 2, 3, or 5 years). Only non-dominated scenarios included in table	A natural history of disease model which incorporates secular trends in breast cancer risk, screening use, and treatment dissemination. By simulating the individual life histories of women aged 20 years or older who were born in 1891 through 1980 in proportion to their prevalence in the U.S. population and aggregating the outcomes, the model can replicate population-level U.S. cancer surveillance data corresponding to calendar years 1975 through 2000. Main data sources include the National Center for Health Statistics, Surveillance, Epidemiology, and End Results (SEER) and costs from the literature. 3% discounting of effects and costs. 2000 USD.	No screening (mean cost and mean QALYs presented)	188 billion	945.8 million QALY	-	Screening in general did not appear to be cost-effective under a $50,000/QALY threshold. The cost per QALY is substantially higher when screening younger women age 40-45. Increasing screening frequency from 5 to 3 years or 3 years to 2 years, incurs a cost per QALY that is on the verge of being cost-effective. Going from screening every 2 years to 1 year is substantially more costly for every QALY gained.
				Screen age 55-70 every 3 years vs Screen age 55-70, every 5 years. (Contribution of screening frequency – 3 vs 5 year intervals)	237 billion	946.8 million QALY	56,667/ QALY	
				Screen age 45-75 every 3 years vs Screen age 50-75 every 3 years. (Contribution of age 45-50)	274 billion	947.4 million QALY	60,000/ QALY	
				Screen age 45-75 every 2 years vs Screen age 50-75 every 2 years (Contribution of age 45-50)	310 billion	948.0 million QALY	63,333 / QALY	
				Screen age 45-75 every year vs Screen age 45-80 every year (Contribution of age 75-80)	406 billion	949.1 million QALY	Dominated	
				Screen age 40-80 every year vs Screen age 45-80 every year at 3% (Contribution of age 40-45)	459 billion	949.6 million QALY	140,000 / QALY	
				Screen age 45-75 every2 years vs Screen age 45-75 every year (Contribution of screening frequency 1 year vs 2 year interval)	381 billion	949.1 million QALY	90,000/ QALY	

				Screen age 45-75 every 3 years vs Screen age 45-75 2 years. (Contribution of screening frequency 2 year vs 3 year interval)	318 billion	984.0 million/QALY	58,333/QALY	
Tosteson et al., 2008	CUA Societal and Medicare perspectives Lifetime United States Women age 40 years and older, screening frequency 1 or 2 years	Discrete-event simulation model All-film mammography (done in all women) vs targeted digital mammography (age-targeted <50, and age- and density-targeted <50 or >=50) vs all-digital mammography (all women)	Computer-based model simulates the life histories of women by using four interacting processes: breast cancer natural history, breast cancer detection, breast cancer treatment, and competing-cause mortality. Simulation incorporates actual age-specific U.S. screening patterns, observed secular trends in cancer risk, and dissemination of adjuvant treatment from 1975 to 2000. Mammography performance and resource use data from the Digital Mammography Imaging Screening Trial. Quality of life weights from the Medical Expenditure Panel Survey which used EuroQoL EQ-5D. 3% discounting of effects and costs.	-	-	-	-	Study assesses differences in screening modalities. Does not provide enough detail to assess contribution of age and/or screening frequency to cost-effectiveness of mammography.

Wong et al., 2007	CEA & CUA Societal perspective Lifetime Hong Kong Chinese women age 40 and older	State-transition Markov model Compared the results from 2-view film mammography every 2 years, beginning at ages 40 years or 50 years and ending at ages 69 years or 79 years, with the results from no screening. Results shown for two models: single-cohort and multiple-cohort	Simulates biennial mammography, breast cancer diagnosis, and treatment in a hypothetical, population-based cohort. Natural history of disease component. Model incorporates cancer incidence treatment, risk of cancer, mortality, stage distributions, and direct medical costs related to screening, diagnostic work-up, initial treatment, and terminal care. Quality of life weights included for health states of Healthy, DCIS, stage I, stage II, stage III, and stage IV cancer. Use of probabilistic sensitivity analysis (parameter uncertainty) around clinical parameters and multi-cohort approach (structural uncertainty). 3% discounting of effects and costs.	No screening (single-cohort)	68.6 million	2.4 million LY	-	None of the screening program options are cost-effective at WTP $50,000 / LYG. Small marginal impact of screening in 70-79 age group and high cost. Greater impact with screening women age 40-49.
				Sc 1: Biennial screening age 50-69 vs no screening (single-cohort)	225.3 million	2.4 million LY	98,540/LYG (Dominated)	
				Sc2: Biennial screening age 50-79 vs Biennial screening age 50-69 (single-cohort)	267.3 million	2.4 million LY	349,833/LYG (Dominated)	
				Sc 3: Biennial screening age 40-69 vs Biennial screening age 50-69 (single-cohort)	353.5 million	2.4 million LY	54,248/LYG	
				Sc 4: Biennial screening age 40-79 vs Biennial screening age 40-69 (single cohort)	395.4 million	2.4 million LY	349,750 / LYG	
				No screening (multiple cohort)	60.84 million	1.9 million LY	-	
				Sc 5: Biennial screening age 50-69 vs no screening (multiple cohort)	192.3 million	1.9 million LY	418,726/LYG Dominated	
				Sc 6: Biennial screening age 40-69 vs Biennial screening age 50-69 (multiple cohort)	232.1 million	1.9 million LY	68,185/LYG	
				Sc 7: Biennial screening age 50-79 vs Biennial screening age 40-69 (multiple cohort)	255.1million	1.9 million LY	Dominated	
				Sc 8: Biennial screening age 40-79 vs Biennial screening age 40-69 (multiple cohort)	294.9 million	1.9 million LY	201,025 /LYG	

Table 5 Summary of included studies - special populations

Reference	Design, Perspective, Time Horizon, Country, Population	Model type, Comparators/ Program designs	Brief methodology	Mammography screening scenario comparison	Cost CAD 2013	Effectiveness	ICER ($/LYG or /QALY)	Findings
Lee J. et al., 2010	CUA Societal perspective Lifetime United States Women age 25 and older with genetic predisposition (BracA1 gene mutation carriers)	Markov Monte Carlo simulation model Three annual screening strategies starting at age 25: (a) screen-film mammography, (b) MR imaging, and (c) combined mammography and MR imaging (combined screening)	Model composed of three linked modules: (a) breast cancer development and detection, (b) treatment and follow-up, and (c) screening. Individual women entered the breast cancer development and detection module at the beginning of the simulation. Use of logarithmic growth model. Use of data from randomized controlled trials for sensitivity and specificity of screening, and costs from administrative data. Quality of life weights derived from medical literature.	No screening (clinical surveillance)	120,544 per woman	44.21 QALY per woman	-	Screening for younger women with genetic predisposition is with film mammography is cost-effective compared to no screening, but not cost-effective when screening is with MRI or combination of MRI and mammogram.
				Sc1: Annual screen film mammography vs No screening	125,935 per woman	44.46 QALY per woman	21,561 /QALY	
				Sc2: Annual MRI vs Annual film screen	136,358 per woman	44.5 QALY per woman	Dominated	
				Sc3: Annual combined mammogram and MRI screening vs Annual film screen mammography	139,285 per woman	44.624 QALY per woman	81,405 /QALY	
Wong et al., 2008	CEA Health care payer perspective Lifetime Australia	Deterministic Markov model Annual breast cancer screening using mammography vs no screening for all	Simulates the natural history of breast cancer in a hypothetical cohort of women on dialysis therapy over time. Model includes benefits, and harms of breast cancer screening across the different health states. Clinical data from Australian	No screening	4,926 per woman	5.9769 LY per woman	-	No screening program option is cost-effective at WTP $50,000 /LYG. Any benefits of cancer screening may be eradicated by the
				Sc 1: Annual mammography of all women (with and without diabetes) age 50+ on dialysis vs no screening	5,377 per woman	5.9732 LY per woman	130,670 / LYG	

	Women age 50-69 on dialysis therapy	women and women on dialysis therapy with a starting age of 50 years.	and New Zealand Dialysis and Transplant Registry and published literature. Direct costs from administrative data. 5% discount rate.	Sc 2: Annual mammography of women with diabetes age 50+ vs no screening	5,371 per woman	5.9773 LY per woman	1,112,700 / LYG	excess mortality and morbidities associated with end-stage kidney disease.
				Sc 3: Annual mammography of women with diabetes on dialysis age 50+ vs no screening	Not specified	Not specified	131,788 / LYG	

Analytic Models and Data Sources

All included studies used either a microsimulation or cohort simulation model. All included studies incorporated a natural history of disease component. However, health states including stage of breast cancer were defined differently across studies. For example, studies by de Gelder et al.(2009) and Okonkwo et al.(2008) used the MIcro-simulation SCreening ANalysis Model (MISCAN) which models tumour development as a progression through successive invasive disease stages T1A, T1B, T1C and T2+. Lee et al.(2010) and Arveux et al.(2003) used invasive TNM stages, plus non-invasive DCIS (ductal carcinoma in situ), while the studies by Mandelblatt et al. (2005), Rojnik et al. (2008), and Stout et al. (2006) categorized stages as DCIS, local, regional, and distant breast cancer. Neeser et al. (2007) defines malignant breast cancer states within their model as DCIS, invasive ductal carcinoma or invasive lobular carcinoma.Wong et al. (2007) used staging employed by Surveillance, Epidemiology, and End Results (SEER) which are categorized as stages I, II, III, and IV. Wong et al. (2008) also used the same SEER staging as well as DCIS. Tosteson et al. (2008) used a model developed at the University of Wisconson as part of the U.S. National Cancer Institute`s Cancer Intervention and Surveillance Modeling Network (CISNET). This model is described to be an adaptation of the Shwartz model which models tumour growth using a logarithm with tumours entering the model at less than 0.2 cm. (Wisconsin-Harvard Model, 2013).

The data used to parameterize the models can be categorized into baseline data (e.g. data establishing the baseline mortality, and incidence rates) and exposure-response information (e.g. data linking interventions with their elemental impacts). Sources of effectiveness and cost data included same country-specific screening program data, other country screening program data, cancer surveillance data, efficacy estimates from randomized trials, effectiveness data from other quasi-experimental studies, and administrative data.

Analytic Perspective and Time Horizon

An analytic perspective is the point-of-view by which the analysis is taken and the results are interpreted. Typical analytic perspectives used in health economic evaluations the health care system or statutory health insurer perspective, the societal perspective, or both, or the

perspective was unclear (or not stated). The studies by Lee et al. (2010), Mandelblatt et al. (2005), and Wong G et al. (2008) took a societal perspective. Tosteson et al. (2008) also took a societal perspective to their analysis in addition to a Medicare perspective. Rojnik et al. (2008) and Neeser et al. (2007) also took a health care system/statutory health care insurer. Wong I et al. (2007) used the health care system perspective which, for the context of the Australian setting, could mean the public (Medicare) and/or private care sector. The studies by Arveux et al. (2003) de Gelder et al. (2009), Okonkwo et al. (2008), and Stout et al. (2006) did not specify their analytical perspective. In health economic evaluations, the analytic perspective has an impact only on costs. The societal perspective takes into account both direct (health care) and indirect costs (e.g. caregiver costs, short-term and long-term disability, productivity, etc.) where as the payer perspective usually includes only direct costs. Costs are therefore generally higher when taking a societal perspective and as a result ICERs would also be higher.

The time horizons adopted across studies were either 20 years (Arveux et al., 2003; De Gelder et al., 2009) or a lifetime (Mandelblatt et al., 2005; Wong G et al., 2008; Wong I et al., 2007; Lee et al., 2010; Rojnik et al., 2008; Stout et al., 2006; Tosteson et al., 2008; Okonkwo et al., 2008; Neeser et al., 2007). The use of a lifetime time horizon is preferred since it allows for the tracking of health benefits and costs over the maximum time and a more accurate account of the long-term impact of screening mammography on life expectancy.

Cost-Effectiveness

The comparative cost-effectiveness of the different screening programs examined in each study is summarized in table 1. The incremental cost-effectiveness ratios of the screening options evaluated varied substantially across studies even after adjusting to 2013 CAD. If one were to use a $50,000 per life year saved or quality adjusted life year saved as a threshold for incremental cost-effectiveness, only seven of eleven included studies reported screening scenarios that were cost-effective (Arveux; De Gelder; Lee; Oknkwo; Rojnik;Tosteson). The study by Arveux et al. (2003) compared the cost-effectiveness of no screening with the current screening program in France. Although the authors specify the screening age of 50 to 65, they do not specify the screening interval. They found that the

current screening program was cost-effective compared to no screening with a cost of $30,223 per LYG with no discounting, and $38,536 per LYG with 5% discounting of costs.

The most cost-effective screening program option reported in the study by De Gelder et al. (2009) was 60% biennial mammogram program screening and 20% opportunistic screening of women age 50 to 69, at a cost of $9,823 per LYG or $10,524 per QALY, compared to 80% biennial mammogram program screening at 3% discounting of costs and effects. Biennial screening mammography for women age 40 to 60 was the most cost-effective screening option in the study by Okonkwo et al. (2008) with a cost of $36,606 per LYG versus annual clinical breast examination with 3% discounting. Rojnik et al. (2008) found screening mammography for women age 40 to 80 every three years more effective than screening women age 40 to 75 every three years at a cost of $21,279 per LYG or $28,879 per QALY with 3% discounting. Age-targeted digital mammography for women age 40 to 50 was found to be more cost effective than all-film mammography for women age 40 and older in the study by Tosteson et al. (2008), at a cost of $32,530 per QALY with 3% discounting. Stout et al. (2006) found that the most cost-effective screening option was screening mammography for women ages 55 to 70 every 5 years at a cost of $45,714 per QALY compared to no screening with 3% discounting. Lee et al. (2010) found annual screen film mammography was cost-effective for women age 25 and older with genetic predisposition (BrcA1 gene mutation carriers) compared to no screening at a cost of $21,561 per QALY. None of the screening options assessed in the studies by Mandelblatt et al. (2005), Neeser et al. (2007), Wong et al. (2008), and Wong et al. (2007) were found to be cost-effective.

Age, Screening Frequency, and Special Populations

It was observed that for the standard age of screening (50-69) and for younger and older age ranges (40 to 45 or 40-49, and 70-75 or 75-80), 2- or 3-year screening intervals were cost-effective. In general, annual screening for any population did not appear to be cost-effective. In terms of screening in special populations, in one study it was found that early screening (starting at age 25) for women with genetic risk for breast cancer was cost-effective compared to no screening (with film mammography), but not when MRI was used instead of film mammography or in combination with film mammography. In the one study that looked

at screening for women on dialysis, screening was not found to be cost-effective because any the benefits of cancer screening is probably offset by the excess mortality and morbidities associated with end-stage kidney disease in this population.

Table 6 summarizes the final results of the review. Based on the reviewed health economic studies, the cost-effectiveness results of only a small number of studies (7 studies) could be compared i.e. they provided enough information on costs and effects and screening program design in terms of age and screening interval or frequency. Based on this review, of the studies that found screening mammography to be cost-effective, there was only 1 study that could support screening for each of the specific screening policies as outlined in table 6. Therefore, based on the limited number of health economic studies, any recommendations to support specific screening mammography policy options based on the health economic literature would be weak due to the low number of supporting studies.

Table 6 Review Summary Results of Included Studies

Number of studies with no comparable results:	4 studies
Number of students where screening mammography was not found to be cost-effective:	4 studies
Characteristics of optimal screening policies –	
Age 25+, 1-year interval:	1 study*
Age 40-49, 2-year interval:	1 study
Age 40-49, 3-year interval:	1 study†
Age 50-64, 3-year interval:	1 study†
Age 65-80, 3-year interval:	1 study†

Notes: * Genetically predisposed population. † Study by Rojnik et al. (2006)

Lessons Learned for a Cost-Effectiveness Study for Canada

Several of the examined studies used limited and what appeared to be incomplete analyses. For example, the study by de Gelder et al. (2009) conducted a fairly simplistic sensitivity analysis that only assessed the impact of variations in false-negative test rates. They could have explored the impact of adjusted values for other and perhaps more policy-relevant parameters including different costs estimate and discount rates, and variations of age eligibility, screening sensitivity, and screening frequency/interval. Mandelblatt et al. (2005) assumed Medicare reimbursements were closely approximated to societal costs within the context of their screening program. It is unclear whether in fact societal costs which would

include indirect productivity costs and patient costs such as transportation and caregiver costs would be covered under that specific national health insurance program. These are considerations that would need to be included in any health economic analysis – costs in relation to the analytic perspective or viewpoint taken. Neeser et al. (2007) only considered the most explicit aspects of breast cancer in their model and did not use a more complex model compared to other studies. These other studies employed a natural history of disease approach e.g. tumour size or stage of cancer progression. A more macro-approach may under- or over-estimate the true impacts of screening.

Some of the included studies omitted important details in their methodology, which was a challenge in comparing and contrasting the individual studies against each other. Neeser et al. (2007) did not specify the screening frequencies explored in their analysis thus leaving the reader to assume that an annual screening interval only was used. In addition, the methods used in the study by Rojnik et al. (2008) to estimate and adjust for health-related quality of life was not well-explained, making it difficult to validate the QALY estimates. The study by Stout et al. (2006) did not include costs related to improving uptake of screening mammography (recruitment costs). As well, the model details in the study by Tosteson et al. (2008) seemed lacking in regards to the sensitivity analysis methods. In the Wong et al. (2008) study, the attendance rate used in the sensitivity analysis could have been adjusted to below 70% and detection rates could have been reduced, which are reasonable assumptions since it could be argued that screening mammography programs in India may not be as effective as it is in the Netherlands where they have been established for many years.

Some of the included studies had potential validity issues. For example, in the study by Arveux et al. (2003), it was difficult to discern the impact of a no screening option since the presence of any screening in the population would no doubt influence opportunistic screening uptake. There was a possibility of length-bias[7], lead time bias[8], and healthy

[7] The overestimation of survival benefit due to the detection of slowly growing lesions by screening tests; including lesions that will never cause mortality. (Institute of Medicine (US) and National Research Council (US) Committee on New Approaches to Early Detection and Diagnosis of Breast Cancer, 2005)
[8] The overestimation of survival time because of the backward shift in the starting point for the measurement of survival as a result of early detection. (Institute of Medicine (US) and National Research Council (US) Committee on New Approaches to Early Detection and Diagnosis of Breast Cancer, 2005)

volunteer bias[9], where screeners would have a lower inherent risk for breast cancer than non-screeners. External validity within the study by Rojnik et al. (2008) may be compromised since they used data from other counties in their model which may not have been generalizable to their target population. The authors did not address this potential bias. Similarly, Wong et al. (2007) used data from the US which may not be generalizable to the Chinese population. For example, base incidence rates or risk of breast cancer are much higher in the US than in China. The study results of Okonkwo et al. (2008) may not be valid as they observed larger health benefit from clinical breast examination (CBE) compared to screening mammography in Indian women which is most likely due to their use of the same sensitivity estimate for mammography as for CBE. If CBE and mammography truly had the same sensitivity rate, the lack of improved sensitivity via mammography may be a reflection of a limited availability of capacity or technical expertise to do and/or read a mammogram. Nevertheless, screening mammography should still yield an improved sensitivity estimate over CBE and at the least; Okonkwo et al. (2008) could have explored the impact of differential sensitivity by modality in a sensitivity analysis.

Discussion

Based on the included studies in this review, it was challenging to determine which screening program design (in terms of appropriate age eligibility and screening frequency/interval) is the most cost-effective since the studies varied across several characteristics. Depending on the context (baseline risk of breast cancer, disease incidence) and the screening program design (age eligibility, screening frequency/interval, screening modality/technology); breast cancer screening in general may or may not be cost-effective. For developed countries, it appears that annual screening with mammography is never cost-effective, unless it is for high-risk population (genetic predisposition). Screening every two years or more may be cost-effective. In terms of age eligibility, screening women ages 50 up to 80 seems to be the most cost-effective age group to screen. However, as mentioned, very few studies in this review supported this finding. In terms of modalities, screening mammography was in general more cost-effective than digital mammography or

[9] When the participants are healthier than the general population. (Froom et al., 1999)

mammography and clinical breast examination combined. In the case of digital mammography, a mixed program where digital mammography was offered to younger women with dense breasts (where film mammography would be not as effective) seemed to be more effective and cost-effective compared to film mammography. Screening was more cost-effective in countries where breast cancer incidence is relatively high, such as in France, Switzerland, Australia, the Netherlands, and the US. Study findings from these countries may be comparable to the Canadian context based on similar demographic characteristics including population risk of breast cancer.

There were several limitations of this review. Firstly, the search strategy may be considered somewhat limited in scope which may have biased the study findings. For example, the review excluded published studies in languages other than English, as well as grey literature (reports and other documents that were not peer reviewed or published in an academic tradition). It has been reported that cost-effectiveness studies tend to show publication bias, though this relates usually to industry-sponsored studies (such as studies evaluating the cost-effectiveness of drugs, technology, and medical devices) (Bell et al., 2006). When published ratios cluster around a proposed threshold, bias may exist, and health policies that are then based on these values may be erroneous (Bell et al., 2006). Although it is possible that this review may have missed some unpublished cost-effectiveness studies, based on a preliminary review of the grey literature using a general internet search, no traditionally published studies that could have been included were found. Further, health economic analysis of screening or other public health interventions are less likely to be amongst the grey literature (unpublished) as these types of studies are unlikely to be sponsored by industry. It is therefore unlikely that the exclusion of grey literature would have biased the results of this review.

One of the more important limitations of this review is that it may be considered outdated. This review was completed in 2010 and since this time, several years of health economic research evaluating screening mammography have been published. A preliminary review of the more recently published literature found at least 30 potentially included studies for review. An update to this review with literature published after 2010 would be required to ensure greater relevance of the review findings.

Another potential limitation is the exclusion of screening modalities other than mammography in this review. In certain circumstances, magnetic resonance imaging is used for screening; however it is not standard practice in Canada which is why we chose to limit the review to mammography only. Further, more recently Canada has almost completely phased-out the use of film mammography and replaced older technology with digital mammography. It is likely that newer studies cost-effectiveness studies will evaluate the health economic impact of this more sensitive and efficient technology. Several provinces in Canada employ the use of mobile units in order to provide access to screening for women in hard-to-reach regions, particularly in the very northern regions of Canada. This review did not include mobile units as a primary intervention for evaluation which may have provided additional information on cost-effectiveness of mobile units for these populations.

Screening mammography has long been accepted in the Western world as an effective public health intervention in the secondary prevention of breast cancer (Public Health Agency of Canada, 2014). Since its inception, screening mammography for masses has received little opposition from the general public. However, discussions surrounding the true benefits and harms of screening mammography have emerged over time. The advancement of treatment, the reduction in other risk factors such as hormone replacement therapy has resulted in improved breast cancer survival (Canadian Cancer Statistics, 2014).

The reported effectiveness of screening mammography on cancer mortality appears to vary across studies depending on study design and other contextual factors. For example, a study by Kalager et al. (2010) reported only a 10% reduction in breast cancer mortality attributable to screening mammography in Norway. In a more recent study by Broeders et al. (2012) that reviewed twelve European observational studies, among studies that reported on annual percentage changes, breast cancer mortality reductions attributable to population-based screening mammography were estimated to be in the range of 1% to 9%, annually. In studies that compared post- and pre-screening periods, the mortality reduction was estimated at 28% to 36%. In the same review, among incidence-based studies (prospective cohort-like studies), the pooled mortality reduction was estimated to be 25% among women who were invited for screening and 38% among women who were actually screened. Among case-control studies, the pooled mortality reduction was estimated to be 31%. Other researchers

have publicly denounced population-based screening mammography based on claims of harms outweighing benefits, including excessive use of lumpectomies, mastectomies, and radiotherapy, high rate of false positive tests, and over-diagnosis (Wright, 1995; Olsen and Gotzsche, 2001; Horton, 2001; Gotzsche and Olsen, 2000).

In 2009, the U.S. Preventive Task Force updated their screening mammography guidelines by advising screening on a biennial basis for women aged 50-74 only (U.S. Preventive Services Task Force, 2009). This garnered much displeasure among women's groups who have argued that women aged 40-49 should also be screened, despite a lack of evidence for success or cost-effectiveness to support screening for this age group (Rosenquist and Lindfors, 1994). Whether screening for the masses is worth the potential harms and costs is still a debated issue. Trade-offs between the benefits, harms, and costs associated with various screening guidelines should be considered when making recommendations for routine screening. Mammography screening younger women (under the age of 50) has not been found to be as cost-effective as screening older women (Rosenquist and Lindfors, 1994). There is also an issue of resource capacity. Recommendations that involve increasing the frequency or proportion of women screened risk causing backlog and with consequent increases in wait-times for all women, including those who are at increased risk.

The lack of consensus in regards to the cost-effectiveness of screening mammography is most likely due to the lack of strong evidence available to support current practices in terms of effectiveness and efficiency. Other considerations include the assessment of the potential impact of longer screening intervals for women of moderate risk or the impact of tailored screening for women at high-risk according to age and family history or genetic predisposition, and the use of more sensitive screening modalities (e.g. magnetic resonance imaging) for women under the age of 50 in this group We must also consider the impact of screening vulnerable sub-groups of the population, including women with mental and physical disabilities who face challenges with not only accessing screening, but accessing the health care system in general. The evidence is vastly lacking on the cost-effectiveness of population-based screening mammography in Canada.

The studies included in this review provide useful guidance in developing a model for Canada. Currently, Canadian women participate in screening either through an organized program, or are screened opportunistically (Public Health Agency of Canada, 2008). Over time there have been a number of important shifts in the way women are screened within organized programs in Canada, and these policies vary regionally. In addition, some provinces are phasing out the use of analog or film mammography for digital mammography, which has been found to be more sensitive in detecting true cancers as opposed to false positives (positive after a screen but negative at diagnosis) (Pisano et al., 2005). The varying policies have significant impacts on a number of outcomes, including the ability for a program to obtain adequate coverage of the at-risk population, wait-times, and costs related to screening, diagnosis, and treatment (Gunes et al., 2004). Within the context of a publically funded health care system, decisions regarding which services should or can be funded, and by how much, are particularly difficult to make. Health economic evaluations are meant to guide these decisions in a way that makes sense for the context in which they must be made. It is important, however, that health economic evaluation studies are explicit in their methodology and assumptions, and that the evidence from them is reliable and valid if they are to be used in practice and policy.

Chapter 4 (manuscript 2): The impact of age and screening interval on the cost-effectiveness of screening mammography in Canada

Authors:

Ngoc-Thy Dinh[1], Kevin Brand[2], Doug Coyle[3], Howard Morrison[4], Catherine Deri-Armstrong[5]

Affiliations:

1. Health Economics, Public Policy Division, Conference Board of Canada

2. Telfer School of Management, University of Ottawa

3. Department of Epidemiology & Community Medicine, University of Ottawa

4. Department of Epidemiology & Community Medicine, University of Ottawa

5. Department of Economics, University of Ottawa

Abstract and Key Words

Background**:** Canadian public health guidelines recommend screening mammography every 2 to 3 years for women age 50 to 74 years. More than 60% of Canadian women participate in routine screening. Population-based or mass screening mammography is a debated issue as clinicians, researchers, policy-makers, and the general public are still uncertain and divided over the trade-offs between health benefits, health risks, and costs, of screening. Although there are Canadian studies that address the health benefits of screening mammography in terms of improved survival from breast cancer, there is a lack of cost-effectiveness analyses comparing the costs and health benefits attributable to mass screening. To add to the body of knowledge, the objective of this study was to identify the optimal screening mammography program options for Canada characterized by age eligibility and screening frequency.

Methods: A Canadian microsimulation model was used to examine the health and cost impact of different screening policy options in a hypothetical female cohort over a lifetime. The impact of screening mammography options on breast cancer outcomes, including mortality and costs associated with screening and treatment was assessed. The modeled screening options were created by altering screening age, with possible start age of 40 and end age of 79, and screening frequency, with possible annual and biennial screening intervals. The costs and benefits (measured in life-years) were estimated for 11 screening mammography program options, including no screening. The base case analysis compared the expected direct costs and life-years across screening options. Costs included screening and treatment costs and were adjusted to 2013 Canadian dollars. Health benefits and costs were discounted at an annual rate of 5%. The optimal screening policy options were determined based on sequential analyses and a willingness-to-pay (WTP) threshold of $50,000 per life-year gained (LYG). Univariate sensitivity analyses using low and high estimates for the parameters of discount rate, cost per screen, initial participation rate, and test sensitivity and specificity were conducted, as well as using low and high WTP thresholds. Several additional screening policy options were added to the analysis to examine the impact of offering less frequent screening (3-year intervals).

Results: For the base case analysis the optimal screening policy option under a WTP of $50,000 per LYG was option biennial screening for women ages 40 to 79. Cost-effectiveness

of screening was sensitive to the screen costs and discount rate and WTP threshold. If WTP was $30,000 the optimal screening policy option would have been biennial screening for women ages 50 to 69.

Conclusions: Cost-effective screening is dependent on the payer's willingness-to-pay threshold which may not be the standard $50,000. It has been suggested that a more conservative (lower) WTP threshold should be employed for cost-effectiveness studies. Decisions for screening should consider other factors including quality of life, individual level risks, and other costs associated with disease including out-of-pocket patient costs and lost productivity

MeSH: screening; mammography; model; breast; cancer; breast neoplasms; cost-effectiveness

Introduction

Breast cancer is the second leading incident cancer and the second leading cause of cancer death among Canadian women (Canadian Cancer Society, 2014). It was estimated in 2014 there were about 24,400 new cases of breast cancer and 5,000 breast cancer deaths in Canada (Canadian Cancer Society, 2014). The estimated lifetime odds of developing breast cancer for a Canadian woman are 1 in 9, with 1 in 28 Canadian women eventually dying from it (Canadian Cancer Society, 2014). The current estimated average 5-year survival rate for breast cancer is relatively high at 88% (Navaneelan and Janz, 2011). Between 1979 and 1999, the incidence of breast cancer in Canada steadily increased but since 1999 it has been declining by a rate of about 1.7% per year (Canadian Cancer Society, 2010). Breast cancer costs the Canadian economy almost $500 million a year in health care costs and lost productivity (lost wages) in terms of short- and long-term disability and premature mortality (Economic Burden of Illness in Canada, 2014). In comparison with other female cancers, the cost of breast cancer is almost 20-times higher than that of cervical cancer (Economic Burden of Illness of Canada, 2014). It has been suggested that the recent decline in breast cancer mortality may be attributable to participation in screening mammography and advancements in diagnostic technology and treatment, while decline in disease incidence may be attributable to changes in the population prevalence of risk factors, such as the decrease in the use of hormone replacement therapy in the population (Canadian Cancer Society, 2014).

The Population Health Model (POHEM) is a microsimulation model of disease and risk factors developed by Statistics Canada. The model generates and ages a large cohort that is representative of Canada, one individual at a time, and tracks the life-trajectory that is guided by exposure to varying life events of each individual until death. The model includes data from various Canadian sources, including large cross-sectional and longitudinal surveys, disease registries, health administrative databases, vital statistics, Census, and treatment cost data estimates extracted from the published literature. The model has been adapted for several chronic diseases including osteoarthritis, acute myocardial infarction, and lung and colorectal cancer. POHEM with the added breast cancer module (POHEM-BCS) was used in a study by Will et al. (1999) to project the lifetime treatment costs attributable to breast

cancer. The updated model includes a breast cancer screening component that offers the capacity to simulate the impacts of changes to various input parameters, including cohort size, age, screening frequency, participation rate, test sensitivity and specificity, and baseline disease incidence. This study uses POHEM-BCS to study the impact of age eligibility and screening frequency on the cost-effectiveness of screening mammography in the secondary prevention of breast cancer in Canada.

Methods

Study Design

This study is a cost-effectiveness analysis (CEA) study examining the impact of age eligibility and screening frequency on the cost-effectiveness of screening mammography in the secondary prevention of breast cancer. Cost-effectiveness is determined based on a measure of the incremental cost per life-year gained across several screening mammography policy options characterized by age and screening frequency. The tested screening policies include a "no screening" option.

Population

In this analysis the population of interest are Canadian women with screening mammography offered to women as young as 40 years and as old as 79 years, depending on the screening option tested in the analyses. Currently, Canadian women can access screening mammography either opportunistically through a referral from her family physician or through self-referral to an organized regional program. For the most part, provincial and territorial breast cancer screening programs actively recruit women ages 50 to 69 every two years for screening. It is recommended that women with elevated risk of breast cancer compared with the general population be screened more frequently. There are certain regions in Canada that also accept women age 40-49 years and/or 70-79 years to participate in their organized screening programs. The selection of a lower limit of 40 and an upper limit of 79 is in-line with the lowest and highest age limits within eligibility screen criteria across provincial/territorial programs in Canada. Women starting at age 40 years are accepted into screening mammography programs in the following regions: Northwest Territories, Yukon Territories, British Columbia, Alberta, Nova Scotia, and Prince Edward Island. Women are

continued to be invited for screening up to age 79 in only British Columbia. Few regions provide screening or follow-up with women ages 80 and older except in the Northwest Territories and Yukon Territories. (Canadian Partnership Against Cancer, 2013)

The debate surrounding population-based breast cancer screening is mostly regarding whether the health benefits of screening women in different age groups is worth the costs and potential harms. There are ongoing debates about the effectiveness of screening mammography which point out that much of the evidence to support screening is based on old data; evidence which too frequently fails to reflect the improvements in treatment, and greater public awareness about the disease.

Population Health Model-Breast Cancer Screening Module (POHEM-BCS)

This study uses a Canadian microsimulation model developed by Statistics Canada called the Population Health Model - Breast Cancer Screening Module (POHEM-BCS). The model allows for the comparison of alternative options or policies ("what if" options) in terms of health and economic outcomes including burden of diseases, life-years lived, mortality, diagnosis and treatment costs. POHEM-BCS is parameterized with Canadian demographic and surveillance data. It is a longitudinal, discrete-event[10], and stochastic[11] model that simulates the life-paths or health biographies for individuals over the life course. The base model infrastructure (generally referred to as POHEM) has been used to model cancers of the breast, colon, and lung, as well as for osteoarthritis (Houle et al., 1997; Will et al., 1999; Kopec et al., 2010; Berthelot et al., 2000; Flanagan et al., 2003; Flanagan et al, 2006; Nadeau et al., 2013). Most recently the POHEM has been used to estimate the impact of physical activity on prevalence of chronic conditions, mortality, and quality of life in the Canadian population. (Flanagan et al., 2013)

POHEM–BCS was used in this study to estimate the life-time impact of a variety of screening program options characterized by age eligibility and screening frequency (1 or 2 year intervals). In the model, life histories are applied to a hypothetical birth cohort of

[10] Each event occurs at a particular instant in time and marks a change of state in the system. - Stewart Robinson (2004). Simulation – The practice of model development and use. Wiley

[11] Of or pertaining to a process involving a randomly determined sequence of observations each of which is considered as a sample of one element from a probability distribution. - stochastic. (n.d.). *Dictionary.com Unabridged*. Retrieved March 16, 2014, from Dictionary.com website: http://dictionary.reference.com/browse/stochastic

Canadian women. A population of 6 million women was used in the model to approximate the number of women age 40 and older in Canada. The model is not equipped to predict tumour size (as are some models) and thus does not link its transition probabilities to tumour size. Rather the model uses breast cancer stage distribution estimates in conjunction with stage-specific transition rates. A stage distribution representing pre- and post- screening conditions as well as for the transition rates (i.e., representing the stage-specific survival probabilities) are imbedded in the model. An individual woman modeled by POHEM-BCS can transition through various (some progressive) health states. All women start out as healthy and as they age are subject to various age-specific state transition probabilities. These probabilities determine events ranging from participation in screening to transitioning from full-health to disease, and from disease to death. The health states of interest to this work include no breast cancer, breast cancer, treatment (initial, local recurrence, distant recurrence, or terminal care), and death (from breast cancer or other cause). The TNM classification of breast cancer staging was used to split cancers into stage I, II node negative, II node positive and stages III and IV combined. The original developers of the POHEM-BCS model (Flanagan et al, 2006) chose this staging for two reasons. They argued first that survival differs significantly between these stages and second it was considered to be more amenable to modeling of treatment impact in future work.

POHEM-BCS generates several different population-level health and economic outcomes. The outcomes of interest in this analysis, which are all available by age group, include:

- Survival in person-years;
- Number (and outcomes) of screens; and
- Diagnosis and treatment costs

Data Sources and Model Structure

The majority of the input data used in the model were pre-existing as part of the POHEM-BCS infrastructure. This version of the model (available for public use) was not easily amenable to change or update for the majority of input parameters. Pre-existing data included breast cancer incidence rates in the absence of screening, breast cancer incidence rates in the presence of screening, the stage distribution of clinically detected cancers, the stage distribution of screen-detected cancer, survival rates, and diagnostic and treatment

costs. These data and their sources are provided in appendix1. Breast cancer incidence rates in the absence of screening in POHEM were originally estimated from the National Cancer Incidence Reporting System database (a Canadian database). In earlier work by Will et al. (2000), the breast cancer incidence rates used in POHEM-BCS were obtained by adjusting rates that were derived for the period 1982-1985 from the Saskatchewan Cancer Registry in a manner that reproduced the age-specific rates in the absence of screening. It was assumed that screening in Canada began in 1986 therefore pre-1986 incidence rates were used to determine breast cancer incidence rates that would be representative of the no-screening scenario. The stage distribution of cancers detected in screening programs was obtained from the Canadian Breast Cancer Screening Database (CBCSD) for 1997-98.

Screening participation rates were specified in POHEM-BCS as being age-invariant, with fixed rate of 70% being applied as representing the probability that a eligible woman (who has to date not yet participated in screening) would volunteer for a first screen and a fixed rate of 90% being used to represent the probability of participating in subsequent screens (Public Health Agency of Canada, 2008). These participation rates were chosen based on Canadian target values. (Public Health Agency of Canada, 2008) Based on data from the 2008 Canadian Community Health Survey (CCHS), self-reported routine mammography within the last two years among females age 50 to 69 was a little under 50%.[12]

The sensitivity and specificity of organized screening mammography were estimated from the Canadian Breast Cancer Screening Database (CBCSD) maintained by the Public Health Agency of Canada (PHAC) for the screen year 2003-2004. PHAC receives updated program screening data from each participating province and territory. These screening estimates ought to be representative of all provinces and territories in Canada with the one exception being Nunavut which does not have an organized screening program (and thus does not contribute data to the CBCSD). In building the model, clinically detected cancers (when the cancer is caught outside of screening, usually symptomatically) were simulated based on the incidence rates and were assumed to have an average sojourn[13] time (Flanagan et al., 2006).

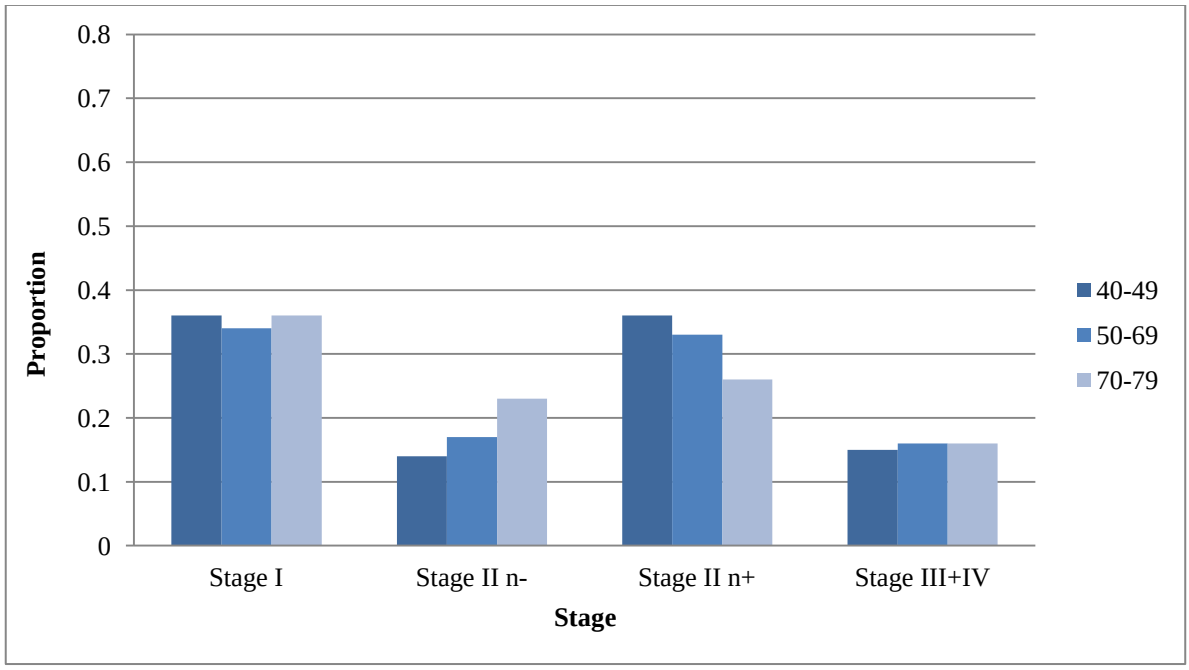
[12] Statistics Canada. *Table 105-0243 - Mammogram obtained within last 2 years, by age group, females aged 50 to 69 years, Canada, provinces, territories, health regions (June 2003 boundaries) and peer groups, every 2 years,* CANSIM (database). (accessed: 2014-03-17)

[13] Average sojourn time is the duration of a disease before clinical symptoms become apparent but during which it is detectable by a screening test (Duffy, 2005)

When mammography for pre-clinical cancers (screen-detected) was simulated, the sensitivity estimate was used to determine the proportion of screen-detected cases (true positives) and those that are missed (false negatives). When no cancer is present over the sojourn period, estimates of specificity were used to determine the proportion of cancers that were identified correctly (true negatives) and those that were incorrectly identified as cancer (false positives) (Flanagan et al., 2006). We followed the approach of Flanagan et al, 2006, using the updated (age-stratified) estimates of sensitivity and specificity derived from the CBCSD as described above.

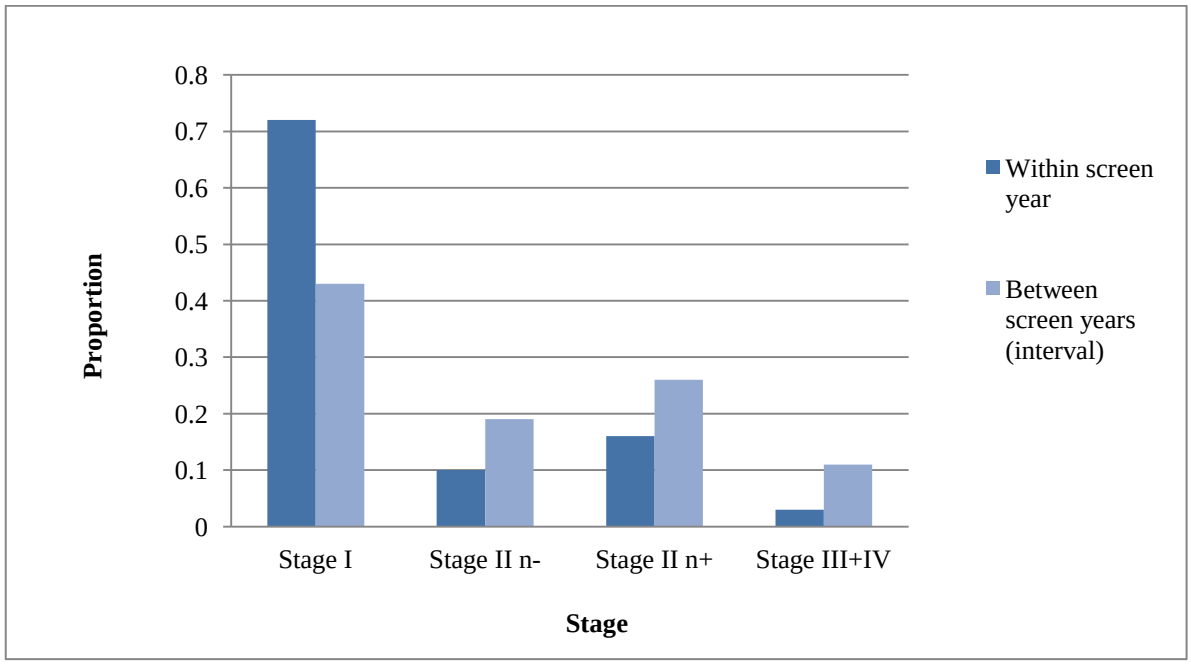
The stage distribution of clinically detected cancers by age group and related survival rates in the absence of screening used in POHEM-BCS were also estimated from the Saskatchewan Cancer Registry from 1982-1985 so that the effects of screening (which occurred mostly after 1986) were avoided. When a cancer is detected in the presence of screening, a stage is assigned according to how the cancer was detected (first/initial screen, at subsequent screen, or between screens/interval) from the estimated stage distribution. Survival benefit from early detection is conferred through the improved stage distribution (stage shift). Figures 10 and 11 show the pre-screening and post-screening stage distributions used in POHEM-BCS. The model does not capture the survival benefits within stage shift, such as moving from late to early stage II, for example.

Figure 10 Stage distribution of clinically detected (pre-screening) breast cancers by age group



Source: Flanagan et al. (2006). Evaluating the impact of breast cancer screening in Canada using POHEM – A demonstration project. (Unpublished work)
Note: IIA = node negative; IIB = node positive

Figure 11 Stage distribution of breast cancers among women age 40-79 participating in organized breast cancer screening programs



Source: Flanagan et al. (2006). Evaluating the impact of breast cancer screening in Canada using POHEM – A demonstration project. (Unpublished work)
Notes: IIA = node negative; IIB = node positive

Direct costs associated with screening, diagnosis, and treatment, were included in the costing component of the analysis. Cost of diagnosis and treatment were modeled in POHEM-BCS while cost of screening was applied external to the model in the study's analyses. Recognizing that the cost of screening varies regionally, a request was made in 2010 to all provincial/territorial representatives on the Canadian Breast Cancer Screening Initiative (CBCSI). Only British Columbia (BC) and Quebec provided published screening costs estimates for use in this study. For the purposes of the base case analysis, the cost per screen from BC for the fiscal year 2008/2009 of $70 inflated to 2013 CAD ($75) was used (Warren Burhenne et al., 2007). The BC estimate was used in this analysis as it is a published estimate with detailed breakdown of included costs, and considered to be a more representative proxy for a Canada wide average than the estimate from Quebec (higher at $125) (Lefrançois et al., 2007). The BC estimate is also comparable to the unpublished cost per screen estimated for Ontario in 2009.[14] This estimate encompasses several costs, including: central services, other operating costs, professional reading fees, and capital allocation. The BC capital allocation costs include capital differential allocated to privately administered centres in their annual operating budget and amortization of equipment purchased through the BC Cancer Agency, Provincial Health Services Authority. Capital allocation does not include capital expenditures capitalized and amortized through host hospitals. The cost per screen estimated for the province of Quebec was used in the sensitivity analysis to represent a high cost per screen. POHEM-BCS includes the initial diagnosis costs which include initial workup of breast cancer, staging, and treatment costs (neo-adjuvant chemotherapy, breast surgery, hospitalization, radiation therapy, hormone therapy, and follow-up costs). The costs associated with metastatic cancer include costs for care of terminal breast cancer including hospital stay, other medical, and palliative radiation therapy. Details of these costs are included in appendix 1. Total costs were calculated as the

[14] Presentation by Verna Mai at the International Meeting on Breast Cancer Screening Rio de Janeiro on April 16-17, 2009. "Economic Aspects of Breast Cancer Screening-The Ontairo Experience" Accessed on March 18, 2014 from http://bvsms.saude.gov.br/bvs/palestras/cancer/economic_aspects_breast_cancer_screening.pdf

sum of total screen and total treatment costs which were all adjusted to 2013 Canadian dollars based on the Bank of Canada's consumer price index.

First Base Case Analysis (Base Case 1)

POHEM-BCS is ideally suited to examine "what-if" scenarios including the potential impacts of different screening mammography policy options on health and economic outcomes for the Canadian population. This study aimed to assess the impact on life expectancy and costs (screening, diagnosis, and treatment) when two design attributes of screening mammography programs were changed: namely the age eligibility criterion and screening frequency which is also known as, screening interval, the duration of time between screens such as two years, usually prescribed to patients in an organized mammography program). Eleven screening mammography policy options including no screening were constructed in POHEM-BCS.

In table 7, the 11 policy options assessed are described by row, with three columns to identify eligible age groups. The screening interval is described within each cell, where 0, 2, and 1 represent no eligible screening for that age-group, eligible for screening once every two years, and eligible for screening once per year, respectively. The standard recommendation (status quo) across Canada is for women age 50 to 69 years participate in screening mammography once every two years, which is represented in screening policy A in this analysis.

Table 7 Characteristics of screening mammography policies – intervals in years by age group

Screening policy	Low Age Group (40-49 years)	Medium Age Group (50-69 years)	High Age Group (70-79 years)
No screen	0	0	0
A	0	2	0
B	1	2	0
C	0	2	2
D	1	2	2
E	0	1	0
F	1	1	0
G	0	1	1
H	1	1	1
I	0	2	2
J	2	2	2

Notes: Screening intervals are 1 year (annual) or 2 years (biennial), and represent the duration of time between screens. 0 denotes no screening in that age group.

For each screening policy option, we used POHEM-BCS to track the cohort of women from birth as described in the previous section detailing the characteristics of the model. The outcomes of interest for the analysis included the total number of: screens, life-years, and costs related to breast cancer diagnosis and treatment contributed by the cohort. To calculate screening costs, the total number of screens recorded by the model was used as a multiplier of the cost per screen. The screening costs were then added to the diagnosis and treatment costs recorded in the model to obtain total costs. These total costs were recorded for each screening policy along with an estimate of the health benefits measured in life years, which were also obtained from POHEM-BCS. The cost and benefits recorded for each policy were then used to form incremental cost-effectiveness ratios, for various pairs of programs (alternative versus comparator).

An incremental cost-effectiveness ratio (ICER) is a measure of excess cost to achieve an extra unit of health benefit conferred by the alternative program (Braithwaite et al., 2008). For each policy comparison, the difference in total cost was divided by the difference in the total number of life-years (see equation 1).

Equation 1

$$ICER = \frac{\Delta C}{\Delta E} = \frac{(C_2 - C_1)}{(E_2 - E_1)}$$

For the purposes of this study, C denotes total cost of screening, diagnosis, and treatment, and E denotes effectiveness in life-years. Subscripts 1 and 2 denote the screening options being compared, for example the cost-effectiveness of option 2 versus option 1.

Incremental cost-effectiveness ratios (ICER) were used to compare the screening options the principle of extended dominance was applied in a sequential analysis. The options were first ranked according to their effectiveness on the basis of securing maximum effect, neglecting cost. Sequentially, each option was compared to the previous in terms of incremental cost-effectiveness. Dominated options were excluded from consideration. An option is considered 'dominated' if there is an alternative screening option which achieves at least the same benefit but at a lower cost or which is at worst as costly but with greater benefit. We also applied principles of extended dominance whereby a screening policy is subject to extended

dominance when there is a more costly screening policy with a lower incremental cost-effectiveness ratio compared to the base. Non-dominated options were compared with each other in a sequential fashion starting from the lowest benefit option and sequentially identifying whether the next most beneficial option has an ICER that still meets the WTP criterion; stopping at the option that precedes the first breech of the WTP criterion. This option, which has the greatest benefit in terms of life-years gained (LYG) within the established willingness-to-pay (WTP) threshold of $50,000 per LYG, was identified as the optimal screening policy option.

This cost-effectiveness analysis takes the perspective or viewpoint of the Ministry of Health (from the standpoint of what the Ministry reimburses or pays). Both costs and health benefits were discounted at the annual rate of 5%, by convention (Canadian Agency for Drugs and Technologies in Health, 2006).

Sensitivity Analyses

Multiple sensitivity analysis was used to determine whether conclusions regarding cost-effective screening would change from the base case depending on costs per screen, discount rate, willingness to pay threshold, participation rate, mammography test sensitivity, mammography test specificity, and screening interval. The sensitivity of the results to plausible low and high estimates for cost per screen and discount rate (costs and health benefits) were explored. A low cost per screen at $48 based on an estimated 5% annual decline in costs based on BC's provincial trend was used (BC Cancer Agency, 2013). A high cost per screen estimate of $125 was used as it is the highest estimated cost per screen from another province (Quebec cost per screen in 2007 inflated to 2013 CAD) (Lefrançois et al., 2007). A low discount rate of 0% to reflect an analysis with no discount and high discount rate of 10% was used. These rates have been used in other health economic evaluation studies. (Mitton et al.m 1998; Willich et al., 2006; Hoeflma and Hanewinkel, 2007). We also assessed cost-effectiveness using a low and high WTP threshold of $30,000 and $100,000 per LYG, respectively. The impact of using 50% and 60% initial participation rates to reflect lower self-reported participation rates in Canada was examined. When assessing the impact of test sensitivity and specificity, a separate analysis was conducted, setting each parameter in the model to 100% to reflect the benefits of screening with perfect effectiveness in terms

of a picking up true cancers and not picking up inexistent cancers. Since sensitivity and specificity were already high in the base-case analysis (see appendix 1, table 4.3), an increase to 100% could be considered a minimal adjustment. The results of these sensitivity analyses were discussed in the context of the results of the base case 1 analysis.

Adjusted Base Case Analysis (Base Case 2)

To assess the impact of longer screening intervals (or less frequently screening), six additional screening policy options were constructed to include three-year screening intervals for different age groups. The base-case analysis was repeated to include these new screening options resulting in a comparison of 17 total screening options including no screening. Table 8 describes the characteristics of each of these additional screening options in terms of age eligibility and screening intervals.

Table 8 Characteristics of additional screening mammography policies – intervals in years by age group

Screening policy	Low Age Group (40-49 years)	Medium Age Group (50-69 years)	High Age Group (70-79 years)
K	3	3	3
L	0	3	0
M	3	0	0
N	3	2	2
O	2	3	3
P	1	2	3

Notes: Screening intervals are 1 year (annual), 2 years (biennial), or 3 years (triennial), and represent the duration of time between screens. 0 denotes no screening in that age group.

Results

Base Case Analysis 1

Table 9 describes the summary results of the base case analysis and demonstrates the variability in the average life-time costs per woman (screening, diagnosis, and treatment) and health benefits (life-days gained) across the screening policy options. Other results relating to detection (false positive cases, true positive cases, false negative cases, true negative cases) are included in appendix 2 (table 1).

When comparing each of the screening policies against no screening, the average life-years gained (LYG) per 1,000 woman (total population 6 million) from birth across screening options varied from 9.1 (option A) to 18.4 (option H). As screening becomes more intensive (more screening as a result of expanded age eligibility and/or shorter screening interval), the LYGs increases, however those screening policy options that screen women age 40-49 had larger LYGs compared to those policy options that screen women age 50-79, despite comparable number of lifetime screens.

Table 9 Summary of outcomes (cost and life-year benefits) projected for each of the 11 screening scenarios (base case 1)

Screening Option*	Age (Interval)	Average number of lifetime screens per 1,000 women	Average lifetime cost of screening per 1,000 women ($ CAD)	Average lifetime diagnosis and treatment cost per 1,000 women ($ CAD)	Average lifetime total cost per 1,000 women ($ CAD)	Average number of life-years gained per 1,000 women
No screen	-		-	-	-	-
A	50-69 (2)	8,300	51,116	-2,296	48,820	1.5
C	50-69 (2) 70-79 (2)	11,500	70,783	-2,555	68,228	1.7
I	40-49 (2) 50-69 (2)	12,700	77,813	-3,212	74,601	2.5
E	50-69 (1)	15,700	96,324	-2,391	93,932	1.6
J	40-49 (2) 50-69 (2) 70-79 (2)	15,900	97,255	-3,455	93,800	2.7
B	40-49 (1) 50-69 (2)	16,800	103,208	-3,429	99,779	2.8
D	40-49 (1) 50-69 (2) 70-79 (2)	20,000	122,868	-3,718	119,150	3.0
G	50-69 (1) 70-79 (1)	22,000	135,131	-1,938	133,193	1.8
F	40-49 (1) 50-69 (1)	24,200	148,445	-3,450	144,995	2.9
H	40-49 (1) 50-69 (1) 70-79 (1)	30,600	187,797	-2,976	184,821	3.1
Per cent difference between high and low†		73%	73%	48%	74%	48 %

* Screening options are ordered according to average number of lifetime screens per woman.
† Per cent difference is used to demonstrate the spread in values.
Notes: Costs and life-days gained discounted annually at 5%. Per 1,000 woman estimates are in relation to a total female population of 6 million.

The estimated average lifetime cost of screening per 1,000 women, compared to no screening, ranged from $51,116 (option A) to $187,797 (option H). Recall that screening costs were estimated outside of POHEM-BCS by applying an estimated cost per screen to the total number of screens for each policy option. The estimated reduction of lifetime cost of diagnosis and treatment per 1,000 women compared to no screening ranged from $1,938 (option G) to $3,718 (option D). The lower cost in diagnostic workup and treatment can be explained by stage shifting from a later, more costly stage of cancer to an earlier, less costly stage of cancer that is because of screening. These cost differences are evident in the input estimates for the POHEM-BCS parameters: cost of initial work-up, cost of initial

hospitalizations, cost of initial follow-up, cost of chemotherapy for breast surgery, cost of radiation therapy for breast surgery, and hormonal therapy (see appendix 1, tables 1.1-1.17). Further detail around the diagnostic and therapeutic approaches to breast cancer in Canada and their associated costs can be found in Will et al. (1999;2000). It is somewhat unclear if a relationship exists between stage of diagnosis and cost of treatment. In our analysis, the diagnosis and treatment cost reductions were low in comparison to the increases in screening costs across screening policy options. The resulting total costs (screening, diagnosis, and treatment) across screening policy options are therefore approximately proportional to the number of screens.[15] The estimated average increase in lifetime costs per 1,000 women ranged from $48,829 (option A) to $184,821 (option H).

The most effective policy was option H: annual screening for women age 40 to 79 (3.1 LYG per 1,000 women) but it was also the most costly ($184,821 per 1,000 women), compared to no screening. The least effective policy was option A: biennial screening for women age 50 to 69 (1.5 LYG per 1,000 women) but it was also the least costly option ($48,820 per 1,000 women), compared to no screening. A sequential analysis is essential in determining optimal policy options based on a comparison across several options using incremental cost-effectiveness ratios (ICER). In the context of this research, we use "optimal" to describe a policy option that is non-dominated and which incurs the most life-years gained within the willingness-to-pay threshold (WTP). The results of the first base case sequential analysis are presented in table 10.

Screening policy options A, C, J, D, and H lie on the cost-effective efficiency frontier (non-dominated). Of these policy options, the optimal screening policy was J: biennial screening for women ages 40 to 70 years because it was the most effective, non-dominated policy (16,421 LYGs compared to no screening) that met the ICER constraint of meeting the $50,000 per LYG WTP criterion. If WTP was $30,000 per LYG, the optimal screening policy option would be A: biennial screening for women ages 50 to 69 (status quo). If WTP was $20,000 or less per LYG no screening policy option would have been optimal under the base case analysis.

[15] In a plot of lifetime total costs per 1,000 women versus total screens per 1,000 women, R^2=0.998; y=1.0182x +2579.2.

Table 10 Results of the sequential analysis of screening mammography policy options (base case 1)

Screening option*	Age group (screen interval in years)	Total LYGs	Total LYGs per 1,000 women†	Total excess (vs no screen) health care cost ($ million CAD)	Total excess health care cost ($CAD) per 1,000 woman†	ICER vs no screening (cost in $ CAD per LYG)	Sequential ICER (cost in $ CAD per LYG)
A	50-69 (2)	9,104	1.5	199	33,170	21,860	21,859
C	50-69 (2) 70-79 (2)	10,287	1.7	235.5	39,240	22,889	30,854
J	40-49 (2) 50-69 (2) 70-79 (2)	16,421	2.7	441.5	73,580	26,884	33,583
D	40-49 (1) 50-69 (2) 70-79 (2)	17,824	3.0	631	105,170	35,402	135,068
H	40-49 (1) 50-69 (1) 70-79 (1)	18,370	3.1	845	140,830	45,998	391,941
I	40-49 (2) 50-69 (2)	15,248	2.5	405.5	67,580	26,592	Extended dominance through J and A
B	40-49 (1) 50-69 (2)	16,630	2.8	594.6	99,100	35,755	Extended dominance through D and A
E	50-69 (1)	9,698	1.6	373.7	62,280	38,530	Dominated by C
G	50-69 (1) 70-79 (1)	10,810	1.8	446.1	74,350	41,266	Dominated by I and J
F	40-49 (1) 50-69 (1)	17,218	2.9	770.5	128,420	44,751	Dominated by D

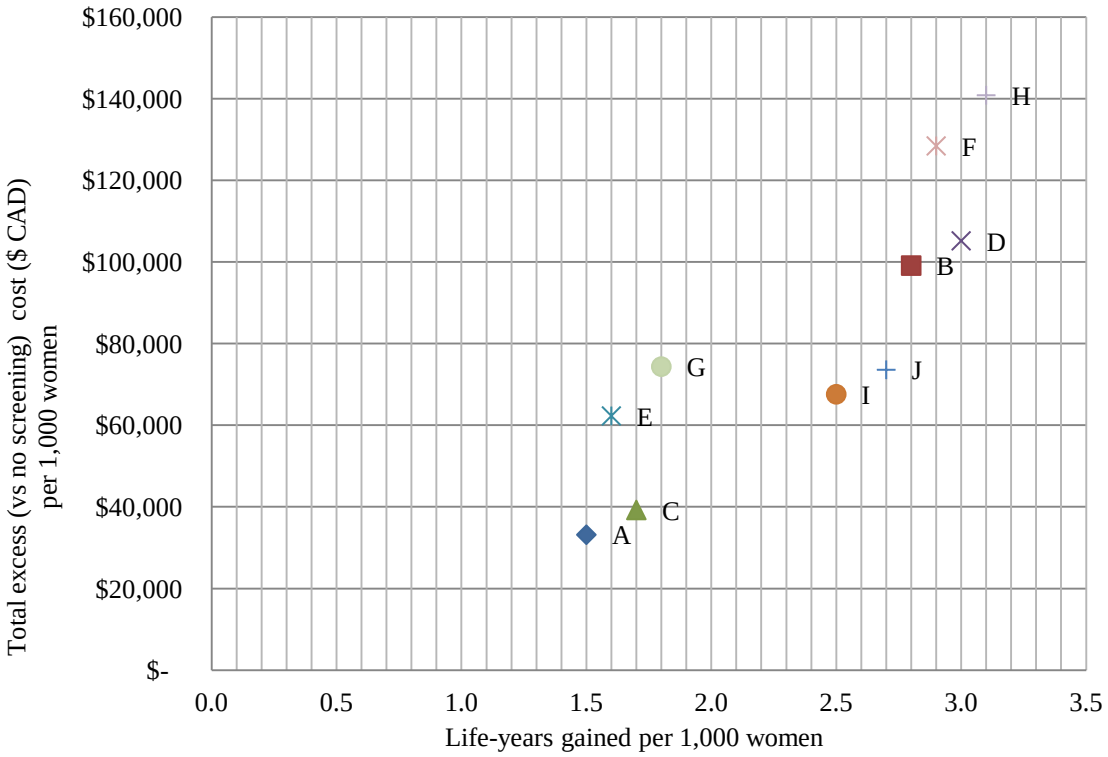
* Screening option ordered according to total excess cost compared to no screening

† Total population of 6 million women followed in the model.

Δ = difference (gain); LYG = life years gained; ICER= incremental cost-effectiveness ratio

Notes: Discount rate of 5% applied to health benefits and costs. An optimal screening policy option incurs the greatest number of LYGs with an ICER < $50,000 per LYG (in this case, option J).

Figure 12 Cost-Effectiveness Plane – Incremental cost versus incremental life-years (base case analysis



Sensitivity Analyses

Table 11identifes those screening policy options deemed optimal in the base case analysis (base case 1), the sensitivity analysis, and the adjusted base case analysis (base case 2).

At a lower WTP of $30,000 per LYG, in the base case analysis, option A: biennial screening for women ages 50 to 69 was the optimal screening policy option, while option J: biennial screening for women ages 40 to 79 were optimal policies at a low cost per screen and low discount rate, respectively. No screening policy option was optimal when a high cost per screen or high discount rate was used.

At a higher WTP of $100,000, option J: biennial screening for women ages 40 to 79 was optimal in the base case analysis, and when high cost per screen was used and high discount rate was used, whereas option D: annual screening for women age 40 to 49 and biennial screening for women age 50 to 79 was optimal when a low cost per screen and low discount rate was used.

These results are intuitive because as the cost per screen, discount rate, and WTP become more conservative the optimal screening policy option, if any, is characterized by less frequent screening or no screening compared with higher cost per screen, higher discount rate, and higher WTP. Optimal screening is therefore dependent upon these parameters.

Compared to the base case analysis, cost-effectiveness of screening was sensitive to low and high per screen cost ($48 and $125 CAD, respectively) and to low and high discount rate (0% and 10%, respectively). The detailed results of the sequential analyses are included in appendix 2, tables 2 through 10.

At a WTP of $50,000 per LYG, option J: biennial screening for women ages 40 to 79 years was the optimal screening policy option when using a low cost per screen. When a high cost per screen or high discount rate (per screen cost at nominal value of $75) was used, the optimal screening policy option was A: biennial screening for women ages 50 to 69. For a low discount rate the optimal screening policy option was D: annual screening for women age 40 to 49 and biennial screening for women age 50 to 79.

To reflect current screening practices we observed the impact of lower participation rates in the population using initial participation rates of 50% and 60%. When a 50% or 60% participation rate was used, option J: biennial screening for women ages 40 to 79 was the optimal screening policy option

Table 11 Summary results – optimal screening intervals in years by WTP and age group

Scenario	Optimal Screening Policy (Screening interval in years by WTP and age group)								
	WTP: $30,000/LYG			WTP: $50,000/LYG			WTP: $100,000/LYG		
	Age: 40-49 years	Age: 50-69 years	Age: 70-79 years	Age: 40-49 years	Age: 50-69 years	Age: 70-79 years	Age: 40-49 years	Age: 50-69 years	Age: 70-79 years
Base Case 1*	-	2	-	2	2	2	2	2	2
Cost per screen = $48	2	2	2	2	2	2	1	2	2
Cost per screen = $125	-	-	-	-	2	-	2	2	2
Discount rate = 0%	2	2	2	1	2	2	1	2	2
Discount rate = 10%	-	-	-	-	2	-	2	2	2
Participation rate = 50%	-	2	2	2	2	2	1	1	-
Participation rate = 60%	-	2	2	2	2	2	2	2	2
Test sensitivity = 100%	-	2	2	2	2	2	1	2	2
Test specificity = 100%	-	2	2	2	2	2	2	2	2
Base Case 2†	-	2	-	2	2	2	2	2	2

* Base case 1 denotes the original base case analysis which compared screening policy options that included 1 and 2-year screening intervals.

†Base case 2denotes the original base case analysis replicated with additional screening policy options which include 3-year screening intervals for specific age groups.

No optimal screening	*Low-volume screening*	*Medium-volume screening*	*High-volume screening*

Base Case Analysis 2

The results of the replicated base case analysis that included 6 new screening policy options with three-year screening intervals for some or all of the ages found no change to the original base case analysis in terms of optimal screening (see table 4 and chart 4). Optimal screening was still option J: biennial screening for women ages 40 to 79 years. To get a sense of the impact on the effectiveness and cost impact of longer screening intervals, we compared selected screening policy options, isolating specific age groups (table 5). When we compared the impact of 2- and 3-year screening against annual screening for women ages 40 to 79, we observed a reduction in life-years gained by a factor of 0.9 and 0.6, respectively. Costs reduced by a factor of 0.5 when going from annual to 2-year and to 3-year screening. When age was isolated, it was found that going from 2- to 3-year screening had a different impact by age group particularly in the costs. Increasing screening intervals from 2 to 3 years for women ages 40-49, 50-69, and

70-79, appeared to reduce life years gained by a factor of 0.7, 08, and 0.9, respectively. Costs decreased by a factor of 0.8 for women ages 40-49, but increased by a factor of 1.4 and 1.2 for women ages 50-69 and 70-79, respectively.

Table 12 Cost-Effectiveness Plane – Incremental cost versus incremental life-years (base case analysis 2)

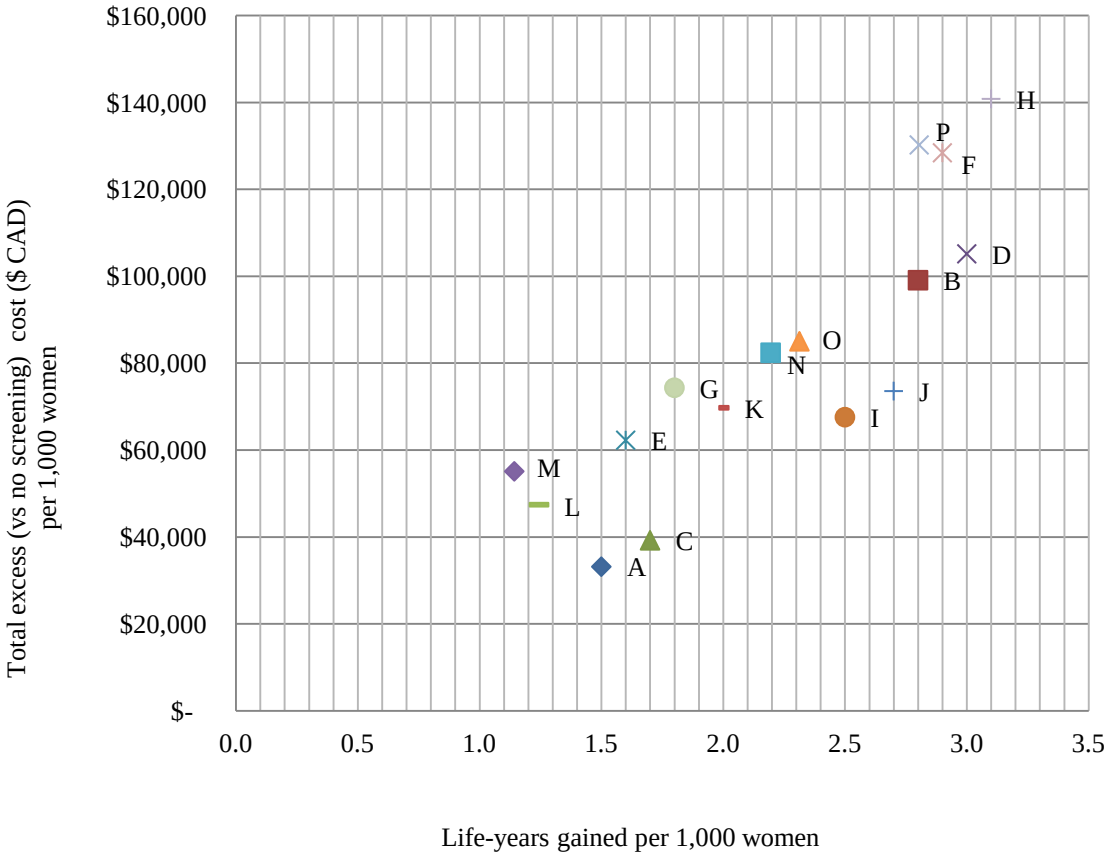


Table 13 Examination of the impact of screening frequency on life-years gained and costs

Description	Screening policy option	Age group in years (screening interval in years)	Life-years gained ratio	Total excess cost ratio
Increasing screening intervals (decreasing screening frequency) from 1 to 2 to 3 years in women ages 40 to 79 years on life-years gained and costs.	H*	40-49 (1) 50-69 (1) 70-79 (1)	-	-
	J	40-49 (2) 50-69 (2) 70-79 (2)	0.9	0.5
	K	40-49 (3) 50-69 (3) 70-79 (3)	0.6	0.5
Increasing screening intervals (decreasing screening frequency) from 2 to 3 years in women age 40 to 49 years on life-years gained and costs.	J*	40-49 (2) 50-69 (2) 70-79 (2)	-	-
	N	40-49 (3) 50-69 (2) 70-79 (2)	0.7	0.8
Increasing screening intervals (decreasing screening frequency) from 2 to 3 years in women ages 50 to 69 years on life-years gained and costs.	A*	50-69 (2)	-	-
	L	50-69 (3)	0.8	1.4
Increasing screening intervals (decreasing screening frequency) from 2 to 3 years in women ages 70 to 79 years on life-years gained and costs.	D*	40-49 (1) 50-69 (2) 70-79 (2)	-	-
	P	40-49 (1) 50-69 (2) 70-79 (3)	0.9	1.2

* Reference

Discussion

Overall, this study found that for the Canadian context, population-based annual and biennial screening mammography for women age 40 to 79 years may or may not be deemed cost-effective, depending on input choices. In the base case analyses (1 and 2), biennial screening for women ages 40 to 79 was determined to be optimal (most cost-effective and non-dominated). Annual screening for women age 40 to 49 years was only optimal when the screening program also included biennial screening for women age 50 to 79 years and if no discount rate was used or when cost per screen was low or WTP was high. Annual screening for this age group was also deemed optimal when test sensitivity was high (100%) and WTP was high. The only scenario that showed annual screening for women age 50-69 as part of an optimal screening policy was when annual screening was also offered to women ages 40-49, when participation rate was 50%, and when the WTP was $100,000/LYG. At a low WTP ($30,000/LYG), optimal screening would never include women ages 40-49 under any of the conditions we tested. In both base case analyses it was found that no screening policy option would be optimal if the WTP was equal to or lower than $21,000 per LYG.

Depending on the province or territory, current screening practices in organized programs accept women ages between 40 and 79 years for annual and/or biennial screening, but current breast cancer screening guidelines recommend screening for average-risk women between the ages of 50 to 65 years, every 2 to 3 years. Base case analysis 1 examined only 1 and/or 2-year screening intervals for different age ranges. In the second base case analysis several screening options that included 3-year screening intervals were included. Although results of base care analysis 2 found no change in the selection of optimal screening compared to the original base case, some interesting relationships emerged when the impact of decreasing screening frequency on LYGs and costs by age group were isolated. As expected, when screening frequency decreased (going from 1 to 2 to 3 year intervals), LYGs decreased for all women ages 40 to 79. This is a result of screening intervals that are longer than the estimated sojourn (the duration of time a cancer is detectable before it becomes symptomatic). In POHEM, sojourn times were set at 1 year for women ages 40-49 and 2 years for women ages 50-79. The impact of screening frequency on total costs is not as intuitive. Increasing screening intervals, thereby decreasing screening frequency, will

decrease screening costs but it may result in higher treatment costs because more cancers will be detected at a later, more costly stage. The reason for possibly higher treatment costs can also be explained by screening intervals that are longer than sojourn times. The results of this analysis showed that that going from 2 to 3 year screening intervals for women ages 50-69 years and women ages 70 to 79 years not only reduced the number life-years gained, but it also increased costs. Going from 2 to 3 year screening intervals still reduced total costs in women ages 40 to 49, which might mean screening in this age group screen costs are not as greatly offset by savings in treatment costs as they are for the older age groups.

Statistics Canada's POHEM-BCS microsimulation model was chosen for this analysis because it is accessible, uses Canadian data, and has a track record of application in Canada. Although POHEM-BCS is grounded in theory from many disciplines, including epidemiology, statistics, demography, medicine, and health economics, it is not based on biological theories, such as those models used by the Cancer Intervention and Surveillance Modeling Network (CISNET) out of the US. These models are based on the biological theory of tumour growth (Cancer Intervention and Surveillance Modeling Network (CISNET) Breast Cancer Collaborators, 2006) whereas POHEM-BCS models impact on mortality via health-state distributions.

One of the limitations in this study is the limited data on the validity (accuracy) and reliability (generalizability) of the POHEM-BCS. The relationships between all model variables as well as the parameter estimates themselves come from a variety of sources that may or may not reflect true life. These data come from theory, empirically collected data, surveillance data, and registry data (Flanagan et al. 2006). Previous analysis has shown that breast cancer incidence rates generated by the model were able to reproduce actual incidence for the 40 to 79 year age group between years 1986 and 2011 (Flanagan et al., 2006). As the model has been used for nearly two decades, been developed by modeling experts at Statistics Canada, and used in many other published studies, it is reasonable to assume that the model is valid and reliable for the purposes of this study (Houle et al., 1997; Will et al., 1999; Berthelot et al., 2000; Flanagan et al., 2003; Flanagan et al, 2006; Kopec et al., 2010;Nadeau et al., 2013).

Indirect costs were excluded in the analysis, including the costs of home care, out-of-pocket oral medications, other out-of-pocket expenses resulting from visits to healthcare institutions, production losses and additional costs due to other diseases in the life-years gained. Excluding such costs is consistent with the provincial/territorial payer (department of health) perspective chosen. Adding-in the indirect costs would have enabled a societal perspective to be pursued. It has been estimated that the inclusion of indirect costs could more than double the cost-effectiveness ratio of screening (Koning, 1991).

There was intent to quantify effectiveness in terms of quality-adjusted life-years (QALY) in this analysis but this was not possible due to the output required for such a calculation was more detailed than that provided by POHEM-BCS. It was not possible to link health utilities (health-related quality of life measures) to the LYGs generated by the model. Incorporating health-related quality of life could have accounted for anxiety, pain, and discomfort related to undergoing the mammography exam and definitive diagnoses, the wait for and receipt of results, false positive and false negative tests, and treatment into the measure of effectiveness. One of the limitations of a cost-effectiveness analysis compared to cost-utility analyses is that LYGs do not to incorporate the quality of life benefits from early detection. Although health-related quality of life adjustments to LYGs would reduce the overall health benefits across all screening policy options, including no screening, the application of health utilities to LYGs in this study could potentially alter the relative cost-effectiveness of the screening policy options and result in a different optimal screening policy options chosen. There is some evidence to suggest a quality of life gradient across breast cancer TMN stages, age groups, and socioeconomic and cultural groups whereby lower quality of life is experienced in later stages of cancer, older age groups, and more socially vulnerable or disadvantaged populations (Chopra and Kamal, 2012).

The impact (if at all) of having one or more false positive screening tests on quality of life was not assessed in the current study. In the study by Bonomi et al. (2008), female subjects was asked to rate on a scale utility/preference scale of 0 (least preferred) to 100 (most preferred) routine, population-based screening mammography options. Those screening options that would have resulted in having a high negative predictive value but were likely to result in a certain number of false positive tests were on average rated at 80 or above suggesting the subjects perceived the risk of having a false positive test result was not

deleterious (Bonomi et al, 2009). On the other hand, another study found that women who had experienced a false positive test result were more likely to experience concerns regarding their breast health compared to women who had a normal (true negative) test result. In addition, these women were also more likely to use health care services within one year of the mammogram, potentially as a result of this heightened concern (Barton et al, 2001). The subjects in the Bonomi study found the following experiences much lower on the utility scale compared to experiencing a false positive test: cancer treatments, palliation/end-of-life care, recurrence of breast cancer, receiving news of a new breast cancer diagnosis, and receiving delayed news of breast cancer diagnosis (ratings of 33 to 49 on the scale of 0 to 100). These experiences are not modeled in POHEM-BCS which would further limit the ability to accurately capture the full impact of screening and interrelated services and treatments on quality of life.

Discounting was applied to both health benefits and costs from birth. Since benefits are not expected to materialize until at least the age of screening (earliest age of benefit is 40 in this study), it is likely that the benefits of screening in terms of extended life-years are underestimated. A more meaningful measure may be the amount of extended life from a later age such as from age 65 as reported in national statistics.[16] The approach we used for discounting is included in appendix 3.

Several international studies have assessed the cost-effectiveness of screening mammography for breast cancer over the past decade (Arveux et al., 2003; Mandelblatt et al., 2005; Neeser et al., 2007; Okonkwo et al., 2008; Wong et al., 2008; De Gelder et al., 2009; Wong et al., 2007; Lee et al., 2010; Stout et al., 2006; Tosteson et al., 2008). These studies vary in the screening options examined and the methodology used to assess cost-effectiveness, namely the analytical and modeling approach, the screening options being compared, and the setting. These studies have been systematically reviewed in the chapter 2 of this dissertation and are briefly discussed here to place the results of the current study in context of the existing literature. De Gelder et al. (2009) performed a cost-effectiveness analysis and cost-utility analysis of biennial screening for women age 50-69 years over 20-years in Switzerland. They reported a cost-effectiveness ratio of screening versus no

[16] An example of such life expectancy statistics is available from Statistics Canada, CANSIM, table 102-0512.

screening of about $17,000 (CAD 2013) per life-year gained which is higher than our estimate of almost $10,000 (CAD 2013) per LYG without discounting and lower than our $20,000 per LYG with 5% discounting. The authors did not state the perspective adopted which may limit comparability.

Arveux et al. (2003) reported a cost-effectiveness ratio of $28,000 per LYG (CAD 2013), with costs discounted at 5%, when comparing the cost-effectiveness of screening versus no screening in the female population age 50 to 65 years over a 20-year period in France. They used a Markov-based decision model that incorporated regional screening program data, morbidity and mortality data, demographic data, and direct costs pertaining to screening, diagnosis, initial treatment, and breast cancer surveillance. The cost-effectiveness analysis did not specify the screening interval or frequency, nor did they specify the analytic perspective which compromises the comparability of their findings to those of this current study. However, there were several screening policy options in our study that were close to $28,000 per LYG with discounting at 5%. The current study found that biennial screening for ages 40 to 79 (option J) and biennial screening for ages 40 to 69 had cost-effectiveness ratios of almost $27,000 per LYG, compared to no screening.

As survival improves in late-stage breast cancer due to improvements in treatment effectiveness and accessibility increases (minimizes screening effectiveness) and as health care budget constraints tighten (lowers WTP), screening mammography eligibility might need to become more narrow in order to stay within the WTP. Although a WTP of $50,000 per LYG is most commonly used in the health economic evaluations literature (Winquist et al., 2012) , it can be argued that a lower threshold of $30,000 per LYG may be more appropriate (Grima et al., 2014; Burger et al., 2014). Given concerns in relation to reining in health care budgets, a lower threshold may be more appropriate, depending on the payer's WTP or what they believe is the appropriate cost per unit benefit based on multiple factors including budget constraints, uncertainty regarding risks and benefits, and competing priorities. Further, the National Institute for Health and Clinical Excellence (NICE) suggests that the WTP threshold for new health technologies is approximately $30,000/QALY from the societal viewpoint (National Institute for Health and Clinical Excellence, 2007). In addition to using a WTP threshold of $30,000 per LYG and the standard WTP threshold of

$50,000 per LYG, we also selected a high WTP threshold of $100,000 which has been used in other studies (Pataky et al., 2013; Sen et al., 2014; Burger et al., 2014).

The way in which effectiveness of screening is captured in POHEM is based on observed rates of breast cancer incidence in a time when screening was implemented in the population versus when it was not. The shortcomings of this approach are that it is based on older surveillance data and at a time when screening was less effective. Inaccurately modeling stage distribution (a new cancer's probability of being diagnosed in a specific stage) can have an important impact on predicting screening effectiveness in terms of breast cancer mortality. The impact of screening on breast cancer mortality is modeled in POHEM as a function of breast cancer incidence, stage distribution, and stage-specific survival. For the screening policy option characterized by annual screening for women ages 40-49 years and biennial screening for women ages 50-79, POHEM predicted that breast cancer screening would reduce breast cancer mortality in all women by 10% over a lifetime. This appears to be an underestimate or overestimate of the impact of screening compared with other studies. For example, the National Cancer Institute references the studies by Nelson (2009), Moss et al. (2006) and Miller et al. (2014) when summarizing the benefits of mammography in terms of impact on breast cancer mortality. Based on a review of randomized controlled trials, Nelson et al. (2009) estimated that mammography was associated with a 15-20% relative reduction in breast cancer mortality among women age 40 -74 years. The study by Moss et al. (2006) reported an absolute mortality benefit for women screened annually over 10 years of about 1%. In the Canadian study by Miller et al. (2014), it was reported that annual breast cancer screening had no impact on breast cancer mortality among women ages 40-59 based on 25-year trial data. Other studies have reported much larger benefits. In a review study by Broeders et al. (2012), it was estimated a reduction in breast cancer mortality attributable to screening of 25-31% for all women. This estimate was based on a review of observational studies including prospective and case-controlled studies from Europe. There is an obvious discrepancy in the literature on the measured impact of screening on breast cancer mortality which may be explained by differences in study design and screening contexts, such as differences in population demographics, variability in the way in which screening is administered, and the quality of reporting for cause of death. It is therefore difficult to

decipher whether the estimated effectiveness of screening in terms of reducing breast cancer mortality in POHEM would be an overestimate or underestimate of the true impact.

This study contributes to the current body of knowledge on the impact of age and screening frequency on the cost-effectiveness of screening mammography in Canada. To date there have been no other published Canadian studies examining the cost-effectiveness of population-wide screening mammography. Future research could examine the cost-effectiveness of more frequent screening for higher risk sub-populations, including women with family history of breast cancer. The analyses in this study could also be replicated with updates to the several of POHEM's input parameters including stage distribution, stage-specific survival, and diagnosis and treatment costs. Future studies could also take a broader analytic perspective by including indirect costs or take into account quality of life in a cost-utility analysis.

Chapter 5 (Manuscript 3): Cost-effectiveness of screening mammography for Canadian women with elevated risk of breast cancer.

Authors:

Ngoc-Thy Dinh[1], Kevin Brand[2], Doug Coyle[3], Howard Morrison[4], Catherine Deri-Armstrong[5]

Affiliations:

1. Health Economics, Public Policy Division, Conference Board of Canada

2. Telfer School of Management, University of Ottawa

3. Department of Epidemiology and Community Medicine, University of Ottawa

4. Department of Epidemiology and Community Medicine, University of Ottawa

5. Department of Economics, University of Ottawa

Abstract and Key Words

Title: Cost-effectiveness of screening mammography for Canadian women with elevated risk of breast cancer.

Objectives: This study aimed to determine the impact of age eligibility and screening frequency on the cost-effectiveness of screening mammography in women at higher risk for breast cancer compared to the general female population.

Methods: A cohort microsimulation model was used to estimate the impact of disease risk and screening sensitivity on the cost-effectiveness of screening mammography in Canada. The analysis involved changes in baseline breast cancer incidence that approximated the increased risk of breast cancer associated with family history of the disease. Family history was described in terms of having an affected second degree relative (SDR), an affected first degree relative (FDR), two or more FDR and/or SDR, or having extremely dense breasts density (BIRADS 4). The impact of screening for high-risk populations on total direct costs (screening, diagnosis, and treatment) and life-years gained (improvement in survival) were estimated for ten different screening options characterized by varying screening age eligibility (40 to 79 years) and screening frequency (1-and/or 2-year intervals between screens). Costs and life-years gained were discounted at 5% annually. Costs were adjusted to 2013 Canadian dollars. Analyses were conducted from the perspective of the third party payer and a lifetime time horizon. Incremental costs and life-years gained for each screening scenario were compared against no screening and each other. Sequential analysis allowed identification of the optimal screening strategy based on a willingness-to-pay threshold of $50,000.

Results: In the previous chapter's base case analyses, the optimal screening policy for the general female population was biennial screening for women ages 40 to 79 years. In comparison to the base case, screening women with elevated risk of breast cancer, specifically women with family history or extremely dense breasts, was less costly but more effective in terms of extended life-years. As breast cancer risks increase, LYGs increase and total costs decrease. Screening women age 40 to 49 years annually and women age 50 to 79 years biennially was found to be optimal for the highest-risk group examined in this

analysis: having a family history of two or more SDR or FDR. Similar to the base case, biennial screening for women ages 40 to 79 was found to be optimal for the high-risk populations: family history SDR, family history FDR, and extreme breast density. The results of our research may suggest that risk-based screening may be more efficient and effective than uniform recommendations for the general risk population.

Conclusions: Screening recommendations for women with elevated risk of breast cancer may include more frequent screening, particularly among the 40 to 49 age group. Future research considerations include the impact of more sensitive screening technology such as the use of digital mammography and magnetic resonance imaging, treatment innovations, and other types of preventive measures such as lifestyle modification, drug therapy, and preventive mastectomy.

MeSH: screening; mammography; model; breast; cancer; breast neoplasms; cost-effectiveness

Introduction

Screening mammography has been widely used as a population-based public health measure in the secondary prevention of breast cancer in many developed countries, including Canada, the US, the UK, Australia, New Zealand, Belgium, the Netherlands, Finland, Norway, Japan, and the Republic of Korea (Saika and Sobue, 2011). Guidelines for screening mammography have been focused on the appropriate age to start screening and the appropriate screening frequency for the average-risk population. Less focus is paid to the question of whether customized screening programs might be appropriate for high-risk groups in Canada. For example should high-risk groups such as those with family history, be subject to a different screening program. There have been numerous international studies published on the impact of screen age and frequency on the cost-effectiveness of breast cancer screening, as well as studies looking at the effectiveness of screening in high-risk women (Arveux et al., 2003; Mandelblatt et al., 2005; Neeser et al., 2007; Okonkwo et al., 2008; Wong et al., 2008, Lee et al., 2010; Rojnik et al., 2008; Stout et al., 2006; Tosteson et al., 2008; De Gelder et al., 2009; Wong et al., 2007). In the context of assessing the impact of screening for high-risk sub-groups of the female population, it is unclear whether more vigorous screening for these sub-groups is cost-effective. Screening is more effective in terms of mortality reduction and efficient in terms of program costs when the burden of disease is adequately high (Miller, 1996). Preferred screening may include tailoring screening programs for high-risk sub-groups; perhaps pursuing more vigorous (earlier and more frequent) screening programs for these groups.

In a previous study by the authors, a cost-effectiveness analysis was conducted to examine the impact of age eligibility and screening frequency on the cost-effectiveness of screening mammography in the general female Canadian population through the examination of 11 screening policy options including no screening. These policy options varied in screening start and end age (40 to 79) and screening frequency or interval (1, 2, or 3 years). The results of the base case analyses identified biennial screening for women ages 40 to 79, as the optimal screening policy option with WTP= $50,000, a 5% discount rate, and $75 CAD cost per screen. If WTP was $30,000 the optimal screening policy option would have been biennial screening for women ages 50 to 69 (the prevailing base recommendation or status

quo). If WTP was $20,000 then none of the screening policy options would have been considered cost-effective. Cost-effectiveness of screening was sensitive to the screen costs, discount rate, and WTP.

The Canadian Cancer Society recommends that women with higher risk for breast cancer, characterized by family history (an affected first or second degree relative), certain gene mutations such as BRCA1 or BRCA2, or previous breast biopsies showing changes in the breast tissue, should be screened at earlier age and/or more frequently (The Canadian Cancer Society, 2014). It is uncertain however how much of an increase in risk is required to result in a meaningful change to the results of our base case analysis which aimed to identify optimal screening in the general female population age 40 to 79 (Dinh et al., 2015).

The degree of risk associated with several factors for breast cancer including family history of disease and extreme breast density were examined in the review article by Nelson et al. (2012). The family history and breast density risk factors were cited as being related to increased breast cancer risk that range from 70% to 200% increased risk of disease in women with the risk factor compared to women without. Personal factors such as race or ethnicity, body mass index, physical activity, alcohol consumption, smoking use and status were not identified as being significantly associated with breast cancer risk in those studies examined in the review.

The objective of this study was to determine the optimal screening mammography policy (characterized by age eligibility and screening frequency) when restricting attention exclusively to high-risk groups (targeted screening). It is possible, for example, that more rigorous screening programs may be more appropriate for high-risk groups. This study examines this question, using cost-effectiveness analysis as means for examining whether different screening policies might be prescribed for these high-risk groups. It was hypothesized that more vigorous screening amounting to an earlier start age and later end-age and/or more frequent screening will be relatively more cost-effective in high-risk populations compared to the general population. This study examines the impact of screening that is targeted to women with high-risk of breast cancer based on their family history of disease and breast density. These two risk factors were singled-out because of

their comparatively high association with disease onset and progression as indentified in the literature, and because they have and continue to be used in clinical practice as markers for elevated risk.

Methods

The Population Health Model - Breast Cancer Screening Module (POHEM-BCS)

The Population Health Model-Breast Cancer Screening Module (POHEM-BCS) was used to identify the optimal screening mammography policy option for women based on their risk profile. The screening policy options under study are varied by screening start and end age and frequency. POHEM breast cancer module is a microsimulation model built and maintained by Statistics Canada which uses Canadian data from a variety of sources on breast cancer incidence, mortality rates, population demographics, and interventions to predict survival and direct health care costs. Details of POHEM-BCS's structure, data sources, parameter values, assumptions, and validity have been described elsewhere and in the authors' previous study (Wolfson, 1994; Will et al., 1998; Will et al., 2000; Will et al., 2001; Kopec, et al., 2010; Dinh et al., 2015).

Screening Policy Options

The focus of this study's analysis was to quantify the impact of age eligibility and frequency of screening mammography on direct health care costs and life-years for sub-populations at elevated risk for breast cancer in order to identify optimal screening strategies for these sub-populations (Dinh et al., 2015). Given the low prevalence of the risk factors in the general population, the results for the total population will approximate the results for the population without elevated risk – although the incremental cost-effectiveness ratios will be likely slightly higher for this population but insufficient to change the study's interpretation. The 10 screening option scenarios examined in this study considered variations of age eligibility and screening frequency as summarized in Table 13. Taken together with a no screening policy option, this amounts to a total of 11 screening policy options for analysis. The screening policy options were constructed in POHEM-BCS by varying the values for the following input parameters: screen inclusion criteria, screen frequency, and participation rates. Hypothetical life histories were generated through model simulation for five synthetic

birth cohorts of women, each representing Canadian women with high-risk for breast cancer, specifically having strong family history of disease or extreme breast density. There was an attempt to characterize each cohort through an adjustment in baseline breast cancer incidence according to age-adjusted relative risk estimates and estimated female population prevalence of the underlying risk factor identified in the scientific literature. Apart from this adjustment, the model demanded a change in the mammography sensitivity[17] in women with extreme breast density. It has been estimated that mammography sensitivity in the female population with extreme breast density level of BIRADS[18] category 4 is approximately 44% compared to a sensitivity rate of about 81% in the average-risk female population (Kolb et la., 2002). It was assumed that all other model parameters were the same for higher-risk women compared to the general female population within the same age range. For each simulation the health profiles including breast cancer cases (detected and undetected via mammography), mortality from breast cancer and other causes, person years lived, and other outcomes including costs of diagnosis and treatment, were generated and tracked until the death of each individual. We ran a total of 55 simulations (11 scenarios for 4 high-risk populations) with a 6 million population for each cohort.

[17] According to the Dictionary of Public Health by Last (2007), sensitivity is the property of testing positive for the highest possible proportion of individuals who have, or might be suspected of having, the condition for with the test is performed. A sensitive test should have a zero or very low proportion of false-negative results (a negative result occurring in a person who actually has the disease or possesses the attribute for which the test is done).

[18] Breast Imaging-Reporting and Data System (BIRADS) is a standardized approach to describing mammogram findings. BIRADS level 4 represents a suspicious abnormality in the breast that warrants further investigation and potentially biopsy for diagnosis. This category is also often used to describe very high breast density.

Table 14 Characteristics of the 11 screening mammography policies, intervals by age group.

Screening policy	Low Age Group (40-49 years)	Medium Age Group (50-69 years)	High Age Group (70-79 years)
No screen	0	0	0
A	0	2	0
B	1	2	0
C	0	2	2
D	1	2	2
E	0	1	0
F	1	1	0
G	0	1	1
H	1	1	1
I	0	2	2
J	2	2	2

Notes: The age eligibility is divided into three categories across the columns, and the screening programs are labelled by row. A zero appearing within an age category denotes no screening. "2" denotes the duration of 2 years between screens or biennial screening. "1" denotes the duration of 1 year between screens or annual screening.

Risk Factors and Population Prevalence

Having a family history of breast cancer is often described as having one or more first- and/or second-degree relative with current or past diagnosis of breast cancer. A first-degree relative (FDR) is defined as the parents, siblings, or children of an individual (National Cancer Institute, 2014). A second-degree relative (SDR) is defined as the aunts, uncles, grandparents, grandchildren, nieces, nephews, or half-siblings of an individual (National Cancer Institute, 2014). Pharaoh et al. (1997) systematically reviewed the literature published between 1966 and 1996 examining the risk associated with a family history of breast cancer and found that the degree of risk was related to whether the relative affected is first-degree or second-degree, the number of relatives affected, and the age at which the relative was diagnosed with the disease. Family history is hypothesized to be associated with greater risk of breast cancer because a proportion of disease cases have been associated with genetic mutations, the most common being BRCA1 and BRCA2 inherited mutations. The population prevalence of first-degree or second-degree family history of breast cancer in US has been estimated at about 1 per cent, and population prevalence of women with two or more affected SDR and/or FDR at about 0.4 per cent (Hall et al., 2008; Collaborative Group on Hormonal Factors in Breast Cancer, 2001). About 20 per cent of breast cancer is associated with family history of disease while 5 to 10 per cent of disease is due to genetics

(Carroll et al., 2008). According to the review by Nelson et al. (2012), the increased risk (incidence rate ratio) of breast cancer among women with family history of disease compared with women without is about 1.90 for having an affected FDR and about1.70 for having an affected SDR.

The risk of disease increases with an increasing number of affected relatives. The Canadian study by Martin et al. (2010) reported almost 2.5 times the likelihood of breast cancer (lifetime probability) when a woman has two or more affected SDR or FDR. Family history is considered one of the strongest known risk factors for breast cancer and is the most likely risk factor to contribute to a women's decision to participate in screening mammography.

Breast density is described by the appearance of the breast tissue as visualized in a mammographic image (Berne and Spernak, 2005). The visual appearance of breast tissue on a mammogram has been classified both qualitatively and quantitatively as a way to describe breast density, and has typically been described using a proportion estimate; capturing the proportion of the breast area on the mammogram that is comprised of the dense stromal and epithelial tissue (Pharaoh et al., 1997). Although the evidence is still inconclusive it is hypothesized that breast density represents levels of endogenous hormones (Pharaoh et al., 1997). Breast Imaging Reporting and Data System (BIRADS) categories for breast density include category 1 and 2 representing less-than-average risk for breast cancer, and categories 3 and 4 representing higher-than-average risk. Breast tissue that is categorized as BIRADS 4 is considered extremely dense. In this category the breast contains greater than 75 per cent glandular and fibrous tissue which substantially reduces the sensitivity of mammography to about 44 per cent (Kerlikowske et al., 1996). The relative risk (incidence ratio) of breast cancer for women with extreme breast density (BIRADS 4) is 2.0 compared to women not in this category. The prevalence of women in these different breast density categories is not reported for Canada but the prevalence of BIRADS 4 among US women as been reported to be about 2 per cent (Eberl et al., 2006).

Risk Adjustment

The relative risk (RR) estimates in Table 14 represent the ratio of risk as represented by disease incidence between the exposed versus unexposed groups (those with the risk factor

versus those without). This ratio would be helpful if one wanted to scale up the breast cancer incidence rate of an unexposed population to the rate ratio implied for the exposed. However, in this study the baseline breast cancer risk needed to scale-up actually represents the risk among the general population (not among an unexposed population). We therefore needed to adjust the original RRs (exposed/unexposed) so that they reflect the ratio of risk in the exposed to the risk in the general population. This adjustment can be achieved by dividing the conventional RR in Table 1 by the adjustment represented in the following equation:

$$(1 - p) + RR * p$$

Note that when prevalence (p) is zero, this amounts to no correction. As the prevalence increases, the adjustment makes more of a material difference, provided the RR deviates materially from 1.

Table 15 Prevalence of risk factors for breast cancer from literature

Risk Factor	Adjusted relative risk for breast cancer	Estimated female population prevalence
Family history of disease, affected second-degree relative (SDR)	1.7	0.01
Family history of disease, affected first-degree relative (FDR)	1.9	0.01
Extreme breast density (BIRADS 4)	2.0	0.02
Family history of disease, two or more SDR and/or FDR (2+ FDR/SDR)	2.5	0.004

Analysis

All costs are reported in 2013 Canadian dollars. Both life-years and costs were discounted annually at 5%.

The analyses were carried out from the perspective of the third party payer i.e. provincial/territorial government as the main payer in a universal health care system and a lifetime time horizon. The impacts of each screening scenario for each high-risk population were measured as incremental cost and incremental benefits (life-years gained) compared to no screening.

For each screening policy option, we used POHEM-BCS to track the cohort of women from birth as described in a previous study (Dinh et al., 2015). The outcomes of interest for our analysis included the total number of: screens, breast cancer cases, life-years, and costs related to breast cancer diagnosis and treatment contributed by the cohort. To calculate screening costs, the total number of screens recorded by the model was used as a multiplier of the cost per screen. The screening costs were then coupled with the diagnosis and treatment costs recorded in the model to obtain total costs. These total costs were recorded for each screening policy along with an estimate of the health benefits measured in life years, which were also obtained from POHEM-BCS. The cost and benefits recorded for each policy were then used to form incremental cost-effectiveness ratios, for various pairs of programs (alternative versus comparator).

An incremental cost-effectiveness ratio (ICER) is a measure of excess cost to achieve an extra unit of health benefit conferred by the alternative program (Braithwaite et al., 2008). For each policy comparison, the difference in total cost was divided by the difference in the total number of life-years using the following equation:

$$ICER = \frac{\Delta C}{\Delta E} = \frac{(C_2 - C_1)}{(E_2 - E_1)}$$

For the purposes of this study, C denotes total cost of screening, diagnosis, and treatment, and E denotes effectiveness in life-years. Subscripts 1 and 2 denote the screening options being compared, for example the cost-effectiveness of option 2 versus option 1.

Incremental cost-effectiveness ratios (ICER) were used to compare the screening options and the principle of extended dominance was applied in a sequential analysis. The options were first ranked according to their effectiveness on the basis of securing maximum effect, neglecting cost. Sequentially, each option was compared to the previous in terms of incremental cost-effectiveness. The details of this analytic approach have been reported elsewhere (Dinh et al., 2015).

Also provided is a descriptive comparison between the results of each high-risk population's sequential analyses and the base case analysis results from the previous study (Dinh et al.,

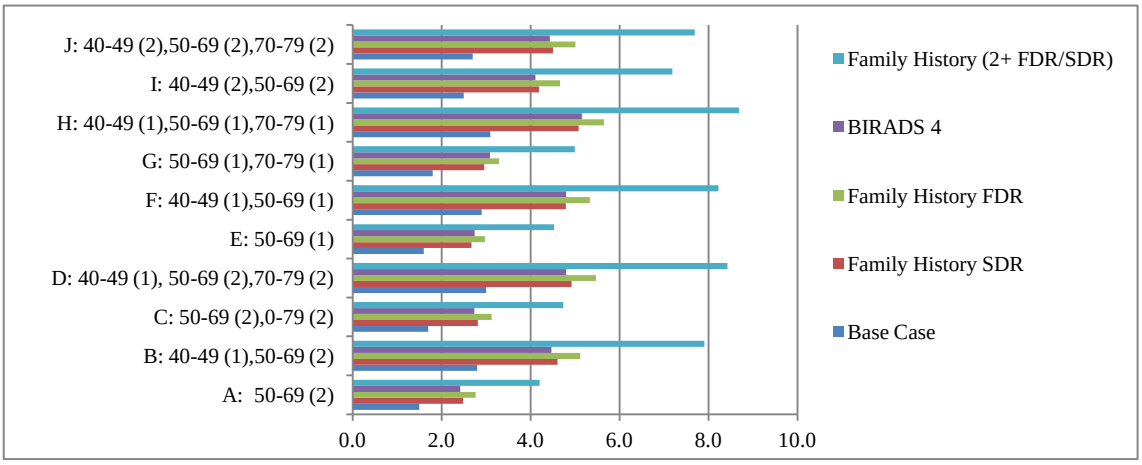
215). The previous study's base case analysis reflects estimated cost-effectiveness of screening mammography in the general or average-risk population. In this base case analysis, disease risk (age-specific incidence) of breast cancer was lower than that of the high-risk groups. All other model parameters were the same for the base case as for the high-risk analysis, with the exception of lower test sensitivity in the population with extreme breast density (BIRADS 4) in comparison to the general population.

Results

Life-Years Gained

We measured health benefits of screening in terms of the number of accumulated life-years gained (LYG) for each cohort by screening policy option. For each risk group, the cumulative total life-years for each screening policy option were compared against each other in a sequential analysis. The number of total LYGs across all screening options increased with breast cancer risk (see Figure 13). However, despite high-risk of disease in the group with extreme breast density (BIRADS 4), the life expectancy gains were muted because of the low test sensitivity of mammography in this group. Substantially greater health benefits of screening were seen for the high-risk group characterized by two or more affected FDR and/or SDR compared to all other populations. Screening in the base case was the least effective. When comparing across screening policy options however, screening health benefits in terms of LYGs is not so dramatic. Across all populations, it was observed that annual screening for ages 40 to 79 years (option H) resulted in the most LYGs across populations, and was most marked for the group with two or more affected SDR and/or FDR (8.7 LYGs from birth per 1,000 women or 3.2 days gained from birth per woman). Screening women age 50 to 69 every two years (option A; status quo) resulted in the least number of LYGs in comparison to no screening across all populations.

Figure 13 Total life-years gained per 1,000 women by screening policy and population, annual discounting at 5%



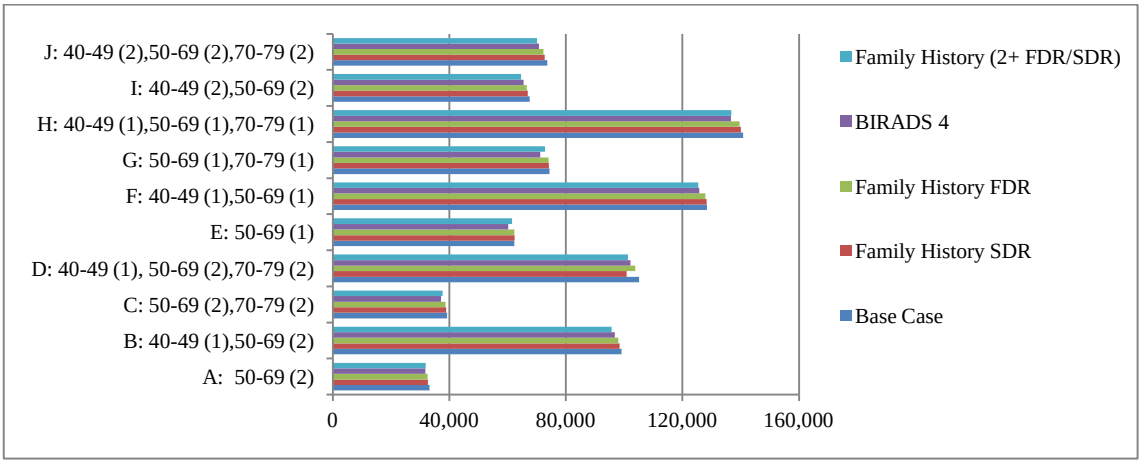
Costs

Screening costs were not expected to vary significantly across populations since in the model the number of screens is independent of the breast cancer risk and screening participation rates for each population were not adjusted. The only cost component that would depend on breast cancer incidence (and thus upon the categories of subpopulations) were diagnosis and treatment costs, but these costs represent a relatively small proportion of total (per capita) costs. Similar to the base case analysis in the previous study, total costs across all high-risk populations increased with the number of total screens. The relationship between total costs and disease risk is rather muted because the predominant portion of total costs (screening costs) is unrelated to breast cancer incidence in this model.

All costs are expressed in per capita (more specifically per 1,000 women) terms and represent excess costs, namely the costs of the program minus the costs incurred under the no screening option. The label "total" in relation to costs refers to the sum of screening, diagnosis, and treatment costs. The term "treatment cost" as a label was used for the sum of diagnosis and treatment (exclusive of screening) costs. In relation to total costs, it was observed that the highest value in the base case analysis and the lowest value in the high-risk analysis for population with two or more affected SDR and/or FDR (see Figure 14). This group had the highest risk for breast cancer among the modeled high-risk populations.

Again, in relation to total cost, but this time within population, it was observed that the highest total cost was for annual screening for women age 40 to 70 years (screening policy option H). The cost for this screening policy option (H) ranged from $136,792 in excess cost per 1,000 women in the highest-risk population (two or more SDR and/or FDR) to $140,830 in excess cost per 1,000 women in the base case. Across all high-risk populations, the lowest total cost was observed for biennial screening for women age 50 to 69 years (screening policy option A), which is also considered the status quo.

Figure 14 Total excess cost per 1,000 women by screening policy option and population ($ CAD 2013), annual discounting at 5%



Sequential Analysis

The results of the sequential analyses are described in Tables and Figures 15-19. We describe the results of the analyses for the high-risk populations in order of lowest to highest risk (incidence) of breast cancer.

Under a WTP of $50,000 per LYG, the optimal screening policy option for population with one SDR was biennial screening for women ages 40 to 79 years (J). This screening policy option was estimated to cost $16,146 per life-year gained compared to no screening (see Table 2). The same screening policy option was identified as optimal under the base case. As shown in Figure 3, screening for SDR compared to its base case counterpart (general population) manifests in a shift in LYGs (rightward along the LYG axis), but no discernible shift in the total cost (the vertical axis). This pattern of no discernible shift in total cost is consistent with observations of muted shifts in costs made in relation to Figure 2 earlier.

The results of the sequential analysis for one FDR also found biennial screening for women ages 40 to 79 years (J) as the optimal screening policy option at a WTP of $50,000 per LYG, with an incremental cost-effectiveness ratio of $14,432 per LYG compared to no screening (see Table 16). This result is similar to that found in the base case. In addition, screening in this risk group generates more LYGs compared to the base case across all screening policy options (see Figure 15). As in the case of the one SDR high-risk group, the one FDR group relative to its base case option counterpart manifests a rightward shift along the LYG axis, with no discernible shift in the cost (vertical axis). The rightward shift (increase) along the LYG axis is slightly more pronounced in the case of FDR than was observed for SDR, which is consistent with our emerging understanding of the influence of upward shifts in breast cancer incidence.

Table 16 Sequential Analysis Results –Family History with an Affected Second Degree Relative (SDR)

Scenario	Age (Interval)	LYG	LYG per 1,000 women	Total excess (vs no screen) health care cost ($ millions CAD)	Total excess health care cost ($CAD) per 1,000 women	ICER vs no screening (cost in $ CAD per LYG)	Sequential ICER (cost in $ CAD per LYG)
A	50-69 (2)	14,914	2.5	196.2	32,707	13,158	13,158
C	50-69 (2) 70-79 (2)	16,891	2.8	233.4	38,896	13,816	18,780
J	**40-49 (2) 50-69 (2) 70-79 (2)**	**27,023**	**4.5**	**436.3**	**72,719**	**16,146**	**20,030**
D	40-49 (1) 50-69 (2) 70-79 (2)	29,526	4.9	605.3	100,884	20,501	67,524
H	40-49 (1) 50-69 (1) 70-79 (1)	30,486	5.1	840.4	140,070	27,568	245,019
I	40-49 (2) 50-69 (2)	25,158	4.2	401.6	66,926	15,961	Extended dominance through J and A
E	50-69 (1)	16,043	2.7	374.3	62,385	23,332	Dominated by C, I, J, B, and D
G	50-69 (1) 70-79 (1)	17,759	3.0	445.3	74,215	25,074	Dominated by I, J, B, and D
B	40-49 (1) 50-69 (2)	27,625	4.6	590.3	98,380	21,367	Dominated by D
F	40-49 (1) 50-69 (1)	28,785	4.8	769.6	128,268	26,737	Dominated by D

Notes: Total population of 6 million women in each population was modeled in POHEM-BCS. LYG = Life-years gained; ICER = incremental cost-effectiveness ratio. Interval labels of 1 and 2 denote the duration in years between screens where 1 denotes annuals screening and 2 denotes biennial screening. Costs and benefits (LYGs) have been discounted by 5%. 6 million women were assumed to face an identical amplification of breast cancer incidence appropriate to their high-risk category, and thus represent a fiction n which all 6 million Canadians fall into the requisite high-risk category.

Figure 15 Cost-Effectiveness Ratios for the Base Case and Family History (Affected Second-Degree Relative) Analyses (Costs per 1,000 women versus Life-Years Gained per 1,000 women)

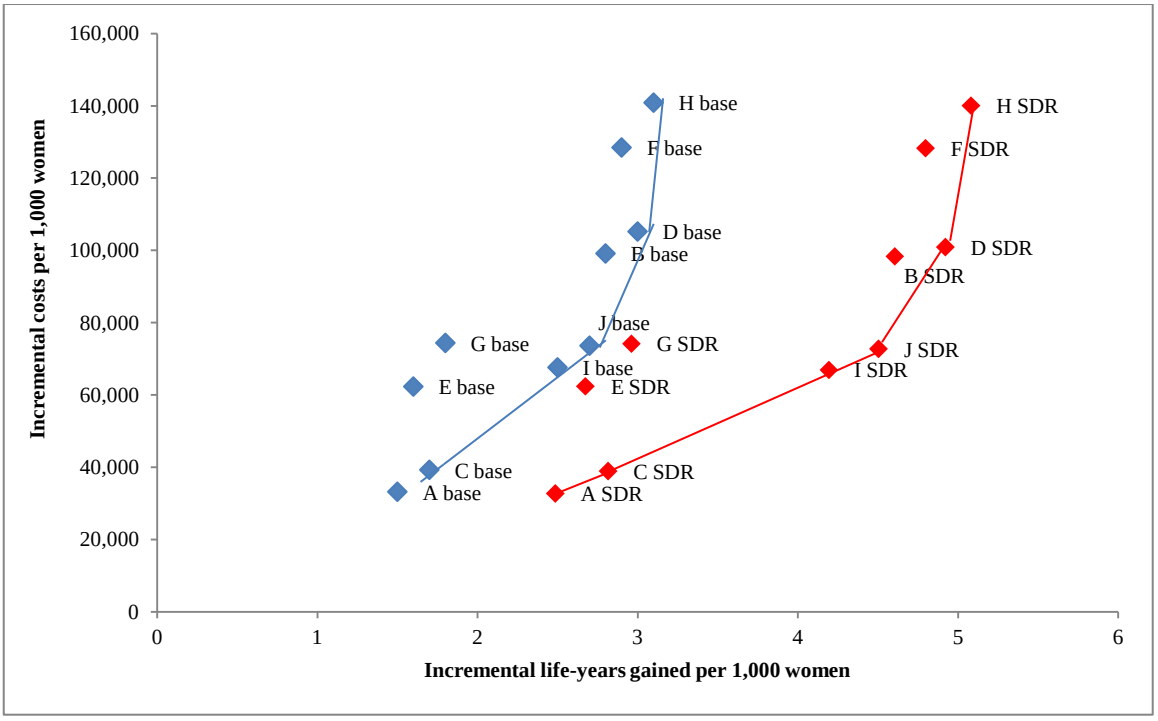
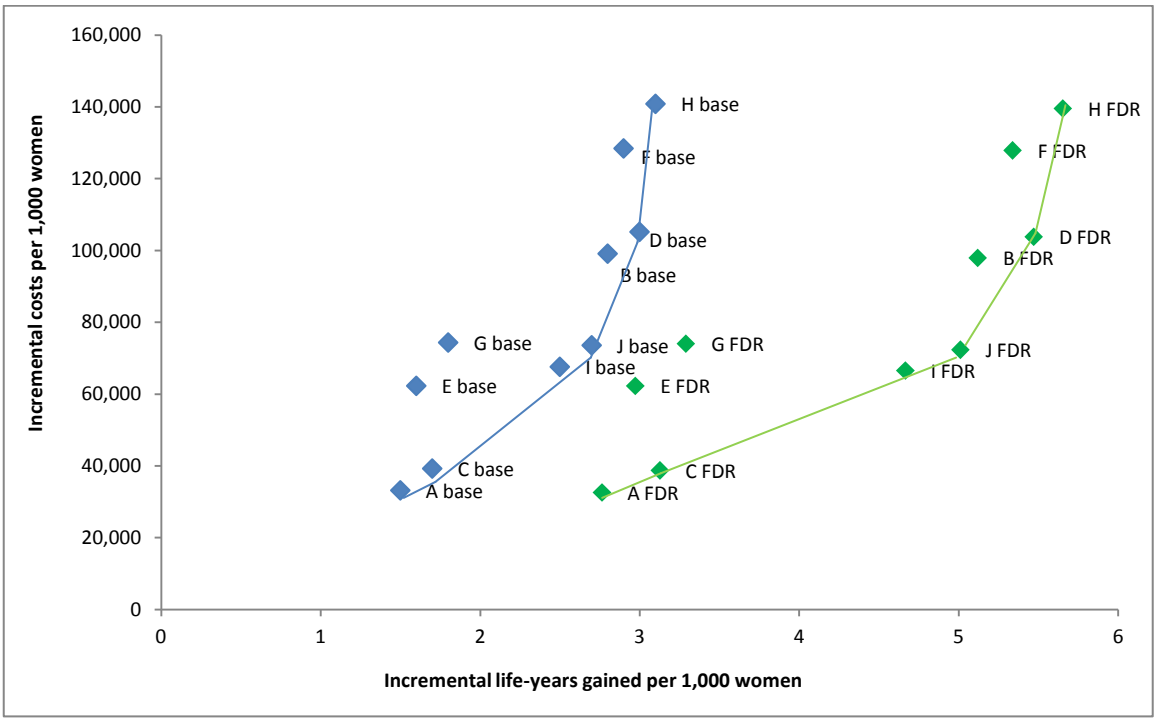


Table 17 Sequential Analysis Results – Family History with an Affected First Degree Relative (FDR)

Scenario	Age (Interval)	LYG	LYG per 1,000 women	Total excess (vs no screen) health care cost ($ millions CAD)	Total excess health care cost ($CAD) per 1,000 women	ICER vs no screening (cost in $ CAD per LYG)	Sequential ICER (cost in $ CAD per LYG)
A	50-69 (2)	16,584	2.8	195.5	32,579	11,787	11,787
C	50-69 (2) 70-79 (2)	18,764	3.1	232.3	38,715	12,380	16,889
J	**40-49 (2)** **50-69 (2)** **70-79 (2)**	**30,059**	**5.0**	**433.8**	**72,305**	**14,432**	**17,842**
D	40-49 (1) 50-69 (2) 70-79 (2)	32,821	5.5	622.8	103,807	18,977	68,455
H	40-49 (1) 50-69 (1) 70-79 (1)	33,913	5.7	837.5	139,586	24,696	196,578
I	40-49 (2) 50-69 (2)	27,999	4.7	399.4	66,560	14,264	Extended dominance through J and A
B	40-49 (1) 50-69 (2)	30,713	5.1	587.8	97,965	19,138	Extended dominance through D and A
E	50-69 (1)	17,836	3.0	373.7	62,278	20,950	Dominated by C
G	50-69 (1) 70-79 (1)	19,736	3.3	444.3	74,047	22,511	Dominated by I and J
F	40-49 (1) 50-69 (1)	32,023	5.3	767.1	127,850	23,955	Dominated by D

Notes: Total population of 6 million women in each population was modeled in POHEM-BCS. LYG = Life-years gained; ICER = incremental cost-effectiveness ratio. Interval labels of 1 and 2 denote the duration in years between screens where 1 denotes annuals screening and 2 denotes biennial screening. Costs and benefits (LYGs) have been discounted by 5%. 6 million women were assumed to face an identical amplification of breast cancer incidence appropriate to their high-risk category, and thus represent a fiction n which all 6 million Canadians fall into the requisite high-risk category.

Figure 16 Cost-Effectiveness Ratios for the Base Case and Family History (Affected First-Degree Relative) Analyses (Costs per 1,000 women Versus Life-Years Gained per 1,000 women)



For the high-risk group with extreme breast density (BIRADS 4), the optimal screening policy option was also biennial for women ages 40 to 79 years (J) with a cost-effectiveness ratio of $15,957 per LYG compared to no screening (see Table 18). Compared to the base case, screening in this high-risk group resulted in greater LYGs and slightly lower costs across all screening policy options (see Figure17). As in the case of the one SDR and one FDR high-risk groups, the extreme breast density group relative to its base case option counterpart also manifests a rightward shift along the LYG axis, with no discernible shift in the cost (vertical axis). The rightward shift (increase) along the LYG axis is not as pronounced in relation to our understanding of the influence of upward shifts in breast cancer incidence. This muted shift in the LYG axis may be due to lower test sensitivity in this high-risk group.

For the highest-risk group of two or more SDR and/or FDR, the optimal screening policy option was annual screening for women ages 40 to 79 years and biennial screening for

women ages 50 to 79 years (D). The cost-effectiveness ratio for this screening option compared with no screening was $12,019 per LYG (see Table 19). In contrast to the base case, more frequent screening for the age group 40 to 49 was more cost-effective under a WTP of $50,000. Also in comparison with the base case, screening in this risk group generated more LYGs across all policy options (see Figure 18) As in the case of the other high-risk group, the 2+SDR/FDR group relative to its base case option counterpart manifests a rightward shift along the LYG axis, with no discernible shift in the cost (vertical axis). The rightward shift (increase) along the LYG axis is more pronounced in the case in this high-risk group than was observed for FDR, SDR, or extreme breast density. This is consistent with our understanding of the influence of upward shifts in breast cancer incidence.

Table 18 Sequential Analysis Results – Extreme Breast Density (BIRADS 4)

Scenario	Age (Interval)	LYG	LYG per 1,000 women	Total excess (vs no screen) health care cost ($ millions CAD)	Total excess health care cost ($CAD) per 1,000 women	ICER vs no screening (cost in $ CAD per LYG)	Sequential ICER (cost in $ CAD per LYG)
A	50-69 (2)	14,538	2.4	190,818,477	31,803	13,126	13,126
C	50-69 (2) 70-79 (2)	16,418	2.7	222,966,205	37,161	13,581	17,096
J	**40-49 (2)** **50-69 (2)** **70-79 (2)**	**26,624**	**4.4**	**424,827,568**	**70,805**	**15,957**	**19,779**
D	40-49 (1) 50-69 (2) 70-79 (2)	28,794	4.8	612,803,012	102,134	21,282	86,612
H	40-49 (1) 50-69 (1) 70-79 (1)	30,920	5.2	819,903,219	136,651	26,517	97,435
E	50-69 (1)	16,471	2.7	361,383,892	60,231	21,941	Extended dominance through J and A
I	40-49 (2) 50-69 (2)	24,665	4.1	393,196,112	65,533	15,942	Extended dominance through J and A
B	40-49 (1) 50-69 (2)	26,831	4.5	580,673,453	96,779	21,642	Extended dominance through D and A
F	40-49 (1) 50-69 (1)	28,827	4.8	754,465,352	125,744	26,172	Extended dominance through H and A
G	50-69 (1) 70-79 (1)	18,553	3.1	426,932,433	71,155	23,011	Dominated by I

Notes: Total population of 6 million women in each population was modeled in POHEM-BCS. LYG = Life-years gained; ICER = incremental cost-effectiveness ratio. Interval labels of 1 and 2 denote the duration in years between screens where 1 denotes annuals screening and 2 denotes biennial screening. Costs and benefits (LYGs) have been discounted by 5 per cent. Six million women were assumed to face an identical amplification of breast cancer incidence appropriate to their high-risk category, and thus represent a fiction n which all 6 million Canadians fall into the requisite high-risk category.

Figure 17 Cost-Effectiveness Ratios for the Base Case and Extreme Breast Density (BIRADS 4) Analyses (Costs per 1,000 women Versus Life-Years Gained per 1,000 women)

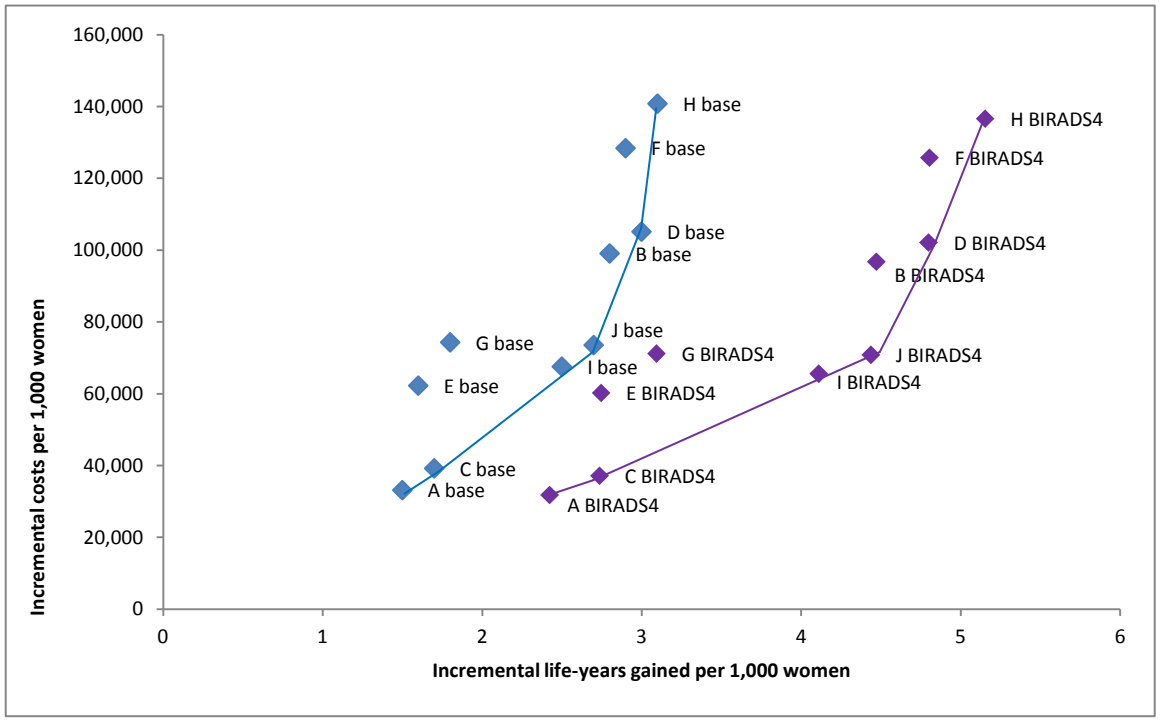
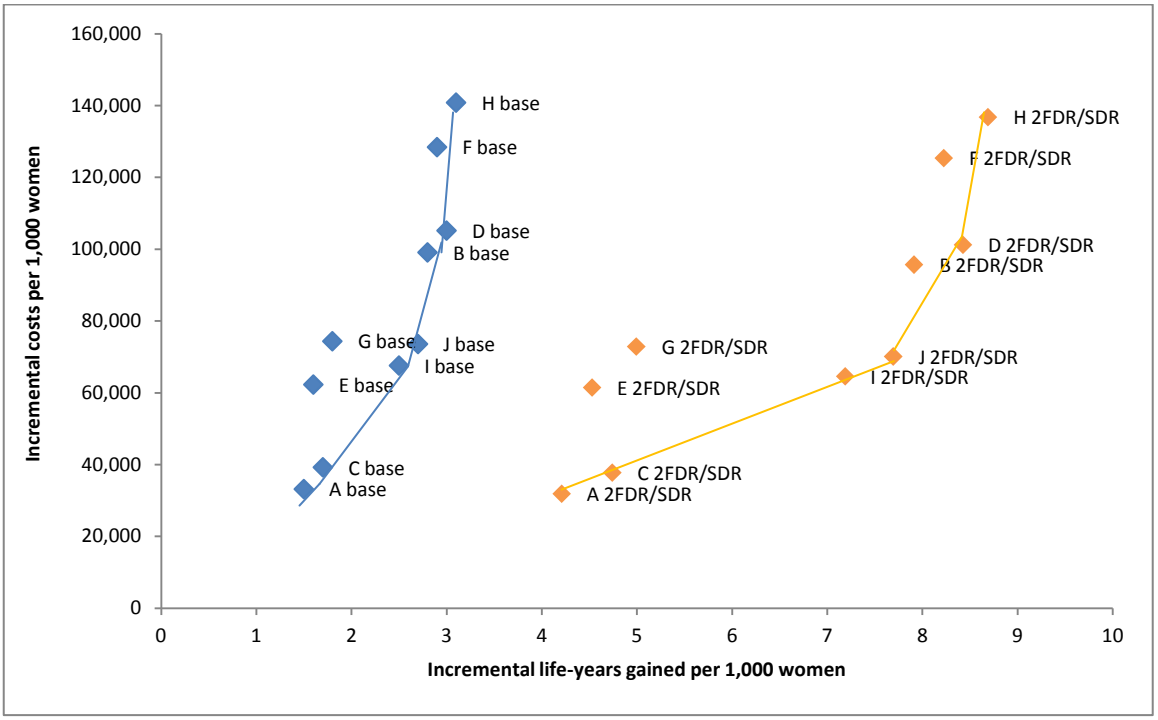


Table 19 Sequential Analysis Results – Family History with Two or More Affected First and/or Second Degree Relatives (2+FDR/SDR)

Scenario	Age (Interval)	LYG	LYG per 1,000 women	Total excess (vs no screen) health care cost ($ millions CAD)	Total excess health care cost ($CAD) per 1,000 women	ICER vs no screening (cost in $ CAD per LYG)	Sequential ICER (cost in $ CAD per LYG)
A	50-69 (2)	25,249	4.2	191.2	31,873	7,574	7,574
J	40-49 (2) 50-69 (2) 70-79 (2)	46,159	7.7	420.5	70,088	9,110	10,966
D	**40-49 (1)** **50-69 (2)** **70-79 (2)**	**50,547**	**8.4**	**607.5**	**101,257**	**12,019**	**42,619**
H	40-49 (1) 50-69 (1) 70-79 (1)	52,131	8.7	820.7	136,792	15,744	134,606
C	50-69 (2) 70-79 (2)	28,444	4.7	226.5	37,757	7,965	Extended dominance through J and A
I	40-49 (2) 50-69 (2)	43,131	7.2	387.5	64,590	8,985	Extended dominance through J and A
E	50-69 (1)	27,160	4.5	369.0	61,501	13,586	Dominated by C
G	50-69 (1) 70-79 (1)	29,959	5.0	437.4	72,900	14,600	Dominated by I
B	40-49 (1) 50-69 (2)	47,467	7.9	574.0	95,663	12,092	Dominated by D
F	40-49 (1) 50-69 (1)	49,345	8.2	752.5	125,426	15,251	Dominated by D

Notes: Total population of 6 million women in each population was modeled in POHEM-BCS. LYG = Life-years gained; ICER = incremental cost-effectiveness ratio. Interval labels of 1 and 2 denote the duration in years between screens where 1 denotes annuals screening and 2 denotes biennial screening. Costs and benefits (LYGs) have been discounted by 5%. 6 million women were assumed to face an identical amplification of breast cancer incidence appropriate to their high-risk category, and thus represent a fiction n which all 6 million Canadians fall into the requisite high-risk category.

Figure 18 Cost-Effectiveness Ratios for the Base Case and Family History (Two or More Affected First- and/or Second-Degree Relative) Analyses (Costs per 1,000 women Versus Life-Years Gained per 1,000 women)



Discussion

Interpretation of Results

Family history of breast cancer and extreme breast density are considered important risk factors for breast cancer, with relative risks for disease ranging from 1.7 to 2.5 in this study. They are also considered risk factors that are relatively easily identifiable by women or by a clinician. Our results show that, in comparison to the base case (average risk or general population), screening women with elevated risk of breast cancer, specifically women with family history or extremely dense breasts, was less costly and more effective in terms of extending total life-years. As breast cancer risks increases, LYGs increase and total costs decrease. Screening women age 40 to 49 years annually and women age 50 to 79 years biennially was found to be optimal for the highest-risk group in this study (having two or more affected SDR and/or FDR). For all other high-risk groups (one SDR, one FDR, or BIRADS 4) and for the base case (screening in the average-risk population), it would be optimal to screen women age 40 to 79 biennially. Despite high-risk for disease among the

population with high breast density, more frequent screening for this group was not found to be cost-effective, probably due to low test sensitivity in these women. In the previous study it was found that optimal screening for the average-risk population was biennial screening for women age 40 to 79 years. The results of our research may suggest that risk-based screening may be more efficient and effective than uniform recommendations for the general risk population.

The cost-utility analysis study by Schousboe et al. (2011) estimated the cost-effectiveness of mammography stratified by breast density and family history of breast cancer for women of varying ages using microsimulation modeling. Their modeling approach was similar to the one used in this study as they employed epidemiological data as well as estimates from the medical literature to determine impact of screening. In contrast to this study, they applied quality of life to their health benefits and found that women with high breast density and family history should be recommended biennial screening for women age 40 to 79 years. They did not find annual screening to be cost-effective for any age group. If quality of life was applied to the LYGs in this study, similar results may have been found. Quality of life takes into account certain harms of vigorous screening including pain and anxiety associated with mammography, waiting for results, and false-positive tests.

The more recent study by Vilaprinyo et al. (2014) found that risk-based screening strategies were more cost-effective and have better harm-benefit ratios than screening recommendations for the general risk population. The researchers also recommended against using a “one-size-fits-all” approach to breast cancer screening and a shift towards more personalized recommendations, noting that risk-based strategies can reduce costs as well as potential harms of screening (over-diagnosis, radiation exposure, anxiety, etc). They concluded that optimal screening is characterized by quinquennial (every 5 years) or triennial (every 3 years) screening for the population with low or moderate risk of breast cancer and annual screening for the high-risk populations. They also recommended the development of more accurate measures of individual risk for breast cancer which may result in improvements in risk-based organized screening programs and screening guidelines.

Limitations

There are several model assumptions that should be considered when interpreting the results of this study. Although these assumptions helped provide some structure to the analysis, they potentially limit the generalizability or external validity of these results. These limitations and their impact on the interpretation of model results are included in the previous study (Dinh et al., 2015).

One of this study's limitations was the reliance on the published literature for relative risk estimates in terms of incidence of disease among the population with the risk factor versus the population without, and prevalence of the risk factors in the population. It is believe that the estimates obtained from the published literature are reasonably generalizable to the Canadian female population. Although the relative risk of disease estimates came from a fairly recent systematic review, the individual studies that were included in that review are more than 10 years old which could influence validity. However, since the main interest of this study is the comparative performance of different screening policy options, its primary findings may be robust to small shifts in the true values of the rate or risk ratios.

Another limitation is that with one exception (the subpopulation with dense breasts) it was assumed that the only difference between the high-risk populations examined in this study and the general population examined in the previous study is breast cancer incidence (with screening sensitivity being considered as well as in the aforementioned exception). The model used in this study was adjusted for baseline breast cancer incidence using relative risk of disease and estimated population prevalence of the risk-factor in question in order to observed the changes in life-years and costs for different screening policy options when disease risks increases incrementally. It is possible that these high-risk populations may be different from the average-risk population and from each other, in other ways that may alter the impact of screening mammography on disease mortality.

For each analysis, a single rate ratio estimate as a multiplier for breast cancer incidence equally across all age strata was used since either age-stratified rate ratio estimates or age-stratified prevalence estimates could not be found. It is suspected that for strong risk factors (high rate ratios) such as for the risk factors of having family history of disease and extreme

breast density, one would likely see higher rate ratios in younger age groups which would likely contribute to greater life-years lost. Further, as there are no national level data on the true prevalence of family history of breast cancer and breast density in Canada, this study relied on prevalence estimates from the published literature. It is believed however that the populations in the studies from which these estimates were taken is reasonably comparable to the Canadian population. Risk factors for extreme breast density include late age at first birth and increased circulating hormone levels related to hormone replacement therapy use in postmenopausal (Titus-Ernstoff et al., 2006). An increase in these factors may be expected to increase the prevalence of women falling in the high breast density category.

Health benefits accumulated from birth were reported in this study. Since the benefits of screening do not materialize until at least the age of screening (earliest age of benefit is 40 years in this study), the framing of the analysis discounts benefits by an amount requisite to at least 40 years. This has the expected result of heavily discounting benefits, in part as an artefact of the study's analytical design. While the absolute values of these benefits may be instructive to individuals at birth, they are not instructive to women of age 35 or 40, for example, who might be weighing their choice to participate in screening. In the case of a 40-year old woman, her prospective benefits accruing in the next 10 years would be more prominent if benefits were only discounted within the time horizon of 0 to 10 years, yet this study's analysis effectively discounts their benefits not within this time horizon but one spanning between 40 to 50 years. This results in a substantially greater diminution of benefits. These similar concerns also apply to the costs. For the highest risk group – family history with two or more affected SDR and/or FDR, the most effective screening policy option generated 3.2 days of extended life from birth compared with no screening. A more meaningful measure in this case may consider changes in residual life expectancy.[19]

The data used in POHEM-BCS came from several different sources, many of which are quite old. The breast cancer screening module of POHEM-BCS includes data that dates back as far as the 1980's with most recent data from the early 2000's. Breast cancer incidence data from the 1980's was used in the model to predict rates of the disease before screening programs were implemented in Canada. However, were observed a close match between the

[19] An example of such life expectancy statistics is available from Statistics Canada, CANSIM, Table 102-0512.

historical trends in baseline breast cancer incidence rates in POHEM-BCS with more recent trends. Over the past decade there have not be significant changes in breast cancer incidence rate in Canada therefore we believe the breast cancer incidence data in POHEM-BCS is appropriate. Screening sensitivity and specificity values were estimated from more recent data from the Canadian Breast Cancer Screening Database and therefore more accurately reflects current experience.

There was an attempt to adjust the 1995 treatment costs in POHEM-BCS to reflect present day value by simply inflating the cost estimates to 2013 Canadian dollars. As the original process to estimating treatment costs required a separate study to identify current patterns of health care utilization from the time of breast cancer diagnosis, it was not possible to more accurately reflect any changes in treatment regimes and changes in costs that may be attributable to higher prices for new innovations or lower prices for older technologies. Nor did the estimation of costs reflect any changes in provider fees. This study was unable to predict whether treatment cost is an under- or over-estimate of the true cost of breast cancer treatments. The cost per screen was also inflated from the year 2008 to the year 2013. The cost per screen of about $75 from the Ministry of Health in British Columbia was used in this study to represent the average cost per screen for the country. Since health care provision is predominately a provincial/territorial responsibility, screening costs can vary significantly across provinces and territories due to differences in component costs such as adjustments in salaries and technology costs. It is uncertain whether the cost per screen used in this study is an under- or over-estimate of the average national cost of screening.

Finally, more targeted screening based on disease risk may incur a higher cost per screen. Since population prevalence is relatively low for risk factors examined in this study, it is reasonable to assume that annual screening volumes would decrease dramatically. Although the total cost of screening would hypothetically reduce, the per unit cost of screening (cost per screen) may actually be higher than current program screening that is defined by age because of lower total volume screening as result of more a more targeted screening program. A higher cost per screen with lower total volume screening is based on the theory of “economies of scale” or operational efficiency which assumes cost advantages that an administrating organization would obtain due to high volume output or large scale operation.

Under these circumstances, the cost per unit of output would decrease with increasing scale because fixed costs or overhead costs would be spread out more evenly with more units of output. In addition, there would be a cost to identifying women who would be considered as SDR, FDR, or BIRADS 4. A process of pre-screening would be required to identify these women which could be conducted via qualitative or biomedical means. These approaches would incur a range of different costs that would need to be captured in the full cost of screening. Since the cost-effectiveness of screening is sensitive to the cost per screening and of course the total cost of screening, underestimating the true cost could result in the erroneous selection of an optimal screening policy option.

Research and Policy Implications

Future epidemiologic studies on population prevalence of breast cancer risk factors and their contributions to disease risk by age group would be beneficial in increasing the accuracy of the results of this study and future studies using similar methodology. Future studies may also assess the health and economic impact of different screening options for other populations for which more recent evidence show high-risk for breast cancer. For example, the prevalence of high body mass index (BMI), which has been found in some studies to be positively associated with post-menopausal breast cancer risk but higher screening sensitivity, has been increasing steadily over time especially in women. The systematic review by Nelson et al. (2012) using older trial data reported an increased breast cancer risk among woman in the upper BMI quartiles versus the lower quartile with relative risk (RR) of 1.28. While the estimate was not statistically significant (RR 95 per cent confidence interval: 0.98-1.66), more follow-up may provide a more precise estimate. The results of our study may support guidelines for more frequent and earlier screening mammography of women with increased risk of breast cancer.

Future studies may also consider the impact of longer intervals such as screening every three years and younger age eligibility (before age 40). Also future studies may also take into account the impact of other preventive measures such as lifestyle modification and treatments such as tamoxifen and preventive mastectomy, as well as broadening the scope of measured benefits for health economic evaluation including the measurement of quality of life impacts through a cost-utility analysis. Costs could also include indirect costs such as

productivity measures (short and long term disability, premature mortality), caregiver costs, home and community care, and other out of pocket costs. The inclusion of indirect costs would change the perspective of the health economic analysis to a societal perspective.

More recent studies have looked at the cost-effectiveness of using more sensitive but potentially more costly modalities for breast cancer screening including digital mammography and magnetic resonance imaging (MRI). Future research may include a cost-effectiveness or cost-utility analysis to determine optimal screening for high-risk populations using these more sensitive modalities.

In terms of policy implications, the findings of this study seem to suggest that women with significantly elevated risk of breast cancer compared to the general female population with the same policy recommendations. Optimal screening is characterized by biennial screening for women ages 40 to 79 for the general population as well as for women with family history SDR or FDR as well as for women with extreme breast density. Optimal screening for the highest-risk group in this study, women with family history of two or more SDR and/or FDR, was characterized by annual screening of women age 40 to 49 and biennial screening for women ages 50 to 79. It should be noted that should Canada operationalize a multi-tier screening strategy; it is possible that the cost per screen as well as diagnosis and treatment costs that we applied in our model may not apply. In our analysis we assumed a "one size fits all" approach to screening and this was reflected in a use of an average costing by screen and by stage of cancer (treatment). It is plausible that more tailored screening for high-risk populations may incur higher per screen costs per patient (depending on the balance between capital and operational costs of screening). Although our analysis showed greater efficiency in screening high-risk populations compared to the general population, greater consideration or more refined analysis in regards to more accurately estimating costs (screening, diagnosis, and treatment) should be made.

The Canadian Task Force on Preventive Health Care (2011) recommends that women at average risk of breast cancer be screened from age 50 to 74 every 2 to 3 years with mammography however does not provide recommendations for screening in subpopulations with elevated risk of disease. The Task Force's guidelines are currently in place to help

clinicians make recommendations to their patients and to help women to decide on when they should start screening for breast cancer and how often. National guidelines could be amended to include evidence-based screening protocols for women with significant elevated risk for breast cancer. The findings of this study along with the evidence from other similar studies could contribute to national guidelines for screening for specific subgroups of the population.

Chapter 6: General Discussion

Preamble

Breast cancer continues to be one of the most common cancers and cause of cancer-related deaths among Canadian women. The goal of screening is to reduce breast cancer mortality through early detection at a point of time during disease progression where treatment is more effective than if the disease was detected clinically or symptomatically. Regular screening mammography (every two to three years) has been recommended for women of moderate risk based on age (usually starting at age 50 until 70). In practice, women as young as 40 have been recommended to screen every year, as well as women age 70 and older are offered screening every two years. Initial randomized controlled trials from over twenty years ago established the effectiveness of screening mammography in reducing breast cancer mortality. However, in the past several years the benefits of breast cancer screening have been questioned. These concerns are influenced by increased survival rates, even in late-stage breast cancer, which may be attributable to greater awareness of the disease as well as improvements in diagnosis and treatment over time. Also, mammography is not without risks. For example, mammography can cause anxiety and unnecessary work-ups and treatment in the case of a false-positive test. Further, wide-scale screening of the female population characterized by age alone is resource-intensive and costly. In the Canadian context, there have been very few studies that examine the cost-effectiveness of screening mammography.

This dissertation has six research objectives. The first objective, which was addressed in the introductory chapter, was to provide a comprehensive review of the current issues pertaining to breast cancer and breast cancer screening. The second objective was to provide an overview of quantitative models used in the health and economic evaluation of screening mammography. The third objective was to appraise the health economic evaluation literature on the impact of age-eligibility, screening frequency or interval, and breast cancer risk, on the cost-effectiveness of population-based screening mammography in order to inform a cost-effectiveness study for Canada. The fourth objective was to determine the optimal screening policy option, characterised by age-eligibility and screening frequency, using a

cost-effectiveness analysis approach. The fifth objective was to determine optimal screening, characterised by age-eligibility and screening frequency, for women with elevated risk of breast cancer including with family history of disease and extreme breast density.

This final section of the dissertation provides a summary of the research findings including a discussion of the overall strengths and limitations of the research. Also included is a commentary on other important issues that were not addressed in the analyses as well as the potential research, practice, and policy implications of the research.

Summary of the Analytic Findings

Three of the six analytic chapters of this dissertation were individual studies that will be prepared and submitted for publication. The findings of each of these studies are summarized in this section.

Systematic Review of Health Economic Studies on the Cost-Effectiveness of Mammography: The Impact of Age Eligibility and Screening Frequency

The objective of the third chapter of this dissertation was to review the health economic evaluation literature on the impact of age eligibility, screening frequency, and breast cancer risk on the cost-effectiveness of screening mammography in order to inform a cost-effectiveness study for the Canadian context. At the time of when the review was conducted, there were no satisfactory Canadian health economic studies. The review was conducted in 2010 and included health economic evaluation studies published in the last ten years. After screening titles, abstracts, and full-text reviews for relevance and quality, a total of 11 studies were retained for data abstraction and synthesis. The included studies were compared using a narrative review approach on several components including analytic perspective, time horizon, population, characteristics of the screening programs or policies under study, measurement of health benefits and costs, discount rate used, and results (health benefits and costs converted to CAD 2013). The results of the included studies were evaluated to provide insight on the contributions of different ages and screening frequencies on screening mammography cost-effectiveness. In terms of age ranges, we were specifically interested in the contribution of younger and older age ranges outside of 50 to 69 (the

standard recommended age for screening). We also wanted to see whether more or less frequent screening in comparison to biennial screening made a material difference on cost-effectiveness. In addition, insights into the benefits of targeted screening for special populations were also of interest to the review. Based on the findings of the review, we determined that screening every 2 or 3 years was cost-effective for age groups ranging from 40 to 80. Annual screening was found to often incur substantially more costs than life-years. We also realized that screening in high-risk populations was often considered cost-effective with mammography but women with comorbidities may not see the benefits of screening due to competing health risks that may have a large impact on mortality. We also learned from this review that a Canadian cost-effectiveness study that specifically assesses the contributions of age, screening frequency, and risk to the cost-effectiveness of screening mammography would be adding new knowledge to the existing literature.

Cost-Effectiveness of Screening Mammography in Canada: Impact of Age Eligibility and Screening Frequency in the General Female Population

The objective of the fourth chapter of the dissertation was to identify the optimal screening mammography policy for the general Canadian population, based on several design options related to age eligibility and screening frequency (1, 2, or 3 year intervals). Optimal screening is described as a screening scenario that incurs maximum benefits at a low incremental cost compared to another screening option and whereby the cost per additional life-year gained (LYG) is below the willingness-to-pay threshold (WTP). The Statistics Canada's Population Health Model- Breast Cancer Screening Module (POHEM-BCS) was used to evaluate the trade-offs between health benefits in total life-years gained and total costs, including screening, diagnosis, and treatment over the lifetime of a hypothetical cohort of Canadian women. This study was conducted from the perspective of the third-party payer (provincial/territorial government that oversees the funding and operation of breast cancer screening programs, an annual discount rate of 5 per cent applied to both future health benefits and costs, and a WTP threshold of $50,000 per LYG was used in the base-case analysis. The study also conducted a series of one-way sensitivity analyses using varying estimates (usually high and low) for cost per screen, discount rate, initial participation rate, mammography test sensitivity and specificity, WTP thresholds.

The results of these analysis showed that it is optimal to screen women ages 40 to 79 every two years when the WTP threshold is $50,000 per LYG under the base case analyses. This observation also held true under several conditions (sensitivity analyses): lower cost per screen, lower participation rates, and perfect test sensitivity or specificity – all with a WTP of $50,000 per LYG. The only conditions under which annual screening (more frequent screening) was optimal for any one of the age groups (annual screening for age 40 to 49 years and biennial screening for age 50 to 79 years) was with no discounting under a WTP of $50,000 per LYG. Less frequent screening (no screening for ages 40 to 49 years and 70 to 79 years and biennial screening for age 50 to 69 years) was optimal under the condition where cost per screen or discount rate was high under a WTP of $50,000 per LYG. Although $50,000 per LYG is the standard WTP in health economic analyses, it has been suggested that practically, a lower WTP would be appropriate. When a WTP of $30,000 per LYG was used, only biennial screening for ages 50 to 69 years was optimal under the base case conditions; no screening option was optimal under the conditions of high cost per screen or high discount rate; annual screening for any age group was not optimal under any condition; and more frequent screening outside of the standard 50 to 69 year age group (biennial screening for age 40 to 49 and/or 70 to 79) was optimal under the conditions of low cost per screen, low discount rate, lower participation rate, and perfect test sensitivity or specificity. In general, these analyses show that at the broad population level, that population-based screening based on age eligibility alone can be costly at an incremental health benefit over the long-term. Additional analysis (as is explored in the subsequent chapter) could assess the cost-effectiveness of more targeted screening for women of higher-risk for breast cancer which could serve as a less costly, more effective endeavour at the population-level.

Cost-Effectiveness of Screening Mammography in Canada: Impact of Age Eligibility and Screening Frequency in High-Risk Women

The objective of the fifth chapter (and third manuscript), was to identify optimal screening in high-risk populations including women age 40 to 79 who have family history of disease in a one or more first-degree relative (FDR) and/or one or more second-degree relative (SDR), or women in that age range with extreme breast density (BIRADS 4). This study applied the same methodology, microsimulation model, and base-case assumptions as in the previous

analyses (chapter four). The risk profile of the hypothetical cohort in the model was created by adjusted the baseline breast cancer incidence by age in POHEM-BCS. The impact of various screening policy options in terms of costs (screening, diagnosis, and treatment) and health benefits (LYGs) in each of the high-risk populations was compared to determine optimal screening based on risk. The screened populations assessed in this study in order of lowest to highest risk for breast cancer were as follows: one SDR, one FDR, BIRADS 4, and two or more SDR and/or FDR. Similar to the previous chapter, optimal screening was described as a screening scenario that incurs maximum benefits at a low incremental cost compared to another screening option and whereby the cost per LYG is below the WTP. The results of the analysis for each high-risk population were also compared with the base case results (impact in the general risk female population) from chapter four.

The results of this study showed that, as expected, the higher the risk for breast cancer in the screened population and the higher the volume of screening, the higher the number of total LYGs. In contrast, the total costs did not appear to change dramatically across screening populations (risk groups). Similar to the previous chapter's base case analysis, total costs were most affected volume of screening therefore the most costly screening programs were those that included more age groups (40 to 49 years, 50 to 69 years, and/or 70 to 79 years) and more frequent screening (screening every year versus every two years within any one or more age groups). Biennial screening for women ages 40 to 79 was the optimal screening policy option for the SDR population, FDR population, and BIRADS 4 population. The optimal screening policy option for the highest risk population (two or more SDR and/or FDR) was annual screening for women ages 40 to 49 and biennial screening for women ages 50 to 79.

Since the population prevalence of these risk factors is relatively small (varies from 0.4 to 2 per cent), the cost of screening may be underestimated. The per screen costs for a targeted screening program for higher-risk women compared to the current programs defined by age may be higher based on "economies of scale" theory or operational efficiency. In addition, costing of targeted screening programs would need to include the costs related to determining which women would be considered high-risk. Risk-profiling women could be conducted via qualitative or biomedical approaches which would incur a range of different

costs. Future analyses may explore the actual per screen related to lower-volume screening as well as the cost of risk-profiling individual women. The cost-effectiveness of screening appeared to be sensitive to the cost per screening and an underestimate of the true cost of targeted screening based on risk may result in the erroneous selection of an optimal screening policy option.

General Research Strengths and Limitations

The research conducted within this dissertation adds to the current body of literature. At the time of writing, there were no published systematic reviews of health economic evaluations of screening mammography. It was also found that there were very few comprehensive cost-effectiveness studies for breast cancer screening within the Canadian context and that perhaps a greater role for the cost-effectiveness evidence in creating or supplementing recommendations for screening in Canada. Further, there were no published studies that addressed the cost-effectiveness of targeted screening for high-risk populations characterized by family history of disease and extreme breast density for Canada. The studies' results within this dissertation work were consistent with previous literature on the effectiveness and cost-effectiveness of screening mammography in that optimal screening for the general population includes screening women every two years and starting at the age of 50, with more frequent screening starting at the age of 40 may be warranted for higher-risk sub-populations including women with strong family history of disease.

There were several issues that were not adequately addressed in this dissertation but which are important considerations in research, practice, and policy. We address some of these issues in relation to impact of screening on health equity, impact of screening on quality of life, biases in the measurement of screening effectiveness, film versus digital mammography, impact of other screening modalities, and real-world resource allocation decision-making in the health care system.

Detailed discussions regarding the strengths and limitations for each of the analyses conducted can be found in the each of the studies' chapters in this dissertation.

Equity Considerations

One of the limitations of this research is the lack of analysis on the impact of screening on traditionally marginalized populations which would be consistent with traditional population health research (the inclusion of equity-focussed analyses). In contrast, vulnerability was characterized in this dissertation as sub-populations with elevated risk of breast cancer which does not necessarily link neatly to sociodemographic factors. However, there may be a muted equity issue in regards to breast cancer incidence. As noted in the introductory chapter of this dissertation, breast cancer risk appears to be higher among more advantaged female populations, including women of higher socio-economic status. However, disparity in breast cancer mortality has been linked to differences in accessibility and use or quality of health care services.

As part of this general discussion, a commentary regarding some of the equity issues that would be relevant to the conversation in regards to breast cancer screening is included. One issue that should be addressed is variability in timely access to health care services including screening and treatment which could affect the ultimate outcome of premature breast cancer mortality.

In terms of access to health care services and breast cancer screening in particular, there is some evidence that shows lower utilization of screening mammography in Canada among immigrant and lower income populations compared to the rest of the population. Glazier (2004) identified marked variation in mammography rates by area characterized by immigrant density, where lowest mammography use was found in areas with lowest income and highest immigrant density. The disparity in health and health care utilization has been well documented in the US. The main hypothesis being greater socio-economic gradients in that country compared to other OECD countries results in disparities in access to health services and therefore health outcomes. For example, Barnes et al. (2006) reported refugee women in the US do not receive adequate women's health care, including mammography, compared to the better-off counterparts. Access to preventive services may result in delayed treatment and increased mortality (Carrière et al, 2013). Access to treatment and other underlying conditions (comorbidity) in the individual are likely to be strongest drivers of disparity in mortality in vulnerable subgroups in Canada. For example, in Canada, life

expectancies among aboriginal peoples are consistently lower compared to other Canadians. They experience many social and health issues in addition to lack of access to health care services, including higher rates of diabetes and depression, comorbidity, and they are more likely to be socially and economically disadvantaged (Loppie and Wien, 2009).

There is a large body of Canadian research on the facilitators and barriers to screening among vulnerable populations as well as in the general population. For example, studies by Ahmad et al. (2012; 2013) identified eight clusters of barriers preventing women of South Asian ethnicity from participating in screening mammography including: 1) dependence on family, 2) ease of access to a mammography centre; 3) language and transportation; 4) access to a physician (including preferences of type of provider); 5) fear; 6) reliance on self-care; 7) cultural beliefs and practices; and 8) knowledge. These barriers may be generalizable to other ethnic minority groups. Schoueri-Mychasiw et al. (2013) noted that the levels of in breast cancer screening uptake by ethnicity is variable, and that immigrant groups are not homogenous, which may speak to the other stronger predictors of screening uptake including one or more of the eight barriers identified by Ahmad et al. (2012;2013). Research by Vahabi (2011 a, 2011 b) support these findings, noting that interventions to promote breast cancer screening must be culturally sensitive, arguing that much of the existing cancer-related information was developed primarily for Caucasian women and fails to address other ethno-cultural and religious beliefs. In the review by Hanson et al. (2009), it was reported that in the general female Canadian population, barriers to screening mammography include: past and present health actions, socio-economic status, personal attributes, limited physician access, lack of a screening recommendation, lack of health care provider support, an intimidating relationship with the health care provider, limited social support, and fearful descriptions of breast cancer in the media. This indicates that the major differences between the barriers experienced by vulnerable populations and those experienced by the general population are knowledge and socio-economic or socio-cultural factors.

There is also literature that shows some differences in ethnicity in the distribution of breast cancer histological types, regardless of country of residence (Carrière, 2013). A recognized example is the fact that there is a greater incidence of basal type breast cancer within African-American and West African populations (Carrière, 2013). It is seen that in general,

women of visible minority have a lower incidence of breast cancer compared to Caucasian women; however they also have a higher prevalence of advanced breast cancer, poorer five-year survival, and higher rates of breast cancer mortality (Vahabi, 2010). Luo et al. (2004) noted that risk of cancer of the immigrant tends to match the population risk of cancer of the new country – specifically, it was found that Canadian-born Chinese individuals living in Alberta experienced higher all cancer incidence rates than comparable populations of Chinese immigrants living in Canada, which had rates that appeared more closely matching cancer incidence rates of Chinese individuals living in Shanghai. This observation may be explained by variation in environmental factors that contribute to the onset of disease including environmental exposures and diet, and it may also be predictive of future increases in rates of cancer among ethnic minority populations as they become more assimilated to Canadian culture.

Yavari et al. (2010) noted that in parts of the world where healthcare is socialized, economic status does not have an effect on treatment use, and also, because Canadian residents receive universal health care, economic status should not have an effect on treatment. However, in parts of the world where healthcare is privatized, there have been documented disparities in treatment due to economic status. Specifically, It has been reported that breast cancer patients in the United States that were of African American, Mexican and Puerto Rican descent were less likely to receive the standard of care, whereas breast cancer patients in the United States that were of Asian and Pacific Islanders descent were more likely to receive the gold standard of care. In Canada, disparities in health care utilization across different ethnicities were also noted although it is uncertain whether variability in use is related to socioeconomic or sociocultural disadvantage or to other factors. According to records of the British Columbia (BC) Cancer Registry from three periods spanning the years 1980 - 2006, it was observed that women of Chinese descent residing in BC with stage I or II breast cancer were significantly less likely to receive surgery with radiation, and were significantly more likely to receive surgery without radiation, when indicated (Hislop, 2007). A later finding indicated that Asian women in general, with ductal carcinoma in situ of the breast, were more likely to undergo mastectomy rather than lumpectomy, when indicated (Yavari, 2010). It is observed that differences in treatment preferences could perhaps stem from cultural preferences or personal preferences – the latter of which is shaped by the

individual's personal knowledge base, rather than socioeconomic/sociocultural disadvantage. One of the key messages from the current body of research is that it is critical that all breast cancer patients are well-informed about treatment practices that recognized as the best practice, as well as the benefits and harms of varying procedures and treatments for their conditions.

Carrière (2013) noted that an individual's breast cancer mortality risk can vary depending on their birth country as well as time lived post-immigration in the new country. Ethnic and visible minority status has been associated with poorer survival compared with their counterparts (Luckett, 2011). A systematic review and meta-analysis on the topic of psychological morbidity and quality of life of ethnic and visible minority populations with cancer conducted in the US noted that Hispanic patients reported significantly worse distress, depression, and health-related quality of life than majority patients (Luckett,2011). From a Canadian perspective, a recent study by Tjepkerma et al. (2013) aimed to assess cause-specific mortality rates by income adequacy among Canadian adults using a large, population-based sample with income information linked to almost 16 years of mortality data. This study found no mortality gradient by income for breast cancer. However, as mentioned previously, despite higher rates of breast cancer in women of higher socio-economic status, it appears that women of lower socio-economic status and of other vulnerable subgroups who have breast cancer are more likely to be of advanced disease at the time of diagnosis and treatment (Vahabi, 2010). Therefore, although there may be no difference in mortality, there may be differences in quality of life as well as treatment costs in relation to the differences in stage distributions between higher and lower socio-economic status populations.

It has been reported that it is often difficult to differentiate whether low screening participation rates are attributed to low socio-economic status or high volume of immigration. As recent immigrants tend to settle in low socio-economic neighbourhood it is often difficult to disentangle the impact of each (Borkhoff et al, 2013; Lofters et al, 2013; Glazier et al, 2004). In the studies by Tammemagi et al. (2007) and Yavari et al. (2010), and Hislop (2007), ethnicity status was determined by individuals' surnames, which may be subject to misclassification bias, which adds to the difficulties in identifying important sociodemographic risk factors for utilization or access to services. Many of the Canadian

studies examining ethnic disparities in screening uptake are limited to Ontario and BC. Because healthcare is regionally (provinces and territories) funded, governed, and administered, similar studies based in other provinces and territories should be conducted or conducted on a more national scale would contribute to the understanding of ethnic and socio-economic disparities, if any, in screening uptake. Further, certain ethnic/visible minority groups are underrepresented in the literature. It was noted that many of the articles focused on South Asian (especially those articles based in Ontario) or Asian (especially those articles based in British Columbia), which overtime may be less considered as ethnic minorities or disadvantage d populations. Certinaly, the experiences of these particular populations may not be the generalizable to the experiences of other ethnic minority groups in Canada.

In assessing urban versus rural access to mammography, it has been reported that women residing in rural areas of Canada are significantly less likely to undergo screening mammography compared with women residing in census metropolitan areas, even after adjusting for confounding factors including having a family doctor, education, family income, marital status, and age (odds ratio = 0.58) (McDonald and Sherman, 2010). It was hypothesized that one of the important factors that could explain this result are differences in attitude about the importance of screening as opposed to accessibility. Based on a survey of participants on reasons for not having undergone mammography, barriers to timely access was the least frequently chosen reason across all regions although only slightly more frequently selected among rural regions compared with more urban regions. Only 1 in 20 women in the most rural regions identified access barriers as the reason for not getting a mammogram. In the same study, the researchers found that other socio-economic factors were significantly related to lower mammography utilization including being a recent immigrant, not being able to speak English or French, and not having a family doctor.

As noted earlier, the effectiveness and cost-effectiveness of screening mammography on mortality show that the benefits of screening are not as large as they were shown previously in the older randomized controlled trials. There is a lack of understanding whether there is an equity dimension to these types of analyses as there are very few effectiveness and cost-effectiveness studies providing stratified analysis by vulnerable population as stratification is

often reserved for groups with high-risk for breast cancer. Future studies should include an equity lens to effectiveness and cost-effectiveness analyses of screening.

Quality of Life

Quality of life has been defined as "an individual's perception of their position in life in the context of the culture and value system in which they live and in relation to their goals, expectations and standards and concerns." (WHO, 1995) It is a broad concept that is highly, subjective and sensitive to multiple factors including an individual's physical health, psychological state, and level of independence, social relationships, and their relationship to characteristics of their environment. One of the limitations of the research in this dissertation is that there was no assessment of the impact of screening on health-related quality of life. If the impact of screening mammography on mortality is muted due to higher survivability of late-stage breast cancer cases, it could be that screening may have a benefit on health-related quality of life via stage shifting, assuming that the screening test itself poses little to no quality-of-life detriment and that there is a quality-of-life improvement in early stages of breast cancer. It was previously noted that disparities in the impact of screening, if any, in the general female population compared to sub-populations including women at higher-risk of disease as well as more vulnerable populations, may be due to differences in quality of life in addition to or as opposed to mortality. Having a health-related quality of life component added to the POHEM-BCS would allow the re-weighting of LYGs and therefore allow for the estimation of the relative differences in health-adjusted life-years across different screening policy options.

Offsetting Harms with Benefits of Screening

Breast cancer screening is assumed to improve breast cancer survival (or reduce breast cancer mortality) through early detection and treatment, resulting in improved prognosis. A review by the Independent UK Panel on Breast Cancer Screening examined the benefits of breast cancer screening in reducing mortality were weighted against the risk of over-diagnosis – higher rates of screen-detected cancers that would not become clinically apparent in the woman's lifetime in the absence of screening and/or which would not lead to death

(Marmot et al., 2012). The panel reported the best evidence for the relative benefit of screening on breast cancer mortality is a meta-analysis of 11 randomized controlled trials with 13 years of follow-up. It was estimated that a 20 per cent reduction in breast cancer mortality was attributable to screening based on the review. However, several concerns regarding the methodology that may compromise the validity of the estimates were reported by the review panel. Firstly, the 95 per cent confidence interval around the point estimate was relatively wide, ranging from 11 per cent to 27 per cent. Secondly, bias may be inherent due to suboptimal randomization and the lack of adjustment in the cause of the death in the individual studies. Thirdly, the panel notes that the trials used in the meta-analysis are old and therefore may not reflect changes over time that may affect the impact of screening on breast cancer mortality. The trials in the meta-analysis are 2 to 3 decades old and because the panel did not elect to include observational studies, which make up all modern studies evaluating the effectiveness of screening, none of this more current data were included in their review. Finally, there was observed quite a bit of variance around the estimate of absolute mortality benefit from screening. This variance ranges from one breast cancer death prevented for every 2000 women screened to one breast cancer death prevented for every 100 woman screened across studies –a 20-fold difference. Based on a relative risk reduction of 20 per cent, the panel estimated for every one breast cancer death prevented for every 235 women invited for screening or one breast cancer death prevented for every 180 women actually screened. In aggregate, there is an estimated 1,300 breast cancer deaths prevented every year due to screening in the UK or approximately 22,000 life-years saved.

In terms of over-diagnosis, the panel aimed to estimate the magnitude of over-diagnosis in breast cancer screening. Citing few sources of reliable data, the panel reported not being able to directly address this objective and instead reported provisional estimates instead. They estimated over-diagnosis using two different denominators and one numerator. These were assessed using two different perspectives: 1) population perspective: the number of breast cancers, both invasive and ductal carcinoma in situ (DCIS), diagnosed throughout the rest of a woman's lifetime after the age that screening begins; 2) perspective of a woman invited to screening: using the total number of breast cancers diagnosed during the screening period. The panel used data from three randomized controlled trails that did not systematically screen the control group at the end of the screening period and followed these women for

several more years. They estimated that the frequency of over-diagnosis in breast cancer screening was of the order of 11 per cent from population perspective, and about 19 per cent from the perspective of a woman invited to screening. Again, by limiting the estimate of over-diagnosis to data from only randomized controlled trials resulted in the use of older data. A limitation cited was that it is not currently possible to distinguish whether cases of over-diagnosis are actually screened cancer that would lead to death or whether they are cases that simply would not have come to attention if there was no breast cancer screening. Since screening is already disseminated in the population, it is difficult to assess reliable estimates of over-diagnosis without conducting a new randomized controlled trial. The negative consequences of over-diagnosis include women being unnecessarily treated for breast cancer and the related consequences including the detrimental impacts on quality of life, psychological wellbeing, productivity, and public and private costs. The panel however concluded that the potential harms of screening are minimal compared to the mortality benefit, therefore screening should continue despite uncertainty around the true benefit.

In general, women do not have a clear understanding of whether breast cancer screening will be more or less beneficial or harmful to them. The messages that women receive including from their physicians, can vary greatly. Decision-aids have been developed to help women decide whether and when they should be screened for breast cancer, however physician recommendations still remain the greatest influence on a woman's decision to participate in screening mammography. At the individual level, some women feel that the benefits of breast cancer screening are worth the risks (Hersch et al., 2013), however there is some research showing that women may not be adequately aware of the actual benefits and risks of screening (Domenighetti et al, 2003; Chamot et al., 2005). The study by Domenighetti et al. (2003) reported that in US and European countries a high proportion of surveyed women overestimated the benefits that can be expected from screening mammography. Over the 10 plus years since the publication of the study, it is expected that women would be more informed however more research is required in this area to make this conclusion.

At the population-level, concerns have been raised in regards to whether it is appropriate to place so many resources on mass screening with mammography (age-based screening) on the basis of the minimal benefit at some level of risk for some women. One of the challenges for governments is to decide whether and by how much screening programs should be scaled

back given the strong support for routine screening among breast cancer by the general population. A recently published article discussed the debate in Switzerland over its breast cancer screening program (Arie, 2014). This article discussed the controversial recommendation by the Swiss Medical Board to suspend the country's screening mammography programme because it was leading to too many unnecessary tests and treatments. The Board made this recommendation in February 2014, noting that while systematic screening mammography for breast cancer saves 1 to 2 women's lives for every 1,000 women screened, it leads to unnecessary investigations and treatment for about 100 women in every 1000 screened (Swiss Medical Board, 2013). In the same month, another Canadian study by Miller et al. (2014) reported a muted benefit of screening based on 25-year follow-up data of the Canadian National Breast Screening Study. This study compared breast cancer incidence and mortality up to 25 years in women age 40 to 59 years who did or did not participate in screening mammography using follow-up randomized screening trial data using data linkage to cancer registries and vital statistics data. They compared breast cancer incidence and mortality rates among women who had five annual mammography screens versus women who did not participate in mammography, with both groups receiving physical examination every year. The study found that annual mammography in this age group did not reduce mortality from breast cancer and that the benefits to mortality are explained through physical examination or usual care when adjuvant treatment is highly accessible. Further, they estimated that 22 per cent of screen-detected invasive breast cancers were cases of over-diagnosis that would not have become cancer if left undetected. This represented 1 case of over-diagnosis for every 424 women screened. Both the Swiss Board recommendations and the Canadian study received their share of backlash from breast cancer groups (research, lobby, health care providers, etc.) who maintain the benefits of screening still outweigh the harms.

Advancements in Technologies and Treatments

The studies in this dissertation were limited to the analog screening mammography as the intervention being modeled in POHEM-BCS in the context of the screening test sensitivity and specificity. At the start of the main study (chapter four), the majority of screening data being collected was for analog mammography – the printing of x-ray images of the breast on

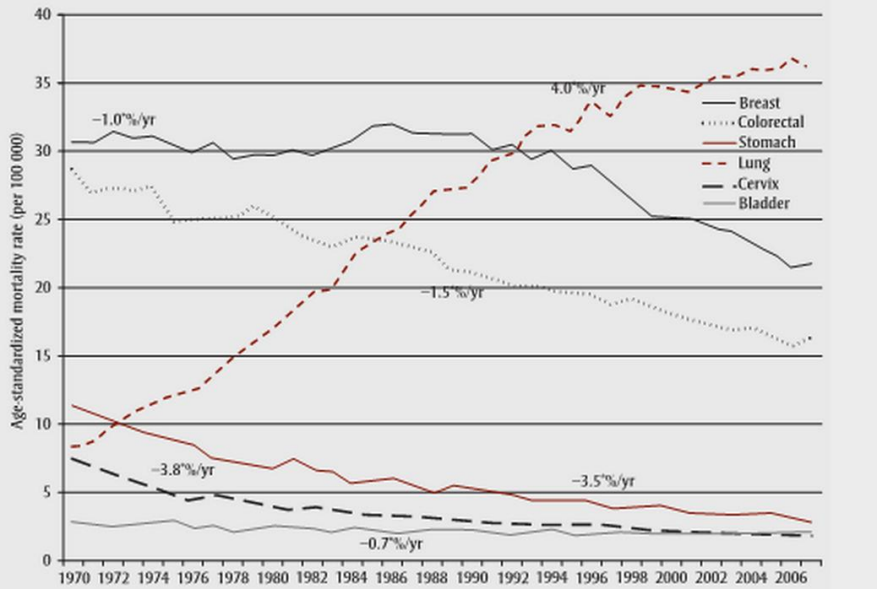
film. Over the past decade, digital mammography has been more widely used across the country and is now considered the standard technology used in screening mammography. Digital mammography is a much more sensitive than analog at detecting cancers among women with high breast density; it is also much more efficient, allowing for the generation of multiple high-quality images of the breast in a short period of time. There was an attempt to adjust sensitivity and specificity of mammography in the sensitivity analysis in chapter four to account for advancements in technology. In a study by Pisano et al. (2008) digital mammography was twice as sensitive than analog in detecting breast cancers in women with high breast density (0.59 vs 0.27). The Canadian Agency for Drugs and Technologies in Health (CADTH) reported that although digital mammography is more sensitive for younger women (under age 50), who have dense breasts, and who are pre- or peri-menopausal, the effectiveness of digital mammography in detecting cancers in terms of overall accuracy is similar to analog mammography (CADTH, 2006). They also note that the costs of digital mammography are much higher than analog, however this may be due to high initial/acquisition costs and that over time, costs should level.

Currently, many provinces have already transitioned fully to digital mammography with some provinces also using breast ultrasound in combination with mammography or magnetic resonance imaging (MRI) in certain cases. MRI is considered a newer technology for the use of detecting smaller, earlier breast cancers in higher-risk women with denser breasts. It is however not considered to replace mammography. There are currently randomized controlled trials taking place to assess the effectiveness of MRI in the detection of breast cancer. Future technologies may include digital breast tomosynthesis, breast computed tomography, and biomarker imaging (Canadian Breast Cancer Foundation, 2014). There has been a very gradual shift towards using technology that is more sensitive and efficient in examining the breast for abnormalities. Coupled with higher reading volumes for radiologists, overall sensitivity of screening programs to detect cancers has improved resulting in perhaps a large increase in the number of incident breast cancer cases over the past few decades.

In a recent study by Kachur et al. (2013), observed trends in age-standardized breast cancer mortality in Canada between 1970 and 2007 showed statistically significant decreases over time (see figure 1). According to Surveillance, Epidemiology, and End Results (SEER) data,

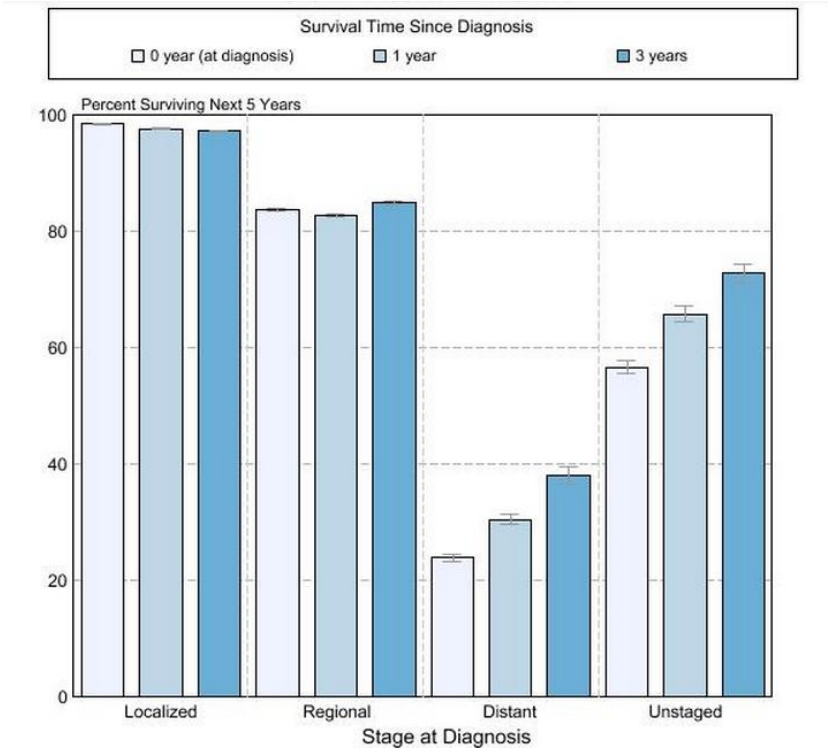
the 5-year relative survival in the general female population in the US is 99% for localized disease (stage I), 85 per cent for regional disease, and 24 percent for distant stage disease (Howlader et al., 2014). The same data show that conditional relative survival is significantly higher among the distant breast cancer cases representing increased 5-year survival among women diagnosed with late stage cancer after one and three years of having already survived with the disease (see figure 2). Part of the improvements in breast cancer cannot be solely explained by screening since it is known that treatments have also become more effective and accessible over time. As mentioned earlier in the review of the CISNET microsimulation models (chapter two) for breast cancer screening, treatment has been established to a significant role, and potentially more important role than screening, in determining breast cancer survival.

Figure 19 Average annual percent change in age-standardized mortality rates for selected cancers in females, Canada, 1970–2007. (Excludes Quebec)



Source: Kachur et al. (2013).

Figure 20 Cancer of the female breast, 5-year SEER conditional relative survival and 95 per cent confidence intervals by stage, U.S., 1998-2010.



Source: Howlander et al. (2014)

Making Resource Allocation Decisions in Health Care: A Population Health Issue

Making a case to adjusting screening programs or to dismantle population-based screening all together is difficult to make when individual and population-level health impacts, both beneficial and harmful, are so marginal. What is not so marginal is cost. Despite the high costs of screening and unnecessary treatment of those over-diagnosis cases, it is still not enough to scaling back screening to the public and physicians who may believe that a price tag cannot be put on a life. However, regional governments are always faced with the challenge of balancing budgets with higher health care costs and lower revenues – which may be attributable to an aging population and lower health care transfers from the federal government. Over the past few years economists have warned provincial and territorial

governments to reduce health care spending or raise revenues (through taxes) or else continue to fund health care at the expense of cutting funding to other social programs including education and housing, economic initiatives such as job creation and pensions, and infrastructure. This dilemma shows how resource allocation decision-making has important impacts on population health. The research in this dissertation aimed to address questions of which screening program is optimal in terms of age to screen and how often to screen? The real difficult questions that governments and administrators face is how to pay for all the programs that are currently in the big basket of covered health care services and programs for Canadians, which new programs should be covered, and which should be scaled back or eliminated. This applies to not only health care services and treatments, but also to public health programs.

Implications for Policy

Analysis conducted at Statistics Canada found that increases in breast cancer prevalence proportions were primarily due to population aging. The trends in age-standardized prevalence of breast cancer were not found to be significant over time (Ellison and Wilkins, 2012). These data show trends in breast cancer incidence and prevalence can be better predicted by using forecasts of age-structure. As a result, the prevalence of those disease risk factors that were discussed in chapter five may not drive disease rates in the future as much as population aging.

In real terms, it is not clear how much screening mammography costs in Canada. It was reported in the U.S. that $8 billion a year in health care expenditures are for mammography. The U.S. also happens to be one of the countries where, despite national guidelines to the contrary, in practice annual screening has been highly encouraged for the age 40 and older population while most other countries, including Canada recommend screening to start age 50 years and at less frequent intervals (screening every 2 to 3 years) (Elmore and Kramer, 2014). The cost-effectiveness analysis study in this dissertation (chapter four) estimated total cost differences across the examined screening policy options (characterized by age eligibility and screening frequency) were driven solely by screening costs as diagnosis and treatment costs across policy options were not substantially different across these options.

Given survival gains across policy options are also modest across screening policy options; governments may find greater cost-savings by re-evaluating screening policies and programs and making the appropriate policy changes.

The cost-effectiveness analyses in this dissertation estimated incremental cost-effectiveness ratios to compare screening policies against each other, but this metric may not be so meaningful for policy makers. A WTP threshold for the payer needs to be established in order to identify an optimal policy option. Further, in reality, the payer is required to compare multiple and often unrelated options to make resource allocation or rather resource de-allocation decisions. As mentioned previously, sometimes the decision is not around the optimal screening scenario but rather how much more or less money should be allocated to specific programs including cancer screening, drug coverage, and seniors care, to name a few. Although effectiveness and cost-effectiveness data and evidence are important in helping decision-makers make these resource allocation decisions, other factors including competing (sector) budgets, political factors, and public perceptions, are also strong influencers on these decisions.

References

Advisory Committee on Breast Cancer Screening. (2006). *Screening for breast cancer in England: past and future.* NHSBSP Publication no.61. Sheffield, UK: NHS Cancer Screening Programmes.

Ahmad, F., Jandu, B., Albagli, A., Angus, J. E., & Ginsburg, O. (2013). Exploring ways to overcome barriers to mammography uptake and retention among South Asian immigrant women. *Health & social care in the community*, *21*(1), 88-97.

Ahmad, F., Mahmood, S., Pietkiewicz, I., McDonald, L., & Ginsburg, O. (2012). Concept mapping with South Asian immigrant women: barriers to mammography and solutions. *Journal of Immigrant and Minority Health, 14*(2), 242-250.

Anderson, R. 2010. Systematic reviews of economic evaluations: Utility or Futility? *Health Econ., 19* (3), 350-364.

Armstrong, K., Moye, E., Williams, S., Berlin, J.A., & Reynolds, E.E. (2007). Screening mammography in women 40 to 49 years of age: a systematic review for the American College of Physicians. *Annals of Internal Medicine, 146*(7), 516-526.

Arozullah, A.M., Calhoun, E.A., Wolf, M., et al. (2004). The financial burden of cancer: estimates from a study of insured women with breast cancer. *Journal of Supportive Oncology, 2*(3), 271-278.

Arveux, P., & Schaffer, P. (2013). Building a model to determine the cost-effectiveness of breast cancer screening in France. *European Journal of Cancer Care, 12*, 143-153.

Barnes, D. M., & Harrison, C. L. (2004). Refugee women's reproductive health in early resettlement. *Journal of Obstetric, Gynecologic, & Neonatal Nursing, 33*(6), 723-728.

BC Cancer Agency. (2013). *Screening Mammography Program 2013 Annual Report.* Retrieved March 28, 2014, from http://www.screeningbc.ca/NR/rdonles/8CD1608D-BE23-41EC-A5E6-8ADE5119F6E4/68074/SMPAnnualReport2013Rev14January2014.pdf

Bell, C.M., Urbach, D.R., Ray, J.G., et al. (2006). Bias in published cost effectiveness studies: systematic review. *British Medical Journal, 332,* 699-703.

Berthelot, J.M., Will, B.P., Evans, W.K., Coyle, D., Earle, C.C., & Bordeleau, L. (2000). Decision framework for chemotherapeutic interventions for metastatic non-small cell lung cancer. *Journal of the National Cancer Institute, 92*, 1321-1329.

Boer, R., Plevritis, S., & Clarke, L. (2004). Diversity of model approaches for breast cancer screening: A review of model assumptions by the cancer intervention and surveillance

network (CISNET) breast cancer groups. *Statistical Methods in Medical Research, 13*, 25-538.

Bonomi, A.E., Boudreau, D.M., Fishman, P.A., Ludman, E., Mohelnitzky, A., Cannon, E.A., & Seger D. (2008). Quality of life valuations of screening mammography. *Quality of Life Research, 17*(5), 801-814.

Borkhoff, C. M., Saskin, R., Rabeneck, L., Baxter, N. N., Liu, Y., Tinmouth, J., & Paszat, L. F. (2013). Disparities in Receipt of Screening Tests for Cancer, Diabetes and High Cholesterol in Ontario, Canada: A Population-based Study Using Area-based Methods. *Canadian Journal of Public Health*, *104*(4).

Boyd, N.F., Martin, L.J., Yaffe, M.J., et al. (2006). Mammographic density: a hormonally responsive risk factor for breast cancer. *J Br Menopause Soc, 12*(4), 186–193.

Braithwaite, R.S., Meltzer, D.O., King, J.T., Leslie, D., & Roberts, M.S. (2008). What does the value of modern medicine say about the $50,000 per quality-adjusted life-year decision rule? *Med Care, 46,* 349 –356.

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. *Journal of Medical Screening,19,*14-25.

Brown, M.L., & Fintor, L. (1993). Cost-effectiveness of breast cancer screening: preliminary results of a systematic review of the literature. *Breast Cancer Research and Treatment, 25*(2), 113-118.

Burger, E.A., Sy, S., Nygård, M., Kristiansen, I.S., & Kim, J.J. (2014). Prevention of HPV-Related Cancers in Norway: Cost-Effectiveness of Expanding the HPV Vaccination Program to Include Pre-Adolescent Boys. *PLOS ONE 9*(3), e89974. doi:10.1371/journal.pone.0089974.

Byrne, C., & Spernak, S. (2005). What is Breast Density? *Breast Cancer Online, 8,* e51.

Cady, B. (1998). How to prevent invasive breast cancer: detect and excise duct carcinoma in situ. *Journal of Surgical Oncology, 69,* 60–62.

Canadian Agency for Drugs and Technologies in Health. (2006) *Guidelines for the economic evaluation of health technologies*: Canada [3rd Edition]. Ottawa. Retrieved March 17, 2014 from http://www.inahta.org/upload/HTA_resources/AboutHTA_Guidelines_for_the_Economic_Evaluation_of_Health_Technologies.pdf

Canadian Cancer Society. (2014). *Hormone receptor status testing*. Retrieved May 3, 2015 from http://www.cancer.ca/en/cancer-information/diagnosis-and-treatment/tests-and-procedures/hormone-receptor-testing/?region=on

Canadian Cancer Society. (2014). *Screening for Breast Cancer*. Retrieved March 30, 2014 from http://www.cancer.ca/en/cancer-information/cancer-type/breast/screening/?region=on.

Canadian Cancer Society. (2014). *Benefits and risks of screening for breast cancer*. Retrieved November 19, 2014 from http://www.cancer.ca/en/cancer-information/cancer-type/breast/screening/?region=on#Benefits

Canadian Cancer Society. (2014). *Staging and grading for breast cancer*. Retrieved November 19, 2014 from http://www.cancer.ca/en/cancer-information/cancer-type/breast/staging/?region=on

Canadian Cancer Society. (2014). *What is breast cancer?* Retrieved November 19, 2014 from http://www.cancer.ca/en/cancer-information/cancer-type/breast/breast-cancer/?region=on

Canadian Cancer Statistics. (2014). *Canadian Cancer Society*. Retrieved on October 5, 2014 from: http://www.cancer.ca/

Canadian Partnership Against Cancer. (2003). *Organized Breast Cancer Screening Programs in Canada: Report on Program Performance in 2007 and 2008*. Toronto: Canadian Partnership Against Cancer.

Canadian Task Force on Preventive Health Care. (2011). *Screening for Breast Cancer*. Retrieved on August 23, 2013 from http://canadiantaskforce.ca/guidelines/2011-breast-cancer/.

Cancer Care Ontario. (2004). *First Nations cancer research and surveillance priorities for Canada: Workshop Report September 23-24, 2003*. Ottawa: Cancer Care Ontario.

Cancer Intervention and Surveillance Modeling Network (CISNET) Breast Cancer Collaborators. (2006). The Impact of Mammography and Adjuvant Therapy on U.S. Breast Cancer Mortality (1975-2000): Collective Results from the Cancer Intervention and Surveillance Modeling Network. *Journal of the National Cancer Institute Monographs, 36,* 1-126.

Cancer Quality Council of Ontario. *Reporting of Cancer Stage at Diagnosis*. February 2013. Retrieved on March 25, 2014 from http://www.csqi.on.ca/by_patient_journey/diagnosis/reporting_of_cancer_stage_at_diagnosis/#.UzHQpvldWSo

Carrière, G. M., Sanmartin, C., Bryant, H., & Lockwood, G. (2013). Rates of Cancer Incidence Across Terciles of the Foreign-born Population in Canada From 2001-2006. *Canadian Journal of Public Health, 104*(7), e443-e449.

Carroll J.C., Cremin C., Allanson J., et al. (2008). Hereditary breast and ovarian cancers. *Canadian Family Physician, 54*(12), 1691-1692.

Chakraborty R., Little M.P., & Sankaranarayanan K. (1998). Cancer predisposition, radiosensitivity and the risk of radiation-induced cancers. IV. prediction of risks in relatives of cancer-predisposed individuals. *Radiation Research, 149,* 493-507.

Chamot, E., et al. (2005). Variability in women's desire for information about mammography screening: implications for informed consent. *European Journal of Cancer Prevention, 14*(4), 413-418.

Champion, V. Health Behaviour Constructs: Theory, Measurement, & Research. Cancer Control and Population Sciences. National Cancer Institute.

Chang, P. (1993). *A Simulation Study of Breast Cancer Epidemiology and Detection since 1982: The Case for Limited Malignant Potential Lesions* [Ph.D.]

Chart, P.L., & Franssen, E. (1997). Management of women at increased risk for breast cancer: Preliminary results from a new program. *Canadian Medical Association Journal, 157*, 1235-1242.

Chlebowski, R.T., Chen, Z., Anderson, G.L., et al. Ethnicity and breast cancer: Factors influencing differences in incidence and outcome. *Journal of the National Cancer Institute, 97*, 439-448.

Chopra, I., & Kamal, K.M. (2012). A systematic review of quality of life instruments in long-term breast cancer survivors. *Health and Quality of Life Outcomes, 10*, 14.

Cicchetti, D.V., & Conn, H.O. (1976). A statistical analysis of reviewer agreement and bias in evaluating medical abstracts. *The Yale Journal of Biology and Medicine, 49*, 373-383.

CISNET Breast Cancer Model Profiles. (2013). Retrieved May 1, 2014 from http://cisnet.cancer.gov/breast/profiles.html

Clavel-Chapelon, F. (2002). E3N-EPIC Group. Differential effects of reproductive factors on the risk of pre- and postmenopausal breast cancer: Results from a large cohort of French women. *British Journal of Cancer, 86*, 723–727.

Collaborative Group on Hormonal Factors in Breast Cancer. (2002). Breast cancer and breastfeeding: Collaborative reanalysis of individual data from 47 epidemiological studies in

30 countries, including 50302 women with breast cancer and 96973 women without the disease. *The Lancet, 360,* 187-195.

Collaborative Group on Hormonal Factors in Breast Cancer. (2001). Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 without the disease. *The Lancet, 358,* 1389-1399.

De Gelder, R., Bulliard, J-L., de Wolf, C., et al. (2009). Cost-effectiveness of opportunistic versus organised screening mammography in Switzerland. *European Journal of Cancer, 45,* 127-138.

Dinh, N-T., Brand. K., Coyle, D., Morrison, H., Deri-Armstrong, C., & Kasmani, A. (n.d.). Impact of screening age and frequency on the cost-effectiveness of screening mammography in Canada, 2014. [Unpublished study]

Dinh, N-T, Brand, K., Coyle, D., Morrison, H., & Deri-Armstrong, C. (n.d.). A systematic review of the cost-effectiveness of Screening Mammography, 2014. [Unpublished study]

Domchek, S.M., & Weber, B.L. (2002). Recent advances in breast cancer biology. *Current Opinion in Oncology, 14,* 589-593.

Domenighetti, G., et al. (2003). Women's perception of the benefits of mammography screening: population-based survey in four countries. *International Journal of Epidemiology, 32*(5), 816-821.

Drummond, M.F., Sculpher, M., Torrance, G.W., O'Brien, B.J., & Stoddart, G.L. (2005). *Methods for the Economic Evaluation of Health Care Programmes, 3rd edition*. Oxford: Oxford University Press.

Duffy, S.W., Tabar, L., & Smith, R.A. (n.d.). Screening for breast cancer with mammography. *The Lancet, 358,* 2166-2166.

Duffy, S.W. (2005). Screening, Sojourn Time. In *Encyclopedia of Biostatistics*. Volume 7. John Wiley & Sons, Ltd.

Dupont, W., Parl, F., Hartmann, W., et al. (1993). Breast cancer risk associated with proliferative breast disease and atypical hyperplasia. *Cancer, 71,* 1258-1265.

Earle, C., Coyle, D., Wells, G., & Papadimitropoulos, E. (1999). A needs based assessment of breast cancer in Canada: the economic burden of illness. *Annu Meet Int Soc Technol Assess Health Care Int Soc Technol Assess Health Care Meet, 15,* 87.

Eberl, M.M., Fox, H.F., Edge, S.B., et al. (2006). BIRADS Classification for Management of Abnormal Mammograms. *American Board of Family Medicine, 19*(2), 161-164.

Ellison, L., & Haber, D. (1998). Hereditary breast cancer. *Annual Review of Medicine, 49,* 425-436.

Ellison, L.F., & Wilkins, K. (2012). Health Reports, Vol. 23, no. 1, March 2012. Statistics Canada, Catalogue no. 82-003-XPE.

Elmore, J.G., Kramer, B.S. (2014). Breast cancer screening: Toward informed decisions. *Journal of the American Medical Association, 311*(13), 1298-1299.

Endogenous Hormones and Breast Cancer Collaborative Group. (2002). Endogenous sex hormones and breast cancer in postmenopausal women: Reanalysis of nine prospective studies. *Journal of the National Cancer Institute, 94,* 606-616.

Endogenous Hormones and Breast Cancer Collaborative Group. (2003). Body mass index, serum sex hormones, and breast cancer risk in postmenopausal women. *Journal of the National Cancer Institute, 95,* 1218-1226.

Espinoza, M.A., Manca, A., Claxton, K., & Sculpher, M.J. (2012). The value of heterogeneity (VoH) for cost-effectiveness subgroup analysis: Theoretical framework and application. Centre for Health Economics, University of York. Retrieved February 13, 2012 from http://www.ispor.org/awards/14euro/HG4-THE-VALUE-OF-HETEROGENEITY-FOR-COST-EFFECTIVENESS-SUBGROUP-ANALYSIS.pdf..

Flanagan, W., Le Petit, C., Berthelot, J-M., Gaudette, L., Gao, R-N., & Onysko, J. (2006) Evaluating the impact of breast cancer screening in Canada using POHEM – A demonstration project. (Unpublished)

Flanagan, W., Le Petit, C., Berthelot, J-M., White, K.J., Coombs, B.A., & Jones-McLean, E. (2003). Potential impact of population-based colorectal cancer screening in Canada. *Chronic Disease Canada, 24,* 81-88.

Flanagan, W., Manuel, D., Nadeau, C., Oderkirk, J., Tremblay, M.S., Wall, R., & Wong, S. (2013). Development of a population-based microsimulation model of physical activity in Canada. Statistics Canada, Catalogue no. 82-003-X, *Health Reports, 24*(10), 11-19.

Fletcher, S.W., Black, W., Harris, R., Rimer, B.K., & Shapiro, S. (1993). Report of the international workshop on screening for breast cancer. *Journal of the National Cancer Institute, 85,* 1644-1656.

Froom, P., Melamed, S., Kristal-Boneh, E., et al. (1999). Healthy volunteer effect in industrial workers. *Journal of Clinical Epidemiology, 52,* 731–735.

Gaudette, L.A., Altmayer, C.A., Nobrega, & K.M., Lee, J. (1996). Trends in mammography utilization, 1981 to 1994. *Health Reports, 8,* 17-27(Eng); 19-30(Fre).

Given-Wilson, R., Layer, G., Warren, M., & Gazettn, J.-C. (1997). False negative mammography: causes and consequences. *The Breast, 6*(6), 361-366.

Glazier, R.H., Creatore, M.I., Gozdyra, F.I., Steele, L.S., & Boyle, E. (2004) Geographic methods for understanding and responding to disparities in mammography use in Toronto, Canada. *Journal of General Internal Medicine, 19*(9), 952–961.

Gordon, L., Scuffham, P., Hayes, S., & Newman, B. (2007). Exploring the economic impact of breast cancers during the 18 months following diagnosis. *Psychooncology, 16* (12), 1130-1139.

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

Greene, M. (1997). Genetics of breast cancer. *Mayo Clinic Proceedings, 72*, 54-65.

Grima, D.T., Brown, S.T., Bainey, K.R., Goeree, R., Oh, P., Ramanathan, K., & Goodman, S.G. (2014). Cost-effectiveness of ticagrelor versus clopidogrel in patients with acute coronary syndromes in Canada. *ClinicoEconomics and Outcomes Research, 6*, 49–62. Retrieved April 1, 2014 from http://www.dovepress.com/getfile.php?fileID=18802

Güneş, E.D., Chick, S.E., & Akşin, O.Z. (2004). Breast cancer screening services: trade-offs in quality, capacity, outreach, and centralization. *Health Care Management Science, 7*(4), 291-303.

Habbema, J.D., Van Oortmarssen, G.J., Lubbe, J.T., & Van der Maas, P.J. (1985). The MISCAN simulation program for the evaluation of screening for disease. *Computer Methods and Programs in Biomedicine, 20*(1):79-93.

Hailey, D. (2006). *Digital mammography: an update [Issues in emerging health technologies issue 91]*. Ottawa: Canadian Agency for Drugs and Technologies in Health.

Hall, E.J. (2000). Radiation, the two-edged sword: Cancer risks at high and low doses. *The Cancer Journal, 6,* 343-350.

Hamilton, N., & Bhatti, T. (1995). *Population Health Promotion: An Integrated Model of Population Health and Health Promotion*. Working Paper, Health Promotion Development Division, Health Canada.

Hankinson, S., Colditz, G., & Willett, W. (2004). Towards an integrated model for breast cancer etiology: The lifelong interplay of genes, lifestyle, and hormones. *Breast Cancer Research, 6*, 213-218. Retrieved from: http://breast-cancer-research.com/content/6/5/213.

Harris, R., & Kingsinger, L.S. (2006). *Principles of screening for cancer. In: Oncology: An Evidence-Based Approach*. New York: Springer.

Hersch, J., Jansen, J., Barratt, A., Irwig, L., Houssami, N., & Howard, K. et al. (2013). Women's views on overdiagnosis in breast cancer screening: a qualitative study. *British Medical Journal, 346,* f158.

Hislop, T. G., Bajdik, C. D., Regier, M. D., & Barroetavena, M. C. (2007). Ethnic differences in survival for female cancers of the breast, cervix and colorectum in British Columbia, Canada. *Asian Pacific Journal of Cancer Prevention, 8*(2), 209.

Hochbaum, G. M. (1958). *Public participation in medical screening programs: a socio-psychological study*. Washington, D.C.: U.S. Dept. of Health, Education, and Welfare.

Hoeflma, D., & Hanewinkel, R. (2008). Do school-based tobacco prevention programmes pay off? The cost-effectiveness of the 'Smoke-free Class Competition.' *Public Health, 122(*1): 34-41.

Horton, R. (2001). Screening mammography -- an overview revisited. *The Lancet, 358,* 1284-1285.

Houle, C., Will, B.P., Berthelot, J-M., & Evans, W.K. (1997). Use of POHEM to Estimate Direct Medical Costs of Current Practice and New Treatments Associated with Lung Cancer in Canada. Statistics Canada. 11F0019MPE No. 99. Retrieved March 28, 2014 from http://www.statcan.gc.ca/pub/11f0019m/11f0019m1997099-eng.pdf.

Howlader, N., Noone, A.M., Krapcho, M., Garshell, J., Miller, D., Altekruse, S.F., Kosary, C.L., Yu, M., Ruhl, J., Tatalovich, Z., Mariotto, A., Lewis, D.R., Chen, H.S., Feuer, E.J., & Cronin, K.A. (eds). (2013). SEER Cancer Statistics Review, 1975-2011, *National Cancer Institute. Bethesda, MD.* Retrieved April 2014 from http://seer.cancer.gov/csr/1975_2011/

Humphrey, L.L., Helfand, M., Chan, B.K.S., et al. (2002). Breast cancer screening: a summary of the evidence for the U.S. Preventive Services Task Force. *Annals of Internal Medicine, 137*(5), 347-360.

Hunt, K.A., & Sickles, E.A. (2000). Effect of obesity on screening mammography: outcomes analysis of 88,346 consecutive examinations. *American Journal of Roentgenology, 174*(5):1251-1255.

Hunter, D.J., Drake, S.M., Shortt, S.E., Dorland, J.L., & Tran, N. Simulation modeling of change to breast cancer detection age eligibility recommendations in Ontario, 2002-2021. *Cancer Detection and Prevention, 28,* 453-460.

Hunter DJ, Drake SM, Shortt SE, Dorland JL, Tran N. Simulation modeling of change to breast cancer detection age eligibility recommendations in Ontario, 2002-2021. Cancer Detect Prev. 2004;28:453-460.

Hurley, S.F., & Kaldor, J.M. (1992). The benefits and risks of mammographic screening for breast cancer. *Epidemiologic Review, 14*, 101-130.

Institute of Medicine (US) and National Research Council (US) Committee on New Approaches to Early Detection and Diagnosis of Breast Cancer; Joy, J.E., Penhoet, E.E., Petitti, D.B., (eds). (2005). *Saving Women's Lives: Strategies for Improving Breast Cancer Detection and Diagnosis*. Washington, DC: National Academies Press (US). Glossary. Available from: http://www.ncbi.nlm.nih.gov/books/NBK22321/

Jacobs, T.W., Byrne, C., Colditz, G., Connolly, J.L., & Schnitt, S.J. (2001). Pathologic features of breast cancers in women with previous benign breast disease. *American Journal of Clinical Pathology, 115*(3), 362-369.

Kachuri, L., De, P., Ellison, L.F., Semenciw, R., & Advisory Committee on Canadian Cancer Statistics. (2013). Cancer incidence, mortality and survival trends in Canada, 1970–2007. *Chronic Diseases and Injuries in Canada, 33*(2).

Kalager, M., et al. (2010). Effect of Screening Mammography on Breast-Cancer Mortality in Norway. *New England Journal of Medicine, 363*, 1203-1210.

Kawachi, I., Subramanian, S.V., & Almeida-Filho, N. (2002). *Journal of Epidemiology Community Health*, *56*, 647–652.

Kelsey, J.L., Gammon, M.D., & John, E.M. (1993). Reproductive factors and breast cancer. *Epidemiology Review, 15*, 36-47.

Kerlikowske, K., Grady, D., Barclay, J., Sickles, E.A., & Ernster, V. (1996). Effect of age, breast density, and family history on the sensitivity of first screening mammography. *The Journal of the American Medical Association*, *276*(1), 33–38.

Kerlikowske, K., Grady, D., Rubin, S.M., Sandrock, C., & Ernster, V.L. (1995). Efficacy of screening mammography. A meta-analysis. *The Journal of the American Medical Association*, *273*, 149-154.

Kerlikowske, K., Salzmann, P., Phillips, K.A., Cauley, J.A., & Cummings, S.R. (1999). Continuing screening mammography in women aged 70 to 79 years: impact on life expectancy and cost-effectiveness. *The Journal of the American Medical Association*, *282*, 2156-2163.

Kolata, G. (2010). Mammograms' Value in Cancer Fight at Issue. *New York Times, Health*, Retrieved February 15, 2011 from http://www.nytimes.com/2010

Kolb, T.M., Lichy, J., Newhous, J.H. (2002). Comparison of the Performance of Screening Mammography, Physical Examination, and Breast US and Evaluation of Factors that Influence Them: An Analysis of 27,825 Patient Evaluations. *Radiology*, *225*, 165-175.

Koning, H.J.D., et al. (1991). Breast cancer screening and cost-effectiveness; policy alternatives, quality of life considerations and the possible impact of uncertain factors. *International Journal of Cancer, 49* , 531.

Koning, H.J.D., Boer, R., Warmerdam, P.G., Beemsterboer, P.M.M., & van P.J.dM. (1995). Quantitative interpretation of age-specific mortality reductions from the swedish breast cancer-screening trials. *Journal of the National Cancer Institute, 87,* 1217-1223.

Kopec, J.A., Sae, E.C., Berthelot, J.M., et al. (2010). Development of a computer simulation model of osteoarthritis in Canada. *Osteoarthritis Cartilage, 18,* 303-311.

Kopec, J.A., Finès, P., Manuel, D.G., Buckeridge, D.L., Flanagan, W.M., Oderkirk, J., et al. (2010). Validation of population-based disease simulation models: a review of concepts and methods. *BMC Public Health, 10,* 710.

Kosters, J. P., & Gotzsche, P. C. (2003). Regular self-examination or clinical examination for early detection of breast cancer. *Cochrane Database of Systematic Reviews, 2,* Article ID CD003373.

Last, J.M., ed. (2001). *A Dictionary of Epidemiology*. Fourth ed. New York: Oxford University Press.

Last, J.M. (2007). *A Dictionary of Public Health*. New York: Oxford UP Print.

Lauzier, S., Maunsell, E., & Drolet, M. et al. (2008). Wage Losses in the Year After Breast Cancer: Extent and Determinants Among Canadian Women. *Journal of the National Cancer Institute, 100,* 321-332.

Lee, J.M., McMahon, P.M., Kong, C.Y., & et al. (2010). Cost-effectiveness of Breast MR Imaging and Screen-Film Mammography for Screening BRCA1 Gene Mutation Carriers. *Radiology, 254,* 793-799.

Lee, S., & Zelen, M. (2003). Modelling the early detection of breast cancer. *Annals of Oncology, 14*(8), 1199-1202.

Lee, K., & Mills, A. (1983). *The Economics of Health in Developing Countries: a Critical Review, In Kenneth Lee and Anne Mills (Eds.).* The Economics of Health in Developing Countries. Oxford: Oxford University Press.

Lefrançois, S., Pelletier, E., Thiboutot, P., Major, D., Simpson, A. (2007). *Évaluation des coûts du Programme québécois de dépistage du cancer du sein : développement des indicateurs coût par participante et coût par cancer détecté*. Institut National de Santé Publique du Québec. Retrieved March 28, 2014 from http://www.inspq.qc.ca/pdf/publications/693_depistage_cancer_sein.pdf.

Lehman, C.D., Gatsonis, C., Kuhl, C.K., Hendrick, R.E., Pisano, E.D., Hanna, L., Peacock, S., Smazal, S.F., Maki, D.D., Julian, T.B., dePeri, E.R., Bluemke, D.A., & Schnall, M.D. (2007). ACRIN Trial 6667 Investigators Group. MRI evaluation of the contralateral breast in women with recently diagnosed breast cancer. *New England Journal of Medicine, 356*(13), 1295-1303.

Lidgren, M., Wilking, N., Jonsson, B., & Rehnberg, C. (2007). Resource use and costs associated with different states of breast cancer. *International Journal of Technology Assessment in Health Care, 23*(2), 223-231.

Lofters, A.K., Gozdyra, P., & Lobb, R. (2013). Using geographic methods to inform cancer screening interventions for South Asians in Ontario, Canada. *BMC public health, 13*(1), 395.

Loppie, C., Wein, F. (2009). *Health inequalities and social determinants of Aboriginal people's health*. The National Collaborating Centre for Aboriginal Health. Retrieved May 5, 2014 from http://www.nccah-ccnsa.ca/docs/social%20determinates/nccah-loppie-wien_report.pdf.

Luckett, T., Goldstein, D., Butow, P. N., Gebski, V., Aldridge, L. J., McGrane, J., ... & King, M. T. (2011). Psychological morbidity and quality of life of ethnic minority patients with cancer: a systematic review and meta-analysis. *The Lancet Oncology, 12*(13), 1240-1248.

Luo, W., Birkett, N. J., Ugnat, A. M., & Mao, Y. (2004). Cancer incidence patterns among Chinese immigrant populations in Alberta. *Journal of Immigrant Health, 6*(1), 41-48.

Maddux, J. F., Ingram, J. M., & Desmond, D. P. (1995). Reliability of two brief questionnaires for drug abuse treatment evaluation. *American Journal of Drug and Alcohol Abuse, 21*(2), 209-221.

Mandelblatt, J., Saha, S., Teutsch, S., Hoerger, T., Siu, A.L., Atkins, D., Klein, J., & Helfand, M. (2003). The Cost-Effectiveness of Screening Mammography beyond Age 65 Years: A Systematic Review for the U.S. Preventive Services Task Force. *Annals of Internal Medicine, 139*(10), 835-842.

Mandelblatt, J.S., Schecter, C.B., Yabroff, K.R., et al. (2005). Toward Optimal Screening Strategies for Older Women. *Journal of General Internal Medicine, 20,* 487-496.

Marmot, M., Altman, D., Cameron, D., Dewar, J., Thompson, S., & Wilcox, M. (2012). The benefits and harms of breast cancer screening: an independent review. *The Lancet, 380*(9855), 1778-1786.

Marshall, L., Hunter, D., Connolly, J., et al. (1997). Risk of breast cancer associated with atypical hyperplasia of lobular and ductal types. *Cancer Epidemiology Biomarkers Prev, 6,* 297-301.

Martin, L.J., Meinichouk, O., Guo, H., et al. (2010). Family History, Mammographic Density, and Risk of Breast Cancer. *Cancer Epidemiology Biomarkers Prev, 19*(2), 456-463.

McDonald, J.T., & Sherman, A. (2010). Determinants of mammography use in rural and urban regions of Canada. *Canadian Journal of Rural Medicine, 15*(2).

Mcpherson, K., Steel, C,M., & Dixon, J.M. (2000). ABC of breast diseases. Breast cancer-epidemiology, risk factors, and genetics. *British Medical Journal, 321*, 624-628.

Miller, A., Baines, C.J., Sun, P. et al. (2014). Twenty five year follow-up for breast cancer incidence and mortality of the Canadian National Breast Screening Study: randomised screening trial. *British Medical Journal, 348*, g366.

Miller, A.B. (1996). *Fundamental issues in screening for cancer.* In: Schottenfeld, D., Fraumeni, J.F., (eds). 2nd ed. New York: Oxford University Press. p.1433.

Miller, C.J., Kramer, B.S., Kreimer, A.R., et al. (2009). Annals Journal Club: Cumulative Incidence of False-Positive Results in Repeated, Multimodal Cancer Screening. *The Annals of Family Medicine, 7*, 212-222.

Mitton, C. R., Rose, S., & Sheldon, R.S. (1998). *A cost-utility analysis of pacemakers for the treatment of neurally mediated syncope*. Edmonton, AB, Canada: Institute of Pharmaco-Economics. [Working paper]. 98-95.

Mittra, I., Baum, M., Thornton, H., & Houghton, J. (2000). Is clinical breast examination an acceptable alternative to mammographic screening? *British Medical Journal,321*, 1071–1073.

Moss, S.M., Cuckle, H., Evans, A., et al. (2006). Effect of mammographic screening from age 40 years on breast cancer mortality at 10 years' follow-up: a randomised controlled trial. *The Lancet, 368*(9552), 2053-2060.

Mushlin, A.I., Kouides, R.W., & Shapiro, D.E. (1998). Estimating the accuracy of screening mammography: a meta-analysis. *American Journal of Preventive Medicine, 14*(2), 143–153.

Nadeau, C., Wong, S.L., Flanagan, W.M., Oderkirk, J., Manuel, D., Wall, R., & Tremblay, M.S. (2013). Development of a population-based microsimulation model of physical activity in Canada. Statistics Canada, Catalogue no. 82-003-X . *Health Reports, 24*(10), 11-19.

National Cancer Institute at the National Institutes of Health. Dictionary of Cancer Terms. http://www.cancer.gov/Common/PopUps/popDefinition.aspx?id=460150&version=HealthProfessional&language=English

National Cancer Institute. (2014) *Breast Cancer Screening (PDQ®)*. Retrieved October 5, 2014 from http://www.cancer.gov/cancertopics/pdq/screening/breast/healthprofessional/page1.

National Institute for Health and Clinical Excellence. (2007). *Briefing paper for the Methods Working Party on the cost effectiveness threshold*. London: National Institute for Health and Clinical Excellence.

Nature.com reviews glossary. Retrieved on July 29, 2007 from http://www.nature.com/nrg/journal/v2/n4/glossary/nrg0401_292a_glossary.html.

Navaneelan, T., & Janz, T. (2011). Cancer in Canada: Focus on Lung, Colorectal, Breast and Prostate. Health at a Glance, Statistics Canada. (Catalogue no. 82-624-X).

Neeser, K., Szucs, T., Bulliard, J-L., et al. (2007). Cost-Effectiveness Analysis of a Quality-Controlled Mammography Screening Program from the Swiss Statutory Health-Care Perspective: Quantitative Assessment of the Most Influential Factors. *Value Health, 10,* 42-53.

Nelson, H.D., Tyne, K., Naik, A., Bougatsos, C., Chan, B.K., & Humphrey, L. (2009). U.S. Preventive Services Task Force. Screening for breast cancer: an update for the U.S. Preventive Services Task Force. *Annals of Internal Medicine 151*(10), 727-37, W237-42.

Nelson, H.D., Zakher, B., Cantor, A., 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.

Okonkwo, Q.L., Draisma, G., der Kinderen, A., et al. (2008). Breast Cancer Screening Policies in Developing Countries: A Cost-Effectiveness Analysis for India. *Journal of the National Cancer Institute, 100,* 1290-1300.

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

Pataky, R., Armstrong, L., Chia, S., Coldman, A., Kim-Sin, C., McGillivray, B., Scott, J., Wilson, C., & Peacock, S. (2013). Cost-effectiveness f MRI for breast cancer screening in BRCA1/2 mutation carriers. BMC Cancer 2013, 13:339. Retrieved April 1, 2014 from http://www.biomedcentral.com/content/pdf/1471-2407-13-339.pdf.

Patra, J., Popova, S., Rehm, J., et al. (2007). *Economic cost of chronic disease in Canada, 1995-2003.* Ontario Chronic Disease Prevention Alliance and the Ontario Public Health Association.

Pharoah, P.D.P., Day, N.E., Duffy, S., et al. (1997). Family history and the risk of breast cancer: a systematic review and meta-analysis. *International Journal of Cancer, 71*, 800–809.

Phillips, C. (2009). *What is cost-effectiveness?* Second Edition. Retrieved from: http://www.medicine.ox.ac.uk/bandolier/painres/download/whatis/Cost-effect.pdf

Phillips et al. (2006). Good Practice Guidelines for Decision-Analytic Modelling in Health Technology Assessment. *Pharmacoeconomics, 24*(4), 355-371.

Phipps, S.M.O, Woodfin, W.K., & Tollefsbol, T.O. (2005). The epigenetics of breast carcinogenesis and metastasis. *Current Genomics, 6,* 129. Retrieved from: http://www.ingentaconnect.com/content/ben/cg/2005/00000006/00000003/art00002.

Pisano, E.D., Gatsonis, C., Hendrick, E., Yaffe, M., Baum, J.K., Acharyya, S., Conant, E.F., Fajardo, L.L., Bassett, L., D'Orsi, C., Jong, R., & Rebner, M. (2005). Digital Mammographic Imaging Screening Trial (DMIST) Investigators Group. Diagnostic performance of digital versus film mammography for breast-cancer screening. *New England Journal of Medicine, 353*(17), 1773-1783.

Pisano et al. (2008). Diagnostic accuracy of digital versus film mammography: exploratory analysis of selected population subgroups in DMIST. *Radiology, 246*(2), 376-83. doi: 10.1148/radiol.2461070200.

Public Health Agency of Canada. (2014) Economic Burden of Illness in Canada, 2005-2008. Retrieved October 5, 2014 from http://www.phac-aspc.gc.ca/publicat/ebic-femc/2005-2008/assets/pdf/ebic-femc-2005-2008-eng.pdf.

Public Health Agency of Canada. (n.d.). *The Economic Burden of Illness in Canada 2005-2008 Report*. Retrieved August 15, 2014 from http://www.phac-aspc.gc.ca/publicat/ebic-femc/2005-2008/assets/pdf/ebic-femc-2005-2008-eng.pdf.

Public Health Agency of Canada. (2007). *Guidelines for monitoring breast screening program performance*: Report from the evaluation indicators working group: 2nd Edition. Ottawa.

Public Health Agency of Canada. (n.d.). *Organized breast cancer screening programs in Canada, 2003 and 2004 report*. Ottawa: Minister of Public Works and Government Services Canada; 2008. Retrieved August 15, 2014 from: http://www.phac-aspc.gc.ca/publicat/obcsp-podcs01/pdf/Breast-En_2003-2004.pdf.

Public Health Agency of Canada. (2011). *Organized Breast Cancer Screening Programs in Canada: Report on Program Performance in 2005 and 2006*. Ottawa.

Public Health Agency of Canada. (n.d.). *What is a population health approach?* Retrieved August 15, 2014 from http://www.phac-aspc.gc.ca/ph-sp/approach-approche/index-eng.php.

Quan, H., Fong, A., De Coster, C., et al. (2006). Variation in health services utilization among ethnic populations. *Canadian Medical Association Journal, 174*, 787-791.

Richardson, L. (2005). Re: Ethnicity and breast cancer: Factors influencing differences in incidence and outcome. *Journal of the National Cancer Institute, 97*, 1619-1620.

Rojnik, K., Naversnik, K., Mateovic-Rojnik, T., et al. (2008). Probabilistic Cost-Effectiveness Modeling of Different Breast Cancer Screening Policies in Slovenia. *Value Health, 11*, 139-148.

Rosenquist, C.J. & Lindfors, K.K. (1994). Screening mammography in women aged 40-49 years: analysis of cost-effectiveness. *Radiology 191*, 647-650.

Saika, K., & Sobue, T. Time Trends in Breast Cancer Screening Rates in the OECD Countries. *Japanese Journal of Clinical Oncology, 41*(4), 591-592.

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.

Samuelson, P.A. (1976): *Economics*. New York: McGraw-Hill.

Schifter, D. E., & Ajzen, I. (1985). Intention, perceived control, and weight loss: An application of the theory of planned behavior. *Journal of Personality and Social Psychology, 49*(3), 843-851.

Schousboe, J.T., Kerlikowske, K., Loh, A., & Cummings, S.R. (2011). Personalizing mammography by breast density and other risk factors for breast cancer: analysis of health benefits and cost-effectiveness. *Annals of Internal Medicine, 155*(1), 10-20.

Schueler, K.M., Chu, P.W., & Smith-Bindman, R. (2008). *Journal of Women's Health, 17*(9), 1477-1498. doi:10.1089/jwh.2007.0603.

Schwartz, L.M., Woloshin, S., Fowler, Jr .F.J, et al. (2004). Enthusiasm for cancer screening in the United States. *Journal of the American Medical Association, 291*(1), 71–78.

Sen, S., Wang, S.Y., Soulos, P.R., Frick, K.D., Long, J.B., Roberts, K.B., Yu, J.B., Evans, S.B., Chagpar, A.B., & Gross, C.P. (2014). Examining the cost-effectiveness of radiation therapy among older women with favorable-risk breast cancer. *Journal of the National Cancer Institute, 106*(3), dju008. doi: 10.1093/jnci/dju008. Epub 2014 Mar 5.

Shen, Y., Yang, Y., Inoue, L.Y.T., Munsell, M.F., Miller, A.B., & Berry, D.A. (2005). Role of detection method in predicting breast cancer survival: Analysis of randomized screening trials. *Journal of the National Cancer Institute, 97,* 1195-1203.

Simulation Technology for Applied Research. Glossary of Terms. Retrieved from http://www.pophealthmodels.ca/index.php?option=com_content&view=article&id=70&Itemid=57

Stone, P. (1998). Commentary on "A brief smoking cessation intervention for adults in hospital was cost effective." Evidence-Based Nursing1999Apr;2:58. Comment on: Meenan RT, Stevens VJ, Hornbrook MC, et al. Cost-effectiveness of a hospital-based smoking cessation intervention. *Medical Care, 36,* 670–678.

Stout, N.K., Rosenberg, M.A., Trentham-Dietz, A., et al. (2006). Retrospective Cost-effectiveness Analysis of Screening Mammography. *Journal of the National Cancer Institute, 98,* 774-782.

Tammemagi, C. M. (2007). Racial/ethnic disparities in breast and gynecologic cancer treatment and outcomes. *Current Opinion in Obstetrics and Gynecology,19*(1), 31-36.

The Canadian Breast Cancer Foundation. (2014). *Breast Screening Technology*. Retrieved May 6, 2014 from https://www.cbcf.org/central/AboutBreastHealth/EarlyDetection/Pages/Breast-Screening-Technology.aspx.

Thomas, H.V., Key, T.J., Allen, D.S., et al. (1997). Reversal of relation between body mass and endogenous estrogen concentrations with menopausal status. *Journal of the National Cancer Institute, 89,* 396-398.

Titus-Ernstoff, L., et al. (2006). Breast cancer risk factors in relation to breast density (United States). *Cancer Causes Control, 17*(10), 1281-1290.

Tjepkerma, M., Wilkins, R., & Long, A. (2013). Cause-sepcfici mortality by income adequacy in Canada: A 16-year follow-up study. *Health Reports, 24*(7), 14-22. Statistics Canada, Catalogue no. 82-003-X

Tosteson, A.N., Stout, N.K., Fryback, D.G., et al. (n.d.). Cost-Effectiveness of Digital Mammography.

US Preventive Services Task Force. Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. *Annals of Internal Medicine, 151*(10), 716-726. W-236. doi: 10.7326/0003-4819-151-10-200911170-00008.

Vahabi, M. (2011 a). Breast cancer and screening information needs and preferred communication medium among Iranian immigrant women in Toronto. *Health & Social Care in the Community, 19*(6), 626-635.

Vahabi, M. (2011 b). Knowledge of breast cancer and screening practices among Iranian immigrant women in Toronto. *Journal of community health, 36*(2), 265-273.

Vainshtein, J. (2008). Disparities in breast cancer incidence across racial/ethnic strata and socioeconomic status: A systematic review. *Journal of the National Medical Association, 100*, 833-839.

Van den Brandt, P.A., Spiegelman, D., Yaun, S., et al. (2000). Pooled analysis of prospective cohort studies on height, weight, and breast cancer risk. *American Journal of Epidemiology, 152*, 514-527.

Van Oortmarssen, G.J., Boer, R., & Habbema, J.D. (1995). Modeling issues in cancer screening. *Statistical Methods in Medical Research, 4*, 33-54.

Velicer, W. F., DiClemente, C. C., Prochaska, J. O., & Brandenburg, N. (1985). Decisional balance measure for assessing and predicting smoking status. *Journal of Personality and Social Psychology, 48*(5), 1279-1289.

Vilaprinyo, E., et al. (2014). Cost-effectiveness and harm-benefit analyses of risk-based screening strategies for breast cancer. *PLoS ONE 9*(2): e86858.

Warren Burhenne, L.J., Coldman, A.J., & Kan, L. (2007) Organized Breast Cancer Screening in British Columbia: The Screening Mammography Program of British Columbia. *Seminars in Breast Disease, 10*(2): 83-88.

Weitzner, M.A., Meyers, C.A., Stuebing, K.K., & Saleeba, A.K. (1997) Relationship between quality of life and mood in long-term survivors of breast cancer treated with mastectomy. *Support Care Cancer, 5*, 241-248.

Welch, H.G. & Black, W.C. (1997). Using Autopsy Series To Estimate the Disease "Reservoir" for Ductal Carcinoma in Situ of the Breast: How Much More Breast Cancer Can We Find? *Annals of Internal Medicine, 127*(11), 1023-1028.

Westbrook, K., & Stearns, V. (2013). Pharmacogenomics of breast cancer therapy: an update. *Pharmacology and Therapeutics, 139*(1), 1-11. doi: 10.1016/j.pharmthera.2013.03.001. Epub 2013 Mar 13.

Will, B., Berthelot, J., Petit, C.L., Tomiak, E., Verma, S., & Evans, W. (2000). Estimates of the lifetime costs of breast cancer treatment in Canada. *European Journal of Cancer, 36*, 724-735.

Will, B.P., Berthelot, J.M., Nobrega, K.M., Flanagan, W., Evans, W.K. (2001). Canada's Population Health Model (POHEM): a tool for performing economic evaluations of cancer control interventions. Euro*pean Journal of Cancer, 37*(14), 1797-1804.

Will, B.P., Le Petit, C., Berthelot, J-M., et al. (1999). Diagnostic and therapeutic approaches for nonmetastatic breast cancer in Canada, and their associated costs. *British Journal of Cancer, 9*(10), 1428-1436.

Willich, S.N., Reinhold, T., Selim, D., Jena, S., Brinkhaus, B., & Witt, C.M. (2006). Cost-effectiveness of acupuncture treatment in patients with chronic neck pain. *Pain, 125*(1-2), 107-113.

Winquist, E., Bell, C.M., Clarke, J.T.R., Evans, G., Martin, J., Sabharwal, M., Gadhok, A., Stevenson, H., & Coyle, D. (2012). An evaluation framework for funding drugs for rare disease. *Value in Health, 15*, 982–986.

Winsconson-Harvard Model. CISNET. 2013. Retrieved September 1, 2014 from https://cisnet.flexkb.net/mp/pub/CISNET_ModelProfile_BREAST_UWISC_001_07232013_58567.pdf#pagemode=bookmarks.

Wolfson, M.C. (1994). POHEM--a framework for understanding and modelling the health of human populations. *World Health Stat Q, 47*(3-4), 157-176.

Wong, G., Howard, K., Chapman, J.R., et al. (2008). Cost-Effectiveness of Breast Cancer Screening in Women on Dialysis. *American Journal of Kidney Disease, 52*, 916-929.

Wong, I.O., Kuntz, K.M., Cowling, B.J., et al. (2007). Cost Effectiveness of Screening mammographyfor Chinese Women. *Cancer, 110*, 885-895.

World Health Organization. (2011). *Health Topics: Obesity*. Retrieved December 11, 2011from http://www.who.int/topics/obesity/en/.

World Health Organization. (1995). The World Health Organization Quality of Life assessment (WHOQOL): position paper from the World Health Organization. *Social Science and Medicine, 41*(10), 1403-1409.

Wright, C. (1995). Screening mammography and public health policy: the need for perspective. *The Lancet, 346*, 29.

Yavari, P. et al. (2010). Breast cancer treatment and ethnicity in British Columbia, Canada. *BMC Cancer, 10*,154.

Yen, M.F., et al. (2003). Quantifying the potential problem of overdiagnosis of ductal carcinoma in situ in breast cancer screening. *European Journal of Cancer, 39*,1746–1754.

Young, L.S, Hwa, J.S., Nam, K.Y., et al. (2009). Cost-effective screening mammography in Korea: High incidence of breast cancer in young women. *Cancer Science, 100,* 1105–1111.

Zackrisson, S., Andersson, I., Janzon, L., Manjer, L., Garne, J.P. (2006). Rate of over-diagnosis of breast cancer 15 years after end of Malmo mammographic screening trial: follow-up study. *British Medical Journal, 332*, 689–692.

Appendices

Appendix1: POHEM-BCS input data with data sources where applicable

Note: all values were used for the status quo screening policy option A: biennial screening for women age 50-69.

1. Screening – Recruitment

Table 1.1 Parameter Description: Screen inclusion criteria					
Screen eligibility items					
Recruitment minimum age	Recruitment maximum age	Recruitment start year	Recruitment end year	Year screening program ends	Maximum age to re-screen
40	80	0	9999	9999	80

Screen inclusion criteria are preset and represent the possible range of ages eligible for screening and the years in which screening would start and end. No screening policy options included screening for women age less than 40 or greater than 79. Since this analysis is not dependent on year, we aimed to employ a starting population of 6 million females who would be tracked from birth to death.

2. Screening – Screen Frequency

Table 2.1 Screen frequency				
	Age Group			
Test/Procedure	(min,50)	[50,60)	[60,70)	[70,max)
Screening Test	0	2	2	2

Screen frequency estimates are preset and represent the possible ranges of screening frequency by age group. For the status quo screening policy option A: biennial screening for women age 50-69, we indicated that women age 50-79 would receive screening every two years.

3. Participation rates

Table 3.1 Participation Rates					
Screen invitation status	Sex	Age Group			
		(min,50)	[50,60)	[60,70)	[70,max)
First invitation	Female	0	0.7	0.7	0
Subsequent invitation	Female	0	0.9	0.9	0

Participation rates are preset and represent the proportion of the population who accept screening under each screening policy option. It represents the probability that a women, depending on her age, will participate in screening. For initial screens (first invitation) the target participation rate is 70% and for subsequent screens (subsequent invitation) the target participation rate is 90% (Canadian Partnership Against Cancer, 2013). The participation rates were therefore set to these values in the model.

4. Detection

Table 4.1 Baseline breast cancer incidence rates (applied in POHEM)

Age	Incidence rate	Age	Incidence rate	Age	Incidence rate	Age	Incidence rate	Age	Incidence rate
0	0	23	9.08E-06	46	0.001474274	69	0.002993167	92	0.003975416
1	0	24	1.72E-05	47	0.00149767	70	0.002997702	93	0.003975416
2	0	25	3.48E-05	48	0.00164945	71	0.003248865	94	0.003975416
3	0	26	3.53E-05	49	0.001682116	72	0.003173168	95	0.003975416
4	0	27	6.40E-05	50	0.00178061	73	0.003081548	96	0.003975416
5	0	28	8.83E-05	51	0.001772747	74	0.003191935	97	0.003975416
6	0	29	0.000122861	52	0.00172392	75	0.003294008	98	0.003975416
7	0	30	0.00012332	53	0.001804896	76	0.003184342	99	0.003975416
8	0	31	0.000196222	54	0.001981448	77	0.003508521		
9	0	32	0.000226199	55	0.001955706	78	0.003322676		
10	0	33	0.000274112	56	0.002002354	79	0.003584737		
11	1.40E-06	34	0.00035728	57	0.002223367	80	0.003477437		
12	0	35	0.000386155	58	0.002337462	81	0.003546167		
13	0	36	0.000444819	59	0.002233186	82	0.003375243		
14	1.32E-06	37	0.000540774	60	0.002341559	83	0.003505821		
15	0	38	0.000548474	61	0.002507056	84	0.003694679		
16	2.55E-06	39	0.000731757	62	0.002504307	85	0.003944241		
17	0	40	0.000847718	63	0.002554073	86	0.004137035		
18	1.14E-06	41	0.000911019	64	0.002658824	87	0.003737634		
19	2.15E-06	42	0.001017628	65	0.002607497	88	0.003686608		
20	2.07E-06	43	0.00107327	66	0.002698942	89	0.003885776		
21	4.05E-06	44	0.001219352	67	0.0028419	90	0.003975416		
22	1.01E-05	45	0.00145637	68	0.002877011	91	0.003975416		

Breast cancer incidence is the number of new cases of breast cancer diagnosed by age. The source data came from the Canadian Cancer Registry (1995) and were originally used for the research conducted by Will et al.

(2000). These data were not altered for the present analysis which may under estimate the impact of screening in terms of life-years gained as well as costs of diagnosis and treatment. Age-standardized breast cancer incidence rate has however remained relatively unchanged (about 100 per 100,000 population) between 1995 and 2013 according to the Canadian Cancer Society`s Canadian Cancer Statistics (2013).

Table 4.2 Breast cancer stage					
Screen detection method	**Age Group**	**Stage**			
		I	**II node negative**	**II node positive**	**III +IV**
No screening	(min,50)	0.362	0.135	0.359	0.145
No screening	[50,60)	0.343	0.171	0.329	0.157
No screening	[60,70)	0.356	0.225	0.256	0.163
No screening	[70,max)	0.351	0.181	0.312	0.156
First screen	(min,50)	0.717	0.097	0.159	0.027
First screen	[50,60)	0.717	0.097	0.159	0.027
First screen	[60,70)	0.717	0.097	0.159	0.027
First screen	[70,max)	0.717	0.097	0.159	0.027
Re-screen	(min,50)	0.717	0.097	0.159	0.027
Re-screen	[50,60)	0.717	0.097	0.159	0.027
Re-screen	[60,70)	0.717	0.097	0.159	0.027
Re-screen	[70,max)	0.717	0.097	0.159	0.027
Interval	(min,50)	0.434	0.192	0.261	0.114
Interval	[50,60)	0.434	0.192	0.261	0.114
Interval	[60,70)	0.434	0.192	0.261	0.114
Interval	[70,max)	0.434	0.192	0.261	0.114
Non-participant	(min,50)	0.362	0.135	0.359	0.145
Non-participant	[50,60)	0.343	0.171	0.329	0.157
Non-participant	[60,70)	0.356	0.225	0.256	0.163
Non-participant	[70,max)	0.351	0.181	0.312	0.156

Breast cancer stage distribution represents the proportion of the population diagnosed with breast cancer in each stage by screen detection method and age group. Stage 0 or ductal carcinoma in situ was not included in the model infrastructure therefore the analysis in this study is specific to invasive breast cancer only. Pre-screening (clinically detected) stage distributions were obtained from the Saskatchewan cancer registry, 1982-1985. The stage distribution of cancers detected in screening programs was obtained from the Canadian Breast Cancer Screening Database (CBCSD) for 1997-98. More current stage distributions in the population appear to be comparable to the distributions estimated from the Ontario Cancer Registry, Collaborative Staging Database (2007-2011 data) in a report Cancer Quality Council of Ontario (February 2013).

Table 4.3 Cancer Detection : Sensitivity & Specificity & Sojourn					
Test/Procedure	**Age Group**	**Screening Round Type**	**Accuracy**		
			Sensitivity	**Specificity**	**Sojourn (Years)**
Screening Test	(min,50)	First	0.87	0.85	1
Screening Test	(min,50)	Subsequent	0.72	0.94	1
Screening Test	[50,60)	First	0.88	0.87	2
Screening Test	[50,60)	Subsequent	0.90	0.94	2
Screening Test	[60,70)	First	0.94	0.90	2
Screening Test	[60,70)	Subsequent	0.83	0.94	2
Screening Test	[70,max)	First	0.94	0.91	2
Screening Test	[70,max)	Subsequent	0.67	0.95	2

Screening test sensitivity and specificity data were estimated for each age group and by screening invitation status or round type (first or subsequent screen). Also indicated is the screening sensitivity and specificity by screen interval or frequency indicated by sojourn time in years. The source data for these estimates were from the Canadian Breast Cancer Screening Database 2006 data for which national data was most complete at the time of analysis (year 2009).Clinically detected cancers are simulated through the incidence rate (4.1)s. According to Flanagan et al. (2006), these cancers were assumed to have an average sojourn time over which the cancer could be detected pre-clinically by mammogram. By evaluating incidence from the person's current age forward in time to an amount equal to the sojourn period, the model identifies the presence of cancers before they occur clinically (i.e. before they are symptomatic). Flanagan et al. (2006) simulated mammography for pre-clinical cancerous cases using the sensitivity estimate to determine the proportion of screen-detected (true positives)cases and cases missed (false negatives). Missed cases occuring before the next screen are called interval cancers. In cases where no cancer is present over the sojourn period, Flanagan et al. (2006) used the specificity estimate to determine the proportion correctly identified (true negatives) and those mistakenly identified as cancerous (false positives).

5. Survival

Table 5. 1 Survival parameters				
Inverse piecewise Weibull parameters	**Stage**	**Breast Cancer Progression**		
		Initial diagnosis- local recurrence	**Initial diagnosis-metastasis**	**Initial diagnosis-mortality**
Weibull_ uFixed	I	0	0	0.929287
Weibull_ uFixed	II node negative	0	0	0.956044
Weibull_ uFixed	II node positive	0	0	0.77334
Weibull_ uFixed	III+IV	0	0	0.313713
Weibull_ tFixed	I	0	0	48
Weibull_ tFixed	II node negative	0	0	18
Weibull_ tFixed	II node positive	0	0	42
Weibull_ tFixed	III+IV	0	0	56
Weibull_ lambda1	I	0	0	0.003751
Weibull_ lambda1	II node negative	0	0	0.014055
Weibull_ lambda1	II node positive	0	0	0.00942
Weibull_ lambda1	III+IV	0	0	0.020943
Weibull_ beta1	I	0	0	1.523876

Weibull_ beta1	II node negative	0	0	2.257141
Weibull_ beta1	II node positive	0	0	1.4652
Weibull_ beta1	III+IV	0	0	0.927157
Weibull_ lambda2	I	0	0	0.001903
Weibull_ lambda2	II node negative	0	0	0.002885
Weibull_ lambda2	II node positive	0	0	0.004561
Weibull_ lambda2	III+IV	0	0	0.023308
Weibull_ beta2	I	0	0	1.09174
Weibull_ beta2	II node negative	0	0	1.048778
Weibull_ beta2	II node positive	0	0	0.822094
Weibull_ beta2	III+IV	0	0	0.554807

Survival was modeled using probabilities of mortality from initial diagnosis by breast cancer stage. The source data of the survival probabilities came from several sources including the Northern Alberta Breast Cancer Registry (1971-1988), the Saskatchewan Cancer Foundation – Special Charts Reviews (1985-1992), and the British Columbia Cancer Agency (1989-1994). These data were originally used for the research by Will et al. (2000). Disease-specific survival from initial diagnosis to death was implemented by stage as a Weibull distribution. Survival has undoubtedly improved since the mid-90s from when the most recent data used in the estimation of survival in the model, particularly in the later stages (III+IV). Changes in survival probabilities may be attributable to improvements in treatments. It is therefore possible that this analysis over-estimates the impact of screening on mortality.

6. Treatment

All of the below tables include data used in the therapeutic algorithms at initial diagnosis, follow-up after initial treatment, and diagnosis and treatment of recurrent or metastatic disease. Neo-adjuvant therapy refers to treatment given as an initial treatment to shrink a tumor before the main treatment, which is usually surgery, is given. They include chemotherapy, radiation therapy, and hormone therapy. These estimates all come from the estimates used in the study by Will et al. (2000), who reference source data for the therapeutic algorithms at initial diagnosis including the Saskatchewan Cancer Foundation (1993), Surveys of Canadian Oncologists (1994), and breast cancer experts. Data on follow-up after initial treatment came from Surveys of Canadian Oncologists (1994) and breast cancer experts. Diagnosis and treatment of recurrent or metastatic disease data came from the Saskatchewan Cancer Foundation – Special Chart Reviews (1985-1992) and the Ottawa Regional Cancer Centre – Special Chart Reviews (1996-1997). Will et al. (2000) noted that breast cancer treatment is dependent on the stage of the disease at initial diagnosis.

Table 6.1 Cumulative Rate of Breast Cancer Neo-Adjuvant Chemotherapy (Neo Adj Chemo)	
No Neo-adjuvant chemotherapy	**Neo-adjuvant chemotherapy**
0.85	0.15

Table 6.2 Cumulative Rate of Breast Surgery				
Stage	**Age at treatment**	**Breast Surgery**		
		No Surgery	**Breast-conserving surgery (lumpectomy)**	**Mastectomy**
I	(min,50)	0	0.8	0.2
I	[50,max)	0	0.7	0.3
II node	(min,50)	0	0.6	0.4

negative				
II node negative	[50,max)	0	0.45	0.55
II node positive	(min,50)	0.15	0.4	0.45
II node positive	[50,max)	0.25	0.1	0.65
III+IV	(min,50)	0.5	0.3	0.2
III+IV	[50,max)	0.6	0.2	0.2

Table 6.3 Cumulative Rate of Adjuvant Radiation Therapy (XRT)				
			Radiation Therapy	
Stage	**Age at treatment**	**Breast Surgery**	**No Radiation Therapy**	**Radiation Therapy**
I	(min,50)	No Surgery	1	0
I	(min,50)	Breast-Conserving Surgery	0.1	0.9
I	(min,50)	Mastectomy	1	0
I	[50,max)	No Surgery	1	0
I	[50,max)	Breast-Conserving Surgery	0.2	0.8
I	[50,max)	Mastectomy	1	0
II node negative	(min,50)	No Surgery	1	0
II node negative	(min,50)	Breast-Conserving Surgery	0.05	0.95
II node negative	(min,50)	Mastectomy	0.85	0.15
II node negative	[50,max)	No Surgery	1	0
II node negative	[50,max)	Breast-Conserving Surgery	0.25	0.75
II node negative	[50,max)	Mastectomy	0.85	0.15
II node positive	(min,50)	No Surgery	0.15	0.85
II node positive	(min,50)	Breast-Conserving Surgery	0.3	0.7
II node positive	(min,50)	Mastectomy	0.65	0.35
II node positive	[50,max)	No Surgery	0.55	0.45
II node positive	[50,max)	Breast-Conserving Surgery	0.35	0.65
II node positive	[50,max)	Mastectomy	0.5	0.5
III+IV	(min,50)	No Surgery	0.7	0.3
III+IV	(min,50)	Breast-Conserving Surgery	0.7	0.3
III+IV	(min,50)	Mastectomy	0.7	0.3
III+IV	[50,max)	No Surgery	0.75	0.25
III+IV	[50,max)	Breast-Conserving Surgery	0.75	0.25

III+IV	[50,max)	Mastectomy	0.75	0.25

Table 6.4 Cumulative Rate of Breast Cancer Local Treatment				
	Breast Cancer Urgent Local Treatment			
Age at treatment	**No urgent treatment**	**Urgent surgery**	**Urgent radiotherapy**	**Urgent surgery and radiotherapy**
(min,50)	0.12	0.5104	0.1584	0.2112
[50,max)	0.2	0.496	0.128	0.176

Local treatments are used to remove or destroy the disease within the breast and surrounding regions, such as in the lymph nodes. These treatments include surgery, either mastectomy or lumpectomy and radiation therapy.

Table 6.5 Cumulative Rate of Local and Systematic Treatment					
		Adjuvant therapy			
Age at treatment	**Urgent local treatment**	**No adjuvant therapy**	**Hormone therapy**	**Chemotherapy**	**Hormone therapy and chemotherapy**
(min,50)	No urgent treatment	0.17	0.166	0.664	0
(min,50)	Urgent surgery	0.58	0.2436	0.1344	0.0462
(min,50)	Urgent radiotherapy	0.58	0.2436	0.1344	0.0462
(min,50)	Urgent surgery and radiotherapy	0.58	0.2436	0.1344	0.0462
[50,max)	No urgent treatment	0.17	0.5229	0.2739	0.0332
[50,max)	Urgent surgery	0.59	0.3854	0.0164	0.0082
[50,max)	Urgent radiotherapy	0.59	0.3854	0.0164	0.0082
[50,max)	Urgent surgery and radiotherapy	0.59	0.3854	0.0164	0.0082

Local treatments are used to remove or destroy the disease within the breast and surrounding regions, such as in the lymph nodes. These treatments include surgery, either mastectomy or lumpectomy and radiation therapy. Systemic treatments are used to destroy or control cancer cells all over the body and include chemotherapy drugs that kill cancer cells, hormone therapy (e.g. tamoxifen) drugs that prevent hormones such as estrogen from promoting growth of the disease cells that may remain after surgery, and biological therapy (e.g. herceptin) that target disease cells that have high levels of HER2.

Table 6.6 Cumulative Rate of Metastatic Breast Cancer Local Treatment					
		Local Therapy			
Age at treatment	**Breast Cancer Metastasis Sites**	**No Urgent Treatment**	**Urgent Surgery**	**Urgent Radiation Therapy**	**Urgent Surgery and Radiation Therapy**
(min,50)	soft tissue metastasis	0.75	0	0.2	0.05
(min,50)	bone metastasis	0.25	0.075	0.6375	0.0375
(min,50)	visceral metastasis	0.6	0.06	0.3	0.04
[50,max)	soft tissue metastasis	0.6	0.14	0.2	0.06
[50,max)	bone metastasis	0.45	0.0275	0.4675	0.055
[50,max)	visceral metastasis	0.75	0.075	0.15	0.025

Table 6.7 Cumulative Rate of Metastatic Breast Cancer Systemic Treatment					
		Adjuvant Therapy			
Age at treatment	Breast Cancer Metastasis Sites	No adjuvant therapy	Hormone therapy	Chemotherapy	Hormone therapy and chemotherapy
(min,50)	Soft tissue metastasis	0.15	0.255	0.425	0.17
(min,50)	Bone metastasis	0.2	0.6	0.16	0.04
(min,50)	Visceral metastasis	0.05	0.285	0.38	0.285
[50,max)	Soft tissue metastasis	0.2	0.56	0.16	0.08
[50,max)	Bone metastasis	0.2	0.72	0.04	0.04
[50,max)	Visceral metastasis	0.15	0.4675	0.255	0.1275

Table 6.8 Cumulative Rate of Metastasis By Hospitalization for Breast Cancer			
Age at treatment	Breast Cancer Metastasis Sites	Hospitalization for Breast Cancer	
		No hospital	Hospital
(min,50)	Soft tissue metastasis	0.9	0.1
(min,50)	Bone metastasis	0.9	0.1
(min,50)	Visceral metastasis	0.7	0.3
[50,max)	Soft tissue metastasis	0.9	0.1
[50,max)	Bone metastasis	0.9	0.1
[50,max)	Visceral metastasis	0.7	0.3

Table 6.9 Cumulative Rate of Metastasis by Diagnostic Surgery at Metastasis Site			
		Breast Diagnostic Surgery (Biopsy)	
Age at treatment	Breast Cancer Metastasis Sites	No diagnostic surgery at metastasis site	Diagnostic surgery at metastasis site
(min,50)	Soft tissue metastasis	0.4	0.6
(min,50)	Bone metastasis	0.95	0.05
(min,50)	Visceral metastasis	0.7	0.3
[50,max)	Soft tissue metastasis	0.4	0.6
[50,max)	Bone metastasis	0.95	0.05
[50,max)	Visceral metastasis	0.7	0.3

Table 6.10 Cumulative Rate of Metastasis By Other Procedures at Metastasis Site			
		Other Breast Cancer Procedures	
Age at treatment	Breast Cancer Metastasis Sites	No other procedures at metastasis site	Other procedures at metastasis site
(min,50)	Soft tissue metastasis	1	0
(min,50)	Bone metastasis	1	0
(min,50)	Visceral metastasis	0.85	0.15
[50,max)	Soft tissue metastasis	1	0
[50,max)	Bone metastasis	1	0
[50,max)	Visceral metastasis	0.85	0.15

7. Costs

According to Will et al. (2000), provincial cancer registry data, provincial fee schedules and special costing studies and chart reviews were used to augment the national databases and to determine costs. Based on therapeutic modalities, costs associated with diagnosis, treatment and follow-up of local recurrence were determined from the Canadian Institute for Health Information (CIHI), the Ontario Health Insurance Plan (OHIP), and the Ontario Case Cost Project (OCCP). For costs associated with the diagnosis and treatment of stage IV and metastatic disease, Will et al. (2000) conducted a retrospective review of 500 Saskatchewan charts to determine types and frequency of interventions in breast cancer patients diagnosed with a recurrence and 100 charges extracted by personnel from the Ottawa Regional Cancer Centre to determine information on types of surgery performed with metastatic disease and length of hospital stay. Costs of therapy were then extracted from the same sources.

Will et al. (2000) note that radiotherapy costs were estimated from the study by Earl et al. (1999).[20] *Chemotherapy costs (drugs and administration) came from the Ottawa Civic Hospital (1995) and the Ottawa General Hospital (1995). Fees for physician`s services, diagnostic and surgical tests and procedures came from the Ontario Fee Schedule (1995). Hospital per diem rates by case mix groups and terminal care came from the OCCP (1993-1995). Facility overhead costs came from came from the 1988 National cancer Institute of Canada Clinical trial (was adjusted to CAD 2000). Hormonal therapy costs came from Ottawa pharmacies.*

Costs for ongoing care are inclusive of all treatments initiated 3 months after the diagnosis and treatment of metastatic disease until 3 months prior to death (these last three months are considered to be the terminal care phase of the illness). Ongoing care costs include those for hospitalisation, inpatient and outpatient medical services, and treatment with radiotherapy or chemotherapy. The costs of home care, oral medications, or out of pocket expenses resulting from visits to healthcare institutions are not included. These data came from the Manitoba Medical Service Foundation, Manitoba Cancer Treatment and Research Foundation (MCTRF) (1990), Manitoba Health Service Insurance Plan, Statistics Canada`s National Person-oriented Database of Hospital Discharges (1992-1994), and the OCCP. Further details on how Will et al. (2000) estimated the monthly costs for ongoing care are included in their paper.

Terminal care costs include those health care system costs that are incurred in the three months prior to death from breast cancer. The cost components are similar to those for ongoing care. According to Will et al. (2000), Statistics Canada's 1993-94 hospital discharge data was used to calculate dying breast cancer patients average number of days spent in hospital. The OCCP was used to calculate the average cost for this time period using the cost per encounter for patients who died between June 1993 and March 1995, based on the Case Mix Groups (CMGs) 429±433, and 443±445. Data from the Manitoba database was used to estimate the

[20] Earle C, Coyle D, Smith A, Agboola O, Evans WK. The cost of radiotherapy at an Ontario regional cancer centre. Crit Rev Oncol/Hematol.1999, 32; 87-93.

proportion of patients receiving palliative radiotherapy and the number of fractions per patient at a cost of $138 per fraction (CAD 1990). The cost of consultation, partial assessment and weekly blood work were added to the total cost of radiotherapy treatment.

Table 7.1 Cost of Initial Work-Up of Breast Cancer
328.98

Table 7.2 Cost of Initial Staging of Breast Cancer			
Stage			
I	**II node negative**	**II node positive**	**III+IV**
141.82	256.72	334.43	334.43

Table 7.3 Cost of Initial Neo-Adjuvant Chemo Therapy	
Neo-Adjuvant Chemotherapy	
No Neo-Adjuvant Chemotherapy	**Neo-Adjuvant Chemotherapy**
0	4087.96

Table 7.4 Cost of Initial Breast Surgery				
		Breast Surgery		
Stage	**Age at Treatment**	**No Surgery**	**Breast-Conserving Surgery**	**Mastectomy**
I	(min,50)	0	665.57	706.74
I	[50,max)	0	665.57	706.74
II node negative	(min,50)	0	665.57	706.74
II node negative	[50,max)	0	665.57	706.74
II node positive	(min,50)	0	665.57	706.74
II node positive	[50,max)	0	665.57	706.74
III+IV	(min,50)	0	665.57	706.74
III+IV	[50,max)	0	665.57	706.74

Table 7.5 Cost of Initial Hospitalization by Age at Treatment and Breast Surgery				
		Breast Surgery		
Stage	**Age at Treatment**	**No Surgery**	**Breast-Conserving Surgery**	**Mastectomy**
I	(min,50)	0	3822.39	4217.67
I	[50,max)	0	4446.89	4643.16
II node negative	(min,50)	0	3822.39	4217.67
II node negative	[50,max)	0	4446.89	4643.16
II node positive	(min,50)	0	3822.39	4217.67
II node positive	[50,max)	0	4446.89	4643.16
III+IV	(min,50)	0	3822.39	4217.67
III+IV	[50,max)	0	4446.89	4643.16

Table 7.6 Cost of Initial Radiation Therapy by Breast Surgery			
	Breast Surgery		
Stage	No Surgery	Breast-Conserving Surgery	Mastectomy
I	0	3867.29	2999.16
II node negative	0	3867.29	2999.16
II node positive	3903.46	3903.46	2999.16
III+IV	2999.16	2999.16	2999.16

Table 7.7 Cost of Initial Chemotherapy by Breast Surgery			
	Breast Surgery		
Stage	No Surgery	Breast-Conserving Surgery	Mastectomy
I	0	2375.99	2565.45
II node negative	0	2596.78	2796.69
II node positive	4087.96	3961.9	4087.96
III+IV	0	0	0

Table 7.8 Cost of Initial Hormone Therapy by Age at Treatment		
	Age at Treatment	
Stage	(min,50)	[50,max)
I	35.6	35.6
II node negative	35.6	35.6
II node positive	35.6	35.6
III+IV	0	0

Table 7.9 Cost of Initial Follow-Up by Time (Months)						
	Month Grouping of Initial Follow-Up					
Stage	(min,12)	[12,24)	[24,36)	[36,48)	[48,60)	[60,max)
I	38.92	34.99	28.72	25.52	23.84	0
II node negative	38.92	34.99	28.72	25.52	23.84	0
II node positive	40.18	39.24	34.73	30.35	26.82	0
III+IV	0	0	0	0	0	0

Table 7.10 Cost of Localized Breast Cancer Diagnosis
220.89

Table 7.11 Cost of Localized Breast Cancer Surgery by Age at Treatment	
Age at Treatment	
(min,50)	[50,max)
356.65	356.65

Table 7.12 Cost of Hospitalization for Localized Breast Cancer by Age at Treatment	
Age at Treatment	
(min,50)	**[50,max)**
5592.51	5592.51

Table 7.13 Cost of Radiation Therapy for Localized Breast Cancer by Age at Treatment	
Age at Treatment	
(min,50)	**[50,max)**
3584.6	3446.6

Table 7.14 Cost of Chemotherapy for Localized Breast Cancer by Age at Treatment	
Age at Treatment	
(min,50)	**[50,max)**
4046.33	3969.69

Table 7.15 Cost of Hormone Therapy for Localized Breast Cancer by Age at Treatment	
Age at Treatment	
(min,50)	**[50,max)**
31.14	61.48

Table 7.16 Cost of Follow-up for Localized Breast Cancer by Time of Follow-up (Months)				
Months				
(min,12)	[12,24)	[24,36)	[36,48)	[48,max)
68.95	68.95	51.01	51.01	41.4

Table 7.17 Cost of Care for Terminal Breast Cancer		
Ongoing or Terminal Care		
Hospital Stay	**Other Medical**	**Palliative Radiation Therapy**
15258.9	188	83.92

Appendix 2: Other Results

Table 1. Screening mammography detection rates

Screening option	Age eligibility (screen frequency)	True negatives per 1,000 women	False negatives per 1,000 women	True positives per 1,000 women	False positives per 1,000 women	True Positive Rate	True Negative Rate
A	50-69 (2)	7,289.00	12.73	57.52	569.28	0.82	0.99
B	40-49 (1) 50-69 (2)	15,248.30	18.51	70.58	1,113.89	0.79	1.00
C	50-69 (2) 70-79 (2)	10,331.90	17.50	92.48	732.51	0.84	0.99
D	40-49 (1) 50-69 (2) 70-79 (2)	18,204.30	23.13	104.61	1,270.28	0.82	0.99
E	50-69 (1)	13,938.08	18.54	82.07	1,031.67	0.82	0.99
F	40-49 (1) 50-69 (1)	21,815.74	24.18	93.87	1,570.73	0.80	1.00
G	50-69 (1) 70-79 (1)	19,764.73	24.95	129.31	1,339.06	0.84	0.99
H	40-49 (1) 50-69 (1) 70-79 (1)	27,720.67	30.71	141.59	1,883.73	0.82	0.99
I	40-49 (2) 50-69 (2)	11,402.35	16.13	64.36	867.51	0.80	0.99
J	40-49 (2) 50-69 (2) 70-79 (2)	14,320.04	20.67	97.99	1,021.90	0.83	0.99

Notes: Per 1,000 women refers to a total population of 6 million women and not just the screened population**.**

Table 2 Sensitivity analysis results: sequential analysis using $48 cost per screen

Screening option	Age eligibility (screen frequency in years)	Total LYGs	Total LYGs per 1,000 women†	Total excess (vs no screen) health care cost ($ millions CAD)	Total excess health care cost ($CAD) per 1,000 women†	ICER vs no screening (cost in $ CAD per LYG)	Sequential ICER (cost in $ CAD per LYG)
No screen	-	-		-		-	
A	50-69 (2)	9,104	1.5	129.2	21,540	14,192	14,360
C	50-69 (2) 70-79 (2)	10,287	1.7	153.5	25,580	14,921	20,200
J	**40-49 (2)** **50-69 (2)** **70-79 (2)**	**16,421**	**2.7**	**284.4**	**47,390**	**17,316**	**21,810**
D	40-49 (1) 50-69 (2) 70-79 (2)	17,824	3.0	406.9	67,820	22,831	68,100
H	40-49 (1) 50-69 (1) 70-79 (1)	18,370	3.1	548.3	91,380	29,847	235,600
I	40-49 (2) 50-69 (2)	15,248	2.5	260.4	43,400	17,075	Extended dominance through J and A
B	40-49 (1) 50-69 (2)	16,630	2.8	382.7	63,780	23,012	Extended dominance through D and A
E	50-69 (1)	9,698	1.6	244.1	40,690	25,174	Dominated by C
G	50-69 (1) 70-79 (1)	10,810	1.8	292.6	48,770	27,069	Dominated by I and J
F	40-49 (1) 50-69 (1)	17,218	2.9	498.5	83,080	28,949	Dominated by D

† Total population of 6 million women followed in the model.

Δ = difference (gain) ; LYGs= life-years gained; ICER = incremental cost-effectiveness ratio; discount rate of 5% applied to health benefits and costs **.**

Table 3 Sensitivity analysis results: sequential analysis using $125 cost per screen

Screening option	Age eligibility (screen frequency in years)	Total LYGs	Total LYGs per 1,000 women †	Total excess (vs no screen) health care cost ($ millions CAD)	Total excess health care cost ($CAD) per 1,000 women†	ICER vs no screening (cost in $ CAD per LYG)	Sequential ICER (cost in $ CAD per LYG)
No screen	-	-		-		-	
A	**50-69 (2)**	**9,104**	**1.5**	**332.2**	**55,370**	**36,492**	**36,492**
C	50-69 (2) 70-79 (2)	10,287	1.7	391.9	65,310	38,093	50,414
J	40-49 (2) 50-69 (2) 70-79 (2)	16,421	3.0	741.3	123,550	45,143	56,968
D	40-49 (1) 50-69 (2) 70-79 (2)	17,824	2.5	1,058.6	176,438	59,394	226,180
H	40-49 (1) 50-69 (1) 70-79 (1)	18,370	2.8	1,411.2	235,203	76,822	645,769
I	40-49 (2) 50-69 (2)	15,248	2.7	682.4	113,733	44,753	Extended dominance through J and A
B	40-49 (1) 50-69 (2)	16,630	3.1	999.0	166,505	60,074	Extended dominance through D and A
E	50-69 (1)	9,698	1.6	620.9	103,475	64,018	Dominated by C
G	50-69 (1) 70-79 (1)	10,810	1.8	739.0	123,163	68,361	Dominated by I
F	40-49 (1) 50-69 (1)	17,218	2.9	1,289.8	214,962	74,908	Dominated by D

† Estimated value per capita (i.e. woman) with total population of 6 million women followed in the model.
Δ = difference (gain) ; LYGs= life-years gained; ICER = incremental cost-effectiveness ratio; discount rate of 5% applied to health benefits and costs

Table 4. Sensitivity analysis results: sequential analysis using 0% discount rate

Screening option	Age eligibility (screen frequency in years)	Total LYGs	Total LYGs per 1,000 women†	Total excess (vs no screen) health care cost ($ millions CAD)	Total excess health care cost ($CAD) per 1,000 women†	ICER vs no screening (cost in $ CAD per LYG)	Sequential ICER (cost in $ CAD per LYG)
No screen		-	-	-	-	-	
A	50-69 (2)	366,397	61.1	3,568.5	594,747	9,739	9,739
I	40-49 (2) 50-69 (2)	507,540	84.6	5,453.0	908,825	10,744	13,351
J	40-49 (2) 50-69 (2) 70-79 (2)	588,267	98.0	6,856.3	1,142,722	11,655	17,384
D	**40-49 (1) 50-69 (2) 70-79 (2)**	**626,761**	**104.5**	**8,709.3**	**1,451,545**	**13,896**	**48,136**
H	40-49 (1) 50-69 (1) 70-79 (1)	667,164	111.2	13,509.5	2,251,578	20,249	118,808
C	50-69 (2) 70-79 (2)	447,562	74.6	4,987.1	831,183	11,143	Extended dominance through I and A
E	50-69 (1)	406,161	67.7	6,866.0	1,144,327	16,905	Dominated by C, I, and J
G	50-69 (1) 70-79 (1)	487,370	81.2	9,735.8	1,622,627	19,976	Dominated by I, B, J, and D
B	40-49 (2) 50-69 (2) 70-79 (2)	544,678	90.8	7,293.4	1,215,562	13,390	Dominated by J
F	40-49 (1) 50-69 (2)	584,427	97.4	10,598.4	1,766,398	18,135	Dominated by J and D

† Estimated value per capita (i.e. woman) with total population of 6 million women followed in the model.
Δ = difference (gain) ; LYGs= life-years gained; ICER = incremental cost-effectiveness ratio

Table 5. Sensitivity analysis results: sequential analysis using 10% discount rate

Screening option	Age eligibility (screen frequency in years)	Total LYGs	Total LYGs per 1,000 women†	Total excess (vs no screen) health care cost ($ millions CAD)	Total excess health care cost ($CAD) per 1,000 women†	ICER vs no screening (cost in $ CAD per LYG)	Sequential ICER (cost in $ CAD per LYG)
No screen		-	-	-	-	-	
A	**50-69 (2)**	**271.39**	**0.05**	**11.73**	**1,955**	**43,222**	**43,222**
C	50-69 (2) 70-79 (2)	289.18	0.05	12.66	2,110	43,779	52,277
J	40-49 (2) 50-69 (2) 70-79 (2)	631.19	0.11	35.55	5,925	56,322	66,928
D	40-49 (1) 50-69 (2) 70-79 (2)	696.68	0.12	55.4	9,233	79,520	303,100
H	40-49 (1) 50-69 (1) 70-79 (1)	698.31	0.12	66.33	11,055	94,986	6,705,521
G	50-69 (1) 70-79 (1)	290.71	0.05	23.35	3,892	80,321	Extended dominance through J and A
I	40-49 (2) 50-69 (2)	613.62	0.10	34.64	5,773	56,452	Extended dominance through J and A
B	40-49 (1) 50-69 (2)	678.76	0.11	54.47	9,078	80,249	Extended dominance through D and A
E	50-69 (1)	275.44	0.05	21.53	3,588	78,166	Dominated by C
F	40-49 (1) 50-69 (1)	681.82	0.11	64.4	10,733	94,453	Dominated by D

† Estimated value per capita (i.e. woman) with total population of 6 million women followed in the model.
Δ = difference (gain) ; LYGs= life-years gained; ICER = incremental cost-effectiveness ratio

Table 6. Sensitivity analysis results: sequential analysis using 50% initial participation rate

Scenario	Age eligibility (screen frequency in years)	Total LYGs	Total LYGs per 1,000 women†	Total excess (vs no screen) health care cost ($millions CAD)	Total excess health care cost ($CAD) per 1,000 women†	ICER vs no screening (cost in $ CAD per LYG)	Sequential ICER (cost in $ CAD per LYG)
No screen	-	-	-	-	-	-	-
A	50-69 (2)	8,417	1.4	354.6	59,104	42,133	42,133
C	50-69 (2) 70-79 (2)	9,476	1.6	385.5	64,245	40,680	68,056
J	**40-49 (2)** **50-69 (2)** **70-79 (2)**	**15,262**	**2.5**	**588.3**	**98,045**	**38,544**	**35,046**
F	40-49 (1) 50-69 (1)	16,247	2.7	681.1	113,512	41,919	94,201
D	40-49 (1) 50-69 (2) 70-79 (2)	16,663	2.8	767.0	127,837	46,031	206,877
H	40-49 (1) 50-69 (1) 70-79 (1)	17,215	2.9	983.6	163,936	57,139	392,696
I	40-49 (2) 50-69 (2)	14,281	2.4	555.2	92,526	38,874	Extended dominance through J and A
E	50-69 (1)	9,023	1.5	521.2	86,862	57,764	Dominated by C
B	40-49 (1) 50-69 (2)	15,652	2.6	736.5	122,744	47,054	Dominated by F
G	50-69 (1) 70-79 (1)	9,986	1.7	591.4	98,566	59,221	Dominated by I and J

Table 7. Sensitivity analysis results: sequential analysis using 60% participation rate

Scenario	Age eligibility (screen frequency in years)	Total LYGs	Total LYGs per 1,000 women†	Total excess (vs no screen) health care cost ($ millions CAD)	Total excess health care cost ($CAD) per 1,000 women†	ICER vs no screening (cost in $ CAD per LYG)	Sequential ICER (cost in $ CAD per LYG)
No screen	-	-	-	-	-	-	-
A	50-69 (2)	8,681	1.4	359.3	59,885	41,390	41,390
C	50-69 (2) 70-79 (2)	9,713	1.6	390.1	65,015	40,161	29,826
J	**40-49 (2) 50-69 (2) 70-79 (2)**	**15,501**	**2.6**	**596.0**	**99,325**	**38,445**	**35,566**
D	40-49 (1) 50-69 (2) 70-79 (2)	16,873	2.8	782.2	130,367	46,358	135,780
H	40-49 (1) 50-69 (1) 70-79 (1)	17,423	2.9	998.8	166,464	57,326	393,952
B	40-49 (1) 50-69 (2)	15,863	2.6	751.6	125,264	47,379	Extended dominance through D and A
I	40-49 (2) 50-69 (2)	14,517	2.4	562.9	93,809	38,773	Extended dominance through J and A
E	50-69 (1)	9,293	1.5	530.1	88,347	57,043	Dominated by C
F	40-49 (1) 50-69 (1)	16,456	2.7	928.7	154,783	56,435	Dominated by D
G	50-69 (1) 70-79 (1)	10,250	1.7	600.3	100,055	58,571	Dominated by I and J

Table 8. Sensitivity analysis results: sequential analysis using 100% test sensitivity

Scenario	Age eligibility (screen frequency in years)	Total LYGs	Total LYGs per 1,000 women	Total excess (vs no screen) health care cost ($ millions CAD)	Total excess health care cost ($CAD) per 1,000 women	ICER vs no screening (cost in $ CAD per LYG)	Sequential ICER (cost in $ CAD per LYG)
No screen	-	-	-	-	-	-	-
A	50-69 (2)	9,329	1.6	365.0	60,825	39,121	39,121
C	50-69 (2) 70-79 (2)	10,383	1.7	396.0	65,986	38,130	29,359
I	40-49 (2) 50-69 (2)	15,970	2.7	570.3	95,053	35,713	31,220
J	**40-49 (2) 50-69 (2) 70-79 (2)**	**16,967**	**2.8**	**603.8**	**100,633**	**35,587**	**33,578**
D	40-49 (1) 50-69 (2) 70-79 (2)	18,511	3.1	795.7	132,611	42,984	95,247
H	40-49 (1) 50-69 (1) 70-79 (1)	19,375	3.1	1,014.0	169,001	54,879	252,542
F	40-49 (1) 50-69 (1)	17,583	2.9	707.3	117,878	40,225	Extended dominance through by D and A
E	50-69 (1)	9,541	1.6	540.0	90,007	56,601	Dominated by C
B	40-49 (1) 50-69 (2)	17,504	2.9	764.8	127,460	43,691	Dominated by F
G	50-69 (1) 70-79 (1)	10,418	1.7	610.9	101,824	58,642	Dominated by I and J

Table 9. Sensitivity analysis results: sequential analysis using 100% test specificity

Scenario	Age eligibility (screen frequency in years)	Total LYGs	Total LYGs per 1,000 women	Total excess (vs no screen) health care cost ($ CAD)	Total excess health care cost ($CAD) per 1,000 women	ICER vs no screening (cost in $ CAD per LYG)	Sequential ICER (cost in $ CAD per LYG)
No screen	-	-	-	-	-	-	-
A	50-69 (2)	8,825	1.5	362.8	60,459	41,105	41,105
C	50-69 (2) 70-79 (2)	9,863	1.6	393.4	65,570	39,889	29,549
J	**40-49 (2) 50-69 (2) 70-79 (2)**	**15,626**	**2.6**	**601.7**	**100,288**	**38,509**	**36,147**
D	40-49 (1) 50-69 (2) 70-79 (2)	17,008	2.8	793.0	132,174	46,627	138,353
H	40-49 (1) 50-69 (1) 70-79 (1)	17,803	3.0	1,029.2	171,541	57,814	297,333
I	40-49 (2) 50-69 (2)	14,641	2.4	568.6	94,768	38,837	Extended dominance through J and A
B	40-49 (1) 50-69 (2)	15,993	2.7	762.5	127,075	47,674	Extended dominance through D and A
E	50-69 (1)	9,444	1.6	536.5	89,418	56,807	Dominated by C
F	40-49 (1) 50-69 (1)	16,572	2.8	939.6	156,605	56,700	Dominated by D
G	50-69 (1) 70-79 (1)	10,399	1.7	606.7	101,125	58,345	Dominated by I and J

Table 10. Base case analysis 2 results: sequential analysis using added screening policy options including 3-year screening intervals

Scenario	Age eligibility (screen frequency in years)	Total LYGs	Total LYGs per 1,000 women	Total excess (vs no screen) health care cost ($ million CAD)	Total excess health care cost ($CAD) per 1,000 women	ICER vs no screening (cost in $ CAD per LYG)	Sequential ICER (cost in $ CAD per LYG)
No screen		-	-	-	-	-	-
A	50-69 (2)	9,104	1.5	199.0	33,170	21,860	21,860
C	50-69 (2) 70-79 (2)	10,287	1.7	235.4	39,240	22,889	30,786
J	**40-49 (2) 50-69 (2) 70-79 (2)**	**16,421**	**2.7**	**441.5**	**73,580**	**26,884**	**33,590**
D	40-49 (1) 50-69 (2) 70-79 (2)	17,824	3.0	631.0	105,170	35,402	135,096
H	40-49 (1) 50-69 (1) 70-79 (1)	18,370	3.1	845.0	140,830	45,998	391,868
B	40-49 (1) 50-69 (2)	16,630	2.8	594.6	99,100	35,755	Extended dominance through D and A
I	40-49 (2) 50-69 (2)	15,248	2.5	405.5	67,580	26,592	Extended dominance through J and A
G	50-69 (1) 70-79 (1)	10,810	1.8	446.1	74,350	41,266	Dominated by K
L	50-69 (3)	7,466	1.2	284,6	47,435	38,122	Dominated by A
M	40-49 (3)	6,856	1.1	330.7	55,110	48,227	Dominated by A, C, and L
E	50-69 (1)	9,698	1.6	373.7	62,280	38,530	Dominated by C
F	40-49 (1) 50-69 (1)	17,218	2.9	770.5	128,420	44,751	Dominated by D
P	40-49 (1) 50-69 (2) 70-79 (3)	16,829	2.8	781.4	130,233	46,432	Dominated by F and D
K	40-49 (3) 50-69 (3) 70-79 (3)	11,904	2.0	418.4	69,740	35,151	Dominated by I
N	40-49 (3) 50-69 (2) 70-79 (2)	13,170	2.2	494.4	82,407	37,543	Dominated by I
O	40-49 (2) 50-69 (3) 70-79 (3)	13,879	2.3	510.3	85,049	36,767	Dominated by I and J

Appendix 3: Discounting

Discounting in our analyses was applied to both health benefits (life-years gained) and total costs (combined cost of screening, diagnosis, and treatment) for each screening policy option. We estimated age-stratified discount factors using the following equation:

$$e - r * L$$

Where r represents the discount rate, for example r=0.05 when using a 5% discount rate.
And L represents the midpoint age for any given age range (category).

The age ranges (categories) in POHEM-BCS include:
[min,30[
[30,35[
[35,40[
[40,45[
[45,50[
[50,55[
[55,60[
[60,65[
[65,70[
[70,75[
[75,80[
[80,85[
[85,max]
All

Discounted values of age-stratified health benefits (life-years gained) compared to no screening, used the following simple equation:

$$LYG_i * e - r * L$$

Where *i* represents the age category. Total discounted life-years gained were calculated by summing all the discounted age-stratified health benefits.

Discounted values of age-stratified total excess costs compared to no screening, used the following simple equation:

$$Cost_i * e^{-r*L}$$

Where *i* represents the age category. Total discounted costs were calculated by summing all the discounted age-stratified excess costs.