Report from the Evaluation Indicators Working Group

GUIDELINES FOR MONITORING BREAST CANCER SCREENING PROGRAM PERFORMANCE

THIRD EDITION

Alberta Health Services

Alberta Breast Cancer Screening Program

BC Cancer Agency
CARE + RESEARCH
An agency of the Provincial Health Services Authority
Screening Mammography Program

Breast Screening Program
Programme de dépistage ducancerdu sein

Government of Nunavut

MANITOBA BREAST SCREENING PROGRAM

PROGRAMME MANITOBAIN DE DÉPISTAGE DU CANCER DU SEIN

New Brunswick / Nouveau Brunswick

NOVA SCOTIA BREAST SCREENING PROGRAM

ontario breast screening program
a cancer care ontario program

Prince Edward Island Breast Screening Program

PROGRAMME QUÉBÉCOIS DE DÉPISTAGE DU CANCER DU SEIN

SCREENING PROGRAM FOR BREAST CANCER
A PROGRAM OF THE SASKATCHEWAN CANCER FOUNDATION

Women's Wellness
BREAST SCREENING PROGRAM FOR NEWFOUNDLAND & LABRADOR

Yukon Mammography Program

CANADIAN PARTNERSHIP AGAINST CANCER   PARTENARIAT CANADIEN CONTRE LE CANCER

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Executive Summary

The purpose of the Report from the Evaluation Indicators Working Group: *Guidelines for Monitoring Breast Cancer Screening Program Performance, Third Edition* is to promote consistent calculation of key evaluation indicators for various monitoring and evaluation efforts across programs and over time. Detailed calculation methods and background for each evaluation indicators are outlined in this report. Each indicator and calculation method was chosen on the basis of their utility for assessing program performance and the ability to provide comparability between different organized breast cancer screening programs in Canada. Targets were chosen based on a detailed literature review, analysis of Canadian data and consensus of the evaluation indicators working group. Consistent calculation methods, standardized data collection and reporting, and striving towards established targets can help monitor and improve the quality of breast cancer screening across Canada.

When this report was near completion, the Canadian Task Force on Preventive Health Care released an updated guideline for breast cancer screening in average risk women aged 40 – 74 years (release date: November 21, 2011). Although the updated guideline may have implications for provincial/territorial screening programs in the future, for the purpose of the current report the Evaluation Indicators Working Group collectively decided to continue to assess the existing provincial/territorial screening practices. Targets will apply to women aged 50-69 screened by organized screening programs in Canada.

Background

INTRODUCTION

The principal goal of breast cancer screening is to reduce breast cancer mortality and morbidity. Regular mammography screening for women aged 50 to 69 is estimated to prevent approximately 25% of breast cancer deaths[1]; although recent reports have shown a lesser reduction in mortality rates[2-4]. The benefits of breast cancer screening are gradual and therefore, screening effectiveness cannot only be measured by reductions in mortality rates[5]. Evaluation indicators[†] related to both the benefits and harms that are valid, reliable and feasible to collect within the screening program are required for ongoing evaluation of breast cancer screening. Furthermore, these indicators provide a means to monitor the individual steps throughout the entire screening pathway in order to confirm that the short-term objectives of a successful screening program are met on an ongoing basis. This ensures that screening programs continually strive to increase the benefits of screening while minimizing the harms.

The Report from the Evaluation Indicators Working Group: *Guidelines for Monitoring Breast Cancer Screening Program Performance, Third Edition* will serve as a guide to promote consistent calculation of key evaluation indicators for various monitoring and evaluation efforts across programs and over time. Indicators used for the ongoing evaluation of organized breast cancer screening programs at the national level include participation rate, retention rate, annual screening rate, abnormal call rate, cancer detection rate, diagnostic interval, positive predictive value (PPV) of the screening mammography program, non-malignant biopsy rate, invasive tumour size, nodal status, post screen cancer rate and sensitivity of the screening mammography program. Provincial and territorial programs may compute additional evaluation indicators that are not monitored at the national level. The description of each evaluation indicator includes a definition, the context in which the indicator is relevant (rationale), method(s) of calculation, target objectives, and modification history. The indicators presented in this document were developed on the basis of recognized population screening principles, comparison to international standards, the experiences of professionals working in Canadian breast cancer screening programs, evidence from randomized controlled trials, demonstration projects, and observational studies. Information on aspects of evaluation indicators not measureable in the Canadian Breast Cancer Screening Database such as technical repeat rate, consent, privacy, health promotion/outreach and patient satisfaction, are not included in this document. These are addressed in the report 'Quality Determinants of Organized Breast Cancer Screening Programs in Canada'[6].

A revised edition of the Evaluation Indicators Report is published every five years or when there are significant changes in screening modalities, evaluation or calculation methods. As part of each review, the scientific evidence used to support each of the indicators requires systematic updating. A literature review protocol was designed to facilitate this update of scientific evidence by using both published and grey literature (Appendix A).

---

† An evaluation indicator is defined as: "a measurable variable (or characteristic) that can be used to determine the degree of adherence to a standard or the level of quality achieved."

ORGANIZED BREAST CANCER SCREENING IN CANADA

In 1988 a national workshop, consisting of expert representatives from government as well as key professional and voluntary organizations, recommended women aged 50 to 69 be invited to participate in an early detection program for breast cancer every two years[7]. In Canada, health care delivery is under provincial/territorial jurisdiction; thus, organized screening programs have been developed and implemented independently across the country. The first screening program started in British Columbia, in 1988. Programs have since been established in all provinces and the Yukon and Northwest Territories. Each program varies in their organization, screening modalities, recruitment methods, ages accepted for screening (outside the targeted 50-69 age group), and in the arrangements for diagnostic assessment following an abnormal screen (see Table 1). Organized programs in Canada typically involve the following four steps:

- Identification and invitation of the target population
- Provision of a screening examination
- Follow-up of any abnormalities detected at screening and
- Reminder to return for the next screening episode

To differing degrees, asymptomatic women in most provinces/territories in Canada can also access mammography outside the structure of the organized breast cancer screening programs. Referred to as "opportunistic screening", follow-up data from these women are not routinely monitored or evaluated on a national level. Therefore, reporting of the evaluation indicators outlined in this document is limited to organized screening programs.

TABLE 1
Organized breast cancer screening programs in Canada[a] – usual practices

PROVINCE/ TERRITORY	PROGRAM INCEPTION	CLINICAL BREAST EXAMINATION ON SITE	PROGRAM PRACTICES FOR WOMEN AGE 30+ IN ADDITION TO BIENNIAL MAMMOGRAPHY FOR WOMEN 50 69 YEARS		
			AGE GROUP	ACCEPT[b]	RECALL
NORTHWEST TERRITORIES	2003	No	30-39	No	N/A
			40-49	Yes	Annual
			70+	Yes	Biennial
YUKON TERRITORY	1990	No	30-39	No	N/A
			40-49	Yes	None
			70+	Yes	Biennial
BRITISH COLUMBIA	1988	No	30-39	Accept with physician referral	None
			40-49	Yes	Annual
			70-79	Yes	Biennial
			80+	Accept with physician referral	None
ALBERTA	1990	No	30-39	No	N/A
			40-49	Yes	Annual
			70-74	Yes	Biennial
			75+	Yes	None
SASKATCHEWAN	1990	No	30-39	No	N/A
			40-49	No[c]	N/A
			70-75	Yes	Biennial[d]
			76+	Yes	None
MANITOBA	1995	No	30-39	No	N/A
			40-49	Accept to mobile unit with physician referral	Biennial
			70+	Accept to mobile unit with physician referral	None
ONTARIO	1990	Yes[e]	30-49	Accept high risk women with physician referral who meet the eligibility criteria[f]	Annual
			70-74	Yes	Biennial
			75+	Yes	None

QUÉBEC	1998	No	30-34	No	N/A
			35-49	Accept with physician referral[g]	None
			70+	Accept with physician referral[g]	None
NEW BRUNSWICK	1995	No	30-39	No	N/A
			40-49	Accept with physician referral	None
			70+	Accept with physician referral	None
NOVA SCOTIA	1991	Yes[h]	30-39	No	N/A
			40-49	Yes	Annual
			70+	Yes	None
PRINCE EDWARD ISLAND	1998	Yes[h]	30-39	Accept high risk women with physician referral who meet the eligibility criteria[i]	Annual
			40-49	Yes	Annual
			70-74	Yes	Biennial
			75+	No	N/A
NEWFOUNDLAND AND LABRADOR	1996	Yes[j]	30-49	No	N/A
			70+	Accept if previously enrolled in program	None

a   Nunavut has not developed an organized breast cancer screening program.
b   Accept to program by self or physician referral but do not send out initial invitation letters.
c   Accept age 49 on the mobile if they would be 50 in that calendar year.
d   If previously enrolled in the program.
e   Nurse provides clinical breast examination at 28% of sites (October 2011).
f   High risk women aged 30-49 accepted as of July 2011. Women are considered high risk if they have one of (a) confirmed genetic mutation that increases risk (b) parent, sibling or child with this genetic mutation, (c) family history and ≥ 25% lifetime risk confirmed through genetic assessment, (d) received chest radiation therapy prior to age 30, and at least 8 years previously.
g   Accept with physician referral if done at a program screening centre, but is not officially considered within the program.
h   Modified examination only, performed by technologist at time of mammography.
i   Women aged 30-39 are accepted if mother was diagnosed within 10 years of their age.
j   Nurse.

THE CANADIAN BREAST CANCER SCREENING DATABASE

The Canadian Breast Cancer Screening Database (CBCSD) is the foundation for a national breast screening surveillance system to enable monitoring and evaluation of organized breast cancer screening across Canada. The CBCSD, derived from provincial breast screening program data, was developed in 1993 through a collaborative effort of the federal, provincial and territorial governments through the Canadian Breast Cancer Screening Initiative (CBCSI). It contains the data from all 10 provinces and 1 territory from program inception, and is updated every two years, providing standard, high-quality data for program evaluation. Data from the Yukon are not currently available and Nunavut does not have an organized screening program.

HISTORY OF THE EVALUATION INDICATORS IN CANADA

The Evaluation Indicators Working Group (EIWG) was formed in 1999 under the guidance of the National Committee for the CBCSI. The objective of this working group is to continually assess and develop evaluation indicators and quality measures to fulfill present and future recommendations. The EIWG was comprised of members from both the Quality Determinants Working Group (QDWG) and the National Committee with assistance from the Database Technical Subcommittee. In February 2000, the seven-member working group held a national workshop to assemble a group of knowledgeable stakeholders from the provinces/territories to refine the available indicators and evaluate their applicability in Canada. The efforts of this workshop resulted in the identification of 30 core evaluation indicators, target values for a subset of these indicators, as well as recommendations on practical means to collect and report on these data[8]. The Report from the Evaluation Indicators Working Group: Guidelines for Monitoring Breast Cancer Screening Program Performance documented the first set of guidelines for reporting a key set of "evaluation indicators"[9]. The second edition updated the guidelines in 2007 and was developed by the QDWG and invited guests[10].

EVALUATION INDICATORS WORKING GROUP: 3RD EDITION

A new EIWG was formed in December 2010 to develop the 3rd edition of the Guidelines for Monitoring Breast Cancer Screening Program Performance. This group included members of the QDWG, Database Technical Subcommittee and invited guests. During the first meeting of the EIWG, 24 potential indicators were identified to be evaluated for inclusion in the 3rd edition. Initial research involved the assessment of international guidelines from comparable countries to Canada, including the United States, United Kingdom, European Union, Australia and New Zealand. A more comprehensive literature review was then conducted on each indicator which included analysis of randomized controlled trials, observational studies and meta-analyses. A summary of the literature review process can be found in Appendix A. This review focused on evidentiary support for the evaluation indicators as well as estimates for associated targets. Based on the results of the literature review, 13 key indicators were identified for inclusion in the 3rd edition. Appropriate targets for each indicator were identified and approved by the EIWG.

EVALUATION INDICATOR DEVELOPMENT

In order to achieve reductions in breast cancer mortality and morbidity and to minimize potential harms associated with screening, the delivery of organized screening must be of high quality. The evaluation indicators and associated targets presented in this document were selected on the basis of their utility for assessing program progress toward these goals. The 13 evaluation indicators detailed here generally met the following criteria:

• Data for the indicator were regularly available;

• Data available for the indicator were of high quality;

• Meaningful targets could be defined on an evidentiary basis*;

• Indicators and targets would be useful for national comparison;

• Monitoring on a regular basis would be valuable; and

• Each indicator was widely accepted for use in program evaluation.

* No targets were set for annual screening rate, in situ cancer detection, non-malignant biopsy rate and sensitivity (see Evaluation Indicators under Review in Future Directions).

Indicators added

The following new indicators were included in the 3rd edition:

• Annual screening rate

• Non-malignant biopsy rate

  ▲ Number of non-malignant open and core biopsies per 1,000 screens

  ▲ Percentage of non-malignant biopsies which were open surgical biopsies

• Sensitivity of the screening mammography program

Indicators removed

The following indicators were removed from the 3rd edition:

• Benign to malignant open surgical biopsy ratio

• Benign open surgical biopsy rate

• Benign to malignant core biopsy ratio

• Benign core biopsy rate

The indicators related to separate open and core biopsies were removed given the increasing use of needle core biopsy as an intermediate step or alternative to an open biopsy[11, 12]. The previous indicators were not providing a concise method for estimating 'unnecessary tests' while allowing for comparability between provinces. When determining evaluation indicators the overall national perspective and comparability of the indicator was an important consideration. In particular, when analyzing the number of benign to malignant open biopsies within smaller provinces these ratios become meaningless. The benign to malignant core biopsy ratio and benign core biopsy rate were replaced by a new indicator which measures the rate of biopsy (core + open) with a non-malignant result. The open and core biopsy rate will be analyzed together as this provides a description of the number of biopsies women are exposed to following an abnormal screen. The percentage

of non-malignant open surgical biopsies within the total number of benign biopsies will also be recorded. This allows for a separate evaluation of open biopsy trends within benign cases. More detailed reports of separate core and open rates as well as benign to malignant ratios may be used for quality assurance with each screening program and for individual radiologists.

DATA SOURCES AND COLLECTION

Evaluation indicators are calculated using data from the CBCSD supplemented by routinely available national statistics, and population estimates. Currently, the CBCSD is enabled through the continued collaboration of the provinces and territories and the Centre for Chronic Disease Prevention and Control, Public Health Agency of Canada. Through the CBCSI, the CBCSD is managed by the Database Management Committee and implemented by the Database Technical Subcommittee.

Evaluation indicators span the clinical pathway from screening mammography through to diagnostic imaging, biopsy, and follow-up beyond mammography including diagnostic tests and cancer diagnosis. Collaboration with external practitioners to ensure women obtain appropriate follow-up is part of the services provided by organized breast cancer screening programs. Many, but not all, programs are directly linked to their provincial cancer registries so that cancer outcome data can be obtained. Further complicating the evaluation process, some programs experience delays in obtaining registry data. In addition, analyses have suggested that cancer pathology data vary from one program to another because of the different ways in which breast tumours are assessed, staged and reported[13]. This must be taken into account when the evaluation indicators are compared across programs.

APPLICATION

Through its monitoring and reporting role, the CBCSI produces a routine biennial report: *Organized Breast Cancer Screening in Canada: Report on Program Performance*[12]. The purpose of this report is to provide formal feedback to the programs regarding their relative performance and to assess the national picture. The approach to standardized evaluation indicators established in this document serves as a consistent template for reporting progress over time, as well as providing a set of targets for programs to strive toward.

CONTEXT OF EVALUATION INDICATORS

For the purposes of these guidelines, the target population for evaluation is the same as the national target population for organized screening. This population is defined as asymptomatic women between the ages of 50 and 69 years with no prior diagnosis of breast cancer. However, women screened at ages 40-49 and 70+ should be reported on for surveillance and monitoring purposes. Age is calculated based on the woman's age on the specific screen date unless otherwise indicated. The screening modality evaluated in these guidelines refers to mammography alone and does not include clinical breast examination (CBE). CBE is used by only two Canadian provinces and will not be included in future reports[12, 14]. Many targets are reported separately by initial and subsequent screens. Initial screens are the first screen within the organized screening program and may include women previously examined through opportunistic screening. Therefore, women categorized as initial

screeners may have previously received a mammogram outside the organized program. Subsequent screens include any return to the organized program regardless of compliance with their annual or biennial screening recommendation.

The targets and standards established in this document are intended to apply to the program's performance as a whole. Targets are calculated based on a two year (biennial screening) population unless otherwise specified. Most indicators are measured by screen (not individual women) and may include multiple screens for women on an annual screening recommendation. Indicators are measured per screen as this gives a more meaningful analysis of the screening program as a whole. It is also recognized that for some evaluation purposes it may be appropriate to stratify the target group in terms of demographic characteristics, 5-year age group, screening history, or by screening technology (film vs. digital). When indicators are used for comparison among Canadian programs or with programs in other countries, it may be necessary to age-standardize the results using the appropriate population standard.

Many of the evaluation indicators presented here only provide meaningful measures of program progress when considered in a broader context. In some cases, meeting ideal targets involves achieving a balance rather than continually working to increase or decrease a particular rate or indicator. For example, while increased participation and retention will always be desirable, targets set for indicators such as positive predictive value and abnormal call rate are set with the realization that we must tolerate some false-positive results in order to maximize cancer detection. At the same time, some evaluation indicators and targets should be considered in relation to other relevant data. For instance, the cancer detection rate must be considered in relation to the underlying cancer incidence rate among initial and subsequent screens in specific age groups. An illustration to clarify the relationship between these evaluation indicators is presented in Figure 1.

FIGURE 1
Pathway of a breast cancer screening program

→ WITHIN PROGRAM
⇢ OUTSIDE PROGRAM
✳ RELEVANT EVALUATION INDICATOR

PROGRAM PROMOTION TARGETING ASYMPTOMATIC WOMEN AGED 50-69[a]
Media campaign, Population-based invitations, Physician education,
Personal invitation to screening or recall for subsequent screens

✳ Participation rate,
Retention rate,
Annual screening rate

Program screening visit

✳ Time from screen to
notification of results

NORMAL ← Communicate result to
participant and physician

ABNORMAL
✳ Abnormal call rate,
Time from abnormal screen
to first diagnostic assessment

Non-malignant
biopsy rate ✳

Cancer detected
outside of program

Diagnostic follow-up

✳ Time from abnormal
screen to final diagnosis

✳ Post screen
invasive
cancer rate[c]

Normal/benign[b]

Program Detected
Cancer[b]

✳

Invasive and in situ cancer
detection rates, Screen-detected
invasive tumour size, Proportion
of node negative screen-detected
invasive cancer, Positive
predictive value of the screening
mammography program

a   Some women also undergo screening (opportunistic screening or diagnostic mammograms) and are diagnosed with cancer outside program.
b   Breast screening programs obtain final diagnoses from sources such as physicians, pathology reports, and cancer registries.
c   Cancers detected six-months after a screening event are considered to be post screen cancers at the national level.

Evaluation Indicators

PARTICIPATION RATE

Definition	Percentage of women who have a screening mammogram (within a 30-month period) as a proportion of the target population.
Context	In order for a screening program to reduce mortality in a population, that population must participate in the program in sufficient numbers. A participation rate of 70% and over was achieved in trials reporting substantial mortality reductions[1, 15]. Many factors can influence the participation rate, such as acceptability, accessibility, promotion of screening and the capacity of a screening program. Although women are recommended to screen within 24 months, it may take up to 30-months for women to be screened in many programs[12]. It is important to note that the participation rate does not represent all breast cancer screening in Canada. In most provinces "opportunistic screening" occurs outside the structure of the organized programs. Opportunistic screening varies by province and is not included in the organized screening participation rate. Estimates of 'screening mammography utilization' (organized + opportunistic screening) should also be calculated to demonstrate the total participation in Canada[12, 16].
Calculations	$$\frac{\text{Number of women within the age group as of Dec 31st of the last year, screened within a 30-month period}}{\text{Target population (Estimate of population as of Dec 31st of last year, from census/forecast – prevalent cases)}} \times 100 = \text{Participation Rate (\%)}$$
Details	This calculation method yields a point estimate as of December 31st of the last (most recent) year. The number of women screened (numerator) includes all women within the age group as of December 31st of the last year who were screened at least once within the prior 30 months. The time period in the numerator was changed to 30 months to take into consideration women not seen within the exact biennial (2 year) recommendation. The target population (denominator) should be obtained from the most recent census results and/or population estimates available from Statistics Canada. Ineligible women (previously diagnosed with breast cancer) should also be removed from the denominator using estimates from the Canadian Cancer Registry.
Targets	**Canada** **≥ 70% of the target population within 30 months.** Europe[17] > 70% of invited women age 50-69 within 30 months (acceptable level). United Kingdom[18] ≥ 70% of invited women age 50-70 within 36 months (minimum standard). Australia[19] ≥ 70% of eligible women age 50-69 within 24 months. New Zealand[20] ≥ 70% of eligible women age 50-69 within 24 months.
Evidence	Based on fundamental principles of population screening[1, 21-23], extrapolation from the results of randomized controlled trials[15, 24-26] and comparison to international calculation methods and results[27, 28].
Modification History	Introduced in 2002. Context updated in 2006. Definition, context, calculation and details updated in 2012.

RETENTION RATE

Definition	The estimated percentage of women aged 50-67* who returned for screening within 30 months.
Context	Optimal benefits of screening are achieved by regular participation in the screening program. In Canada, it is recommended that women aged 50-67* attend screening every 2 years. A high retention rate demonstrates that programs are actively recalling and retaining women at the recommended screening intervals. Although women are recommended to screen within 24 months, it may take up to 30 months for women to be screened in many programs. At present there is no indication that the benefits of screening are lost if screening occurs as much as 6 months after the recommended interval[29, 30]. Retention rate is affected by screening recommendations (annual vs. biennial), non-compliance rates, screen sequence, abnormal call rate, invasive diagnostic tests and false positives.
Calculations	Kaplan-Meier Method $s_t = 1 - (p_0 p_1 p_2 ... p_j)$ Multiply the probabilities at each timepoint (i.e. $t_0 ... t_j$) where $\quad p_j = \dfrac{(n^{\char94}_j - e_j)}{n^{\char94}_j}$ $\qquad\qquad n^{\char94}_j = n_j - c_j$ s_t = the estimated cumulative probability of returning to screen from baseline to the end of the study period; p_j = the estimated probability of not returning to screen at time t_j; e_j = the number of women who returned at time t_j; n_j = the number of women present just prior to time t_j; c_j = the number censored (because of death, breast cancer, or age limit > 67 years) at time t_j;
Details	Probability of returning to screen is estimated based on a study interval of 30 months from screen date. In cases of multiple screens per woman the most recent screen is used. *Women aged > 67 are included in the screening population but censored at their index screen as programs may not send recall letters to women outside this age group. Women lost to follow-up are excluded from the calculation.
Targets	**Canada** **≥ 75% screened within 30 months of an initial screen;** **≥ 90% screened within 30 months of a subsequent screen.** Europe[17] > 95% eligible women aged 50-69 are re-invited within the specified screening interval (acceptable level). United Kingdom[18] ≥ 90% women aged 50-70 whose first offered appointment is within 36 months of their previous screen (minimum standard). Australia[19] ≥ 75% initial rescreens within 27 months (age 50-67); ≥ 90% subsequent rescreens within 27 months (age 50-67). New Zealand[20] > 85% women screened in a program round are subsequently (if eligible) re-screened in the next program round (age 50-69); > 75% of women who return for a screen are re-screened between 20-24 months from their previous screen (age 50-69).
Evidence	Related to participation rate[24], sojourn time[29-31], screening adherence studies[32], extrapolation from the results of randomized controlled trials[24, 25] and comparison to international calculation methods and results[27, 28]. The calculation method is based on survival analysis[33].
Modification History	Introduced in 2002. Targets modified in 2006. Calculation and details changed in 2012.

ANNUAL SCREENING RATE

Definition	The estimated percentage of women aged 50-68[*] who returned to screen within 18 months of their previous screen.
Context	Optimal benefits of screening are achieved by regular participation in the screening program (every 2 years). However, women may be recalled on an annual basis due to increased risk of breast cancer (based on patient or screening history), provincial screening policy or other factors. Although women recommended for annual screening are usually recalled within 12 months, any screens that occur up to 18 months are considered 'annual'. It is important to monitor annual screening rates to understand its impact on abnormal screen and cancer detection rates as well as program capacity. Annual screening rates are also affected by non-compliance rates and screening history.
Calculations	Kaplan-Meier Method $s_t = 1 - (p_0 p_1 p_2 ... p_j)$ Multiply the probabilities at each timepoint (i.e. $t_0 ... t_j$) where $\quad p_j = \dfrac{(n\hat{}_j - e_j)}{n\hat{}_j}$ $\quad\quad\quad n\hat{}_j = n_j - c_j$ s_t = the estimated cumulative probability of returning to screen from baseline to the end of the study period; p_j = the estimated probability of not returning to screen at time t_j; e_j = the number of women who returned at time t_j; n_j = the number of women present just prior to time t_j; c_j = the number censored (because of age limit > 68 years) at time t_j;
Details	The probability of returning to screen is estimated based on a study interval of 18 months from screen date. This indicator should be calculated based on one screen year and in cases of multiple screens per woman the most recent screen should be used. [*]Women aged > 68 who returned to screen are included in the population but censored at their index screen as programs may not send recall letters to women outside this age group. Women lost to follow-up or those that did not return for a subsequent screen are excluded from the calculation.

Targets	**Canada** **No target**
	                % screened within 18 months of an initial screen;
	                % screened within 18 months of a subsequent screen.
	                (Surveillance and monitoring purposes only)
	
	Australia[19] ≤ 10% women (aged 50-69) are screened annually.

Evidence	Based on the impact of early recall[34, 35]. Related to sojourn time[29-31], extrapolation from the results of randomized controlled trials[24, 25] and comparison to international calculation methods[27, 28]. Calculation method based on survival analysis[33].
Modification History	Introduced in 2012.

ABNORMAL CALL RATE

Definition	Percentage of mammograms that are identified as abnormal at program screen.
Context	Abnormal call rate is an important indicator of the quality of the mammography image and interpretation. It is most meaningful when considered in the context of positive predictive value, cancer detection rate, post-screen cancer rate and the underlying breast cancer incidence rate. A high abnormal call rate can increase the false positive rate and result in unnecessary tests[36, 37]. Programs should strive to balance the number of abnormal calls with the number of cancers detected. This can be monitored by comparing the number of abnormal screens per extra cancer detected[37]. Programs with extremely low abnormal call rates should also be monitored as this may result in lower cancer detection and higher post-screen cancer rates[36, 38]. Abnormal call rates will generally be higher for first-time screens (which detect prevalent cancers) than for subsequent screens. It may also be affected by recommended screening interval (annual vs. biennial) as well as screening technology (digital vs. film)[39-41]. Further analysis may include stratification by these subgroups. The abnormal call rate can also be used to estimate a false positive rate and specificity as the majority of abnormal calls will be resolved as benign/normal.
Calculations	$$\frac{\text{Number mammograms identified as abnormal}}{\text{Number of screens}} \times 100 = \text{Abnormal Call Rate (\%)}$$
Details	Cases referred by clinical breast exam (CBE) alone will not be included in this calculation.
Targets	**Canada** — **< 10% (initial screen); < 5% (subsequent screens).** Europe[17] — < 7% (initial screen) (age 50-69) (acceptable level); < 5% (subsequent-regular screens) (age 50-69) (acceptable level). United Kingdom[18] — < 10% (initial screen) (age 50-70) (minimum standard); < 7% (subsequent screens) (age 50-70) (minimum standard). Australia[19, 20, 28] — < 10% (initial screen) (age 50-69); < 5% (subsequent screens) (age 50-69). New Zealand[20] — < 10% (initial screen) (age 50-69) (minimum); < 5% (subsequent screens) (age 50-69) (minimum).
Evidence	Based on studies of recall rate[37, 42] and literature reviews[1, 43-47]. Comparison to international calculation methods and results[27, 28, 48].
Modification History	Introduced in 2002. Definition, context and details modified in 2012.

INVASIVE CANCER DETECTION RATE

Definition	Number of invasive cancers detected per 1,000 screens.
Context	The cancer detection rate is important to evaluate how successful the program is at finding invasive cancers. It is most meaningful when considered in relation to the abnormal call rate, post-screen cancer detection rate, and the underlying rate of breast cancer in the eligible population[37]. Programs should strive to achieve the greatest number of cancers detected while limiting unnecessary tests and cancers missed at screen or assessment. Cancer detection rates will generally be higher for initial screens (which detect prevalent cancers) than for subsequent screens. However, women who received previous "opportunistic screening" outside the programs will contribute to a reduction in the invasive cancer detection rate. Established screening programs should have the majority of initial screens occurring among women in the youngest eligible age group. Therefore, it is suggested that further analysis be stratified by 5-year age group as the underlying incidence of breast cancer increases with age[1]. Cancer detection rates may also be affected by annual vs. biennial screening as well as screening technology (digital vs. film)[39-41]. Further analysis may include stratification by these subgroups.
Calculations	$$\frac{\text{Number of invasive cancers detected}}{\text{Number of screens per 1,000 screens}} \times 1{,}000 = \text{Invasive Cancer Detection Rate}$$
Details	Cancers detected by clinical breast exam (CBE) alone will not be included in this calculation. Cancers diagnosed more than 6 months following an abnormal screen are excluded from this indicator (outside the screening episode) and are counted as post-screen cancers. Women lost to follow-up are also excluded from the numerator and denominator. Invasive cancers include those with microinvasion. Once diagnosed with cancer, women are no longer eligible for screening in most programs and are excluded from this indicator. In the case of bilateral cancer, only the highest stage tumour will be counted in the numerator. Cancer detection rate is represented per 1,000 screens to provide a measure of screening program performance comparable to the calculations for other indicators. It is noted that provinces with a high number of annual screens may have individual women counted twice in the denominator when calculated over a two year period. Further analysis could include a calculation of the number of invasive cancers detected per 1,000 women screened.

Targets		
	Canada	**> 5.0 per 1,000 screens (initial screen);** **> 3.0 per 1,000 screens (subsequent screens).**
	Europe[17]	3 X the underlying, expected breast cancer incidence rate in the absence of screening (IR) (initial screens) (age 50-69) (acceptable level); 1.5 X the underlying, expected breast cancer incidence rate in the absence of screening (IR) (subsequent-regular screening) (age 50-69) (acceptable level).
	*United Kingdom[18]	≥ 2.7 per 1,000 screens (initial screen) (age 50-70) (minimum standard); ≥ 3.1 per 1,000 screens (subsequent screens) (age 50-70) (minimum standard).
	Australia[19]	≥ 50 per 10,000 women (initial screen) (age 50-69); ≥ 35 per 10,000 women (subsequent screens) (age 50-69).
	*Excludes microinvasive	

Evidence	Measured in studies of mammography cancer detection rates[5, 45-47, 49]. Targets were developed based on the experience of Canadian and international breast cancer screening programs[12, 18, 19, 27, 28, 48].
Modification History	Introduced in 2002. Context modified in 2006. Details and context updated in 2012.

IN SITU CANCER DETECTION

Definition	(a) Number of ductal carcinoma *in situ* (DCIS) cancers detected per 1,000 screens (b) Percentage of all cancers that are DCIS
Context	*In situ* carcinoma is a heterogeneous disease and not all cases of ductal carcinoma *in situ* will progress to invasive carcinoma. *In situ* cancer detection may be interpreted as an indicator of screening quality when considered in relation to the invasive cancer detection rate and underlying cancer incidence rate in the eligible population. Further analysis may examine DCIS by grade for a more complete classification of the tumour type. It is also suggested that the *in situ* cancer detection trends be measured over time and in association with the implementation of digital mammography[39-41].
Calculations	(a) $$\frac{\text{Number of } \textit{in situ} \text{ cancers detected}}{\text{Number of screens}} \times 1{,}000 = \textit{In situ} \text{ Cancer Detection Rate per 1,000 screens}$$ (b) $$\frac{\text{Number of } \textit{in situ} \text{ cancers detected}}{\text{Number of } \textit{in situ} + \text{invasive cancers detected}} \times 100 = \textit{In situ} \text{ cancers (\%)}$$
Details	Cancers detected by clinical breast exam (CBE) alone will not be included in this calculation. Cancers diagnosed more than 6 months following an abnormal screen are excluded from this indicator (outside the screening episode) and are counted as post-screen cancers. Women lost to follow-up are excluded from the numerator and denominator. Once diagnosed with cancer, women are no longer eligible for screening in most programs and therefore are excluded from this indicator. In the case of bilateral cancer only the highest stage tumour will be counted in the numerator. Cancer detection rate is represented per 1,000 screens to provide an indicator of screening program performance comparable to the calculations for other indicators. It is noted that provinces with a high number of annual screens may have individual women counted twice in the denominator when calculated over 2 year period. Further analysis could include a calculation of the number of *in situ* cancers detected per 1,000 women screened.
Targets	**Canada** — **(a) No target per 1,000 screens (initial screen); per 1,000 screens (subsequent screens). (Surveillance and Monitoring Purposes Only) (b) No target Percentage (initial screen); Percentage (subsequent screens). (Surveillance and Monitoring Purposes Only)** Europe[17] — 10% screen-detected cancers (age 50-69) (acceptable level) *United Kingdom[18] — ≥ 0.4 per 1,000 screens (initial) (age 50-70) (minimum standard); ≥ 0.5 per 1,000 screens (rescreen) (age 50-70) (minimum standard). Australia[19] — ≥ 12 per 10,000 women (initial screen) (age 50-69); ≥ 7 per 10,000 women (subsequent screens) (age 50-69). New Zealand[20] — 10-25% screen-detected cancers (age 50-69). *Includes microinvasive and LCIS*
Evidence	It is inappropriate to set specific targets for DCIS given the heterogeneity of this disease and the current evidence concerning the transition of all forms of DCIS to invasive cancer and the continually increasing sensitivity of screening techniques[50-52]. A target range of DCIS as a proportion of all cancers was used by some jurisdictions as an indicator of both over and under diagnosis of DCIS[17, 20, 53]. The context of this indicator was also defined based on evidence regarding *in situ* cancer detection trends[12, 54].
Modification History	Introduced in 2002. Definition, calculation and details changed in 2012.

DIAGNOSTIC INTERVAL

Definition	(a) Time from screen to notification of screen result. *Among abnormal screens:* (b) Time from abnormal screen to first diagnostic assessment. (c) Time from abnormal screen to definitive diagnosis.
Context	The wait time from screen to resolution is an important indicator of performance across the entire screening episode from index screen to final diagnosis. Programs should strive to achieve case resolution in a timely manner. An abnormal screen result is associated with anxiety and can have a negative psychological impact on a client, even if follow-up is ultimately benign/normal[55-57]. Moreover, excessive delay to diagnosis may worsen prognosis[58-62]. Therefore, work-up should be completed expeditiously. The time from abnormal screen to first diagnostic assessment and final diagnosis is affected by many factors including mammographic suspicion, type of diagnostic test performed, provincial and programmatic capacity, and the final diagnosis. The diagnostic interval can also be improved by patient navigation, 'fast track' or other referral systems[63-65]. However, many Canadian programs do not have integrated diagnostic capabilities, making management of the diagnostic interval more difficult.
Calculations	(a) Time from screen to notification of screen result = (Date notification was sent) – (screen date) $$\frac{\text{Number of notifications within the target time-range}}{\text{Total number of screens}} \times 100 = \text{Notifications within the target time-range (\%)}$$ *Among abnormal screens:* (b) Time from abnormal screen to first diagnostic assessment. (Date of first diagnostic assessment) – (screen date) $$\frac{\text{Number of first diagnostic assessments within the target time-range}}{\text{Total number of abnormal screens}} \times 100 = \text{First diagnostic assessments within the target time-range (\%)}$$ (c) Time from abnormal screen to definitive diagnosis = (Date of definitive diagnosis) – (screen date) $$\frac{\text{Number of definitive diagnoses within the target time-range}}{\text{Total number of abnormal screens}} \times 100 = \text{Definitive diagnoses within the target time range (\%)}$$
Details	The date notification was sent includes the date that a letter was sent or a phone call was made to the client. The time from abnormal screen to first diagnostic assessment includes referrals to primary care provider, surgical consults or any diagnostic test. The date of definitive diagnosis for cancer is the date of the first core or open biopsy to diagnose cancer (DCIS or invasive) or the first definitive fine needle aspiration (FNA) if there was no prior core or open biopsy. The date of definitive diagnosis for benign cases is the last test before a return to screening or before the recommendation for early recall. Time to first diagnostic assessment or definitive diagnosis does not include cases referred by clinical breast exam (CBE) alone. Cancers diagnosed more than 6 months following an abnormal screen are excluded from this indicator (outside the screening episode) and are counted as post-screen cancers. Benign/normal cases that took > 6 months to diagnose are calculated based on the last test prior to 6 months. Women lost to follow-up or with missing date information are excluded from the numerator and denominator. The total duration from abnormal screen to definitive diagnosis is separated into those with and without a tissue test. Cases are considered 'with tissue test' if an open or core biopsy was performed any time prior to 6 months from abnormal screen to definitive diagnosis. Cases without any diagnostic assessment are excluded from the numerator and denominator. Further analysis may also include an indicator of median wait times and the number of weeks needed to achieve 90% completion.

Targets	Canada	(a) ≥ 90% within 2 weeks; (b) ≥ 90% within 3 weeks; (c) ≥ 90% within 5 weeks if no tissue biopsy* performed; ≥ 90% within 7 weeks if tissue biopsy* performed.
	Europe[17]	90% screening mammography and result within ≤ 10 working days (age 50-69) (acceptable level).; 90% symptomatic mammography and result within ≤ 5 working days (age 50-69) (acceptable level); 90% result of screening mammography and offered assessment within ≤ 5 working days (age 50-69) (acceptable level).; 90% result of diagnostic mammogram and offered assessment within ≤ 5 working days (age 50-69) (acceptable level).; 90% assessment and issuing of results within ≤ 5 working days (age 50-69) (acceptable level).
	United Kingdom[18]	≥ 90% are sent their screening results within 2 weeks (age 50-70) (minimum standard); ≥ 90% attend an assessment centre within 3 weeks of their screening mammogram (age 50-70) (minimum standard); ≥ 90% women have a time interval between the decision to refer to a surgeon and surgical assessment of ≤ 1 week (age 50-70) (minimum standard).
	Australia[19,28]	≥ 90% receive a letter informing them of their results within 14 days of screening(age 50-69); ≥ 90% women requiring assessment attend an assessment visit within 28 days of their screening visit(age 50-69); ≥ 95% women attending assessment complete all assessment within a two week period (age 50-69)
	New Zealand[20]	> 90–95% of women can be notified within 10 working days of the screening mammogram (age 50-69); 90% of women are offered an assessment appointment within 15 working days of their final screening mammogram (age 50-69).
		Tissue biopsy does not include fine needle aspiration (FNA).
Evidence		Based on basic principles of screening[1, 66], evaluation of tumour progression and wait times[58-62] and patient quality of care research[55-57, 67-69]. New indicators developed from previous Canadian reports[8, 10].
Modification History		Introduced in 2002. Targets modified in 2006. Definition, context, calculations and details modified in 2012.

POSITIVE PREDICTIVE VALUE OF THE SCREENING MAMMOGRAPHY PROGRAM

Definition	Proportion of abnormal cases diagnosed with breast cancer (invasive or DCIS) after diagnostic work-up.
Context	Positive predictive value (PPV) is an indicator of the predictive validity of screening. The factors that influence cancer detection rate and abnormal call rate must also be taken into consideration when evaluating a program's PPV[70]. PPV tends to improve with subsequent screens because the initial screen establishes a normal baseline. Consequently, PPV tends to be lower among initial screens relative to subsequent screens.
Calculations	$$\frac{\text{Number of screen-detected cancers}}{\text{Number of abnormal screens}} \times 100 = \text{Positive Predictive Value (\%)}$$
Details	Cancers detected by clinical breast exam (CBE) alone will not be included in this calculation. Screen-detected cancers that took > 6 months to diagnose are excluded from this indicator (outside the screening episode) and are counted as post-screen cancers. Cases that were lost to follow-up are excluded from the numerator and denominator. Abnormal screens with benign result can include findings of LCIS, ADH, papilloma, radial scar and phyllodes tumour. Further analysis may include PPV of performed biopsies (% biopsies that resulted in cancer)[47] as well as comparisons between screening technologies (digital vs. film)[39-41, 41, 44].
Targets	**Canada** **≥ 5% (initial screen);** **≥ 6% (subsequent screens).** New Zealand[20] > 9% (aged 50-69) (all screens).
Evidence	Based on screening program evaluation studies and reports[1, 5, 43, 46, 47].
Modification History	Introduced in 2002. Definition, context and details modified in 2012.

NON-MALIGNANT BIOPSY RATE

Definition	(a) Number of non-malignant open and core biopsies per 1,000 screens (b) Percentage of non-malignant biopsies which were open surgical biopsies
Context	The non-malignant biopsy rate provides an indication of the quality of the pre-operative assessment. Programs should strive to limit the number of unnecessary tests while maximizing the screen-detected cancers (invasive and DCIS). Particularly invasive tests (core biopsies) and surgical procedures (open biopsies) should be monitored. This indicator is most meaningful when considered in relation to the abnormal call rate, cancer detection rate and the post-screen cancer rate. Abnormal screens and associated follow-up biopsy rates will generally be higher for initial screens than for subsequent screens. Variation in the use of open biopsy is reflected in the percentage of non-malignant biopsies which were open.
Calculations	(a) $\dfrac{\text{Number of non-malignant open + core biopsies}}{\text{Total number of screens}} \times 1{,}000 = $ Number of biopsies with non-malignant result per 1,000 screens (b) $\dfrac{\text{Number of non-malignant open biopsies}}{\text{Number of non-malignant open + core biopsies}} \times 100 = $ Percentage of non-malignant biopsies which are open surgical biopsies (%)
Details	Cancers detected by clinical breast exam (CBE) alone will not be included in this calculation. Cancers that took > 6 months to diagnose are excluded from this indicator (outside the screening episode). Biopsies that occurred > 6 months are also excluded from this indicator. Open biopsies include cases that went directly to surgical biopsy and those that underwent an inconclusive core biopsy prior to a definitive diagnosis by open surgical biopsy. This indicator includes multiple biopsies per person if applicable. Biopsies with non-malignant result can include benign, indeterminate/equivocal results, high risk lesions (LCIS, ADH, papilloma, radial scar or phyllodes tumour) or non-primary breast cancers (eg. lymphoma). Further analysis may include separation of non-malignant results by type of high risk lesion[71, 72]. Cases that were lost to follow-up or are missing biopsy test results are also excluded from the numerator and denominator.
Targets	**Canada** — **No targets** **per 1,000 screens (initial);** **per 1,000 screens (subsequent screen);** **Percentage open (initial);** **Percentage open (subsequent screen).** **(Surveillance and monitoring purposes only)** United Kingdom[18] — < 3.6 benign open biopsies per 1,000 screens (initial screen) (age 50-70) (minimum standard); < 2.0 benign open biopsies per 1,000 screens (subsequent screens) (age 50-70) (minimum standard). Australia[19] — ≤ 4.0% of women undergoing assessment are found not to have invasive cancer or DCIS after an open biopsy (initial screen) (age 50-69); ≤ 3.2% of women undergoing assessment are found not to have invasive cancer or DCIS after an open biopsy (subsequent screens) (age 50-69). New Zealand[20] — ≤ 3.5 open biopsies performed for benign disease* per 1,000 women (initial screen) (age 50-69); ≤ 1.6 open biopsies performed for benign disease* per 1,000 women (subsequent screens) (age 50-69). *Benign disease may include high risk lesions
Evidence	This indicator is currently for surveillance and monitoring purposes only. Based on methodology in screening program evaluation studies[73] and evidence of changing biopsy patterns[11, 12].
Modification History	Introduced in 2012.

SCREEN-DETECTED INVASIVE TUMOUR SIZE

Definition	Percentage of screen-detected invasive cancers with tumour size ≤ 15 mm in greatest diameter as determined by the best available evidence: 1) pathological, 2) radiological, and/or 3) clinical.
Context	Invasive tumour size is one of the best known prognostic indicators[1]. The purpose of mammography screening is to detect pre-clinical cancers before symptoms are apparent.
Calculations	$$\frac{\text{Number of screen-detected invasive tumours} \leq 15 \text{ mm}}{\text{Total screen-detected invasive cancers where tumour size was accessed}} \times 100 = \text{Invasive tumours} \leq 15 \text{ mm (\%)}$$
Details	Tumour size is measured at the time of diagnosis and excludes measurements taken after neo-adjuvant treatment. Cancers detected by clinical breast exam (CBE) alone will not be included in this calculation. Cancers that took > 6 months to diagnose are excluded from this indicator (outside the screening episode). Cases that were lost to follow-up or are missing tumour size information are also excluded from the numerator and denominator. Invasive tumours include microinvasive cancers. For bilateral cancers or cancers with multiple primary tumours in the same breast (synchronous tumours), the cancer with the highest stage is selected to report on invasive cancer tumour size.
Targets	**Canada** **> 50% screen-detected invasive tumours ≤ 15 mm.** *Europe[17] 50% invasive cancers < 15 mm in size (acceptable level) (age 50- 69). **United Kingdom[18] ≥ 1.5 per 1,000 screens (< 15 mm, initial screen) (minimum standard, based on pathological evidence only) (age 50-70); ≥ 1.7 per 1,000 screens (< 15 mm, subsequent screens) (minimum standard, based on pathological evidence only) (age 50-70). *Australia[19] ≥ 25 cancers per 10,000 women screened (≤ 15 mm) (based on pathological evidence only) (age 50-69). *New Zealand[20] > 50% of invasive cancers < 15 mm or 30.5 per 10,000 women screened are < 15 mm (initial screen) (age 50-69); > 50% of invasive cancers or 17.3 per 10,000 women screened are < 15 mm (subsequent screen) (age 50-69). *Includes cancer cases with unknown tumour size in the denominator **Excludes microinvasive
Evidence	Based on stage-specific prospective studies and trials[74-76] and comparison to international calculation methods and results[27, 28, 48].
Modification History	Introduced in 2002. Calculations and targets modified in 2006. Definition modified in 2012.

PROPORTION OF NODE NEGATIVE SCREEN-DETECTED INVASIVE CANCER

Definition	Proportion of screen-detected invasive cancers in which the cancer has not invaded the axillary lymph nodes as determined by pathological evidence.
Context	The purpose of screening mammography is to detect breast cancer as early as possible. The proportion of node negative invasive cancer is a good indicator of prognosis as it measures whether the cancer has spread to the lymph nodes[1].
Calculations	$$\frac{\text{Number of cases of screen-detected invasive cancer with negative lymph nodes}}{\text{Total number of screen-detected invasive cancer cases in which lymph nodes were assessed}} \times 100 = \text{Percentage with negative lymph nodes (\%)}$$
Details	Nodal status is measured at the time of diagnosis and excludes measurements taken after neo-adjuvant treatment. Cancers detected by clinical breast exam (CBE) alone are not included in this calculation. Cancers that took > 6 months to diagnose are excluded from this indicator (outside the screening episode). Cases that were lost to follow-up or have missing lymph node information are also excluded from the numerator and denominator. Calculations exclude cases in which lymph nodes are not assessed pathologically. For bilateral cancers the cancer with the highest stage is selected to report on nodal status.
Targets	**Canada** **> 70% screen-detected invasive cancers.** *Europe[17] 75% invasive cancers (subsequent-regular screening) (age 50- 69) (acceptable level). New Zealand[20] > 70% invasive cancers (initial screens) (age 50-69); > 75% invasive cancers (subsequent screens) (age 50-69). *Includes cancer cases with unknown nodal status in the denominator.*
Evidence	Based on stage-specific prospective studies and trials[74-77].
Modification History	Modified in 2006. This indicator replaced the "Positive Lymph Nodes in Cases of Invasive Cancer" indicator that was introduced in 2002. Definition, context and details modified in 2012.

POST-SCREEN INVASIVE CANCER RATE

Definition	Number invasive breast cancers found after a normal or benign mammography screening episode within 0 to < 12 and 12-24 months of the screen date.
Context	Post-screen invasive cancer rate is an indicator of the sensitivity of the mammography screening program. This rate is affected by underlying incidence rates, age, sojourn time, opportunistic screening, and screening interval recommendation. A high rate may negatively affect the mortality reduction expected for a successful, organized screening program[78]. The accuracy of this indicator is also dependent on the completeness of cancer registration.

Calculations

$$\frac{\text{Number of invasive cancers detected in the 0 to < 12 month interval after a normal or benign mammography screening episode}}{\text{Total person-years at risk (0 to < 12 months post screen)}} \times 10,000 = \text{0 to < 12-month Post-Screen Invasive Cancer Rate per 10,000 person-years}$$

$$\frac{\text{Number of invasive cancers detected in the 12-24 month interval after a normal or benign mammography screening episode}}{\text{Total person-years at risk (12-24 months post screen)}} \times 10,000 = \text{12-24 month Post-Screen Invasive Cancer Rate per 10,000 person-years}$$

Details	Screen-detected cancers that took > 6 months to diagnose (outside the screening episode) are included as post screen cancers in this indicator. Screen-detected cancers found by clinical breast exam (CBE) alone are also counted as post-screen cancers. Women who were diagnosed with a post-screen cancer between 12-24 months are included regardless of non-compliance with annual screening recommendations. Total person-years at risk includes time from screen (0 months) or 12 months until the end date (next screen, end of reporting period, cancer diagnosis or death) in women with a normal or benign mammography screening episode. Person years at risk includes women undergoing diagnostic assessment as they may still be at risk of developing a post-screen cancer.

Targets

Canada	**< 6 per 10,000 person-years (0 to < 12 months);** **< 12 per 10,000 person-years (12-24 months).**
*Europe[17]	30% of the underlying, expected, breast cancer incidence rate in the absence of screening (0-11 months) (acceptable level) (age 50-69); 50% of the underlying, expected, breast cancer incidence rate in the absence of screening (12-23 months) (acceptable level) (age 50-69).
*United Kingdom[18]	1.2 per 1,000 women (0-24 months) (expected standard) (age 50-70) ; 1.4 per 1,000 women (24-36 months) (expected standard) (age 50-70).
Australia[19]	< 7.5 per 10,000 women (0 to < 12 months) (age 50-69).
*New Zealand[20]	7.1 per 10,000 women screened within one calendar year of previous screen (maximum) (age 50-69); 15.0 per 10,000 women screened within the second calendar year of previous screen (maximum) (age 50-69).
*Includes DCIS	

Evidence	Based on studies of interval cancer[78-80] and previous CBCSD data[12]. Calculations based on person-years method[81, 82] and comparison to international calculation methods and results[27, 28, 83].
Modification History	Introduced in 2002. Modified in 2006. Definition, calculation, details and targets modified in 2012.

SENSITIVITY OF THE SCREENING MAMMOGRAPHY PROGRAM

Definition	Proportion of breast cancer cases (invasive or DCIS) that were correctly identified as having cancer during the screening episode.
Context	Sensitivity is an indicator of how well the screening mammography program detects cancers. This rate is affected by underlying incidence rates, age, rate of disease progression, opportunistic screening, and screening interval recommendation. The accuracy of this indicator is dependent on the completeness of cancer registration.
Calculations	$$\frac{\text{Number of screen-detected cancers}}{\text{Number of screen-detected cancers + Number of post screen cancers detected 0 - <12 months}} \times 100 = \text{Sensitivity (\%)}$$
Details	Calculation includes subsequent screens only as the sensitivity of the initial screen is affected by opportunistic screening practices and small numbers. Calculation excludes post-screen cancers detected ≥12 months due to varying annual screening practices within provinces. Screen-detected cancers that took > 6 months to diagnose are included as post-screen cancers in this indicator (outside the screening episode). Screen-detected cancers found by clinical breast exam (CBE) alone are also counted as post-screen cancers. Invasive and *in situ* (DCIS) cancers are included in both the numerator and denominator.

Targets	**Canada** **No target**
	                         % (Subsequent screens).
	                         (Surveillance and monitoring purposes only)

Evidence	Based on studies and reports of screening program sensitivity[5, 42, 46, 84, 85] and comparison to international calculation methods and results[27, 28, 83].
Modification History	Introduced in 2012.

Future Directions

The review of a set of meaningful evaluation indicators for organized breast cancer screening programs is an ongoing process. The body of research pertaining to organized breast cancer screening is constantly evolving, as is the technology and methodology used to screen, diagnose and treat the disease. The quality of evidence used to support the use of evaluation indicators presented in this document varies greatly from indicator to indicator and is subject to change with the continual introduction of new research evidence. The data used in the calculation of these indicators, and possible future indicators, are still maturing in terms of quality and timely availability. Consequently, certain evaluation indicators and targets remain under review and may be updated in future iterations of this report.

MONITORING EVALUATION INDICATORS

The formal use of these indicators will be in subsequent release of *Organized Breast Cancer Screening Programs in Canada: Report Program Performance in 2007 and 2008*. The EIWG reassessed the 14 previous evaluation indicators, as well as the proposed indicators identified during group discussions. Based on the literature review and professional recommendations, three new indicators were identified and four indicators were removed. Modifications and additions were also made to some of the existing evaluation indicators. Targets were adjusted or redefined by consensus and supported by new research or expert opinion. Changes to the definitions of the indicators and methods of calculation were also considered on the same basis.

EVALUATION INDICATORS UNDER REVIEW

In situ cancer detection

While ductal carcinoma *in situ* (DCIS) is widely accepted as an obligate precursor of invasive disease, the timeframe in which this occurs is not firmly established[50, 52]. The potential for cases of DCIS to remain asymptomatic throughout the individual's natural lifespan suggests a potential for *over diagnosis* with its associated negative consequences[51]. The Evaluation Indicators Working Group will continue to monitor *in situ* cancer detection rates and will consider defining a target under the appropriate circumstances. It has been proposed that the future studies should measure data on low, intermediate and high-grade DCIS, in order to provide more meaningful data for setting targets.

Non-malignant biopsy rate

There were no targets set for this indicator due to the evolving use of core biopsy as an intermediate step or alternative to an open biopsy[11, 12]. The Evaluation Indicators Working Group will continue to report on non-malignant biopsy rates for surveillance and monitoring purposes.
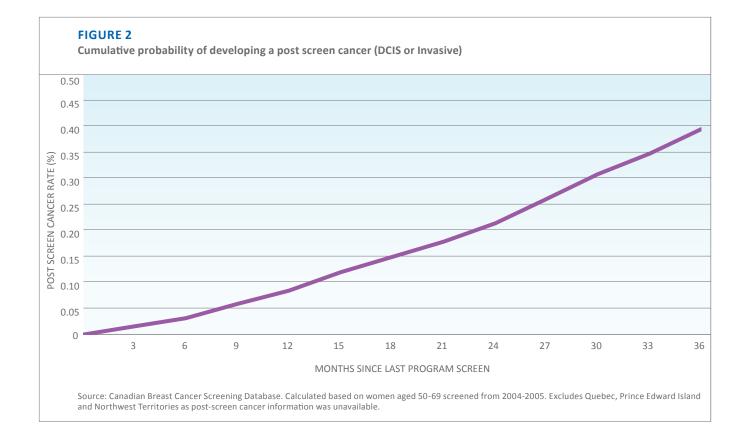
Sensitivity of the screening mammography program

Sensitivity of the screening mammography program is an important indicator of the efficacy of organized screening. In the absence of a gold standard assessment it is impossible to determine sensitivity directly. Defining a target for the current indicator has proven difficult based on the available evidence and variability of Canadian screening data. Sensitivity of the screening mammography program provides an estimate of the proportion of breast cancer cases that were correctly identified as having an abnormality at the time of screening and were correctly identified as breast cancer after completion of diagnostic assessment. However, the calculation of sensitivity has an inherent weakness: true interval cancers cannot be separated from cancer missed at screening or diagnosis which can make it more difficult for programs to report true levels of sensitivity. This indicator uses post-screen cancers as an estimate of false negatives or cases incorrectly identified as benign/normal at screen or diagnostic test. The cumulative probability of a post-screen cancer increases with time since program screen (Figure 2). Therefore, those with annual recommendation will have a lower rate of post-screen cancers. This discrepancy was addressed by only including post screen cancers diagnosed within 12 months. This allows for a greater comparability between provinces but is not consistent with biennial screening recommendations. Further analyses could include calculation of sensitivity within 24 months while attempting to control for annual screeners. Once a baseline for sensitivity of the screening mammography program has been established in subsequent reports on program performance an acceptable target can be determined.

In the current guidelines, sensitivity of the screening mammogram is only presented for subsequent screens. This is due to the presence of opportunistic screening and different age groups targeted across Canada. Initial screens should have a higher sensitivity rate as they identify prevalent cancers. However, analysis of provincial data demonstrates that initial screens may include women who were previously screened outside the organized program. Therefore, sensitivity is more accurately calculated on the subsequent screen as an indicator of the incident cancers correctly identified at the program screen. As programs mature the number of initial screens for women aged 50-69 will also decrease, particularly for provinces where women begin screening at age 40.

Annual screening rate

The EIWG chose to include this indicator due to the impact of annual screening on the current indicators. In provinces with higher rates of annual screening the retention rate and sensitivity may be increased while the cancer detection and post-screen cancer rates may be decreased. Annual screening rates are also an important indicator of program capacity, cost-effectiveness and comparability of results across Canada. In this report the annual screening rate was not assigned a target as recall practices vary widely across jurisdictions. Instead this indicator is meant to provide background information on provincial screening practices that may influence the results of other evaluation indicators.

FIGURE 2
Cumulative probability of developing a post screen cancer (DCIS or Invasive)

Source: Canadian Breast Cancer Screening Database. Calculated based on women aged 50-69 screened from 2004-2005. Excludes Quebec, Prince Edward Island and Northwest Territories as post-screen cancer information was unavailable.

PROPOSED EVALUATION INDICATORS

The following evaluation indicators were investigated during the literature review and may be included in supplementary reports.

- *False positive rate:* Proportion of screens that were incorrectly identified as having an abnormality at the time of screening.
- *Percentage of screen-detected cancers at stage II+:* Proportion of cancers that had a TNM stage greater or equal to II.
- *Incomplete follow-up rate:* Proportion of abnormal screens that were lost to follow-up.
- *Specificity:* Proportion of true negatives (normal screens) that were correctly identified as not having an abnormality at the time of screening.
- *Positive predictive value (PPV) of diagnostic tests:* Positive predictive value of fine needle aspiration (FNA), core or open biopsies.

SUPPLEMENTARY EVALUATION INDICATORS

While the best possible assessment of the morbidity and mortality reducing potential of breast cancer screening was the foremost priority in the selection of these indicators, the availability of high-quality data from the CBCSD was mandatory. Therefore, these criteria do not fully cover the range of evaluation indicators needed to establish comprehensive long-term evaluation plans. From that perspective, factors such as equitable access, waiting time from booking a mammogram, acceptability of services to clients, cost minimization, and program promotion should also be assessed. Factors related to the mammography exam and diagnostic tests such as technical repeat rate and early recall (within 6 months) should also be monitored on a provincial basis. Follow-up of cancers patients including time to treatment, and survival analysis may also be included. In recognition of the need for a more complete inventory of indicators for use in future evaluation initiatives, the QDWG will consider the feasibility of measuring these indicators nationally for inclusion in subsequent editions of this document.

References

1. International Agency for Research on Cancer (IARC). Breast cancer screening. In: H. Vainio, F. Bianchini, editors. IARC handbooks of cancer prevention. Lyon: IARC Press; 2002.

2. Glasziou P, Houssami N. The evidence base for breast cancer screening. Preventive Medicine 2011;53(3):100-2.

3. Gotzsche P, Nielsen M. Screening for breast cancer with mammography (review). Cochrane Database of Systematic Reviews 2011(1).

4. Jorgensen K, Keen J, Gotzsche P. Is mammographic screening justifiable considering its substantial overdiagnosis rate and minor effect on mortality? Radiology 2011;260:621-7.

5. Feig SA. Auditing and benchmarks in screening and diagnostic mammography. Radiol Clin North Am 2007 vi; Sep;45(5):791-800.

6. Health Canada. Quality determinants of organized breast cancer screening programs in Canada. Ottawa: Minister of Public Works and Government Services Canada; 2003.

7. The Workshop Group. Reducing deaths from breast cancer in Canada. Can Med Assoc J 1989;141(3):199-201.

8. Health Canada. Report of the working group on the integration of screening and diagnosis for the Canadian breast cancer screening initiative. Ottawa: Public Works and Government Services Canada; 2000.

9. Health Canada. Evaluation indicators working group: Guidelines for monitoring breast screening program performance. Ottawa: Minister of Public Works and Government Services Canada; 2002.

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

11. Caines J, Schaller G, Iles S, Woods E, Barnes P, Johnson A, Jones G, Borgaonkar J, Rowe J, Topp T, et al. Ten years of breast screening in the Nova Scotia breast screening program, 1991-2001. Experience: Use of an adaptable stereotactic device in the diagnosis of screening-detected abnormalities. Can Assoc Radiol J 2005;56(2):82-93.

12. Public Health Agency of Canada. Organized breast screening programs in Canada: Report on program performance in 2005 and 2006. Ottawa: 2011.

13. Canadian Cancer Society's Steering Committee on Cancer Statistics. Canadian cancer statistics 2011. Toronto: Canadian Cancer Society; 2011.

14. Bancej C, Decker K, Chiarelli A, Harrison M, Turner D, Brisson J. Contribution of clinical breast examination to mammography screening in the early detection of breast cancer. J Med Screen 2003;10(1):16-21.

15. Tabar L, Fagerberg CJ, Gad A, Baldetorp L, Holmberg L, Grontoft O, Ljungquist U, Lundstrum B, Manson J, Eklund G, et al. Reduction in mortality from breast cancer after mass screening with mammography. Lancet 1985; 325(8433):829-32.

16. Doyle G, Major D, Chu C, Stankiewicz A, Harrison M, Pogany L, Mai V, Onysko J. A review of screening mammography participation and utilization in Canada. Chronic Diseases and Injuries in Canada 2011;31(4):152-6.

17. Health and Consumer Protection Directorate-General. European guidelines for quality assurance in breast cancer screening and diagnosis. Luxembourg: European Communities; 2006.

18. NHS Breast Screening Programme. Consolidated guidance on standards for the NHS breast screening programme. Sheffield: NHS Cancer Screening Programmes; 2005. Report nr 60.

19. BreastScreen Australia. National accreditation standards. Canberra: 2008.

20. BreastScreen Aotearoa. National policy and quality standards. Wellington: National Screening Unit; 2008.

21. Elmore JG, Armstrong K, Lehman CD, Fletcher SW. Screening for breast cancer. JAMA 2005 Mar 9;293(10):1245-56.

22. Gotzsche P, Nielsen M. Screening for breast cancer with mammography. Cochrane Database of Systematic Reviews 2006(4).

23. Forrest P. Breast cancer: The decision to screen. London: Nuffield Provincial Hospitals Trust; 1990.

24. Nystrom L, Andersson I, Bjurstam N, Frisell J, Nordenskjold B, Rutqvist LE. Long-term effects of mammography screening: Updated overview of the Swedish randomised trials. Lancet 2002;359:909-19.

25. Feig S. Effect of service screening mammography on population mortality from breast carcinoma. Cancer 2002;95(3):451-7.

26. Shapiro S, Strax P, Venet L. Periodic breast cancer screening in reducing mortality from breast cancer. JAMA 1971;215(11):1777-85.

27. Australian Institute of Health and Welfare. BreastScreen Australia monitoring report 2005-2006. Canberra: AIHW; 2009. Report nr Cancer series no 48.

28. BreastScreen Australia. BreastScreen Australia data dictionary. National Quality Management Committee; 2004.

29. Duffy SW, Chen H, Tabar L, Day N. Estimation of mean sojourn time in breast cancer screening using a Markov chain model of both entry to and exit from the preclinical detectable phase. Statistics in Medicine 1995;14:1531-43.

30. Chen Y, Brock G, Wu D. Estimating key parameters in periodic breast cancer screening - application to the Canadian national breast screening study data. Cancer Epidemiology 2010;34(4):429-33.

31. Chen HH, Duffy SW. A Markov chain method to estimate the tumour progression rate from preclinical to clinical phase, sensitivity and positive predictive value for mammography in breast cancer screening. The Statistician 1996;45(3):307-17.

32. Boudreau DM, Luce CL, Ludman E, Bonomi AE, Fishman PA. Concordance of population-based estimates of mammography screening. Prev Med 2007 Oct;45(4):262-6.

33. Collett D. Modeling survival data in medical research - 2nd edition. Chapman and Hall/CRC; 2003.

34. Ravaioli A, Foca F, Colamartini A, Falcini F, Naldoni C, Finarelli AC, de Bianchi PS, Bucchi L. Incidence, detection, and tumour stage of breast cancer in a cohort of Italian women with negative screening mammography report recommending early (short-interval) rescreen. BMC Medicine 2010;8:11.

35. Ong GJ, Austoker J, Michell M. Early rescreen/recall in UK National Health Service breast screening programme: Epidemiological data. J Med Screen 1998;5:146-55.

36. Otten JD, Karssemeijer N, Hendriks J, Groenewoud J, Fracheboud J, Verbeek A, de Koning H, Holland R. Effect of recall rate on earlier screen detection of breast cancers based on the Dutch performance indicators. J Natl Cancer Inst 2005;97(10):748-54.

37. Gur D, Sumkin JH, Hardesty LA, Clearfield RJ, Cohen CS, Ganott MA, Hakim CM, Harris KM, Poller WR, Shah R, et al. Recall and detection rates in screening mammography. Cancer 2004 Apr 15;100(8):1590-4.

38. Utzon-Frank N, Vejborg I, von Euler-Chelpin M, Lynge E. Balancing sensitivity and specificity: Sixteen year's of experience from the mammography screening programme in Copenhagen, Denmark. Cancer Epidemiology 2011;35:393-8.

39. Skaane P, Skjennald A. Screen-film mammography versus full-field digital mammography with soft-copy reading: Randomized trial in a population-based screening program – the Oslo II study. Radiology 2004 Jul;232(1):197-204.

40. Skaane P, Hofvind S, Skjennald A. Randomized trial of screen-film versus full-field digital mammography with soft-copy reading in population-based screening program: Follow-up and final results of Oslo II study. Radiology 2007;244(3):708-17.

41. Pisano E, Gatsonis C, Hendrick E, Yaffe M, Baum J, Acharyya S, Conant E, Fajarko L, Bassett L, D'Orsi C, et al. Diagnostic performance of digital versus film mammography for breast-cancer screening. N Engl J Med 2005;353(17):1773-83.

42. Schell M, Yankaskas B, Ballard-Barbash R, Qaqish B, Barlow W, Rosenberg R, Smith-Bindman R. Evidence-based target recall rates for screening mammography. Radiology 2007;243(3):681-9.

43. Yankaskas BC, Klabunde CN, Ancelle-Park R, Renner G, Wang H, Fracheboud J, Pou G, Bulliard JL, International Breast Cancer Screening Network. International comparison of performance measures for screening mammography: Can it be done? J Med Screen 2004;11(4):187-93.

44. Vinnicombe S, Pinto Pereira SM, McCormack VA, Shiel S, Perry N, Dos Santos Silva IM. Full-field digital versus screen-film mammography: Comparison within the UK breast screening program and systematic review of published data. Radiology 2009 May;251(2):347-58.

45. Smith-Bindman R, Ballard-Barbash R, Miglioretti DL, Patnick J, Kerlikowske K. Comparing the performance of mammography screening in the USA and the UK. J Med Screen 2005;12(1):50-4.

46. Carney P, Sickles E, Monsees B, Bassett L, Brenner RJ, Feig S, Smith R, Rosenberg R, Bogart TA, Browning S, et al. Identifying minimally acceptable interpretive performance criteria for screening mammography. Radiology 2010;255(2):354-61.

47. Rosenberg RD, Yankaskas BC, Abraham LA, Sickles EA, Lehman CD, Geller BM, Carney PA, Kerlikowske K, Buist DS, Weaver DL, et al. Performance benchmarks for screening mammography. Radiology 2006 Oct;241(1):55-66.

48. NHS Cancer Screening Programmes. NHS breast screening programme annual review 2010: Overcoming barriers. NHSBSP; 2010.

49. Duffy SW, Gabe R. What should the detection rates of cancers be in breast screening programmes? Br J Cancer 2005 Feb 14;92(3):597-600.

50. Evans AJ, Blanks RG. Should breast screening programmes limit their detection of ductal carcinoma in situ?. Clin Radiol 2002 Dec;57(12):1086-9.

51. Yen M, Tabar L, Vitak B, Smith RA, Chen H, Duffy SW. Quantifying the potential problem of overdiagnosis of ductal carcinoma in situ in breast cancer screening. European Journal of Cancer 2003;39(12):1746-54.

52. Kerlikowske K. Epidemiology of ductal carcinoma in situ. J Natl Cancer Inst Monogr 2010;41:139-41.

53. Barchielli A, Federico M, De Lisi V, Bucchi L, Ferretti S, Paci E, Ponti A, Buiatti E. In situ breast cancer: Incidence trend and organised screening programmes in Italy. Eur J Cancer 2005 May;41(7):1045-50.

54. Luke C, Priest K, Roder D. Changes in incidence of in situ and invasive breast cancer by histology type following mammography screening. Asian Pacific J Cancer Prev 2006;7(1):69-74.

55. Hislop TG, Haris S, Jackson J, Thorne S, Rousseau E, Coldman A, Vestrup J, Wright C, Olivotto I. Satisfaction and anxiety for women during investigation of an abnormal screening mammogram. Breast Cancer Res Treat 2002;76:245-54.

56. Lampic C, Thurfjell E, Bergh J, Sjoden P. Short- and long-term anxiety and depression in women recalled after breast cancer screening. Eur J Cancer 2001;37:463-9.

57. Thorne S, Harris S, Hislop TG, Vestrup J. The experience of waiting for diagnosis after an abnormal mammogram. Breast J 1999;5(1):42-51.

58. Olivotto I, Gomi A, Bancej C, Brisson J, Tonita J, Kan L, Mah Z, Harrison M, Shumak R. Influence of delay to diagnosis on prognostic indicators of screen-detected breast carcinoma. Cancer 2002;94(8):2143-50.

59. Wright GP, Wong J, Morgan J, Roy-Chowdhury S, Kazanjian K, Lum S. Time from diagnosis to surgical treatment of breast cancer: Factors influencing delays in initiating treatment. The American Surgeon 2010;76(10):1119-22.

60. Kerlikowske K. Timeliness of follow-up after abnormal screening mammography. Breast Cancer Res Treat 1996;40:53-64.

61. Ganry O, Peng J, Dubreuil A. Influence of abnormal screens on delays and prognostic indicators of screen-detected breast carcinoma. J Med Screen 2004;11(1):28-31.

62. Richards M, Westcombe A, Love S, Littlejohns P, Ramirez A. Influence of delay on survival in patients with breast cancer: A systematic review. The Lancet 1999;353:1119-26.

63. Decker K, Harrison M, Chateau D. Influence of direct referrals on time to diagnosis after an abnormal breast screening result. Cancer Detection and Prevention 2004;28:361-7.

64. Borugian M, Kan L, Chu C, Ceballos K, Gelmon K, Gordon P, Poole B, Tyldesley S, Olivotto I. Facilitated "fast track" referral reduces time from abnormal screening mammogram to diagnosis. Can J Public Health 2008;99(4):252-6.

65. Psooy B, Schreuer D, Borgaonkar J, Caines J. Patient navigation: Improving timeliness in diagnosis of breast abnormalities. Can Assoc Radiol J 2004;55(3):145-50.

66. Strong K, Wald N, Miller A, on behalf of the WHO Consultation Group. Current concepts in screening for noncommunicable disease: World Health Organization consultation group report on methodology of noncommunicable disease screening. J Med Screen 2005;12:12-9.

67. Landercasper J, Linebarger J, Ellis R, Mathiason M, Johnson J, Marcou K, DeMaiffe B, Jago G. A quality review of the timeliness of breast cancer diagnosis and treatment in an integrated breast center. J Am Coll Surg 2010;210:449-55.

68. Rathgeber M, Chan B, MacDermott W, Willoughby K, Cascagnette P. The time it takes: Breast cancer care in Saskatchewan. Saskatoon: Health Quality Council; 2006.

69. Richardson LC, Royalty J, Howe W, Helsel W, Kammerer W, Benard VB. Timeliness of breast cancer diagnosis and initiation of treatment in the national breast and cervical cancer early detection program, 1996-2005. Am J Public Health 2010 Sep;100(9):1769-76.

70. Stewart TR, Mumpower JL. Detection and selection decisions in the practice of screening mammography. Journal of Policy Analysis & Management 2004 Aug;23(4):908-20.

71. Maxwell AJ, Pearson JM, Bishop HM. Crude open biopsy rates for benign screen detected lesions no longer reflect breast screening quality--time to change the standard. J Med Screen 2002;9(2):83-5.

72. Manfrin E, Mariotto R, Remo A, Reghellin D, Falsirollo F, Dalfior D, Bricolo P, Piazzola E, Bonetti F. Benign breast lesions at risk of developing cancer--a challenging problem in breast cancer screening programs: Five years' experience of the breast cancer screening program in verona (1999-2004). Cancer 2009;115(3):499-507.

73. Ghosh K, Melton LJ,3rd, Suman VJ, Grant CS, Sterioff S, Brandt KR, Branch C, Sellers TA, Hartmann LC. Breast biopsy utilization: A population-based study. Arch Intern Med 2005 Jul 25;165(14):1593-8.

74. Michaelson JS, Satja S, Kopans D, Moore R, Silverstein M, Comegno A, Hughes K, Taghian A, Powell S, Smith B. Gauging the impact of breast carcinoma screening in terms of tumor size and death rate. Cancer 2003;98(10):2114-24.

75. Tabar L, Fagerberg G, Duffy SW, Gad A, Grontoft O. Update of the Swedish two-county program of mammographic screening for breast cancer. Radiol Clin North Am 1992;30(1):187-210.

76. Tabar L, Duffy SW, Vitak B, Chen HH, Prevost T. The natural history of breast carcinoma: What have we learned from screening? Cancer 1999;86(3):449-62.

77. Soerjomataram I, Louwman M, Ribot J, Coebergh JW. An overview of prognostic factors for long-term survivors of breast cancer. Breast Cancer Res Treat 2008;107:309-30.

78. Tabar L, Faberberg G, Day NE, Holmberg L. What is the optimum interval between mammographic screening examinations? - An analysis based on the latest results of the Swedish two-county breast cancer screening trial. Br J Cancer 1987;55:547-51.

79. Bulliard JL, Ducros C, Jemelin C, Arzel B, Fioretta G, Levi F. Effectiveness of organised versus opportunistic mammography screening. Annals of Oncology 2009 Jul;20(7):1199-202.

80. Taylor R, Supramaniam R, Rickard M, Estoesta J, Moreira C. Interval breast cancers in New South Wales, Australia and comparisons with trials and other mammographic screening programmes. J Med Screen 2002;9:20-5.

81. Fracheboud J, de Koning H, Beemsterboer P, Boer R, Verbeek A, Hendriks J, van Ineveld B, Broeders M, de Bruyn A, van der Maas P. Interval cancers in the Dutch breast cancer screening programme. British Journal of Cancer 1999;81(5):912-7.

82. Bulliard JL, Sasieni P, Klabunde C, De Landtsheer JP, Yankaskas BC, Fracheboud J. Methodological issues in international comparison of interval breast cancers. International Journal of Cancer 2006 Sep 1;119(5):1158-63.

83. Begg S, Page A, Arnett K, Taylor R. Interval cancers in BreastScreen Aotearoa: 1999-2002. ; 2007.

84. Klabunde CN, Sancho-Garnier H, Taplin S, Thoresen S, Ohuchi N, Ballard-Barbash R, International Breast Cancer Screening Network. Quality assurance in follow-up and initial treatment for screening mammography programs in 22 countries. International Journal for Quality in Health Care 2002 Dec;14(6):449-61.

85. Yankaskas B, Cleveland R, Schell M, Kozar R. Association of recall rates with sensitivity and positive predictive values of screening mammography. Am J Roentgenol 2001;177(3):543-9.

86. Wilson J, Jungner G. Principles and practice of screening for disease. Geneva: World Health Organization; 1968. Report nr 34.

Appendix A

LITERATURE REVIEW PROTOCOL

A new edition of the Evaluation Indicators Report is published every five years or if new evidence becomes available. As part of each update, the scientific evidence used to support each of the evaluation indicators requires systematic review. The following protocol is designed to facilitate the updating of scientific evidence by using both published and grey literature.

1.0 KEY QUESTIONS

Questions that the review will attempt to address for identified indicators:

1. Can meaningful targets for mammography screening program evaluation indicators be defined based on evidence?
2. Based on the available evidence, what is the general consensus on indicator definitions, targets or minimum requirements?
3. How useful is each indicator and each target within a program setting?
4. How can targets be calculated and measured?
5. Which countries/regions use these indicators?
6. What are/is the recommended reporting structures for each target?
7. What is the minimum sample size to be considered valid?

2.0 INCLUSION/EXCLUSION CRITERIAS

2.1 POPULATION(S)

Include: For the purposes of these guidelines for reporting evaluation indicators, the target population for evaluation is the same as the national target population for organized screening. This population is defined as asymptomatic women between the ages of 50 and 69 years with no prior diagnosis of breast cancer.

Exclude: Women outside the ages of 50-69, and those who have already had a diagnosis of breast cancer are outside the target population for evaluation.

2.2 INTERVENTIONS

Include: Articles and reports that evaluate the performance of organized mammography screening programs or discuss proposed evaluation indicators.

Exclude: Articles and reports that do not explicitly evaluate organized mammography screening programs or discuss the evaluation indicators.

2.3 OUTCOMES

This review will focus on evidentiary support for the following evaluation indicators as well as provide estimates for performance targets for the proposed indicators.

2.4 PUBLISHED LITERATURE

Include: Literature published from 2002 to present will be examined. Randomized controlled trials, observational studies, meta-analysis, reviews, international guidelines, and the like from all comparable countries to Canada, including the US, Western European countries, Australia and New Zealand will be included. Only documents in English will be considered.

Exclude: Qualitative reports, methodological studies, editorials and commentaries as well as literature from countries that are less comparable to Canada will be excluded. Documents that are not in English will be excluded.

2.5 GREY LITERATURE

Include: Web pages of international organizations, bilateral agencies, and non-governmental organizations (NGOs) involved in creating, updating, or reporting on mammography screening guidelines and research. Documents that refer to mammography screening performance measurement guidelines will be included or held for review. Reports and articles that were recommended by experts in the field and/or other working group members will be included or reviewed.

Articles published prior to 2002 that are recommended will be included or reviewed.

3.0 METHODS

3.1 STUDY DESIGN

A systematic approach will be used to review relevant published and grey literature.

3.2 LITERATURE SEARCH STRATEGIES

3.2.1 Published literature

The following strategies were used to search the Cochrane library and OVID Medline databases for relevant studies published from 2002 to present. Reference lists of included articles will be scanned for relevant publications

Cochrane Library (Cochrane Reviews)

Access: http://www.library-bibliotheque.hc-sc.gc.ca/eng/services/cisti/health-canada/find-articles.html → Cochrane Library → Cochrane Database of Systematic Reviews (Cochrane Reviews) A broader search was conducted using the Cochrane library in order to ensure large capture. The Cochrane Library is accessible online: http://www.thecochranelibrary.com/view/0/index.html. MeSh trees where applicable will be used.

SEARCH STRATEGY		# IDENTIFIED STUDIES
1	"guidelines **in** Title, Abstract or Keywords **or** standards **in** Title, Abstract or Keywords **and** mammography **in** Title, Abstract **or** Keywords **or** screening **in** Title, Abstract **or** Keywords **in** Cochrane Database of Systematic Reviews"	593

Ovid MEDLINE 1946-January week 2 2011

Access: http://www.library-bibliotheque.hc-sc.gc.ca/eng/services/cisti/health-canada/find-articles.html → Medline

SEARCH STRATEGY		# IDENTIFIED STUDIES
1	mammogra$.mp. or Mammography/ or Breast Neoplasms/	181141
2	guideline$.mp. or Guideline/ or Practice Guideline/	200950
3	standard$.mp. or Reference Standards/	613579
4	2 or 3	788416
5	screen$.mp. or Mass Screening/	383165
6	1 and 5	18155
7	4 and 6	2414
8	limit 7 to (english language and yr="2002 -Current")	1232

3.2.2 Grey literature

The following search strategy was employed to identify the relevant grey literature using keyword searches and webpage scans

• Keyword search ('mammography guidelines, reference standards, practice guidelines') using google.com.

• A search of web pages of international organizations (eg. World Health Agency), bilateral agencies (eg. International Cancer Screening Network), and non-governmental organizations (eg. Canadian Cancer Society) involved in creating, updating, or reporting on mammography screening guidelines and research.

• A search for relevant documents that were referred to in bibliographies of reviews and other reports and articles (scanning of reference lists) as well as relevant websites.

• Reports and articles in press that were recommended by experts in the field and/or other working group members will be included or held for review

- Documents that refer to mammography screening evaluation indicator guidelines will be included or held for review
- Focus on new literature published from 2002 to present however relevant literature prior to 2002 will also be examined.

Website scans

The following websites were scanned for relevant information:

- Breast Cancer Society of Canada – www.bcsc.ca
- Canadian Cancer Society – www.cancer.ca
- Canadian Breast Cancer Network – http://www.cbcn.ca
- British Columbia (BC) Cancer Agency – http://www.bccancer.bc.ca/
- Alberta Health Services – http://www.albertahealthservices.ca/
- Saskatchewan Cancer Agency – http://www.saskcancer.ca/
- Northwest Territories Health and Social Services – http://www.hlthss.gov.nt.ca/
- Cancer Care Manitoba – http://www.cancercare.mb.ca/home/
- Cancer Care Ontario – http://www.cancercare.on.ca/
- Québec Breast Cancer Screening Program – http://www.msss.gouv.qc.ca/
- Institut national de santé publique du Québec – http://www.inspq.qc.ca/
- New Brunswick Breast Cancer Screening Program – http://www.gnb.ca/
- Nova Scotia Breast Screening Program – http://breastscreening.nshealth.ca/
- Prince Edward Island Provincial Breast Screening Program – http://healthpei.ca
- BreastScreen Australia – http://www.cancerscreening.gov.au/internet/screening/publishing.nsf/Content/breastscreen-about
- BreastScreen Aotearoa (New Zealand) – http://www.nsu.govt.nz/current-nsu-programmes/breastscreen-aotearoa.aspx
- European Union (Guidelines for quality assurance in breast cancer screening and diagnosis) – http://ec.europa.eu/health/ph_projects/2002/cancer/fp_cancer_2002_ext_guid_01.pdf
- UK NHS breast screening programme – http://www.cancerscreening.nhs.uk/breastscreen/
- National Cancer Institute (United States) – http://breastscreening.cancer.gov/data/

3.2.3 Literature screening

Articles and documents retrieved from databases and grey literature were screened in two phases by reviewers for relevance based on the inclusion and exclusion criteria. In phase 1, two reviewers separately screened the titles and abstracts of all identified articles. Grey literature was accepted for inclusion based on phase 1. All other articles for which relevance was identified or uncertain based on the title and/or abstract were retained for phase 2. In phase 2 the full-text of the identified articles was examined by at least two members of the EIWG to identify or rule out relevant or irrelevant documents. In an excel workbook, each relevant document was categorized by the indicator(s) for which it provides supporting evidence in terms of employment as breast cancer screening evaluation indicators and performance targets. The reviewers discussed their screening evaluations and come to consensus.

3.2.4 Assessment of Quality

The articles and documents retained following the screening process were assessed and rated by expert reviewers (members of the EIWG) for quality in terms of overall strength for inclusion in the final report. This rating was based on the reviewer's informed opinion and on the article's or document's methodology, setting/context and results. Each document was rated on a 10-point scale (1= low, 10=high). Those documents that are considered of high quality (7 or greater) were retained for inclusion in the final report. This cut-off was adjusted in cases where evidence was sparse for certain indicators.

3.2.5 Data Abstraction

Data were abstracted concurrently with the process of quality assessment (section 2.2.4) for each of the key questions outlined in section 1. The reviewers discussed extracted data and came to consensus on which articles were included in the final document.

4.0 RESULTS REPORTING AND SYNTHESIS

4.1 IDENTIFIED STUDIES

1825 published articles (see flowchart) and 98 grey literature articles, reports and websites were identified. From these, 86 documents (22 published articles and 64 grey literature) were used to address the key questions. All identified studies from each phase of the review were recorded in an excel worksheet. Flowcharts of the process are presented.

4.2 DATA SYNTHESIS

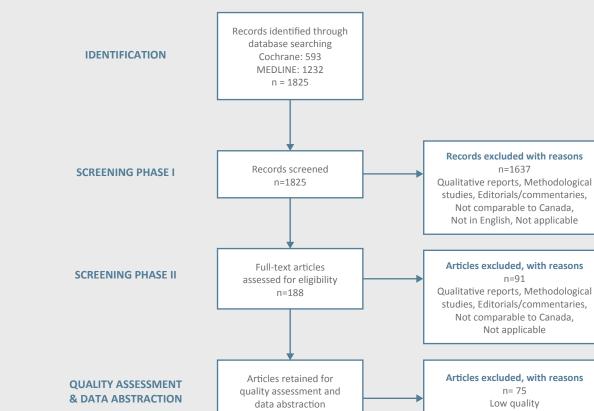The content of the data abstraction worksheet were synthesized to provide the content for the background, target and evidence sections of the report. Data abstraction worksheets included information from each article regarding the key concepts outlined in section 1.

4.3 FINAL CONTENT

Final content to be included in the Report as a result of the literature review was arrived at via consensus of the EIWG.

IDENTIFICATION

Records identified through database searching
Cochrane: 593
MEDLINE: 1232
n = 1825

SCREENING PHASE I

Records screened
n=1825

Records excluded with reasons
n=1637
Qualitative reports, Methodological studies, Editorials/commentaries, Not comparable to Canada, Not in English, Not applicable

SCREENING PHASE II

Full-text articles assessed for eligibility
n=188

Articles excluded, with reasons
n=91
Qualitative reports, Methodological studies, Editorials/commentaries, Not comparable to Canada, Not applicable

QUALITY ASSESSMENT & DATA ABSTRACTION

Articles retained for quality assessment and data abstraction
n=97

Articles excluded, with reasons
n= 75
Low quality
Not applicable

SYNTHESIS

Articles included in synthesis
n=22

Appendix B

CONCEPTUAL FRAMEWORK

The Conceptual Framework is an updated modification of the classic Wilson and Jungner[86] criteria:

- The target cancer should be appropriate for screening.
- The objectives of the screening must be clearly identified.
- There should be an appropriate screening test.
- There should be agreement on the appropriate management of people with positive results on the screening test.
- There must be sound evidence that screening has a favourable impact on its intended objectives.
- Screening should do more good than harm.
- The health care system should be capable of supporting all necessary elements of screening, including diagnosis and treatment.
- Screening should be endorsed only if it is provided in a continuous manner in conjunction with the necessary quality assurance and programmatic elements.

Cancer screening should incorporate all of the essential programmatic elements of the clinical trials that form its evidentiary base. These **Key Elements** include the following:

- Screening must be comprehensive, including recruitment, recall, follow-up, and timely assessment of people with positive screening tests.
- Screening must be supported by public education, including education about primary prevention when applicable.
- Screening must be supported by the education of health care workers.
- All eligible people should have reasonable access to screening, diagnostic assessment and treatment.
- The groups targeted for participation in a screening program should be selected on the basis of a realistic understanding of the harms and benefits of screening and the manner in which health information will be managed.
- All aspects of the screening program must be subject to continuous monitoring and evaluation.
- Screening programs must adopt a culture of continually striving to increase the benefits and minimize the harms of screening.
- Screening programs must have the capacity to modify screening standards, guidelines and best practices on the basis of new scientific evidence.
- The program must have an effective and efficient computerized information system.
- There must be adequate resources (financial, physical, human and informational) to support all aspects of screening.

Screening programs must include a consumer perspective in all aspects of planning and operations.

Appendix C

EVALUATION INDICATORS WORKING GROUP MEMBERS (EIWG)	
Heather Limburg [Chair] Epidemiologist, creening and Early Detection Section	Public Health Agency of Canada 785 Carling Avenue Ottawa, Ontario K1A 0K9
Christina Chu* Biostatistical Analyst, Cancer Surveillance & Outcomes, Population Oncology	BC Cancer Agency 8th Floor, 686 West Broadway Vancouver, British Columbia V5Z 1G1
Dr. Andrew Coldman Vice President, Population Oncology,	BC Cancer Agency 8th Floor, 686 West Broadway Vancouver, British Columbia V5Z 1G1
Thy Dinh* Epidemiologist, Screening and Early Detection Section	Public Health Agency of Canada 785 Carling Avenue Ottawa, Ontario K1A 0K9
Gregory Doyle Coordinator, Breast Screening Program for Newfoundland and Labrador	35 Major's Path, Suite 102 St. John's, Newfoundland A1A 4Z9
Song Gao Team Lead Statistical Specialist, Screening Programs Health Promotion, Disease and Injury Prevention Population and Public Health	Alberta Health Services 2202-2nd Street S.W. Calgary, Alberta T2S 3C1
Vicky (Violet Mary) Majpruz Senior Research Associate, Ontario Breast Screening Program, Screening Unit	Cancer Care Ontario 505 University Avenue Toronto, Ontario M5G 1X3
Gopinath Narasimhan Research Officer, Epidemiology	Saskatchewan Cancer Agency 4-2105 8th Street East Saskatoon, Saskatchewan S7H 0T8
Lisa Pogany Epidemiologist, Screening and Early Detection Section	Public Health Agency of Canada 785 Carling Avenue Ottawa, Ontario K1A 0K9
Sylvie St. Jacques* Agent de planification, de programmation et de recherche Direction de l'analyse et de l'évaluation des systèmes de soins et services	Institut national de santé publique du Québec 945 Wolfe, 5ième étage Québec, Québec G1V 5B3

*formerly

Appendix D

GLOSSARY

Asymptomatic	A woman who does not report symptoms and appears without signs of disease at screening.
Breast cancer	Includes malignant invasive and ductal carcinoma *in situ* (DCIS) of the breast.
Core biopsy	A needle biopsy of the breast used to remove samples of tissue for microscopic evaluation. Most core biopsies are image guided.
Definitive diagnosis	Definitive diagnosis of cancer is the first core or open biopsy that confirms cancer. In rare occasions fine needle biopsy (FNA) may also be used as a definitive diagnosis of cancer. Cancers must be diagnosed within 6 months of the program screen. Definitive diagnosis of benign cases is the last benign test up to 6 months following an abnormal screen.
Ductal carcinoma *in situ* (DCIS)	A non-invasive tumour of the breast, arising from cells that involve only the lining of a breast duct. The cells have not spread outside the duct to other tissues in the breast. DCIS is also referred to as stage 0 cancer.
Fine-needle aspiration biopsy (FNA)	A needle is inserted into the lesion and material withdrawn using a syringe. The material can be stained and the cells examined under microscope in a laboratory to determine whether they are benign or malignant.
Incident cancer	The proportion of new cases of cancer at a given point in time. Refers to new cancers detected during a subsequent screen.
Initial screen	The first screening mammogram provided to a woman by a Canadian organized breast screening program.
Invasive cancer	Cancer cells invading breast tissue beyond the walls of the milk duct or lobule. A ductal carcinoma in situ component may also be present in cases of invasive cancer. Invasive cancer includes stage I-IV.
Normal screening episode	A screening episode that concludes with normal (non-cancer) findings. This includes both a normal screening mammogram and an abnormal screening mammogram with a normal (non-cancer) finding after completion of diagnostic assessment.

Open biopsy	Surgical removal of a breast mass under local or general anesthesia for subsequent microscopic examination by a pathologist.
Post-screen cancer	Cancers that occur outside the screening program following a normal screening episode. This includes women who become symptomatic and develop breast cancer before their next regular screen (interval cancers) and those who did not return for their regular screen and were diagnosed with breast cancer after 24 months from their previous screen date (noncompliant cancers).
Prevalent cancer	The proportion of the population with cancer at a given point in time. Refers to existing cancers detected on the first (initial) screen.
Screen	Two-view bilateral screening mammogram delivered by the organized screening program.
Screening episode (completed)	Defined for normal screens as the date of the program screen; for abnormal screens it is the time from screen to definitive diagnosis. For calculation purposes the screening episode is closed at 6 months following an abnormal screen.
Screen-detected cancer	Cancer detected within 6 months of an abnormal program screen as a result of pathologic confirmation based on diagnostic testing attributed to the mammogram.
Sojourn time	The time interval between the onset of detectable pre-clinical disease and symptomatic disease.
Subsequent screen	Successive screens (screening rounds) after the initial (first) screen under the organized program. This includes women who miss a scheduled round of screening.
Tissue biopsy	A biopsy which provides breast tissue for histopathologic examination (does not refer to fine-needle aspiration biopsy which provides only cells). Includes both core and open biopsies.
Total person-years at risk	Total person-years at risk includes time from screen (0 months) or 12 months until the end date (next screen, end of reporting period, cancer diagnosis or death) in women with a normal or benign mammography screening episode. Person years at risk includes time from screen for women undergoing assessment as they may still be somewhat at risk of developing a post-screen cancer.

CANADIAN PARTNERSHIP AGAINST CANCER   PARTENARIAT CANADIEN CONTRE LE CANCER