Health Canada Santé Canada

Report *from the* Evaluation Indicators Working Group

Guidelines for Monitoring Breast Screening Program Performance

Canada

Our mission is to help the people of Canada
maintain and improve their health

Health Canada

Ce document est aussi offert en français sous le titre :
Rapport du Groupe de travail sur les indicateurs d'évaluation : Lignes directrices pour la surveillance de la performance des programmes de dépistage du cancer du sein

© Her Majesty the Queen in Right of Canada, represented by the Minister of Public Works and Government Services Canada, 2002
Cat. N° H39-617/2002
ISBN 0-662-66515-5

Report from the Evaluation Indicators Working Group: Guidelines for Monitoring Breast Screening Program Performance

Acknowledgments

We would like to acknowledge the following contributors:

Present members of the Evaluation Indicators Working Group:

- Dr. Anna Chiarelli (Chair), Ontario Breast Screening Program;
- Gregory Doyle, Breast Screening Program for Newfoundland and Labrador;
- Lisa Kan, B.C. Cancer Agency;
- Zeva Mah, Alberta Cancer Board;
- Norah Smith, PEI Breast Screening Clinic;
- Christina Bancej, Jay Onysko, Health Canada.

Special thanks go to the following:

- Dr. Paul Jelfs, BreastScreen Australia;
- Dr. Sue Moss, Cancer Screening Evaluation Unit (U.K.);
- Dr. Judy Caines, Nova Scotia Breast Screening Program, for consulting with the Working Group;
- The Database Management Committee (DMC), for its edits and reviews of this report;
- Past members of the Evaluation Indicators Working Group
  - Leslie Gaudette, Health Canada
  - Dr. Françoise Bouchard, Health Canada
  - Gloria Low, Health Canada
  - Veronika Moravan, Cancer Care Ontario
- Dr. Diane Major and colleagues, Institut national de santé publique du Quebec, for reviewing the French version of the report;
- Asako Gomi, Health Canada, for drafting the History of the Evaluation Indicators Working Group;
- Dr. Tammy Horne, WellQuest Consulting Ltd., for facilitating preliminary meetings/drafting initial discussion documents.

Table of Contents

Background

Introduction

Organized breast cancer screening has been established in Canadian provinces and territories. With the Canadian Breast Cancer Screening Database fully implemented, consistent program data are available for evaluation. The purpose of this document is to define an initial set of performance measures for Canadian breast cancer screening programs. The document will serve as a guide to promote consistent calculation of key performance measures for various monitoring and evaluation efforts across programs and over time. The description of each measure includes a definition, the context in which the measure is relevant (rationale), method(s) of calculation, target objectives, and the current status of the measure under evaluation. The measures presented in this document were developed on the basis of recognized population screening principles, evidence from randomized controlled trials, demonstration projects, and observational studies (see Appendix A for a brief framework of screening principles).

Purpose of Performance Measures

The principal goal of breast cancer screening is to reduce breast cancer mortality and morbidity associated with this disease. Regular mammography screening for women aged 50 to 69 is expected to prevent approximately one-third of breast cancer deaths 7 to 12 years after sufficient participation (70% of women in the target group) has been achieved[1]. Because reaching a participation rate of 70% among women aged 50 to 69 will be a gradual process, short-term reductions in mortality rates cannot be used to monitor the effectiveness of breast cancer screening. Instead, performance measures that are valid, reliable and feasible to collect within the screening program are required for interim evaluation of breast cancer screening. Furthermore, these measures provide a means to monitor the individual steps throughout the entire screening process in order to ensure that the objectives of a successful screening program are met. This is the first step in ensuring that screening programs continually strive to increase the benefits of screening while minimizing the negative side effects.

Interim measures used for ongoing evaluation of organized breast cancer screening programs at the national level include participation rate, abnormal call rate, cancer detection rate, rate of advanced cancers, tumour size and nodal status (detailed descriptions to follow). Provincial and territorial programs compute additional measures that are not monitored at the national level. In the past, data from Canadian organized breast cancer screening programs have been collected nationally for comparison against targets set by Sweden[2], the European Union[3,4], the United Kingdom[5,6] and Australia[7]. These comparison are useful, but the uniqueness of Canada's population health context and differences in elements of its programs heighten the need for a set of Canadian core indicators and targets that measure the performance and quality of its organized breast cancer screening programs.

Canada's Unique Population Health Context

Table 1 provides a comparison of selected characteristics of organized breast cancer screening programs in Canada, the United Kingdom, Australia, Sweden and the Netherlands. In Canada, health care delivery is under provincial/territorial jurisdiction; thus, organized screening programs have been developed and implemented independently across the country. Consequently, programs vary in their organization, screening modalities, recruitment methods, ages targeted for screening, and in the arrangements for diagnostic assessment following an abnormal screen. Currently, all 10 provinces and two territories have organized screening programs[8]. To differing degrees, all provinces/territories continue to provide mammography services to asymptomatic women outside the structure of the organized programs. It has been estimated that as much as 80% of bilateral mammography provided in this manner is for screening purposes[9]. Consequently it is referred to as "opportunistic screening".

Table 1
Comparison of Selected Characteristics of Organized Breast Cancer Screening Programs in Canada, the United Kingdom, Australia, Sweden and the Netherlands

Characteristic of National Breast Cancer Screening Program	Canada[8]	U.K.[5,6]	Australia[7]	Sweden[2]	Netherlands[3]
Year of implementation	1988	1988	1991	1986	1989
Number of programs	12	1	1	26	1
Target age group	50-69	50-64	50-69	40-74 (50-69 in some counties)	50-69
Modality of screening	Mammography ± clinical breast examination	Mammography	Mammography	Mammography	Mammography
Number of mammographic views	2-view	1-view	2-view	2-view	2-view
Screening interval	Biennial	Tri/Biennial	Biennial	Biennial	Biennial
Breast cancer incidence in target group prior to implementation of screening, 1982 (age adjusted to the world standard population)[10]	253.36 per 100,000	227.57 per 100,000	210.94 per 100,000 (New South Wales)	248.53 per 100,000	257.11 per 100,000

Organized Screening in Canada

Organized screening in Canada was initiated in 1988 on the recommendation of a national workshop[11]. The workshop consisted of expert representatives from government, and key professional and voluntary organizations. The result was a recommendation that women aged 50 to 69 be offered, and encouraged to participate in, an early detection program offered every 2 years, consisting of mammography, physical examination of the breast by a health care professional, and teaching and monitoring of breast self-examination.

In December 1992, the federal government launched the first phase of the Canadian Breast Cancer Initiative (CBCI), with stable, ongoing funding of $25 million over 5 years. Under the Canadian Breast Cancer Screening component of this initiative, Health Canada enabled a federal/provincial/territorial network to collaborate in the implementation and evaluation of breast cancer screening programs in Canada.

Following the November 1993 National Forum on Breast Cancer, membership of the National Committee on Breast Cancer Screening was expanded, and the group became formally known as the National Committee for the Canadian Breast Cancer Screening Initiative (CBCSI). Its activities included fostering the development of high-quality, organized breast cancer screening programs in Canada with the following essential components: a population-based outcome goal; information about the target population; emphasis on hard-to-reach groups; meticulous quality assurance; outcome data and analysis; information systems and linkages; a woman-centred focus; and excellent coordination with high-quality diagnosis and follow-up. The Canadian Breast Cancer Screening Database, derived from provincial breast screening program data, was developed through a collaborative effort of the National Committee, in 1993. The National Committee for the CBCSI continues its work today as a component of Phase II (1998-2003) of the CBCI (see Future Directions).

The first screening program began in British Columbia in 1988, and programs have since been established in all provinces and in the Yukon and the Northwest Territories. Breast cancer screening in all organized programs includes a bilateral two-view screening mammogram. Manitoba, Ontario, and Newfoundland also provide a clinical breast examination (CBE) carried out by a trained health professional, and Nova Scotia and Prince Edward Island provide a modified CBE carried out by a technologist. In addition, most programs provide information and/or instruction on breast self-examination.

In each province or territory, women of the target age are recruited to the screening program through a letter of invitation from the program, physician referral, or self-referral (except in Quebec for the latter). At the screening facility, which may be a mobile unit or a fixed centre, women receive bilateral two-view mammography. In addition to mammography screening, they undergo clinical breast examination or are encouraged to obtain regular clinical breast exams outside of the program from their family physician.

All programs provide screening results to both the woman and her physician. If the screening result is normal, women are recalled by letter for another routine screen generally after 2 years. However, some women are recalled annually because of their age, mammographic findings, family history, or other factors that vary across programs. Women with abnormal screening results are informed, along with their family physician, of the need for further assessment (which is usually coordinated by the women's physician). Women are also encouraged to consult their physician if symptoms develop in the interval between screening episodes.

Provincial and territorial screening programs obtain information on diagnostic procedures and cancer outcomes through a combination of active follow-up with family physicians, information received from assessment centres, and linkage with provincial medical service databases and cancer registries. Data from the Canadian Breast Cancer Screening Database provide a basis for assessing the performance of organized breast cancer screening in Canada. No comparable data are available from screening activities conducted outside the structure of the screening programs.

History of the Evaluation Indicators Working Group

The necessity for a standardized method of evaluation for all Canadian breast cancer screening programs was first recognized at the 1990 Interchange meeting. This need was identified once more at the 1993 National Forum on Breast Cancer, as well as at the 1997 Workshop on Organized Breast Cancer Screening in Canada[12]. With the Canadian Breast Cancer Screening Database newly established, the infrastructure required to formalize a set of performance measures was in place. The CBCSI's Quality Assurance Working Group, Database Technical Subcommittee, and Database Management Sub-Committee began the process of identifying performance and quality measures and indicators to fulfill these past recommendations. In order to effectively devote time and human resources to the project, the Evaluation Indicators Working Group (EIWG) was formed in 1999.

The process of developing evaluation indicators began with the identification of a set of general categories, which were derived from the results of two surveys of the provincial/territorial programs. Ultimately, nine categories were selected: recruitment and retention, client experience, technical aspects, mammography interpretation, diagnostic assessment and diagnosis, treatment, survival and mortality, data quality assurance, and program management. These categories were then assigned performance and quality indicators gathered through a review of national documents from various countries, published research literature, Canadian federal documents, and Canadian provincial/territorial screening program annual reports. The review focused on indicators that were currently available for breast cancer screening programs in publicly funded health care systems. The EIWG selected indicators from the initial findings on the basis of outcomes, pragmatic considerations, and efficiency.

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 30 core performance and quality indicators, target outcomes for some of these indicators, as well as recommendations on practical means to gather and report these data[13]. Subsequent meetings of the Working Group resulted in the following guidelines for reporting a key set of "performance measures".

Performance Measure Development

In order to achieve reductions in breast cancer mortality and morbidity and to minimize the unwanted effects of screening, the delivery of organized screening must be of high quality. The performance measures and targets presented in this document were selected on the basis of their utility for assessing program progress toward these goals. The 11 performance measures detailed here generally met the following criteria:

- Data for the measure were regularly available.
- Data available for the measure were of high quality.
- Meaningful targets could be defined on an evidentiary basis*.
- Measures and targets would be useful for national comparison.

- Monitoring on an annual basis would be valuable.
- Each measure was widely accepted for use in program evaluation.

**No targets were set for in situ cancer detection rate, given the controversy surrounding the natural history of the condition (see Performance Measures Under Review in Future Directions).*

Application

Through its monitoring and reporting role, the Database Management Sub-Committee (DMC) of the National Committee for the CBCSI produces a routine biennial report: *Organized Breast Cancer Screening in Canada*. The purpose of this report is to provide formal feedback to the programs regarding their relative performance and to assess the national picture. The standardized performance established in this document will serve as a consistent template for reporting progress over time, as well as providing a set of initial targets for programs to strive toward.

Data Sources and Collection

The performance measures are calculated using data from the Canadian Breast Cancer Screening Database (CBCSD) along with routinely available national statistics and population estimates. The CBCSD is a national breast screening surveillance system that permits the monitoring and evaluation of organized breast cancer screening across Canada. Established in 1993, it is operated and maintained through the continued collaboration of the provinces and territories and the Cancer Division at the Centre for Chronic Disease Prevention and Control, Health Canada. The CBCSD currently contains screening information from program inception up to the end of 1998 for the following provinces: British Columbia, Alberta, Saskatchewan, Manitoba, Ontario, Quebec, New Brunswick, Nova Scotia, and Newfoundland. Test data from Prince Edward Island and the Northwest Territories are currently being analyzed. Because the Yukon does not have a computerized information system, its data are not currently available to the CBCSD. Through the Canadian Breast Cancer Screening Initiative, the CBCSD is managed by the Database Management Sub-Committee and implemented by the Database Technical Subcommittee.

The monitoring of screening programs requires reliable, standardized information that is comparable across provinces. Some follow-up data must be obtained from external sources, and this complicates the evaluation process. 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 is that some programs experience delays in obtaining registry data. In addition, analyses have suggested that prognostic data vary from one program to another because of the different ways in which breast tumours are assessed and staged. This must be taken into account when the results of the performance measures across programs are integrated and compared.

Context of Performance Measures

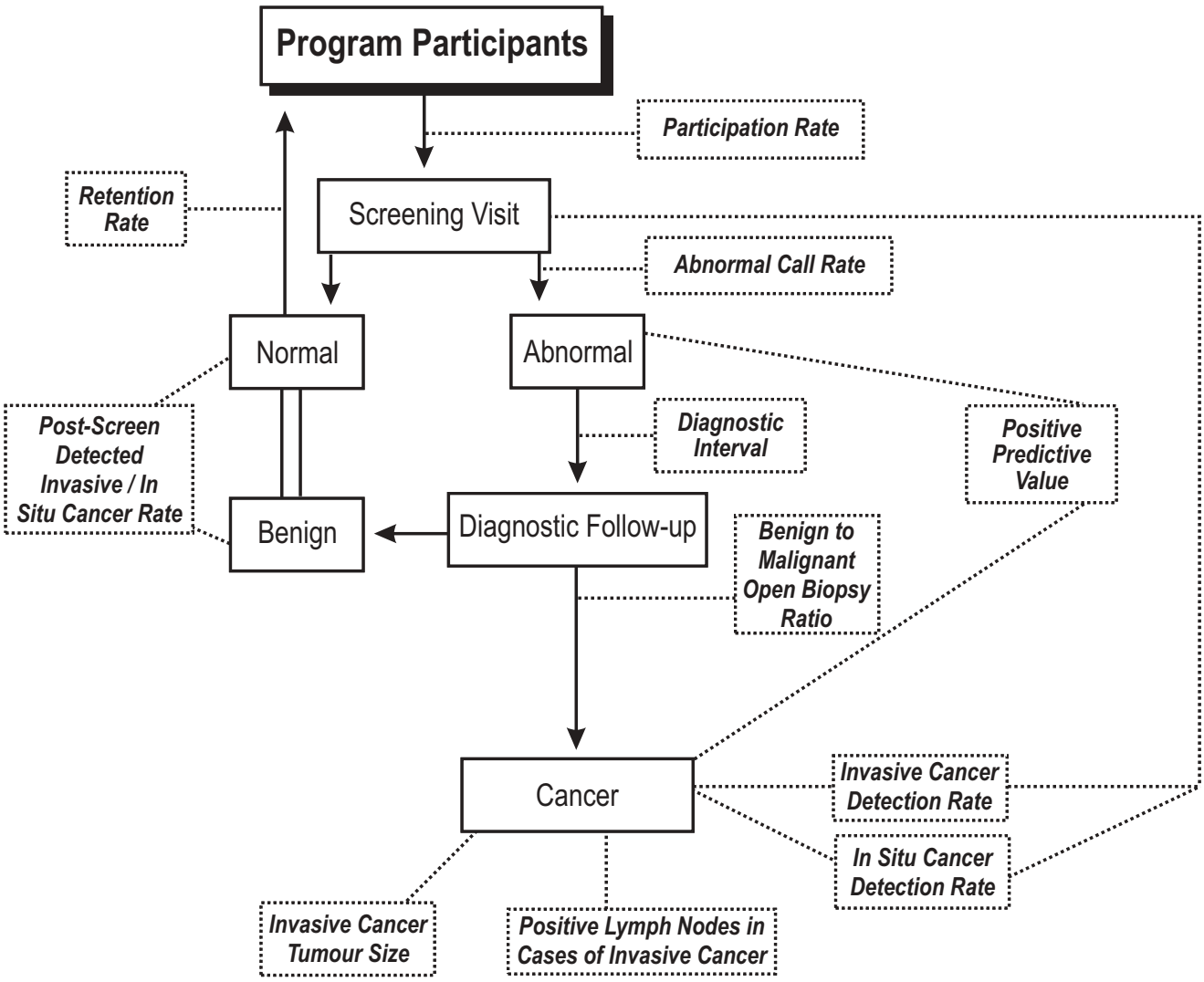
For the purposes of these guidelines for reporting performance measures, the target population for evaluation is the same as the national target population for organized screening. This population is defined as asymptomatic women between the ages of 50 and 69 years with no prior diagnosis of breast cancer.

The targets and standards established in this document are intended to apply to the programs' target group as a whole. It is recognized, however, that for some evaluation purposes it may be appropriate to further stratify the target group in terms of demographic characteristics, screening history, or referral of abnormal result by modality. When measures are used for comparison among Canadian programs or with programs in other countries, it is necessary to age-standardize the results using the appropriate population as the standard.

Five programs (Manitoba, Ontario, Nova Scotia, Prince Edward Island, and Newfoundland) offer clinical breast examination (CBE) in addition to mammography. A special comparison must be made of programs that offer CBE and those offering mammograpy alone. CBE in addition to mammography may confer an advantage in that some palpable breast abnormalities go undetected with mammography. Consequently, programs providing both screening methods may have higher abnormal call rates, making comparisons and the computation of meaningful national figures difficult. An independent research effort to evaluate programs offering CBE is currently under way.

Many of the performance measures presented here provide meaningful measures of program progress only 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 measure. For example, while increased participation and retention will always be desirable, targets for measures such as positive predictive value and biopsy yield ratio are set with the realization that we must tolerate some false-positive results in order to maximize cancer detection. At the same time, performance measures and targets are not necessarily meaningful on their own, and must be considered in relation to each other and (in some circumstances) in relation to other relevant data. For instance, the cancer detection rate must be considered in relation to the underlying cancer incidence rate in the general population before programmatic screening was implemented. An illustration to clarify the relations among the performance measures is presented in Figure 1.

Figure 1


Program Participants
Participation Rate
Retention Rate
Screening Visit
Abnormal Call Rate
Normal
Abnormal
Post-Screen Detected Invasive / In Situ Cancer Rate
Diagnostic Interval
Positive Predictive Value
Benign
Diagnostic Follow-up
Benign to Malignant Open Biopsy Ratio
Cancer
Invasive Cancer Detection Rate
In Situ Cancer Detection Rate
Invasive Cancer Tumour Size
Positive Lymph Nodes in Cases of Invasive Cancer

Program Performance Measures

Participation Rate

Definition:

Percentage of women who have a screening mammogram (calculated biennially) as a proportion of the eligible 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.

Note that program participation rate does not represent all breast cancer screening in Canada. In most provinces screening can also be delivered outside the structure of the program.

Calculations:

$$\frac{\text{Number of women screened at least once (per 2-year period)}}{\text{Target population (1st \& 2nd year populations averaged from census/forecast)}} \times 100 = \text{Participation Rate (\%) (biennial)}$$

Details:

In the case of multiple screens, age at the first screen is the criterion used to determine whether the woman was in the target population.

Target population (denominator) should be obtained from the most recent census results and/or forecasts of population available from Statistics Canada.

Targets:

Canada	**≥ 70% of the eligible population (ages 50-69)**
Europe[3]	60% (ages 50-64)
United Kingdom[5]	70% of women invited (ages 50-64)
Australia[7]	70% (ages 50-69)

Status:

34.3% of Canadian women (age 50-69, 1997-98) received a program screen[8].

Note: From the results of the National Population Health Survey it is estimated that 53.6% of Canadian women received a (program or non-program) mammogram in the 1996/97 2-year period[14].

Evidence:

Based on basic principles of population screening[15,16].

Extrapolation from the results of randomized controlled trials[17,18].

Retention Rate

Definition: The estimated percentage of women who are re-screened within 30 months of their previous screen.

Context: Optimal benefits of screening are brought about by regular participation in the screening program (at least every 2 years). At present there is no indication that the benefits of screening are lost if re-screening occurs up to 6 months after the recommended interval (i.e., 30 month interval).

Calculations: Actuarial Method for Survival Data

$s_t = 1 - (p_0 p_1 p_2 \ldots p_t)$

where $p_t = 1 - q_t$

$q_t = e_t / n^*_t$

$n^*_t = n_t - \frac{1}{2} c_t$

s_t is the estimated cumulative probability of returning from baseline to the end of the study interval that begins at t;

p_t is the estimated probability of not returning during the study interval that begins at time t;

q_t is the estimated probability of women returning during the study interval that begins at time t;

e_t is the number of women returning in the study interval that begins at time t;

n_t is the number of women present at the beginning of the study interval that begins at time t;

c_t is the number censored (because of death, breast cancer, or age limit—68 years) during the interval that begins at time t.

Targets:

Canada	**≥ 75% re-screened within 30 months (age 50-69)**
Australia[7]	75% screened in the previous round (age 50-69); of those re-screened, > 90% to be screened biennially within 27 months.

Status: 79.0% of women (age 50-59, 1997-98) re-screened within 30 months of previous screen

65.0% of women (age 60-69, 1997-98) re-screened within 30 months of previous screen[8].

Evidence: Related to participation rate, sojourn time, screening interval studies[19], and randomized controlled trials[17,18].

Abnormal Call Rate (%)

Definition: Percentage of women screened who are referred for further testing because of abnormalities found with a program screen.

Context: Abnormal call rate is a meaningful indicator when considered in the context of positive predictive value and cancer detection rate. Also, relative to the underlying breast cancer incidence rate, it is an indicator of the quality of mammographic image or interpretation. Abnormal call rate will generally be higher for initial screens (which detect prevalent cancers) than for re-screens.

Calculations:

$$\frac{\text{Number of recalls due to abnormal screens}}{\text{Number of women screened}} \times 100 = \text{Abnormal Call Rate (\%)}$$

Targets:

Canada	**< 10% (initial screen); < 5% (re-screens) (age 50-69)**
Sweden[2]	9% (overall)
Europe[3]	< 7% (initial screen); < 5% (subsequent screens)
United Kingdom[6]	< 7% (initial screen); < 7% (subsequent screens)
Australia[7]	< 10% (initial screen); < 5% (subsequent screens)

Status:

Abnormal recall rates by mode of detection, 1997 and 1998 screen years[8]

Mode of Screening		Age 50-69 (%)
Abnormal by mammography	Initial screen	9.3
	Re-screen	4.9
Abnormal by mammography and/or CBE	Initial screen	11.1
	Re-screen	6.2

Evidence: Measured in randomized controlled trial[17].

Invasive Cancer Detection Rate

Definition:

Number of women detected with invasive cancer during a screening episode per 1,000 women screened.

Context:

Cancer detection rate is meaningful for program evaluation when considered in relation to the abnormal call rate, post-screen cancer detection rate, and the underlying breast cancer incidence rate. The cancer detection rate in an organized screening program should generally exceed the cancer incidence rate in the population prior to organized screening, because some cancers would remain asymptomatic in the absence of screening. Cancer detection rates will generally be higher for initial screens (which detect prevalent cancers) than for re-screens.

Calculations:

$$\frac{\text{Number of invasive cancers detected}}{\text{Number of screens}} \times 1{,}000 = \text{Cancer Detection Rate per } 1{,}000$$

Targets:

Canada	**> 5 per 1,000 on initial screen (age 50-69)** **> 3 per 1,000 on re-screens (age 50-69)**
United Kingdom[6]	> 3.6 per 1,000 first screen (age 50-64); > 4.0 per 1,000 subsequent screens (age 50-64)
Australia[7]	> 5 per 1,000 first screen; > 2 per 1,000 subsequent screens (age 50-69)

Status:

Invasive cancer detection rates per 1,000 screens by mode of detection 1997 and 1998 screen years[8]

Mode of Screening		Age 50-69 (%)
Detected by mammography alone	Initial screen	5.3
	Re-screen	3.5
Detected by mammography and/or CBE	Initial screen	6.7
	Re-screen	4.1

Evidence:

Based on randomized controlled trials[17,18] and the experience of other breast cancer screening programs[5,7].

In Situ Cancer Detection Rate

Definition:	Number of women detected with ductal carcinoma in situ (DCIS) cancer (rather than invasive cancer) during a screening episode per 1,000 women screened.
Context:	While there is no definitive link between in situ cancer and invasive cancer, in situ cancer detection may be interpreted as an indicator of screening quality when considered in relation to the cancer detection rate and underlying cancer incidence rate.
Calculations:	$\frac{\text{Number of in situ cancers detected}}{\text{Number of screens}} \times 1{,}000 = \text{In Situ Cancer Detection Rate per 1,000}$

Targets:	**Canada**	**Surveillance and Monitoring Purposes Only**
	United Kingdom[5]	0.4-0.9 per 1,000 (initial screen) 0.5-1.0 per 1,000 (subsequent screens)
	Australia[7]	10%-20% of cancers detected are in situ

Status:	0.9 per 1,000 in situ cancer (aged 50-69, 1997-98)[8]
Evidence:	It seems inappropriate to set targets for DCIS given the current paucity of evidence concerning the transition of DCIS to invasive cancer and the continually increasing sensitivity of screening techniques[20].

Diagnostic Interval

Definition:	Total duration from abnormal screen to resolution of abnormal screen.
Context:	An abnormal screen result can induce morbidity, given the negative psychological impact it can have on a client, even if follow-up is ultimately negative. Moreover, excessive delay to diagnosis may worsen prognosis. Work-up should therefore be completed expeditiously[21]. Note that some Canadian programs do not have integrated diagnostic capabilities, making measurement of diagnostic interval more difficult.
Calculations:	(date of diagnosis) - (screen date) = Diagnostic Interval $\frac{\text{Number of diagnostic intervals within the target time range}}{\text{Total number of abnormal screens}} \times 100 = \%$ of clients with the target time range
Targets:	**Canada**: **≥ 90% within 5 weeks if no open biopsy (age 50-69)** **≥ 90% within 7 weeks if open biopsy (age 50-69)** Australia[7]: 90% to have first assessment within 10 working days 70% to be provided with definitive diagnosis or recommendation for biopsy within 2 working days of first assessment.
Status:	74.5% within 5 weeks (no open biopsy, age 50-69, 1997-98) 44.9% within 7 weeks (open biopsy, age 50-69, 1997-98)
Evidence:	Based on basic principles of screening[15,16] and screening program evaluation research[22].

Positive Predictive Value

Definition:

Proportion of abnormal cases with completed follow-up found to have breast cancer (invasive or in situ) 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. PPV tends to improve with re-screening because the initial screen establishes a normal baseline. Consequently, PPV tends to be lower among initial screens relative to re-screens.

Calculations:

$$\frac{\text{Number of screen detected cancers}}{\text{Total number of abnormal screens with complete work-up}} \times 100 = \text{Positive Predictive Value (\%)}$$

Note: Includes invasive and in situ cancers

Targets:

Canada **≥ 5% (initial screen) (age 50-69)**
≥ 6% (re-screen) (age 50-69)

Status:

Positive predictive value by mode of detection, 1997 and 1998 screen years[8]

Mode of Screening		Age 50-69 (%)
Detected by mammography alone	Initial screen	4.8
	Re-screen	7.1
Detected by mammography and/or CBE	Initial screen	5.0
	Re-screen	6.1

Evidence:

Based on methodology in screening program evaluation studies[23].

Benign to Malignant Open Biopsy Ratio

Definition: Among open biopsies, the ratio of number of benign cases to the number of malignant cancer cases.

Context: Benign to malignant open biopsy ratios provide an indication of the quality of the presurgical assessment. Diagnostic specificity and sensitivity are reciprocal. Consequently there is a limit to the extent to which biopsy yield ratios can be improved. This indicator is most meaningful when considered in relation to the underlying breast cancer incidence rate and the post-screen detected cancer rate.

Calculations:

$$\frac{\text{Number of benign cases detected by open biopsy}}{\text{Number of malignant cancer cases detected by open biopsy}} : 1$$

Benign to Malignant Open Biopsy Ratio

Note: Each open biopsy performed represents a case. It may be useful to present these figures with confidence intervals when small numbers of cases are observed.

Targets:

Canada	**≤ 2:1 open (initial & re-screen combined) (age 50-69)**
Sweden[2]	3:1 (first & re-screen combined)
Europe[3]	2:1 (first screen), 1:1 (re-screen)
United Kingdom[5]	3:1 (first & re-screen combined)
Australia[7]	2:1 (first screen), 1:1 (re-screen)

Status: 1.6:1 benign to malignant open biopsy ratio (age 50-69, 1997-98)

Evidence: The targets are based on experience from research trials (e.g., Swedish Two County study)[8].

Invasive Cancer Tumour Size

Definition:	Percentage of invasive cancers with tumour size of 10 mm in greatest diameter as determined by the best available evidence: 1) pathological, 2) radiological, 3) clinical.
Context:	Invasive tumour size is the best known prognostic indicator. The purpose of mammographic screening is to detect pre-clinical cancers before symptoms are apparent.
Calculations:	$\frac{\text{Number of invasive tumours} \quad 10\text{mm}}{\text{Total number of invasive tumours}} \times 100 = \%$ of invasive tumours 10mm

Targets:		
	Canada	**> 25% ≤ 10 mm (age 50-69)**
	Sweden[2]	> 50% < 15 mm
	Europe[3]	20% 10 mm (initial screen) 25% 10 mm (subsequent screens)
	United Kingdom[5]	1.5 per 1,000 (< 15 mm; initial screen) 1.65 per 1,000 (< 15 mm; subsequent screens)
	Australia[7]	> 8 per 10,000 (10 mm)

Status:	37.6 % of tumours 10 mm (age 50-69, 1997-98)[8]
Evidence:	Stage-specific prospective studies and trials[2,3,7].

Positive Lymph Nodes in Cases of Invasive Cancer

Definition: Proportion of invasive cancers in which the cancer has invaded the lymph nodes.

Context: Positive lymph nodes are a prognostic indicator. The purpose of mammographic screening is to detect breast cancer as early as possible – before it spreads to the lymph nodes.

Calculations:

$$\frac{\text{Number of cases of invasive cancer with positive lymph nodes}}{\text{Total number of invasive cancer cases in which lymph nodes were assessed}} \times 100 = \%\ \text{cases with positive lymph nodes}$$

Note: Excludes cases in which lymph nodes are not assessed.

Targets:

Canada	**< 30% node positive (age 50-69)**
Sweden[2]	> 70% node negative (age 50-64)

Status: 21.7% node positive in assessed cases of invasive cancer (age 50-69, 1997-98)[8]

Evidence: Stage-specific prospective studies and trials[2,3,7].

Post-screen Detected Invasive Cancer Rate

Definition:

Number of women with a diagnosis of invasive breast cancer after a negative screening episode per 10,000 person-years at risk, within 12 AND 24 months of the screen date.

Context:

Post-screen detected invasive cancer rate is an indicator of the sensitivity of the screening program. This rate is affected by population incidence, age, rate of disease progression, and screening interval recommendation. A high rate may negatively affect the mortality reduction expected for a successful, organized screening program. The accuracy of this measure is dependent on the completeness of cancer registration.

Calculations:

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

$$\frac{\text{Number of cancers detected in the 0-24 month interval after a negative screening episode}}{\text{Total person-years at risk (0-24 months post screen)}} \times 10{,}000 = \text{24-month Post-screen Detected Invasive Cancer Rate per 10,000}$$

Note: Calculations include all cases regardless of screening interval recommendation.

Targets:

Canada	**< 6 per 10,000 person-years (within 12 months, age 50-69)** **< 12 per 10,000 person-years (within 24 months, age 50-69)**
United Kingdom[5]	12 per 10,000 (within 24 months)
Australia[24]	< 6.5 per 10,000 (within 12 months)

Status:

Cancers detected outside of program after normal screening episode among program participants aged 50-69 at screening, 1994 and 1995 screen years[8]

Months After Screening	0-12	0-24
Rate per 10,000 person-years at risk	5.8	8.6

Evidence:

Studies of interval cancer[25,26].

Future Directions

The development of a set of performance measures for organized breast cancer screening programs must be 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 performance measures presented in this document varies greatly from measure to measure and is subject to change with the continual introduction of new research evidence. The data used in the calculation of these measures, and possible future measures, are still maturing in terms of quality and timely availability. Consequently, it is a challenge to establish comprehensive, long-term evaluation plans with valid, reliable performance measures. As a result, a multiphase strategic plan has been proposed with the release of these guidelines.

Phase 1: Monitoring Proposed Performance Measures (Years 1 and 2)

The first formal use of these measures will be in subsequent releases of the biennial report on *Organized Breast Cancer Screening Programs in Canada*[8]. In 2 years' time, a subsequent strategic working group will be formed to examine the outcomes of these monitoring efforts. At that time, the group will reassess the 11 measures in terms of progress made towards achieving the national targets. Targets will be adjusted or redefined by consensus supported by new research or expert opinion. Changes to the definitions of the measures and methods of calculation will also be considered on the same basis. Several measures have already been slated for special review for a number of reasons, as follows.

Performance Measures Under Review

In situ cancer detection rate: While ductal carcinoma in situ (DCIS) is widely accepted as an obligate precursor of invasive disease, the time frame in which this occurs is not firmly established. The potential for cases of DCIS to remain asymptomatic throughout the individual's natural lifespan suggests a potential for *overdiagnosis* with its attending negative consequences. The Working Group will continue to monitor in situ cancer detection rates and will consider defining a target under the appropriate circumstances.

Benign to malignant open biopsy ratio: The precision of this measure needs to be considered in light of the small numbers of cases involved. Therefore it is recommended that open biopsy yield ratios be presented with confidence intervals. These ratios need to be observed over time to assess trend stability and thus the appropriateness of the specified target.

Positive lymph nodes in cases of invasive cancer: In cases of invasive cancer, involvement of the lymph nodes can be difficult to establish. Data for this indicator may therefore be unreliable. New diagnostic techniques (e.g., sentinel node biopsy) are likely to have an impact on this indicator.

Performance by screening modality: Screening modality (mammography alone; mammography and/or CBE) has an impact on program performance, as indicated in the Status sections of the preceding measures. Separate targets will be established by consensus, supported by new research or expert opinion, for each screening modality.

Phase 2: Revisiting Other Performance Measures (Years 3 and 4)

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 measures, the timely availability of high quality data was also an influential factor. Meaningful targets, useful for national comparison through frequent monitoring, were also requisite. These criteria do not, however, fully cover the range of performance measures needed to establish comprehensive long-term evaluation plans. From that perspective, factors such as equitable access, representative participation, acceptability of services to clients, cost minimalization, and program promotion must be assessed. In recognition of the need for a more complete inventory of indicators for use in future evaluation initiatives, a working group will be formed again, after 2 years, to reconsider the feasibility of adding new measures or including measures previously explored (but not published) to a second edition of this document. Previously explored measures with the potential for future use are outlined in Table 2.

Phase 3: Ongoing Evaluation (Year 5)

Having completed phases 1 and 2 of the evaluation indicators strategic plan, the working group will consider the need to continue the process of revisiting this collection of performance measures. A plan will be formed reflecting the contemporary state of breast cancer research and the performance measurement needs of organized screening programs in Canada. This plan will detail the future scope of the working group's activities.

Table 2: Possible Canadian Breast Cancer Screening Measures for Future Use

Proposed Measure	Outcome Area	Level Reportable	Frequency of Reporting
Average wait time	Client experience	Regional, provincial, local	Annual
Duration from abnormal screen to first assessment	Client experience	Regional, provincial, local	Annual
Time from diagnosis to notification of the client	Client experience	Regional, provincial, local	Annual
Total program operating cost per screen	Cost effectiveness	Regional, provincial, local	Infrequent
Total program operating cost per cancer detected	Cost effectiveness	Regional, provincial, local	Infrequent
Level of data completeness	Data quality	Local	Infrequent
Proportion of error for each data field	Data quality	Local	Infrequent
Diagnostic procedures after an abnormal screen	Diagnostic assessment	Regional, provincial, local	Annual
Tumour size TNM stage 2 or worse	Early detection	Regional, provincial, local	Annual
Management practices: policies/procedures	Management quality	Local	Infrequent
Proportion of women lost to follow-up	Management quality	Regional, provincial, local	Annual
Breast cancer relative survival rate	Mortality	Regional, provincial, local	Annual
Overall breast cancer mortality	Mortality	Regional, provincial, local	Annual
Proportion of units CAR* accredited	Technical quality	Regional, provincial, local	Annual
Technical quality: repeated films	Technical quality	Local	Annual
Sensitivity	Validity	Regional, provincial, local	Annual
Specificity	Validity	Regional, provincial, local	Annual

**Canadian Association of Radiologists*

References

1. Kerlikowske K, Grady D, Rubin SM, Sandrock C, Ernster VL. *Efficacy of screening mammography. A meta-analysis*. JAMA 1995;273:149-54.

2. Tabar L, Fagerberg G, Duffy SW, Day NE, Gad A, Grontoft O. *Update of the Swedish two-county program of mammographic screening for breast cancer*. Radiol Clin North Am 1992;30:187-209.

3. Europe Against Cancer. *European guidelines for quality assurance in mammographic screening*. Luxembourg: Office for Official Publications of the European Commission, June 1996.

4. Europe Against Cancer. *European guidelines for quality assurance in mammographic screening*. Luxembourg: Office for Official Publications of the European Commission, January 2001.

5. *Standards for the NHS Breast Screening Programme* (updated August 1998). Sheffield, England: NHSBSP publication.

6. *Informing choice in breast screening: Breast Cancer Screening Program annual review 2001*. Sheffield, England: NHSBSP publication, 2001.

7. Commonwealth Department of Human Services and Health. *National program for the early detection of breast cancer: national accreditation requirements*. Canberra, Australia: Commonwealth Department of Human Services and Health, March 1994.

8. Health Canada. *Organized breast screening programs in Canada 1997 and 1998 report*. Ottawa: Minister of Public Works and Government Services Canada, 2001.

9. Gaudette LA, Altmayer CA, Nobrega KMP, Lee J. *Trends in mammography utilization, 1981 to 1994*. Health Reports 1996;8(3).

10. *Cancer incidence in five continents (volume VII)*. Lyon: IARC Scientific Publication No. 143, 1997.

11. The Workshop Group. *Reducing deaths from breast cancer in Canada*. Can Med Assoc J 1989;141;199-201.

12. Health Canada. *National workshop on organized breast cancer screening programs: Report of the theme discussion group*. Ottawa: Health Canada, 1997.

13. Horne T, Evaluation Indicators Working Group. *Proposed indicators for performance and quality of breast screening programs in Canada*. Ottawa: Health Canada, April 2000.

14. Maxwell CJ, Bancej CM, Snider J. *Predictors of mammography use among Canadian women aged 50-69: findings from the 1996/97 National Population Health Survey*. Can Med Assoc J 2001;164(3):329-34.

15. Wilson JMG, Jungner G. *Principles and practice of screening for disease*. Geneva: World Health Organization, 1968.

16. Day NE, Williams DRR, Khaw KT. *Breast cancer screening programs: the development of a monitoring and evaluation system*. Br J Cancer 1989;59:954-58.

17. Tabar l, Fagerberg G, Chen HH, Duffy SW, Smart CR, Gad A, et al. *Efficacy of breast cancer screening by age: new results from the Swedish two-county trial*. Cancer 1995;75:2507-17.

18. Nyström L, Rutqvist LE, Wall S, Lindgren A, Lindqvist M, Ryden S et al. *Breast cancer screening with mammography: overview of the Swedish randomized trials*. Lancet 1993;341:973-8.

19. Duffy SW, Hill C, Estève J. *Quantitative methods for the evaluation of cancer screening*. London: Arnold, 2001.

20. Evans AJ, Pinder SE, Ellis IO, Wilson ARM. *Screen detected ductal carcinoma in situ (DCIS): overdiagnosis or an obligate precursor of invasive disease?* J Med Screen 2001;8:149-51.

21. Steering Committee on Clinical Practice Guidelines for the Care and Treatment of Breast Cancer. *Clinical practice guidelines for the care and treatment of breast cancer: 1. The palpable breast lump: information and recommendations to assist decision-making when a breast lump is detected*. Can Med Assoc J 1998;158(suppl):3S-8S.

22. Olivotto IA, Bancej C, Goel V, Snider J, McAuley RG, Irvine B, Kan L, Mirsky D, Sabine MJ, McGilly R, Caines JS. *Waiting times from abnormal breast screen to diagnosis in 7 Canadian provinces*. Can Med Assoc J 2001;165(3):303-4.

23. Blanks RG, Moss SM, Wallis MG. *Monitoring and evaluating the UK National Health Service Breast Screening Programme: evaluating the variation in radiological performance between individual programmes using PPV-referral diagrams*. J Med Screen 2001;8:24-8.

24. Draft Breastscreen Australia Monitoring Plan, National Advisory Committee, work in progress.

25. Moss SM, Blanks RG. *Calculating appropriate target cancer detection rates and expected interval cancer rates for the UK NHS Breast Screening Programme*. J Epidemiol Community Health 1998;52:111-15.

26. Rickard MT, Taylor RJ, Fazli MA, El Hassan N. *Interval breast cancers in an Australian mammographic screening program*. Med J Aust 1998;169:184-87.

27. W.H.O. Public Health Papers No. 34. Geneva: World Health Organization, 1968.

Appendix A

Conceptual Framework

The Conceptual Framework is an updated modification of the classic Wilson and Jungner[27] 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 B

Evaluation Indicators Working Group Members

Dr. Anna Chiarelli
(Evaluation Indicators Working Group Chair)
Ontario Breast Screening Program
Cancer Care Ontario
620 University Avenue
Toronto, Ontario M5G 2L7

Gregory Doyle
Breast Screening Program for Newfoundland and Labrador
20 Cordage Place
P.O. Box 13122, Station A
St. John's, Newfoundland A1 B 4A4

Lisa Kan
B.C. Cancer Agency
686 West Broadway, 8th Floor
Vancouver, B.C. V5Z 1G1

Zeva Mah
Screen Test
Alberta Cancer Board
Suite 120, 1040-7th Avenue SW
Calgary, AB T2P 3G9

Norah Smith
PEI Breast Screening Clinic
Queen Elizabeth Hospital
Department of Diagnostic Imaging
1 Riverside Dr., P.O. Box 6600
Charlottetown, PEI C1A 8T5

Christina Bancej
Cancer Division
Centre for Chronic Disease Prevention and Control
Population and Public Health Branch
Health Canada
Rm 2431 Building 6
Tunney's Pasture, PL 0602E2
Ottawa, Ontario, Canada K1A OL2

Jay Onysko
Cancer Division
Centre for Chronic Disease Prevention and Control
Population and Public Health Branch
Health Canada
Rm 2426 Building 6
Tunney's Pasture, PL 0602E2
Ottawa, Ontario, Canada K1A OL2

Appendix C

Glossary

Asymptomatic

⇢ A woman who does not report symptoms and appears without signs of disease at screening.

Cancer

⇢ Includes malignant and ductal carcinoma in situ (DCIS) of the breast.

Diagnosis

⇢ The first pathologic or cytologic diagnosis of cancer, last known biopsy for benign cases, or last intervention prior to a recommendation to return to screening or return for early recall[1].

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.

Fine-needle aspiration biopsy

⇢ A technique used to differentiate cystic from solid lesions in the breast. A needle is inserted into the lesion and material drawn out using a syringe. If the material is solid, it can be stained and the cells examined in a laboratory to determine whether or not they are benign or malignant.

Incident cancer

⇢ Cancers detected by a program screen after the initial screen.

In situ

⇢ Refers specifically to 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.

Initial screen

⇢ The first Canadian screening program screen provided to a woman.

Interval cancer

⇢ Any invasive breast cancer diagnosed in the interval following a “normal” screening result and before the next scheduled screening examination.

Invasive cancer

⇢ Cancer cells invading beyond the basement membrane of the milk duct or lobule. A ductal carcinoma in situ component may also be present in cases of invasive cancer.

Negative screening episode

⇢ A screening episode that concludes with normal findings, including program-initiated work up that did not reveal any cancer.

Open biopsy

⇢ Surgical removal of a breast mass under local anesthesia for subsequent microscopic examination by a pathologist.

Post-screen cancer

⇢ A cancer detected outside the program within 24 months of a negative screening episode.

Prevalent cancer

⇢ The proportion of the population with cancer at a given point in time.

Screen

⇢ Can comprise mammography, or both clinical breast examination and mammography, delivered by a program.

Screening episode (completed)

⇢ Defined for normal screens as the date of the last screen; for abnormal screens, the date of tissue diagnosis if biopsy is performed, the date of the last test before a return to screening or before the recommendation for repeat diagnostic imaging. A “negative screening episode” can include all follow-up, provided that the end result is negative.

Re-screening

⇢ Subsequent screening, according to policy, after initial screening under the program. This includes women who miss a scheduled round of screening.

Screen-detected cancer

⇢ Cancer detected as a result of a positive test with histologic confirmation attributed to the screening findings of the program.

Total person-years at risk

⇢ Within a 12 or 24-month period after a negative screening episode, women are considered at risk for postscreen detected cancer. Women contribute a count in the denominator for each year or fraction of a year within the period of interest before a post-screen detected cancer or the next regular program screen.