Int. J. Cancer: **110,** 756–762 (2004)

Publication of the International Union Against Cancer

MAMMOGRAPHY BENEFIT IN THE CANADIAN NATIONAL BREAST SCREENING STUDY-2: A MODEL EVALUATION

Adriana J. RIJNSBURGER[1]*, Gerrit J. VAN OORTMARSSEN[1], Rob BOER[1,2], Gerrit DRAISMA[1], Teresa TO[3,4], Anthony B. MILLER[3] and Harry J. DE KONING[1]

[1]*Department of Public Health, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands*
[2]*RAND Health, Santa Monica, CA, USA*
[3]*Department of Public Health Sciences, University of Toronto, Toronto, Ontario, Canada*
[4]*Population Health Sciences Program, Research Institute, The Hospital for Sick Children, Toronto, Ontario, Canada*

The CNBSS-2 among women aged 50–59 did not show any significant difference in breast cancer mortality between a control arm screened annually by CBE and a study arm screened by CBE and mammography. Because of this design, the benefit of screening compared to no screening could not be evaluated. We therefore conducted a modeling effort to estimate the benefit of mammography or CBE compared to no screening. We incorporated demographic, epidemiologic and screening characteristics of the CNBSS-2 in MISCAN. Stage-specific sensitivities of CBE, with and without mammography, and breast cancer incidence rate in the trial were estimated by comparing observed trial data with model predictions. We predicted the number of breast cancer deaths for both study arms of the CNBSS-2 and in the absence of screening, assuming improvement in prognosis by early detection. We estimated a 24–29% higher breast cancer incidence rate in the CNBSS-2 than the average Canadian rate. Estimated sensitivity of CBE (control arm) varied from 0.29 to 0.48 for stage T1c and from 0.6 to 0.65 for stage T2+. Estimated sensitivity of CBE supplemented with mammography (study arm) varied from 0.5 to 0.79 for stage T1c and was 0.95 for stage T2+. Expected breast cancer mortality reduction by annual CBE screening is 20.5% compared to no screening. Estimated breast cancer mortality reduction by mammography screening compared to no screening for the CNBSS-2 fell within the range 13.6–34.1%. Enrolled women had above average risk. Screening sensitivity in both arms was high. A benefit of mammography screening is supported by our modeling of the CNBSS-2 results.

Key words: *breast cancer; screening; mortality; mammography; clinical breast examination; Canadian National Breast Screening Study-2*

Several randomized controlled trials have assessed the efficacy of mammography screening at reducing breast cancer mortality. An updated overview of 4 Swedish randomized trials showed breast cancer mortality reductions in invited women of 16% (50–59 years) and 33% (60–69 years).[1] However, in a Cochrane review, the authors found smaller breast cancer mortality reductions in 2 so-called medium-quality trials.[2] The Malmö trial and the CNBSS-2 showed 20% (≥55 years) and 3% (50–59 years) breast cancer mortality reductions after 7 years of follow-up.

The CNBSS-2, however, was designed to evaluate the efficacy of annual mammography over and above annual CBE. The 7- and 13-year follow-up results of this randomized trial did not show a significant difference in mortality from breast cancer between the 2 groups, even though high detection rates were found.[3,4] Because the control arm was screened, no estimations of mammography benefit compared to no screening could be made, as is available from other mammography trials, nor any estimate of the benefit of CBE screening.

Here, we provide a quantitative interpretation of the CNBSS-2 results, using a MISCAN model evaluation of the trial. MISCAN is a microsimulation program, which has been successfully applied in the evaluation of different cancer-screening programs.[5–7] The main purpose of the model evaluation is to show what can be learned about breast cancer screening for women aged 50 and above and about the natural history of the disease from the CNBSS-2 in relation to what we know from other breast cancer–screening trials. The evaluation addresses 2 specific questions: *(i)* Can any explanations, resulting from the model evaluation, be given for the observed equal breast cancer mortality in both study arms of the CNBSS-2? *(ii)* What percentage of breast cancer mortality reduction by mammography screening compared to no screening can we estimate for the CNBSS-2?

MATERIAL AND METHODS

CNBSS-2

The CNBSS-2 is an individually randomized trial in women aged 50–59 years, designed to compare breast cancer mortality following annual screening by both mammography and physical examination of the breasts (MP group) to breast cancer mortality following annual screening by physical examination only (PO group). All participants of both study arms were taught breast self-examination. Women who volunteered for the study with no personal history of breast cancer and no mammograms in the previous 12 months were included. A detailed description of the trial, which started in 1980, and the 7- and 13-year follow-up results have been given elsewhere.[3,4,8] For the present analysis, we used the CNBSS-2 database, containing records for 39,405 women. We used the number of breast cancers detected at the different screening rounds, number of interval cancers, distribution of tumor characteristics, number of breast cancer deaths (11 years of follow-up), and age distribution and attendance of the trial population. There were small differences in numbers of screen-detected and interval cancers between the CNBSS-2 database and the published CNBSS-2 results.[3,4]

Abbreviations: CBE, clinical breast examination; CI, confidence interval; CNBSS-2, Canadian National Breast Screening Study-2; DCIS, ductal carcinoma *in situ;* Dev, deviance; df, degrees of freedom; MISCAN, microsimulation screening analysis; MP, mammography + physical examination; PO, physical examination only; RR, relative risk.

The study sponsor had no role in the conduction of the study or the writing of the report. The authors declare no conflict of interest.

Grant sponsor: Health Care Insurance Board (the Netherlands).

*Correspondence to: Department of Public Health, Erasmus MC, University Medical Center Rotterdam, PO Box 1738, 3000 DR, Rotterdam, the Netherlands. Fax: +31-10-4089449.
E-mail: a.rijnsburger@erasmusmc.nl

Received 17 September 2003; Revised 16 December 2003; Accepted 18 December 2003

DOI 10.1002/ijc.20143
Published online 15 March 2004 in Wiley InterScience (www.interscience.wiley.com).

Implementation of demographic, epidemiologic and screening characteristics of the CNBSS-2 in MISCAN

The MISCAN breast cancer–screening model[6,7] was used to analyze the results of both study arms of the CNBSS-2. A detailed description of the model is given in the Appendix. The PO and MP groups were implemented separately in the model. First, the number of women in each study arm, their age distribution, screening ages, screening interval and attendance rates at successive screening rounds were incorporated. We adjusted the model to the Canadian 1978–1982 clinical age-specific breast cancer incidence data[9] and used the Canadian clinical stage distribution for women aged 50–69 years.[10] There was no indication that age- and stage-specific breast cancer survival in Canada differed from that in the Netherlands, so Dutch survival rates were used (see Appendix). The resulting simulated breast cancer mortality rates from MISCAN were comparable to the 1981 Canadian national mortality data.[11]

Parameter estimation

In MISCAN, the probability of screen detection is dependent on the sensitivity of the screening test and the age-specific mean durations of the screen-detectable preclinical stages. We started our analysis with a base model in which mean durations are based on analyses of data from the Dutch nationwide screening program[12] (see Appendix).

Stage-specific sensitivities of CBE were derived first (model A) by comparing the observed stage-specific number of screen-detected cancers in round 1, rounds 2–5 and interval cancers in years 1–5 in the PO group with model predictions. Interval cancers are defined as those occurring less than 12 months after a screening examination that did not result in a recommendation for diagnostic evaluation. Numerical optimization with the downhill simplex method[13] was used to find CBE sensitivity estimates that minimize the deviance, calculated by subtracting the log likelihood of the estimated model from the log likelihood of the saturated model and multiplying by 2. A χ^2 test applied to the deviance was used as a test of goodness of fit.

The same procedure was performed for the MP group (model A), resulting in estimated stage-specific sensitivities of CBE supplemented with mammography. We assumed sensitivities of the different invasive stages to be ascending, with a maximum of 0.975 for stage T2+. Invasive cancers without known tumor size were not included in the observed stage-specific numbers, while expected stage-specific numbers were adjusted by the proportion of unknowns. Stage T4 cancers were classified under stage T2+ cancers.

Since women participated voluntarily in the trial, their breast cancer risk might have been different from the average Canadian level. Three alternative model variants were developed: breast cancer incidence rate together with the stage-specific sensitivities of CBE and of CBE supplemented with mammography were estimated in model B, the duration of the screen-detectable preclinical phase T2+ in addition to breast cancer incidence rate and stage-specific sensitivities were estimated in model C, and the total duration of all screen-detectable preclinical invasive stages in addition to breast cancer incidence rate and stage-specific sensitivities were estimated in model D. These parameters were estimated jointly by numerical optimization, minimizing the total deviance statistic for agreement between observed and expected numbers of cancers for the PO and MP groups, calculated as the sum of the individual deviances of both study arms.

Lead time, defined as the time between screen detection and clinical diagnosis in the absence of screening, is generated by MISCAN for every screen-detected case in the model. Mean lead time was estimated for both the PO and MP groups.

Breast cancer mortality

For each of the models (A–D) we predicted breast cancer mortality in the absence of screening, given estimated model parameters (breast cancer incidence rate, durations of screen-detectable preclinical invasive stages) and breast cancer survival. We then predicted the number of breast cancer deaths (1–11 years of follow-up) for both the PO and MP groups of the CNBSS-2, using estimated sensitivities from models A–D. We estimated mammography benefit and CBE benefit, both compared to no screening for the CNBSS-2, and compared predicted breast cancer mortality to observed breast cancer mortality in the CNBSS-2.

Improvement in prognosis after screen detection was defined as 1 minus the ratio of the risk of dying of screen-detected breast cancer divided by the risk if the same cancer had been diagnosed in the absence of screening. We assumed this benefit of early detection to be independent of the screening test used. We first used the improvement in prognosis after screen detection based on the results of analyses of the 5 Swedish mammography screening trials, which was estimated to be 1.000, 0.892, 0.814, 0.567 and 0.395, respectively, for cancers screen-detected in stages DCIS, T1a, T1b, T1c and T2+.[7] The effect of alternative assumptions for the improvement in prognosis after screen detection on expected mortality results (of model C) was also explored. We calculated 95% CIs for breast cancer mortality using the Poisson distribution.

RESULTS

Parameter estimates

The upper part of Tables I and II gives estimated stage-specific sensitivities of CBE supplemented with mammography in the MP group and of CBE alone in the PO group for 4 MISCAN models. The lower part shows the observed number of breast cancers in the CNBSS-2 and the expected number of breast cancers according to the 4 models. In model A, in which breast cancer incidence is assumed to be equal to the average Canadian rate, estimated sensitivities of CBE supplemented with mammography lead to a poor fit of the MP group (Dev = 78.8, 10 df) (Table I). The same is true for the PO group, where the expected numbers of breast cancers at screen 1 and interval cancers in years 1–5 are much lower than observed (Dev = 28.2, 10 df) (Table II).

Since the CNBSS is a volunteer trial, breast cancer risk of participants might have differed from the average Canadian level. We estimated breast cancer incidence together with stage-specific sensitivities and found a significantly improved fit [ΔDev (PO + MP) = 49.6, Δdf = 1] for a 40% higher incidence rate (model B). Estimated sensitivities were then, of course, lower compared to model A. However, in model B, there remained a large discrepancy, at both the first and the repeat screenings, between the observed high numbers of screen-detected T2+ cancers in the MP group and the expected numbers, even though the estimated sensitivity of CBE supplemented with mammography is as high as the preset limit (0.975). Furthermore, the estimated CBE sensitivity of 0.97 for stage T2+ could be considered high.

In model C, we therefore also estimated the duration of the screen-detectable preclinical phase T2+ and found a significantly improved fit [ΔDev (PO + MP) = 14.8, Δdf = 1] for a 29% higher incidence rate and a 95% longer duration of the screen-detectable preclinical phase T2+ (and corresponding estimated T2+ sensitivities of CBE and CBE supplemented with mammography of 0.649 and 0.953, respectively). Instead of estimating the duration of the screen-detectable preclinical phase T2+, we finally estimated the total duration of all screen-detectable preclinical invasive stages (model D). This model also resulted in a significantly improved fit [ΔDev (PO + MP) = 21.8, Δdf = 1] with a 24% higher breast cancer incidence rate, a 115% longer duration of all screen-detectable

TABLE I – COMPARISON OF OBSERVED CNBSS-2 RESULTS OF THE MP GROUP WITH MODEL PREDICTIONS FOR DIFFERENT SENSITIVITY ESTIMATES OF CBE SUPPLEMENTED WITH MAMMOGRAPHY, BREAST CANCER INCIDENCE RATES IN THE TRIAL, DURATION OF THE SCREEN-DETECTABLE PRECLINICAL PHASE T2+ AND TOTAL DURATION OF ALL SCREEN-DETECTABLE PRECLINICAL INVASIVE STAGES

		Model A: average Canadian incidence	Model B: 40.0% ↑ incidence	Model C: 29.15% ↑ incidence, 94.82% ↑ duration T2+	Model D: 23.66% ↑ incidence, 115.03% ↑ duration invasive stages
Sensitivity CBE + mammography					
DCIS		0.771	0.555	0.600	0.635
T1a (≤5 mm)		0.600	0.575	0.579	0.322
T1b (6–10 mm)		0.605	0.575	0.586	0.328
T1c (11–20 mm)		0.777	0.720	0.789	0.500
T2+ (>20 mm)		0.975	0.975	0.953	0.938
	CNBSS-2 observation[1]	Model expectation[2]	Model expectation[2]	Model expectation[2]	Model expectation[2]
Screen 1, cases of cancer (*n*)					
DCIS	26	27.7	28.0	27.9	28.2
T1a (≤ 5 mm)	8	2.4	3.2	3.0	3.4
T1b (6–10 mm)	16	10.9	14.7	13.8	14.7
T1c (11–20 mm)	45	24.2	31.7	31.9	35.0
T2+ (>20 mm)	26	12.9	18.2	30.1	24.8
Total	140	88.1	109.4	122.4	121.7
Screens 2–5, cases of cancer (*n*)					
DCIS	43	26.5	38.9	35.8	34.1
T1a (≤5 mm)	10	9.3	12.9	11.9	13.1
T1b (6–10 mm)	39	34.3	47.4	44.1	42.9
T1c (11–20 mm)	56	36.9	52.3	50.1	60.5
T2+ (> 20 mm)	27	9.3	15.0	17.7	20.4
Total	186	123.8	177.1	169.9	182.5
Interval years 1–5, cases of cancer (*n*)					
DCIS	2	1.4	3.3	2.8	2.4
T1a (≤5 mm)	1	1.8	2.6	2.4	2.3
T1b (6–10 mm)	5	5.8	8.5	7.7	7.4
T1c (11–20 mm)	23	15.9	24.8	20.8	21.4
T2+ (> 20 mm)	15	17.4	26.2	17.4	16.1
Total	54	49.7	76.8	60.0	58.1
Deviance[3]		78.8	33.3	21.7	15.3

[1]Invasive cancers without known tumor size (screen 1, 19; screens 2–5, 11; interval cancers years 1–5, 8) are included in the observed total numbers of cancers but not in the observed stage-specific numbers.– [2]Expected stage-specific numbers are adjusted by the proportion of invasive cancers without known tumor size, while the expected total numbers are not.– [3]Deviance is calculated as 2 × (log likelihood of saturated model − log likelihood of estimated model). The likelihood of the estimated model was calculated assuming the observed number in each stage to be Poisson-distributed, the mean being the number predicted by the model.

preclinical invasive stages and estimated sensitivities which were consequently lower.

The average lead time for the MP group was estimated to be 3.6 years (model A), 3.5 years (model B), 4.0 years (model C) and 5.4 years (model D). For the PO group, estimates were 2.3, 2.2, 3.0 and 4.2 years, respectively. According to models C and D, the average additional lead time gained by mammography in the CNBSS-2 was 1.0 and 1.2 years, respectively.

Breast cancer mortality and estimated mortality reduction by screening

In Figure 1*a,b,* the observed cumulative numbers of breast cancer deaths in the MP and PO groups are compared to our model predictions, including the estimate for a no-screening policy for the enrolled women, using model C. The expected cumulative numbers of breast cancer deaths in the MP group correspond quite closely to the observed cumulative numbers during the entire follow-up period (Fig. 1*a*), with 91 observed and 87.9 expected (95% CI 69.5–106.3) breast cancer deaths at 11 years (model A 69.8, model B 101.1, model D 76.1). In a situation without screening, we predict 133.4 (95% CI 110.7–156.0) breast cancer deaths after 11 years of follow-up in each arm, using model C (Fig. 1*a,b*). According to this model, the expected breast cancer mortality reduction by annual mammography screening (for 4 or 5 years), supplemented with annual CBE, is 34.1% at 11 years compared to a situation without screening (model A 34.1%, model B 32.6%, model D 32.9%).

Throughout the 11-year follow-up period, the observed cumulative numbers of deaths from breast cancer in the PO group were lower than the cumulative numbers expected by model C (Fig. 1*b*). At 11 years, the total number of observed breast cancer deaths in the PO group was 83 while the model predicted 106.0 (95% CI 85.8–126.1) breast cancer deaths (model A 82.5, model B 118.7, model D 91.0). Using model C, the expected breast cancer mortality reduction by annual CBE screening is 20.5% at 11 years compared to no screening (model A 22.1%, model B 20.9%, model D 19.7%).

The model outcomes for the MP and PO groups lead to an expected RR (MP/PO) of dying of breast cancer of 0.829 (95% CI 0.625–1.099) after 11 years of follow-up (model A 0.845, model B 0.851, model D 0.835). The observed RR of 1.095 is not significantly different from the model expectation.

Sensitivity analysis

Raising the improvement in prognosis for breast cancers screen-detected at higher stages will probably lower the expected number of breast cancer deaths in the PO group. After changing the improvement in prognosis to 0.575 for all invasive stages, while keeping the total average improvement in prognosis at the level of the combined Swedish trials,[7] the expected (model C) cumulative number of breast cancer deaths in the MP group at 11 years of follow-up stayed consequently at almost the same level (87.3). The model predicted 102.3 breast cancer deaths in the PO group at 11 years of follow-up (95% CI

TABLE II – COMPARISON OF OBSERVED CNBSS-2 RESULTS OF THE PO GROUP WITH MODEL PREDICTIONS FOR DIFFERENT SENSITIVITY ESTIMATES OF CBE, BREAST CANCER INCIDENCE RATES IN THE TRIAL, DURATION OF THE SCREEN-DETECTABLE PRECLINICAL PHASE T2+ AND TOTAL DURATION OF ALL SCREEN-DETECTABLE PRECLINICAL INVASIVE STAGES

		Model A: average Canadian incidence	Model B: 40.0% ↑ incidence	Model C: 29.15% ↑ incidence, 94.82% ↑ duration T2+	Model D: 23,66% ↑ incidence, 115.03% ↑ duration invasive stages
Sensitivity CBE					
DCIS		0.060	0.045	0.053	0.062
T1a (≤5 mm)		0.030	0.010	0.008	0.004
T1b (6–10 mm)		0.170	0.145	0.158	0.083
T1c (11–20 mm)		0.550	0.500	0.478	0.287
T2+ (> 20 mm)		0.975	0.970	0.649	0.598
	CNBSS-2 observation[1]	Model expectation[2]	Model expectation[2]	Model Expectation[2]	Model expectation[2]
Screen 1 cases of cancer (*n*)					
DCIS	4	2.1	2.3	2.4	2.7
T1a (≤ 5 mm)	0	0.13	0.06	0.05	0.05
T1b (6–10 mm)	5	3.5	4.2	4.2	4.2
T1c (11–20 mm)	33	19.6	25.1	22.1	22.8
T2+ (> 20 mm)	20	14.7	20.7	23.4	18.0
Total	65	42.0	54.9	54.7	50.0
Screens 2–5, cases of cancer (*n*)					
DCIS	3	6.4	6.9	7.5	8.2
T1a (≤ 5 mm)	0	0.49	0.22	0.16	0.18
T1b (6–10 mm)	12	12.1	14.8	14.8	14.3
T1c (11–20 mm)	38	46.1	62.2	54.7	56.4
T2+ (>20 mm)	27	17.3	26.7	29.9	27.7
Total	86	88.3	119.0	114.9	114.4
Interval years 1–5, cases of cancer (*n*)					
DCIS	7	6.6	9.5	8.6	8.1
T1a (≤ 5 mm)	2	2.1	3.0	2.8	2.6
T1b (6–10 mm)	15	8.7	12.5	11.4	10.5
T1c (11–20 mm)	27	25.6	38.6	36.2	33.5
T2+ (> 20 mm)	33	23.6	34.9	34.7	32.2
Total	97	76.6	113.6	108.0	100.2
Deviance[3]		28.2	24.1	20.9	20.3

[1]Invasive cancers without known tumor size (screen 1, 3; screens 2–5, 6; interval cancers years 1–5, 13) are included in the observed total numbers of cancers but not in the observed stage-specific numbers.– [2]Expected stage-specific numbers are adjusted by the proportion of invasive cancers without known tumor size, while the expected total numbers are not.– [3]Deviance is calculated as 2 × (log likelihood of saturated model − log likelihood of estimated model). The likelihood of the estimated model was calculated assuming the observed number in each stage to be Poisson-distributed, the mean being the number predicted by the model.

82.5–122.1). Even with these favorable assumptions for CBE screening, model predictions remain clearly higher than the 83 observed breast cancer deaths in this group.

DISCUSSION

We realize that our model analysis and the CNBSS-2 differ in their approaches. The CNBSS-2 was designed to evaluate the efficacy of annual mammography screening over and above annual CBE, while the primary objective of the model evaluation was to make estimations of mammography benefit and CBE benefit, both compared to no screening, for the CNBSS-2.

Our model analysis shows that the sensitivity of CBE screening in the CNBSS-2 is quite high: 29–48% of T1c and 60–65% of T2+ tumors were detected by CBE. Is it known that the CBEs in the CNBSS were performed well. The CBE took 5–10 min and required both visual examination and palpation. It was primarily performed by specially trained nurse-examiners, who were closely monitored during the study.[14,15] High CBE sensitivities for larger tumors are plausible because of this intensive type of screening. However, the quality of CBE in a community setting is likely to be lower.[16,17]

The sensitivity of CBE supplemented with mammography is also high, which contradicts some of the criticisms of mammography quality in the CNBSS.[18] However, part of the observed high detection rate at screening is explained by an estimated 24–29% higher breast cancer incidence rate in the trial compared to the average Canadian rate. Because the CNBSS-2 was not population-based, women participated voluntarily. This may have resulted in selection of women with a higher breast cancer risk. Miller *et al.*[3] identified higher observed breast cancer incidence in the CNBSS-2 than in national data. They calculated the cumulative ratio of observed to expected (based on national data) rates of invasive breast cancer for years 2–5 after entry. In the MP group, this cumulative ratio was 1.28 and in the PO group, 1.18.

The observed breast cancer mortality differential between both study arms (RR = 1.095) is not significantly different from our model expectations (RR = 0.829, 95% CI 0.625–1.099). But our study suggests that annual mammography screening supplemented with annual CBE is expected to reduce breast cancer mortality by 34.1% compared to no screening, while expected breast cancer mortality reduction by annual CBE screening alone is 20.5%. From this we estimate a 13.6% (difference between 34.1% MP group and 20.5% PO group) to 34.1% (difference between MP group and no screening) breast cancer mortality reduction by mammography screening compared to no screening for the CNBSS-2 (50–59 years). While the model predictions of breast cancer mortality correspond quite closely to the observed cumulative numbers of breast cancer deaths in the MP group, the model predicts 28% more breast cancer deaths than observed in the PO group after 11 years of follow-up. Although this difference between observed breast cancer mortality and model predictions for the PO group

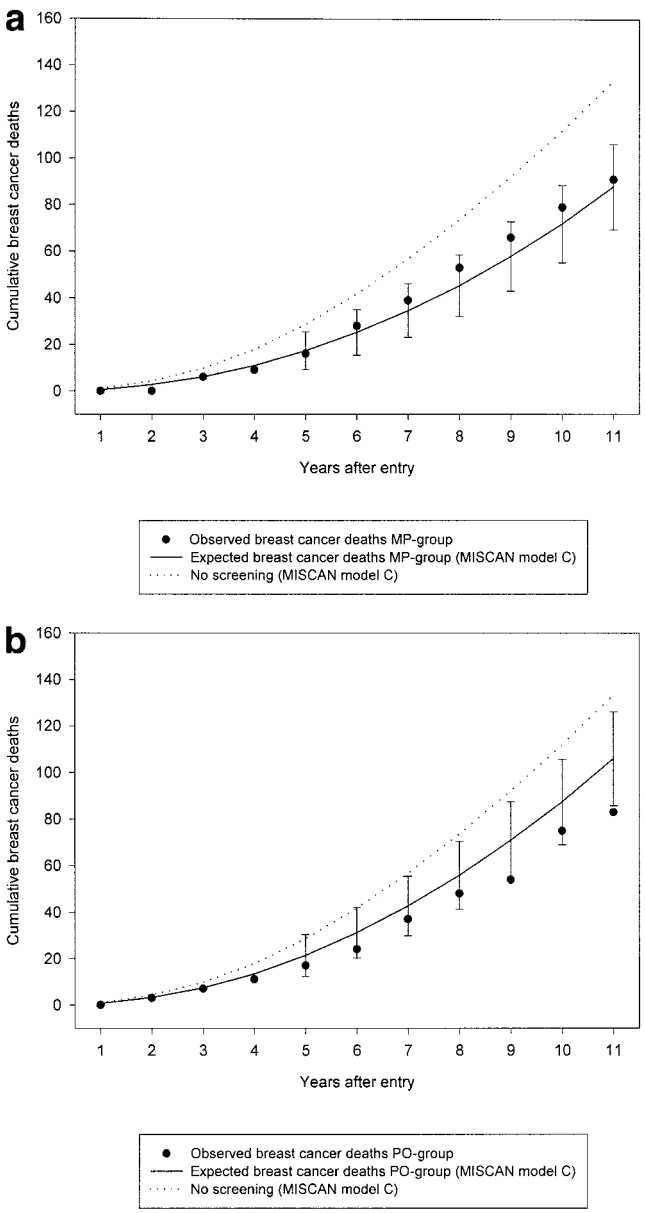


FIGURE 1 – *(a)* Cumulative number of observed (data set*) and expected (model) breast cancer deaths (with vertical bars representing 95% CIs) in the MP group and cumulative number of expected (model) breast cancer deaths in a situation without screening. *(b)* Cumulative number of observed (data set*) and expected (model) breast cancer deaths (with vertical bars representing 95% CIs) in the PO group and cumulative number of expected (model) breast cancer deaths in a situation without screening. (*We used an updated data set with breast cancer deaths known by record linkage with the Canadian National Cancer Registry and National Mortality Database to 31 December 1993 and by active follow-up of breast cancer patients to 30 June 1996.)

is unexpected, all observed cumulative numbers of breast cancer deaths in the PO group are within the 95% CI around the model expectation (except for years 9 and 11).

Doubling the duration of the screen-detectable preclinical phase T2+ (model C) or of all screen-detectable preclinical invasive stages (model D) led to a significantly improved model fit. We presented expected breast cancer mortality using model C instead of model D, though the total deviance of model C was higher. However, model D led to mean lead times for the PO and MP groups that were almost 2 years longer compared to our general estimates from other programs.[12]

We predicted breast cancer mortality using stage-specific improvement in prognosis after screen detection based on the results of analyses of the 5 Swedish randomized controlled trials.[7] Although there has been much discussion about the methodologic quality of several of these trials,[2,19] the general consensus is that their findings are valid.[1,20]

Improvement in prognosis after screen detection was based on breast cancer mortality reductions in women aged 50–69 years.[21] However, CNBSS-2 women were 50–59 years old at randomization, and reported effectiveness for this age group is lower.[1] On the other hand, CNBSS-2 women were screened annually, while the screening interval in the Swedish trials varied from 18 to 33 months.

The Swedish trials were conducted in an era when adjuvant systemic therapy was not used.[22,23] In contrast, all CNBSS-2 participants had access to such therapy, introduced in Canada in the beginning of the 1980s. This could have influenced survival and improvement in prognosis after screen detection, which we did not take into account.

The duration of the screen-detectable preclinical phase appears to be longer in trial participants than in the general population. One possible explanation for this longer duration is length bias. The trial recruitment procedures discouraged the entry of women with "clinically obvious" breast cancer, who were urged to visit their physician instead. Women enrolled were likely therefore to exclude those with rapidly progressive disease. Also, the volunteer trial may have attracted women who had already had minor breast symptoms for some time but had not yet visited a physician. This could have caused a delay in diagnosis among these women, prolonging the mean duration of the screen-detectable preclinical phase compared to the general population. It seems highly unlikely that the biology of breast cancer in Canadian women, or more specifically in the trial women, is different from our model assumptions, implying another natural history of the disease.

Some uncertainty remains in our model results because of lack of fit of the model. However, in the pooling of data with other mammography screening trials, an estimated breast cancer mortality reduction by mammography screening compared to no screening for women aged 50–59 in the CNBSS-2 of at least 13.6% is plausible. Therefore, the CNBSS-2 can be judged as compatible with the results of other trials of mammography screening.

REFERENCES

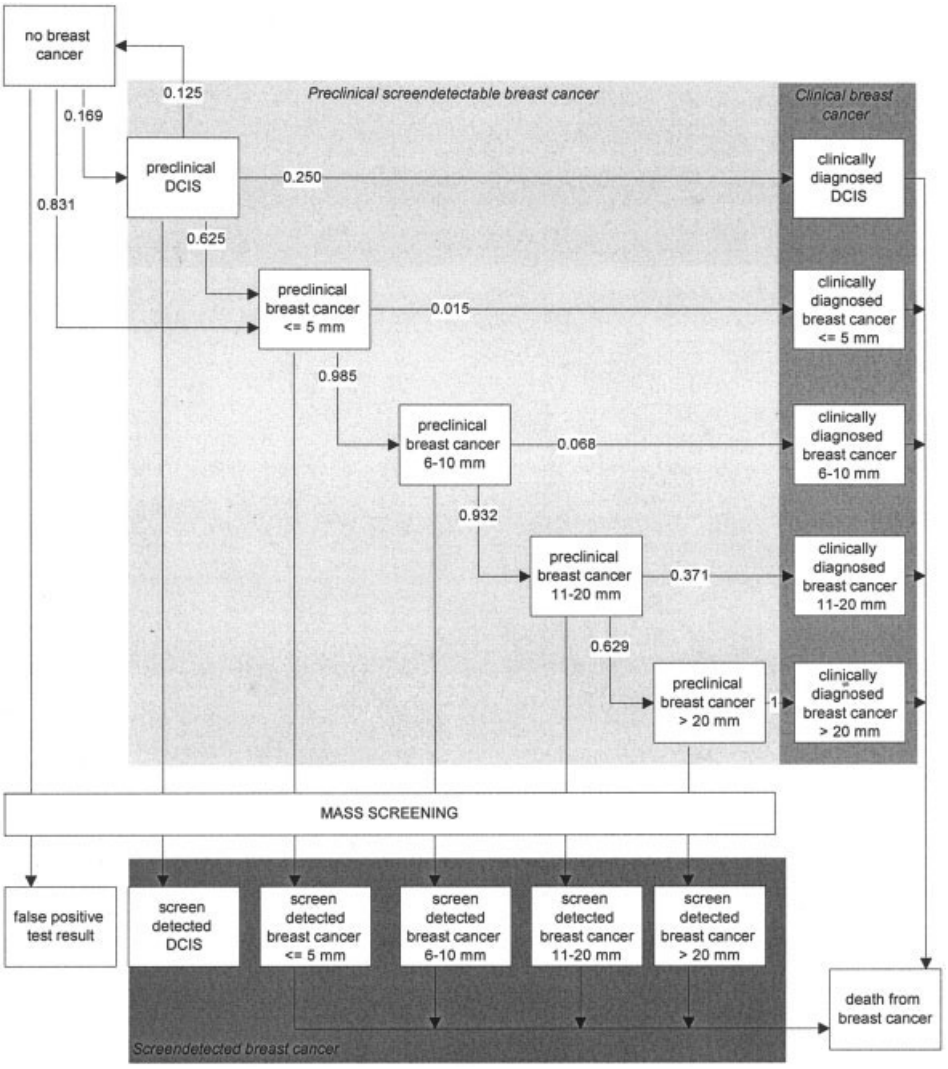
1. Nyström L, Andersson I, Bjurstam N, Frisell J, Nordenskjöld B, Rutqvist LE. Long-term effects of mammography screening: updated overview of the Swedish randomised trials. Lancet 2002; 359:909–19.
2. Olsen O, Gøtzsche PC. Screening for breast cancer with mammography. Cochrane Database Syst Rev 2001:CD001877.
3. Miller AB, Baines CJ, To T, Wall C. Canadian National Breast Screening Study: 2. Breast cancer detection and death rates among women aged 50 to 59 years. CMAJ 1992;147:1477–88.
4. Miller AB, To T, Baines CJ, Wall C. Canadian National Breast Screening Study-2: 13-year results of a randomized trial in women aged 50–59 years. J Natl Cancer Inst 2000;92: 1490–9.
5. Forrest AP, Alexander FE. A question that will not go away: at what age should mammographic screening begin? J Natl Cancer Inst 1995; 87:1195–9.
6. van Oortmarssen GJ, Habbema JDF, van der Maas PJ, de Koning HJ, Collette HJA, Verbeek ALM, Geerts AT, Lubbe JThN. A model for breast cancer screening. Cancer 1990;66:1601–12.

7. de Koning HJ, Boer R, Warmerdam PG, Beemsterboer PMM, van der Maas PJ. Quantitative interpretation of age-specific mortality reductions from the Swedish Breast Cancer-Screening Trials. J Natl Cancer Inst 1995;87:1217–23.
8. Miller AB, Howe GR, Wall C. The National Study of Breast Cancer Screening. Protocol for a Canadian randomized controlled trial of screening for breast cancer in women. Clin Invest Med 1981;4:229–58.
9. Muir C, Waterhouse J, Mack T, Powell J, Whelan S, eds. Cancer incidence in five continents. vol. V. Lyon: IARC, 1987.
10. Gaudette LA. Should population-based screening for breast cancer be extended to Canadian women aged 70–79? A model-based approach. Ottawa: School of Graduate Studies and Research, University of Ottawa, 2001. Dissertation.
11. WHO. World health statistics annual. Geneva: WHO, 1984.
12. de Koning HJ, Fracheboud J, Boer R, Verbeek ALM, Collette HJA, Hendriks JHCL, van Ineveld BM, de Bruyn AE, van der Maas PJ. Nation-wide breast cancer screening in the Netherlands: support for breast-cancer mortality reduction. National Evaluation Team for Breast Cancer Screening (NETB). Int J Cancer 1995;60:777–80.
13. Press WH, Flannery BP, Teukolsky SA, Vetterling WT. Numerical recipes. The art of scientific computing (Fortran version). Cambridge: Cambridge University Press, 1989.
14. Miller AB, Baines CJ, Turnbull C. The role of the nurse-examiner in the National Breast Screening Study. Can J Public Health 1991;82:162–7.
15. Baines CJ, Miller AB, Bassett AA. Physical examination—its role as a single screening modality in the Canadian National Breast Screening Study. Cancer 1989;63:1816–22.
16. Bobo JK, Lee NC, Thames SF. Findings from 752,081 clinical breast examinations reported to a national screening program from 1995 through 1998. J Natl Cancer Inst 2000;92:971–6.
17. Oestreicher N, White E, Lehman CD, Mandelson MT, Porter PL, Taplin SH. Predictors of sensitivity of clinical breast examination (CBE). Breast Cancer Res Treat 2002;76:73–81.
18. Kopans DB, Feig SA. The Canadian National Breast Screening Study: a critical review. Am J Roentgenol 1993;161:755–60.
19. Gøtzsche PC, Olsen O. Is screening for breast cancer with mammography justifiable? Lancet 2000;355:129–34.
20. Vainio H, Bianchini F, eds. IARC handbooks of cancer prevention. vol. 7. Breast cancer screening. Lyon: IARC, 2002.
21. Nyström L, Rutqvist LE, Wall S, Lindgren A, Lindqvist M, Rydén S, Andersson I, Bjurstam N, Fagerberg G, Frisell J, Tabár L, Larsson LG. Breast cancer screening with mammography: overview of Swedish randomised trials. Lancet 1993;341:973–8.
22. Holmberg LH, Adami HO, Tabár L, Bergstrom R. Survival in breast cancer diagnosed between mammographic screening examinations. Lancet 1986;2:27–30.
23. Tabár L, Chen HH, Duffy SW, Krusemo UB. Primary and adjuvant therapy, prognostic factors and survival in 1053 breast cancers diagnosed in a trial of mammography screening. Jpn J Clin Oncol 1999;29:608–16.

APPENDIX

The MISCAN model

In MISCAN, individual life histories are generated as a Markov process of stages and transitions. The natural history of



APPENDIX FIGURE – Structure of the MISCAN model for breast cancer, with transition probabilities for age 55 years.

breast cancer is modeled as a progression through 4 invasive, screen-detectable, preclinical stages according to tumor size: T1a, T1b, T1c and T2+ (≤5, 6–10, 11–20 and >20 mm, respectively). Some of the invasive cancers are assumed to be preceded by screen-detectable DCIS. Without screening, a preclinical cancer may either be diagnosed clinically or progress to the next preclinical stage. The model generates individual life histories for women, incorporating demographic and epidemiologic characteristics of the (trial) population under study, to calculate the clinical incidence and mortality from breast cancer in a situation without screening. For the situation with screening, screening characteristics (screening ages, interval and attendance rate) and performance of screening are added to the model and screening is applied to the individual life histories of women: preclinical cancers may be detected, depending on the sensitivity of the screening test. In the Appendix figure, the structure of the MISCAN model for breast cancer is presented.

Transitions between stages depend on transition probabilities (Appendix figure, age 55 years) and dwelling time distributions. Durations are generated from exponential distributions with stage- and age-dependent means. The same level of sensitivity is assumed at each screening.

Key parameters in the model of the performance of screening are mean durations of the preclinical screen-detectable stages, sensitivity of the screening test and improvement in prognosis after screen detection. The latter is defined as 1 minus the ratio of the risk of dying of screen-detected breast cancer divided by the risk if the same cancer had been diagnosed in the absence of screening. We assumed this benefit of early detection to be independent of the screening test used. The important base model parameters on natural history and screening, from which we started our analysis, are shown in the Appendix table.

The output of the model contains the number of screen-detected cancers and number of interval cancers (both including stage distribution), clinical age-specific breast cancer incidence by stage and age-specific breast cancer mortality (both with and without screening).

APPENDIX TABLE – IMPORTANT BASE MODEL PARAMETERS ON NATURAL HISTORY AND SCREENING IN THE MISCAN BREAST CANCER MODEL, FROM WHICH WE STARTED OUR MODEL ANALYSIS

Mean duration (years) of screen-detectable preclinical stage by age	Age (years)			
Stage	*40*	*50*	*60*	*70*
Preclinical DCIS	5.2	5.2	5.2	5.2
Preclinical T1a (tumor ≤5 mm)	0.1	0.1	0.1	0.2
Preclinical T1b (tumor 6–10 mm)	0.4	0.5	0.7	0.9
Preclinical T1c (tumor 11–20 mm)	0.8	1.0	1.5	1.8
Preclinical T2+ (tumor >20 mm)	0.6	0.8	1.1	1.4

Long-term relative survival by clinical stage and age					
Age (years)	*DCIS*	*T1a*	*T1b*	*T1c*	*T2+*
40	1.000	0.857	0.787	0.628	0.417
50	1.000	0.855	0.785	0.626	0.412
60	1.000	0.831	0.748	0.562	0.312
70	1.000	0.851	0.777	0.612	0.391

Probability of surviving by time since diagnosis and stage				
Time since diagnosis (years)	*T1a*	*T1b*	*T1c*	*T2+*
1	0.935	0.951	0.953	0.907
3	0.745	0.854	0.838	0.698
5	0.601	0.627	0.614	0.521
7	0.497	0.481	0.437	0.399
10	0.386	0.295	0.205	0.304
20	0.201	0.182	0.145	0.132
30	0.124	0.108	0.081	0.072
50	0.000	0.000	0.000	0.000

Sensitivity of mammography by stage and age	Age (years)		
Stage	*40–44*	*45–49*	≥50
Preclinical DCIS	24%	32%	40%
Preclinical T1a (tumor ≤5 mm)	39%	52%	65%
Preclinical T1b (tumor 6–10 mm)	48%	64%	80%
Preclinical T1c (tumor 11–20 mm)	54%	72%	90%
Preclinical T2+ (tumor >20 mm)	57%	76%	95%

Reduction in risk of dying of breast cancer by stage in which (pre)cancer is detected	
Stage	*Reduction in risk*
Preclinical DCIS	100%
Preclinical T1a (tumor ≤5 mm)	89.2%
Preclinical T1b (tumor 6–10 mm)	81.4%
Preclinical T1c (tumor 11–20 mm)	56.7%
Preclinical T2+ (tumor >20 mm)	39.5%