Using Simulation to Model and Validate Invasive Breast Cancer Progression in Women in the Study and Control Groups of the Canadian National Breast Screening Studies I and II

Sharareh Taghipour, PhD, Laurent N. Caudrelier, MSc, Anthony B. Miller, MD, FRCP, FFPH, FACE, Bart Harvey, MD, MSC, PHD, MED, FACPM, FRCPC

Background. *Modeling breast cancer progression and the effect of various risk is helpful in deciding when a woman should start and end screening, and how often the screening should be undertaken.* ***Methods.*** *We modeled the natural progression of breast cancer using a hidden Markov process, and incorporated the effects of covariates. Patients are women aged 50–59 (older) and 40–49 (younger) years from the Canadian National Breast Screening Studies. We included prevalent cancers, estimated the screening sensitivities and rates of over-diagnosis, and validated the models using simulation.* ***Results.*** *We found that older women have a higher rate of transition from a healthy to preclinical state and other causes of death but a lower rate of transition from preclinical to clinical state. Reciprocally, younger women have a lower rate of transition from a healthy to preclinical state and other causes of death but a higher rate of transition from a preclinical to clinical state. Different risk factors were significant for the age groups. The mean sojourn times for older and younger women were 2.53 and 2.96 years, respectively. In the study group, the sensitivities of the initial physical examination and mammography for older and younger women were 0.87 and 0.81, respectively, and the sensitivity of the subsequent screens were 0.78 and 0.53, respectively. In the control groups, the sensitivities of the initial physical examination for older and younger women were 0.769 and 0.671, respectively, and the sensitivity of the subsequent physical examinations for the control group aged 50–59 years was 0.37. The upper-bounds for over-diagnosis in older and younger women were 25% and 27%, respectively.* ***Conclusions.*** *The present work offers a basis for the better modelling of cancer incidence for a population with the inclusion of prevalent cancers.*
Key words: *invasive breast cancer; prevalent cancer; progression modelling; risk factors; over-diagnosis; simulation.* ***(Med Decis Making XXXX;XX:xx–xx)***

Received 23 June 2015 from Ryerson University, Department of Mechanical and Industrial Engineering, Toronto, ON, Canada (ST); University of Toronto, Toronto, ON, Canada (LNC); and University of Toronto, Dalla Lana School of Public Health, Toronto, ON, Canada (ABM, BH). The authors declare no conflict of interest. Revision accepted for publication 27 June 2016.

Address correspondence to Sharareh Taghipour, PhD, Ryerson University, Department of Mechanical and Industrial Engineering, Eric Palin Hall, Room 338A, 87 Gerrard Street East, Toronto, ON M5B 2K3, Canada; Tel: 1-416-979-5000 ext. 7693; e-mail: sharareh@ryerson.ca.

DOI: 10.1177/0272989X16660711

Breast cancer is the most common type of cancer (excluding non-melanoma skin cancer) among Canadian women. In 2014, it is estimated that, on average, 67 new breast cancer cases will occur daily and 14 women will die per day due to breast cancer.[1] Although, breast cancer cannot be eliminated completely, like other diseases, the knowledge about its progression and the effect of different risk factors on the progression, such as age or family history of breast cancer, can be helpful in deciding when a woman should start and end screening, and how often the screening should be undertaken.

In the literature, the progression of many chronic diseases, such as cancer, has been modeled using multi-state Markov processes.[2-4] The main reason

behind this approach is that Markov processes enable the description of diseases with staged progression.[5,6] Markov processes have also been used widely for modeling the progression of breast cancer and estimating the transition rates between progression stages.[2-8] In general, when an appropriate stochastic model is selected, the model is fit to the clinical data, and the progression parameters, such as transition rates, are estimated.

In this paper, we assumed the progression of breast cancer follows a hidden Markov process, and estimated the process's parameters for women aged 40–49 years and 50–59 years in both the study and control groups of the Canadian National Breast Screening Studies (CNBSS). The present paper is an extension of the paper[8] in which a Markov model was fit to the cancer data from only the study groups (the data about women who took annual mammogram and physical examination), and the transition parameters and sensitivities were simultaneously estimated. In the current paper, we first fit a continuous-time hidden Markov multi-state model by maximum likelihood to estimate the parameters of the natural progression based on the study groups of the CNBSS. We then calibrated the parameters (obtained only based on the study groups) to the control groups of the CNBSS, and validated the progression model and its parameters using a simulation model for both the study and control groups. The simulation model generated the events, i.e., breast cancer incidence, progression, and the other-causes of death, as the competing risks to breast cancer incidence[9] for the women in the CNBSS according to a hidden Markov model, and if the progression model and its parameter estimates were valid, the simulation model would generate the same numbers of screen-detected and clinical cancers. The current work is not just a ''goodness of fit'' analysis, because it validates the developed progression model to the control groups' data, which was not initially included in the parameter estimations of the cancer natural progression model.

Moreover, we included the prevalent cancers in our model for both age groups, and estimated the sensitivities at the first screens (after inclusion of prevalent cancers), and the sensitivity of the subsequent physical exams in the control group of women aged 50–59 years. In addition, we estimated the upper bounds for over-diagnosis for both age groups. None of these estimations were performed in the previous investigation.[8]

In short, the new estimates presented in the current paper include the sensitivities of the initial screens in both the study and control groups of the CNBSS, the sensitivity of the annual physical examinations for women aged 50–59 years in the control group of the CNBSS, the upper bounds for over-diagnosis in both age-groups, as well as the calibrated parameter estimates; these estimates provide a validated progression model for both the study and control groups of the CNBSS.

The main motivation of the current study was to validate the breast cancer progression model for both the study and control groups of the CNBSS as well as to model the prevalent cancers. The inclusion of control groups in the analysis can assure us that the parameters are also valid when no screening program is in place. Moreover, the prevalent cancers are a reality that cannot be ignored in the process of decision-making for cancer screening and prevention. The validated model can be used as one of the main components of a prediction model for the cost-effectiveness analysis of various screening policies as well as the no-screening scenario for the Canadian population, and to help policy makers establish the best policies at the population or individual levels. Since we have modeled the effect of several risk factors on the transitions of the progression model, the model for the cost-effectiveness analysis can also be used to evaluate screening policies for sub-populations of Canadian women with certain risk factors (or those who are at higher risk of breast cancer).

METHODS

Study Population and Period

We used data from the Canadian National Breast Screening Studies (CNBSS) I and II.[10-12] The CNBSS consists of two randomized controlled screening trials designed to evaluate screening efficacy in women aged 40–49 and 50–59 years. In the CNBSS, all women who were enrolled in the trial from 1980 to 1985 underwent an initial physical examination and were then randomly assigned to either the study groups or the control groups. The women in the study groups (40–49 and 50–59 years) were offered four or five annual mammograms and physical examinations of the breasts. The women in the control group aged 40–49 years did not take any further screening or physical examination, whereas women in the control group aged 50–59 years received annual physical examinations up to five times. The analytical cohort considered in the current paper is presented in Figure 1.

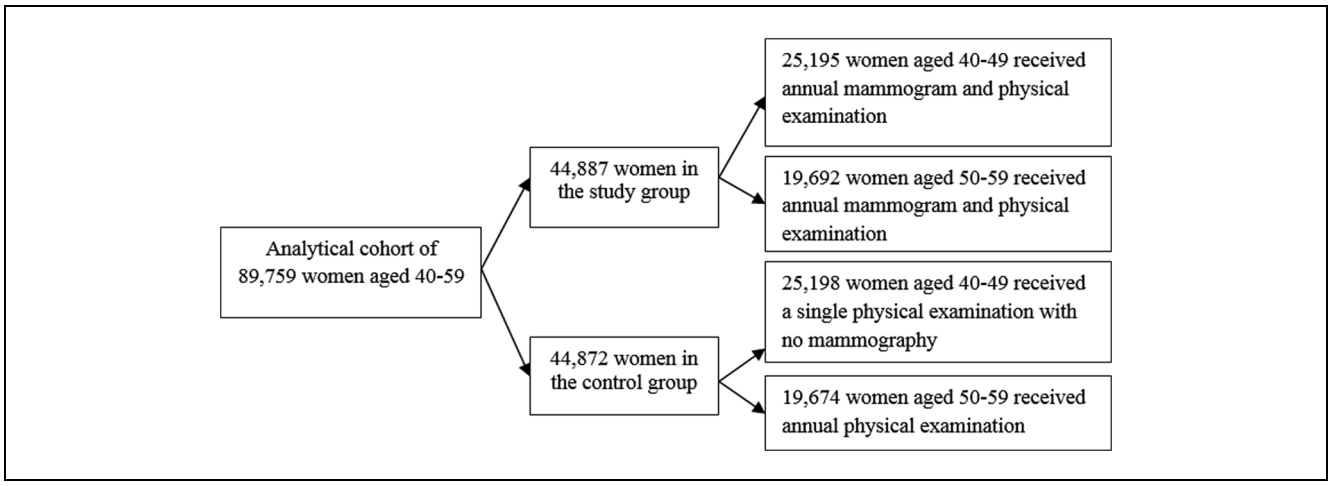

Figure 1 Analytical cohort diagram of women aged 40-49 and 50-59.

All eligible women in the CNBSS completed an enrolment and epidemiology questionnaire to collect information on some of the potential risk factors for breast cancer, such as family history of breast cancer and personal history of breast disease. A complete list of risk factors studied is provided elsewhere.[13]

Prevalent, Screen-detected and Clinical Cancers

In a randomized screening trial, such as the CNBSS,[10-12] women who have been already diagnosed with breast cancer are not eligible to participate. Some women with prevalent cancer may be enrolled into a clinical study if their cancer is revealed at the initial screen (if any) or shortly thereafter. Prevalent cancers are a reality that cannot be ignored in cancer screening decision-making. For example, in the CNBSS, the cancers that were diagnosed within the first year of the screening program accounted for 30% and 37% of all reported cancers in the study groups of the women aged 40–49 and 50–59 years, respectively. Similarly, at the population level, a reasonable proportion of women may be diagnosed with prevalent cancer at their initial screens.

Table 1 shows the number of participants aged 40–49 and 50–59 years for both the study and control groups of the CNBSS, the number of prevalent cancers, and the number of cancers diagnosed at the subsequent screens and the clinical cancers. Women aged 40–49 years were followed-up for breast cancer incidence until the end of 1989. This included cancers that were diagnosed within the screening program (i.e., at or within the four or five annual screens), and cancers that were diagnosed after the screening program through linkage to the breast cancer registry. For instance, a woman who enrolled in 1980 was followed-up for 9 years. However, follow-ups of women aged 50–59 years ceased 5 years after their initial enrolment. As a consequence of these different follow-up times, the total number of cancers in the age groups may not be directly comparable. We chose a shorter follow-up period for women aged 50–59 years because they are likely to be subject to more incident cancers (i.e., cancers diagnosed more than 12 months after the last screen in the screening program), and this may confound our model development and validation.

In Table 1, among the women aged 40–49 years, 80 of 25,195 women are diagnosed with breast cancer at the first screen, with an additional 16 women diagnosed within the first year of their enrolment (i.e., between the first and second screening visits). For women in the study group aged 50–59 years, 104 breast cancers were diagnosed at the initial screen, and 15 clinical cancers were reported within the first year of follow-up. We presume that these 16 women aged 40–49 years (or 15 in the 50–59-year age group) had cancer at the initial screen but were misclassified (false-negative screens).

Sensitivity and Misdiagnosis

When a woman is screened for breast cancer, there is always a chance of misdiagnosis; i.e., a negative screening result (false-negative). Reciprocally,

Table 1 Cancer Cases in the Study and Control Groups of Women Aged 40–49 and 50–59 Years

Study/Control groups	Age Group (Years)	Number of Participants	Number of Prevalent Cancers		Number of Subsequent Cancers		Total
			Detected at the First Screen	Clinical Within the First Year	Detected at Subsequent Screens	Clinical After the First Year	
Study Group	40–49	25,195	80	16	106	120	322
	50–59	19,692	104	15	138	64	321
Control Group	40–49	25,198	53	26	-	193	272
	50–59	19,674	60	18	83	99	260

a positive result may be obtained for a woman with no breast cancer (false-positive). Although frequent, false-positive results have very limited implication, since further diagnostic examinations are performed to confirm the presence /absence of the disease when a result is positive. The risk of misdiagnosis decreases when the tumor is larger, as it is more likely to be detected.[14] Furthermore, as screening technology improves, the risk of false-negative cases may also reduce.

The probability of detecting breast cancer is called the sensitivity of a screening test. In the models widely developed for describing breast cancer progression and screening, although unrealistic, it is assumed that the sensitivity of all screens is fixed;[2,15] i.e., when a woman undertakes several screens over the follow-up period (e.g., four or five screens in the CNBSS), she has the same chance of having cancer detected at each screen if she has the disease. In this analysis, we will base our progression model on this assumption, which allows us to simultaneously estimate the progression parameters as well as the sensitivity of the detection methods.

However, although we consider there to be a constant value for sensitivity of the consecutive screens, the sensitivity of the initial screen is estimated independently. This is because a large number of prevalent cancers are diagnosed at initial screens in both age groups, with much fewer cases detected at subsequent screens (the second, third, etc. screens).

Modeling Cancer Incidence and Progression, Estimating the Sensitivities, and Simulating the Study and Control Groups of the CNBSS

Briefly, the following steps were undertaken for modeling, parameters estimation, and simulation of the CNBSS; additional detail for each step is provided in the subsequent subsections.

1. We first fit a partially observable Markov model to the data from the study groups of the CNBSS, excluding the prevalent cancers, and simultaneously obtained the progression parameters as well the sensitivity of the subsequent screens (those following the initial screen).
2. We then considered the prevalent cancers and estimated numerically the sensitivity of the initial screens for the study groups.
3. Then, we simulated the study groups of the CNBSS, including the prevalent cancers, to validate the parameters' estimates and the sensitivities of the initial and subsequent screens.
4. Finally, we simulated the control groups of the CNBSS and estimated the sensitivities of the initial and subsequent screens, if available.

Modeling the Incidence and Natural Progression of Breast Cancer

Taghipour and others[8] considered a four-state, continuous time, partially observable Markov model to describe the progression of breast cancer from the state of ''healthy/non-detectable cancer'' to ''cancer diagnosis'' (see Figure 2). The states of their model are described as follows:

1. Healthy/non-detectable cancer
2. Preclinical phase (screening detectable cancer)
3. Clinical phase (symptoms are evident)
4. Death due to causes other than breast cancer

In Figure 2, possible transition (hazard) rates between two states are shown. It is assumed that the transition rates of moving from States 1 and 2 to

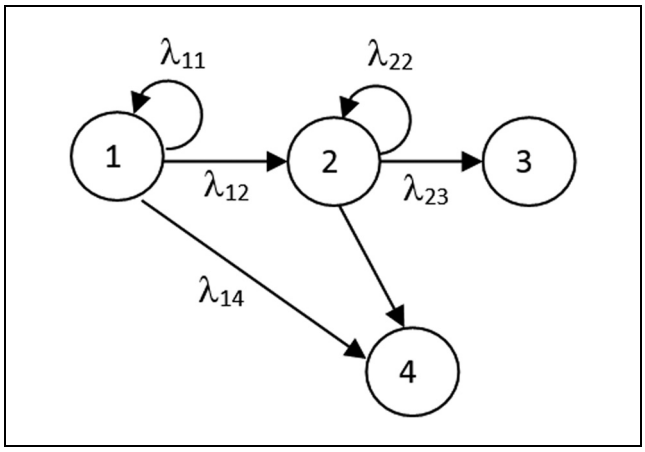

Figure 2 A four-state breast cancer progression model. The loops on States 1 and 2 indicate the possibility that the Markov process continues to occupy the same state in the next period of time. States 3 and 4 are absorbing.

State 4 are the same. In the figure, the loops on States 1 and 2 indicate the possibility that the Markov process continues to occupy the same state in the next period of time. States 3 and 4 are absorbing.

To obtain the progression parameters, Taghipour and others[8] excluded the prevalent cases of breast cancer (the breast cancers detected at the initial screen or during the first six months of the study). From this, we could assume that women were cancer-free or had non-detectable cancer at enrolment. Taghipour and others[8] used only the data from the study groups to develop the progression model, because women in the study groups had several planned screens. They also incorporated the effect of some risk factors, including ''age at enrolment'', ''prior personal history of breast disease'', ''family history of breast cancer'', ''duration of menstruation (years)'', and the ''number of live births'' on the transition rates between the states.[8]

Both States 1 and 2 are not fully observable, i.e., there is a probability of misdiagnosis from breast physical examinations and/or mammograms (i.e., $P(1|2) > 0$). However, the probability of having a false-positive test is neglected, since a woman receives further examinations before a cancer is confirmed (i.e., $P(2|1)=0$). The sensitivity is given by $P(2|2)$.

For the current paper, we first repeated the analysis of Taghipour and others[8] with the data excluding prevalent cancers. In both the current study and in their study, the effect of several risk factors, including ''Family history of breast cancer'', on the transition rates of the progression model were investigated. In the CNBSS, women with a family history of breast cancer could specify up to four female relatives (mother, sister, cousin, etc.) with breast cancer. We classified the relatives into five categories: 'first degree', 'second degree', 'third degree', 'fourth degree' and 'fifth degree and above', and assigned to each woman a single score based on the degrees of relatives with breast Cancer (the method of assigning the scores can be found elsewhere[13]). In the study by Taghipour and others[8], ''sister'' was considered as a ''second degree'' relative, which is incorrect. We fixed this problem in the current study and considered ''sister'' as a ''first degree'' relative. A first degree relative results in a higher score assigned to a woman. Thus, the effect of the risk factor ''Family history of breast cancer'' which is reported in the current study slightly differs from that reported by Taghipour and others.[8]

Based on the study group of the CNBSS, we fit a continuous-time hidden Markov multi-state model by maximum likelihood to estimate the parameters of the natural progression. The likelihood was calculated from the transition probability matrix. The partially observable Markov model was fitted using the package *msm* in R.[16] To fit the partially observable Markov model, the CNBSS data was organized such that every woman had a set of records consisting of the observed state (the states described in Figure 2) and the observation date. The observation dates were the times of transition between states for each woman in the study group, which were known for States 3 and 4, and unknown (censored) for the transition from State 2 to State 3. Thus, the transition rate from State 2 to State 3 was estimated based on the transitions from State 1 to State 2 and from State 1 to State 3. The covariates were also fit to the hidden Markov observation process. A woman can only be found in State 2 after a positive screen, and in State 3 when her cancer becomes revealed clinically within the screening intervals or after the end of their planned screens (four or five screens), but before the end of the follow-up period.

Modeling Prevalent Cancers and Estimating the Sensitivity of the First Screens

After estimating the transition parameters in the natural progression model, we included the possibility of women entering the study with prevalent breast cancer. We first estimated the sensitivity of the initial screens for all groups (i.e., the study and

control groups of women aged 40–49 and 50–59 years) by dividing the total number of breast cancers diagnosed at the initial screen by the total number of cancers (screen-detected and clinical) diagnosed within the first year. For example, the sensitivity of the initial screen for women in the study group aged 40–49 years was estimated to be 0.83, because 80 women were detected at the first screen and the cancer of 16 women became clinical within the first year (80/96~0.83). It should be noted that the sensitivity of the first screen estimated for the study groups incorporates a combination of physical examination and mammography. However, the sensitivity of the first screens of the control groups is only for physical examination, since women in these groups did not have mammography.

For women whose prevalent cancers were not diagnosed at the initial screen (i.e., a false-negative result was obtained), their cancer may or may not become clinical over the follow-up period. The expected number of women with prevalent cancers at the initial screen can also be approximated as the total number of cancers diagnosed within the first year (screen-detected or clinically).

Simulating the Study Groups of the CNBSS

We developed a simulation model based on the natural progression model (partially observable Markov model) combined with the results of the prevalent cancers analysis. Using the simulation model, we first generated the life histories for the same women in the study groups of the CNBSS to compare the outcomes of the simulation model with the actual data, and to validate the progression model parameters and sensitivities. This simulation model was developed in SAS 9.3 TS (2002-2010) (SAS Institute Inc., Cary, NC).

We first validated our model using the basic, initial data from the women in the study groups, which included their age of entry, the number of years of follow-up recorded, and the value of their risk factors. Using the simulation model, we then generated for each woman the probability of having a prevalent cancer, the time to detectable breast cancer (i.e., preclinical cancer), the time from preclinical cancer to clinical cancer (in the absence of screening), and the time from healthy state/preclinical cancer and clinical cancer to other causes of death, given her risk factors. All these events are natural events in the life history of a woman, which may be affected by screening. In the simulation model, we then incorporated screening, which may bring forward the diagnosis of a preclinical cancer with certain probability (sensitivity). For each woman in the study group, we applied exactly the number of annual screens that she took in the real data. If a woman with preclinical cancer was not diagnosed at a screen, her cancer may be diagnosed at subsequent screens, or may become clinical, or may never become evident over the follow-up period. A woman may also die due to causes other than breast cancer.

We compared the results of the simulation to that of the CNBSS to validate the simulation model. It is important to note that the length of the follow-up period varied among the women included in the study. In the simulation model, the follow-up period for a woman is the difference between her date of enrolment and the administrative censoring date (the cut-off date of the data). Moreover, the simulation model may produce different outcomes for the same woman as compared with what was actually observed in the real dataset. For example, a woman who died from other causes in the CNBSS may be diagnosed with breast cancer within the follow-up period. However, what is important, is that the simulation model should produce the total number of events of interest approximating the real data. For example, the number of screen-detected cancers or clinical cancers produced in the simulation model should be a reasonable approximate of the real data. We developed two separate simulation models for women aged 40–49 and 50–59 years because the progression parameters and screening sensitivities are different for these two age groups.

Simulating the Control Groups of the CNBSS

None of the control groups of the CNBSS were used to estimate the parameters of the natural progression. Therefore, we used the natural progression parameters that we obtained from the study groups to simulate the control groups of the CNBSS. To calibrate the estimated parameters for the control group (based on the study group), we considered the ranges for the transition rates parameters, sensitivities, and the coefficient of the covariates. The ranges are the 95% CI for each estimated parameter obtained from the hidden Markov model. All combinations of the parameters' estimates within their 95% CI have been investigated in the simulation model to conclude which combination culminates in the minimum sum of the squares of the residuals. The residuals are the differences between observed

Table 2 Simulation Results for the Study and Control Groups with Prevalent Cancers

Study/Control Group	Age Group	Number of Participants	Detected at the First Screen	Detected at Subsequent Screens	Clinical Cancers
Study Group	40-49 (Real data)	25,195	80	106	136
	40-49 (Simulated data)	25,195	79.96 (8.8)	105.39 (10.06)	134.61 (11.77)
	50-59 (Real data)	19,692	104	138	79
	50-59 (Simulated data)	19,692	103.81 (10.49)	138.22 (11.60)	78.95 (9.04)
Control Group	40-49 (Real data)	25,198	53	-	219
	40-49 (Simulated data)	25,198	52.78 (7.34)	-	219.45 (14.27)
	50-59 (Real data)	19,674	60	83	117
	50-59 (Simulated data)	19,674	60.41 (7.48)	82.56 (9.01)	116.91 (10.36)

(from the CNBSS) and predicted (from the simulation model) numbers of screen-detected and clinical cancers for each age group. We utilized ''lsqnonlin'' function in MATLAB, which solves a nonlinear, least-squares problem to determine the optimal set of parameters (shown in Table 3).

The control group of women aged 40–49 years had only one initial physical examination at enrolment; thus, for this group, we used the sensitivity of the initial physical examination that we had estimated earlier. This initial examination allowed for large prevalent cancers to be reported; however, all other cancers reported after this time point would be considered clinically diagnosed. Women in the control group aged 50–59 years underwent annual physical examinations for up to five times; thus, for this group, we used the sensitivity of the initial and subsequent physical examinations. To estimate the sensitivity of the subsequent physical examinations after the initial examination, we examined different values for the sensitivity by trial and error, and this resulted in the same number of screen-detected and clinical cancers.

Validation of the Progression Model and the Estimates of the Sensitivities

We validated the breast cancer progression model and sensitivity estimates for both the study and control groups at both age ranges. To have a single valid model for both the study and control groups, we had to slightly adjust the progression parameters and the sensitivities that were obtained based on only the study groups. We ran the simulation models 1,000 times. The results of the simulation models and the validations of the models are presented in Table 2. The standard deviations of the predicted number of screen-detected and clinical cancers from the simulation model are calculated based on the results from 1,000 simulation runs.

It should be noted that we also investigated the possibility that the initial screen may have the same sensitivity as the subsequent screens. However, no valid result was obtained and there was a significant departure between the observed and predicted results from the simulation model.

RESULTS

Parameter Estimates in the Natural Progression Model and the Sensitivities of Screens

The parameters of the transition rates, the sensitivity of screenings, and the expected number of prevalent cancers at the initial screens for the study and control groups of both age groups are presented in Table 3. From the number of prevalent cancers at the first screen and the sensitivity of the first screens in the study and control groups, we obtained the expected number of prevalent cancers at the initial screen. For example, 80 women in the

Table 3 Transition Rates in the Partially Observable Markov Model, the Sensitivities and the Expected Number of Prevalent Cancers at the Initial Screens for Women Aged 40–49 and 50–59 Years

Transition Rates	40–49	50–59
Healthy to preclinical, λ_{12}	0.00118	0.00229
Preclinical to clinical, λ_{23}	0.39540	0.33670
Healthy/Preclinical to other causes of death, λ_{14}	0.00062	0.00152
Log-linear effect of covariate on transition from healthy to preclinical		
Age at entry	0.05116	-
Absence of prior personal history of breast disease	–0.43740	-
Family history of breast cancer	0.01580	0.01032
Menstruation length (years)	-	0.02960
Log-linear effect of covariate on transition from healthy or preclinical to other-cause mortality		
Age at entry	0.09212	0.06598
Sensitivity of the initial physical examination and mammography (study group)	0.81	0.87
Expected number of prevalent cancers at the initial physical examination and mammography (study group)	98	120
Sensitivity of the subsequent physical examination and mammography (study group)	0.53	0.78
Sensitivity of the initial physical examination (control group)	0.671	0.769
Expected number of prevalent cancers at the initial physical examination (control group)	78	78
Sensitivity of the subsequent physical examinations (control group)	-	0.37

study group aged 40–49 years were screen-detected at the initial screen, and the sensitivity of the initial screen was 0.81, which implies that 98 women were expected to have prevalent cancers at enrolment.

Estimating an Upper Limit for Over-diagnosis

We can estimate the upper limit for over-diagnosis[17] for both age groups. For the women aged 40–49 years, the difference between the total number of screen-detected cancers in the study and control groups is 80+106-53=133 cancers, and the difference for the clinical cancers is 219-136=83 cancers. Thus, the cancer of 133-83=50 women did not become clinical over the follow-up period, which implies 50/(80+106)=50/186~0.27 as an upper boundary for the over-diagnosis in women aged 40-49. For the women aged 50–59 years, the difference of screen-detected cancers in the study and control groups is (104+138)-(60+83) =99, and the difference of clinical cancers in the study and control groups is 117-79=38. Thus, the cancers of 99-38=61 women did not become clinical over the follow-up period, and 61/(104+138)=61/242~0.25 can be obtained as the upper boundary for the over-diagnosis in women aged 50–59 years.

DISCUSSION

The results presented in this paper describe a new and comprehensive approach to examine in-depth the screening data reported in the Canadian National Breast Screening Study (CNBSS) I and II. The modeling of the progression of breast cancer using both age groups of the CNBSS as well as the inclusion of preclinical prevalent cancers has not been previously performed. The inclusion of the preclinical prevalent cancers (reported in Table 4) in this model brings forward an understanding of the natural development of breast cancer that had been omitted until now.[7,8] These cases represent a visible percentage of all reported cancer cases within the considered follow-up period.[18] Table 4 demonstrates the strong influence of these preclinical prevalent cancers on the overall outcome of the reported cancer cases in the CNBSS.

The results in Table 4 assume that all cancer cases are considered of equal weight. However, it could be argued that more uncertainty exists in the medical outcome of early detected cases since these cases have a higher proportion of node-negative cancers.[19] If a differentiation were to be applied, the contribution of prevalent cancer cases would go up, as they can be assumed to represent more advanced cases of cancer. We did not deem it necessary to go through this extra step, as the data in Table 4 already illustrates their strong contribution.

In previously reported research on the modeling of the CNBSS data, prevalent preclinical cases were not included.[7,8,19] The authors of these studies cite that these cases were excluded because they could

Table 4 Contribution of the Preclinical Prevalent Cancers in the Study and Control Groups of the CNBSS

Age Group (Years)	Reported Prevalent Cases at the Initial Screen and the Clinical Cancers within the First Year (False-negative Cancers at Initial Screen) for Both the Study and Control Groups	Total Reported Preclinical Cases	Percentage of Prevalent Cases in Total Preclinical Cases	Total Number of Reported Cancer Cases	Percentage of Prevalent Cases in Total Cancer Cases
40–49	175 (133 + 42 clinical within 12 months)	281 = 175 + 106 (study group only)	62.3%	594 = 281 + 313 (clinical after 12 months)	29.5%
50–59	197 (164 + 33 clinical within 12 months)	418 = 197 + 221 (138 study group + 83 control group)	47.1%	581 = 418 + 163 (clinical after 12 months)	33.9%

not be used to accurately model the natural progression of the disease. In the current research, we did not use the prevalent preclinical cases to model the natural progression of the disease (as these cases enter the screening study at a detectable state of breast cancer and therefore cannot be used to model the progression from a healthy state to a disease state); instead, they were used to understand the overall effectiveness of screening as well as the influence these cases would have on screening sensitivity.

The assumption that all prevalent cancers are diagnosed at the initial screen or clinically within the first year probably does not completely reflect real life situations. However, because we do not have data on the cancer initiation for prevalent women, we validate and obtain outputs (expected numbers of screen and clinical cancers) based on this assumption, which closely reflects their corresponding numbers in real data. It should be noted that any assumption about diagnoses of unidentified cancer at the initial screen will have an impact on the estimates of the initial and subsequent sensitivities. In other words, a change in our assumption of the initial diagnosis will change the initial and subsequent sensitivities of the data and thus have an impact on the estimated number of screen-detected and clinical cancers.

Our approach allowed us to calculate the screening program sensitivities of different modalities (mammograms and physical examination of the breasts) at different stages of the screening process. We assumed screening sensitivity to be constant over time or age during the time horizon of the CNBSS's screens. Therefore, the values reported represent the average sensitivity over each age cohort. We used different values of sensitivity for the first screen participants in the study.

From the transition rates reported in Table 3, we can obtain the mean sojourn time of the preclinical state using $1/(\lambda_{23}+\lambda_{14})$. For women aged 40–49 years, the mean sojourn time is about 2.53 years, whereas, for the older women (50–59 years), it is about 2.96 years. We compared the reported values of the screening sensitivity as well as the mean sojourn time estimated from the two age groups of the CNBSS with previous reported values from other breast screening studies. Mammography is the only screening modality used in the reported studies with the exception of the HIP study,[19] where both mammography and physical examination of the breasts were utilized. Table 5 outlines these findings.

It is worth noting that we cannot directly compare the reported sensitivity results across studies because of the different screening programs and participants. The CNBSS was not designed to evaluate the effectiveness of the use of mammography as compared with not using mammography, which has been attempted in other studies.[10,11] Rather, the CNBSS-I (i.e., women aged 40–49 years) presents a comparative study to measure the benefits of a group receiving an initial physical examination of the breasts compared with a group receiving an annual combined physical examination and mammogram.[11] In contrast, CNBSS-II evaluated the added-value of a supplementary mammogram to the physical examination at the annual screening, as compared with only receiving annual physical examinations.[10] Some distinction could also be concluded between

Table 5 Estimated Sojourn Time and Sensitivities for Other Clinical Trials

References	Study Population	Mean Sojourn Time	Sensitivity
Duffy et al., 1996[20], 1997[2]; Chen et al., 1997a[21], 1997b[22]	Swedish Two-County, aged 40–49 years	2.46 (95% CI: 2.11, 2.86)	0.83 (95% CI: 0.74, 0.89)
Chen et al., 1997a[21], 1997b[22]; Duffy et al., 1997	Swedish Two-County, aged 50–59 years	3.70 (95% CI: 3.44, 4.17)	1.00
Chen et al., 1998[23]	Swedish Uppsala County, aged 39–49 years	1.52	0.58
Wu et al., 2010[24]	Finnish Mammography program, aged 50–59 years	2.02 (95% CI: 1.64, 2.62)	0.85
Chen et al., 2000[25]	Screening program for high risk women in Taiwan, aged 35–80 years	1.99	0.95
Walter and Day, 1983[26]; Day and Walter, 1984[27]	HIP, aged 40–64 years	1.71	0.82
Bjurstam et al., 1997[28]	The Gothenburg breast screening trial, Sweden, aged 39–49 years	2.21	0.87

*CI: Confidence interval

the effectiveness of a single-view mammogram and two-views mammograms; the latter being used during the CNBSS screening.[10,11,19]

Shen and Zelen[19] estimated the sensitivities of mammography, physical examination, and the combination of both for four screens performed for women aged 40–49 and 50–59 years in the CNBSS. However, they included both in situ and invasive cancers in their analysis, which also occludes any direct comparison of their estimates with ours. Shen and Zelen[19] estimated 0.91 and 0.82 as the average sensitivities for the combined modalities for 40–49 and 50–59 years, respectively, and their estimates are higher than ours (0.53 and 0.78, respectively). This difference may be because we only included invasive cancers in our analysis. Regardless, the higher sensitivity noted for the younger women in their study is surprising, as it would be expected to be lower because of the increased density of breasts in this age group. Moreover, the sensitivities estimated by Shen and Zelen[19] for mammography and physical examinations separately were 0.61 and 0.59, respectively, for younger women, and 0.66 and 0.39, respectively, for older women; these sensitivities are much lower than those estimated for the combined modalities, which is again surprising, since it would be expected that most cancers detectable by physical examination can still be detected by mammography. The estimated sojourn times in their study for the younger and older age groups (1.87 and 3.09 years, respectively), however, are relatively close to our estimates (2.53 and 2.96 years). It should be also noted that we calibrated and validated the progression model for both the study and control groups of the CNBSS, whereas Shen and Zelen's[19] parameters could only be validated for the study groups of the CNBSS. In addition, Shen and Zelen[19] did not include any risk factor (covariates) or prevalent cancers in their analysis.

In the current study, we also examined the influence of covariates that affect modelling the natural progression of breast cancer for the two CNBSS age groups. Indeed, we corrected "sister" as a first-degree relative and not second-degree, as reported by others.[8]

Post-treatment recurrent cancer cases were not considered here, as they do not represent the natural progression of the disease which we are examining in this paper. Likewise, although the current model presents a novel computation of the values of screening sensitivity, it does not include the progression to death from breast cancer (i.e., for participants screened positive, the outcome is not considered further than the clinical cancer stage); extending our model to include death from breast cancer is seen as one possible extension of this paper. Furthermore, the current reported values for the screening sensitivities are based only on the effectiveness of the examined screening modalities (i.e., mammograms and physical examination of the breast). Although the reported cases were verified by biopsy,[10,11] we did not consider the differences in the stage of the detected tumor in this paper. Future work will focus on expanding the follow-up period beyond the 5-year period reported here,

which should allow for a more thorough comparison of the two considered breast screening modalities to support the currently reported conclusions. The difficulty that arises when seeking to extend the follow-up period comes from the inconsistency of the reported cases after the end of the five-year period in the CNBSS. This complicates the possibility of drawing valuable conclusions as to the natural progression of the disease.

The CNBSS database includes information on the detected tumor size, the number of affected lymph nodes as well as biopsy report.[10,11] Therefore, another foreseeable extension of our study would be to better compare the two examined screening modalities by analyzing the differences in the tumors in the two considered screening modalities to see if differences exist in their type or size at detection.[29] Moreover, we could only estimate an upper bound for the over-diagnosis, since the period of follow-up modelled may not have exceeded the lead time of some of the slowly progressive cancers.[30]

Finally, the modelling of death due to breast cancer would allow the reporting of new values for the Quality-Adjusted Life Year (QALY) and economical effectiveness of these findings.[31]

CONCLUSIONS

The work reported in this paper provides a basis with which to better model cancer incidence for a population with the inclusion of prevalent cancers. A better estimation of the incidence rates could also be achieved using a semi-Markov model for the progression between healthy, preclinical and clinical states. In addition, the model can be extended to include mortality due to breast cancer, and be used to evaluate more effective screening and treatment policies.

ACKNOWLEDGMENTS

We acknowledge the Natural Sciences and Engineering Research Council (NSERC) of Canada and the Canadian Institute of Health Research (CIHR) of Canada for their financial support in this collaborative research project.

REFERENCES

1. Canadian Cancer Society. Breast Cancer Statistics. Available from: URL: http://www.cancer.ca/en/cancer-information/cancer-type/breast/statistics/?region=on.

2. Duffy SW, Day NE, Tabár L, Chen HH, Smith RA. Markov models of breast tumor progression: some age-specific results. J Natl Cancer Inst Monogr. 1997;22:93-97.

3. Cowen ME, Chartrand M, Weitzel WF. A Markov model of the natural history of prostate cancer. J Clin Epidemiol. 1994;47:3-21.

4. Newton PK, Mason J, Bethel K, Bazhenova LA, Nieva J, Kuhn P. A Stochastic Markov Chain Model to Describe Lung Cancer Growth and Metastasis. PLoS ONE. 2012;**7**:e34637.

5. Jackson CH, Sharples LD, Thompson SG, Duffy SW, Couto E. Multistate Markov models for disease progression with classification error. Statistician. 2003;52:193-209.

6. Wu HM, Yen MF, Chen HH. SAS macro program for non-homogeneous Markov process in modeling multi-state disease progression. Comput Meth Prog Bio. 2004;75:95-105.

7. Taghipour S, Banjevic D, Montgomery N, Jardine AKS. Modeling Breast Cancer Progression and Evaluating Screening Policies. P Annu Rel Maint Sym. 2013; 28-31 Jan. Orlando, FL, United States.

8. Taghipour S, Banjevic D, Miller AB, Montgomery N, Jardine AKS, Harvey B. Parameter estimates for invasive breast cancer progression in the Canadian National Breast Screening Study. Brit J Cancer. 2013;108:542-8.

9. Taghipour S, Banjevic D, Fernandes J, et al. Predictors of Competing Mortality to Invasive Breast Cancer in the Canadian National Breast Screening Study. BMC Cancer. 2012;12:299.

10. Miller AB, Baines CJ, To T, Wall C. Canadian national breast screening study: 1. *Breast cancer detection and death rates among women age* 40-49 years. Can Med Assoc J. 1992;147:1459-76.

11. Miller AB, Baines CJ, To T, Wall C. Canadian national breast screening study: 2. *Breast cancer detection and death rates among women age* 50-59 years. Can Med Assoc J. 1992;147: 1477-88.

12. Miller AB, To T, Baines CJ, Wall C. The Canadian National Breast Cancer Screening Study-1: breast cancer mortality after 11 to 16 years of follow-up. A randomized screening trial of mammography in women age 40 to 49 years. Ann Intern Med. 2002; 137(5 Part 1):305-12.

13. Taghipour S, Banjevic D, Fernandes J, et al. Incidence of invasive breast cancer in the presence of competing mortality: The Canadian National Breast Screening Study. Breast Cancer Res Treat. 2012;134:839-51.

14. Kolb TM, Lichy J, Newhouse JH. Comparison of the Performance of Screening Mammography, Physical Examination, and Breast US and Evaluation of Factors that Influence Them: An Analysis of 27,825 Patient Evaluations. Radiology. 2002;225: 165-75.

15. Shwartz M. A Mathematical Model Used to Analyze Breast Cancer Screening Strategies. Oper Res. 1978;26:937-55.

16. Jackson CH. Multi-state models for panel data: the msm package for R. J Stat Softw. 2011;38:1-29.

17. Duffy SW, Agbaje O, Tabár L, et al. Overdiagnosis and overtreatment of breast cancer: Estimates of overdiagnosis from two trials of mammographic screening for breast cancer. Breast Cancer Res. 2005;**7**:258-65.

18. Fletcher SW, Black W, Harris R, Rimmer BK, Shapiro S. Report of the International Workshop on Screening for Breast Cancer. J Natl Cancer Inst. 1993;85:1644-56.

19. Shen Y, Zelen M. Screening sensitivity and sojourn time from breast cancer early detection clinical trials: mammograms and physical examinations. J Clin Oncol. 2001;19:3490-99.

20. Duffy SW, Chen HH, Tabar L, Fagerberg G, Paci E. Sojourn time, sensitivity and positive predictive value of mammography screening for breast cancer in women aged 40–49. Int J Epidemiol. 1996;25:1139-45.

21. Chen HH, Duffy SW, Tabar L, Day NE. Markov chain models for progression of breast cancer. Part I: tumour attributes and the preclinical screen-detectable phase. J Epidemiol Biostat. 1997; 2:9-23.

22. Chen HH, Duffy SW, Tabar L, Day NE. Markov chain models for progression of breast cancer. Part II: prediction of outcomes for different screening regimes. J Epidemiol Biostat. 1997;2:25-35.

23. Chen HH, Thurfjel E, Duffy SW, Tabar L. Evaluation by Markov chain models of a non-randomised breast cancer screening programme in women aged under 50 years in Sweden. J Epidemiol Commun. 1998;H52:329-35.

24. Wu JC, Hakam M, Anttila A, et al. Estimation of n5tural history parameters of breast cancer based on non-randomized organized screening data: subsidiary analysis of effects of interscreening interval, sensitivity, and attendance rate on reduction of advanced cancer. Breast Cancer Res Treat. 2010;122:553-66.

25. Chen TH, Kuo HS, Yen MF, Lai MS, Tabar L, Duffy SW. Estimation of sojourn time in chronic disease screening without data on interval cases. Biometrics. 2000;56:167-72.

26. Walter SD, Day NE. Estimation of the duration of a preclinical disease state using screening data. Am J Epidemiol. 1983; 118:865-86.

27. Day NE, Walter SD. Simplified models of screening for chronic disease: estimation procedures from mass screening programmes. Biometrics. 1984;40:1-14.

28. Bjurstam N, Bjorneld L, Duffy SW, et al. The Gothenburg breast screening trial: First results on mortality, incidence, and mode of detection for women ages 39-49 years at randomization. Cancer. 1997;80:2091-9.

29. Shaevitch D, Taghipour S, Miller AB, Montgomery N, Harvey B. Tumour Size Distribution of Invasive Breast Cancers and the Sensitivity of Screening Methods in the Canadian National Breast Screening Study. Cancer Res Thera. 2016; In Press.

30. Duffy SW, Parmar D. Overdiagnosis in breast cancer screening: The importance of length of observation period and lead time. Breast Cancer Res. 2013;15:R41.

31. Gocguna Y, Banjevic D, Taghipour S, et al. Cost-effectiveness of breast cancer screening policies using simulation. The Breast. 2015;24: 440-8.