

Estimation of natural history parameters of breast cancer based on non-randomized organized screening data: subsidiary analysis of effects of inter-screening interval, sensitivity, and attendance rate on reduction of advanced cancer

Jenny Chia-Yun Wu · Matti Hakama · Ahti Anttila ·
Amy Ming-Fang Yen · Nea Malila · Tytti Sarkeala ·
Anssi Auvinen · Sherry Yueh-Hsia Chiu · Hsiu-Hsi Chen

Received: 9 September 2009 / Accepted: 17 December 2009 / Published online: 7 January 2010
© Springer Science+Business Media, LLC. 2010

Abstract Estimating the natural history parameters of breast cancer not only elucidates the disease progression but also make contributions to assessing the impact of inter-screening interval, sensitivity, and attendance rate on reducing advanced breast cancer. We applied three-state and five-state Markov models to data on a two-yearly routine mammography screening in Finland between 1988 and 2000. The mean sojourn time (MST) was computed from estimated transition parameters. Computer simulation was implemented to examine the effect of inter-screening interval, sensitivity, and attendance rate on reducing advanced breast cancers. In three-state model, the MST was 2.02 years, and the sensitivity for detecting preclinical breast cancer was 84.83%. In five-state model, the MST was 2.21 years for localized tumor and 0.82 year for non-localized tumor. Annual, biennial, and triennial screening programs can

reduce 53, 37, and 28% of advanced cancer. The effectiveness of intensive screening with poor attendance is the same as that of infrequent screening with high attendance rate. We demonstrated how to estimate the natural history parameters using a service screening program and applied these parameters to assess the impact of inter-screening interval, sensitivity, and attendance rate on reducing advanced cancer. The proposed method makes contribution to further cost-effectiveness analysis. However, these findings had better be validated by using a further long-term follow-up data.

Keywords Breast cancer service screening · Markov model · Natural history · Sensitivity · Inter-screening interval · Attendance rate

Introduction

Evaluation of breast cancer service screening program has increasingly gained attention after the era of randomized

Electronic supplementary material The online version of this article (doi:10.1007/s10549-009-0701-x) contains supplementary material, which is available to authorized users.

J. C.-Y. Wu · M. Hakama · N. Malila · T. Sarkeala ·
A. Auvinen · S. Y.-H. Chiu · H.-H. Chen
Tampere School of Public Health, University of Tampere,
Tampere, Finland
e-mail: chiayun.wu@gmail.com

M. Hakama
e-mail: matti.hakama@cancer.fi

N. Malila
e-mail: Nea.Malila@cancer.fi

T. Sarkeala
e-mail: Tytti.Sarkeala@uta.fi

A. Auvinen
e-mail: anssi.auvinen@uta.fi

S. Y.-H. Chiu
e-mail: sherrychiu@mail.cgu.edu.tw

M. Hakama · A. Anttila · N. Malila
Mass Screening Registry, Finnish Cancer Registry, Helsinki,
Finland
e-mail: Ahti.Anttila@cancer.fi

A. M.-F. Yen
Division of Biostatistics, College of Public Health, National
Taiwan University, Taipei, Taiwan
e-mail: amyyen@ntu.edu.tw

H.-H. Chen (✉)
Division of Biostatistics, College of Public Health, National
Taiwan University, Room 533, No. 17, Hsu-Chow Road, Taipei
100, Taiwan
e-mail: chenlin@ntu.edu.tw

controlled trials. The comparison of mortality can assess whether screening is effective in reducing mortality but ought to require long-term follow-up and may not throw light on a series of subsidiary questions including the additional benefit by shortening inter-screening interval or improving the sensitivity of mammography. Both considerations may rely on the surrogate endpoint of advanced cancer rate.

These problems depend on the disease natural history of breast cancer, because they are determined by how soon the tumor progresses from the pre-clinical screen-detectable phase (PCDP) to the clinical phase. Annual screening with mammography has been suggested in the young women because they had a more rapid progression than the old women [1–3]. Transition rates of disease natural history of women with family history of breast cancer had provided the reference for health policy-makers to determine inter-screening interval [4]. Similar methods have not been applied to evaluation of population-based organized service screening program from which self-selection bias and lead-time bias are inherent. The natural history of breast cancer seems to be heterogeneous across different populations [5, 6]. Therefore, it may be worthwhile to evaluate each breast cancer service screening program by elucidating the disease natural history to answer a series of subsidiary questions mentioned above.

The aim of this study is to estimate the progression rate of breast cancer using three-state and five-state Markov models, making allowance for measurement errors (sensitivity and specificity) and self-selection bias based on data from non-participants, based on data from one breast cancer service screening program in Finland with 12-year follow-up data on different methods of detecting breast cancers, including screen-detected cancers, interval cancers (cancers diagnosed between screens), cancers arising from non-participants (refuser cases), and cancers diagnosed after last invitations (post-screening cancer). We further applied these transition parameters to assess the influence of inter-screening interval and sensitivity on the reduction in the rate of advanced breast cancer.

Materials and methods

Data source

Registration of the nationwide breast cancer screening program in Finland is centrally maintained at the Mass Screening Registry of the Finnish Cancer Registry. Details of the program implementation and registration have been described in full elsewhere [7, 8]. In brief, data used in this study are derived from the screening center of the Pirkanmaa Cancer Society, Tampere, Finland, in 1988–2000. 95,057 invitations and 84,812 screening visits among

33,375 women who were aged 50–59 years at the time of invitation, 25,834 for 50–54 age group, and 7,541 for 55–59 age group, respectively, were identified and recorded individually by Pirkanmaa center during this screening period. Under the two-yearly screening regime, each woman was invited one to five times during age 50–59 years. The number of screens for the overall 50–59 age group, 50–54 age group, and 55–59 age group were 92,752 (2.78 round per women), 75,707 (2.93 rounds per women), and 17,045 (2.26 per women), respectively. As mammography screening was offered to women aged between 50 and 59 years in Finland between 1988 and 2000 and the upper limit of age for breast cancer screening policy was 60 years, the average numbers of screen offered for women aged 55–59 years (2.53 per women) were less than those for women aged 50–54 years (3.07 per women) given the fixed study period. Therefore, slow-growing breast tumor with long sojourn time (i.e., small but still undetectable by mammography when they were invited to screen) for women aged 55–59 years would not be detected given less rounds of screen offered. The mean sojourn time (MST), particularly in women aged 55–59 years, would be underestimated. In order to solve this truncated problem, we retrieved data on breast cancers from our women cohort occurring after the last invitation and diagnosed after 60 years of age, defined as post-screening cancers (PSC). The total of PSC cases were 247 (see Table 1).

Case definition

In this study, each woman was classified according to their disease status and their detection mode of the disease (see Table 1). All of the breast cancer cases were divided into two categories: screen-detected cases, which are defined as preclinical cancer, and clinically detected cases. The clinically detected cases are those not detected from the service screening program. Among screen-detected cases, the patients detected at her first screen were defined as prevalent cases. The patients detected from subsequent screening were defined as incident cases. Two kinds of clinically detected cases, interval cancer and refuser cases, can be observed during the period the screening program executed. Cancers in women after a negative screening result, but diagnosed before the next screening round, are called interval cancers. Those who rejected to come to screen and diagnosed breast cancer clinically were called refuser cases. Table 1 also shows PSC occurring after last invitation and diagnosed and older than 60 years. Each detection mode encodes information for the estimation of parameters pertaining to three-state natural history model (Table 1).

The detection modes based on five-state natural history model is also presented in Table 1. The definition of localized versus non-localized breast cancer is pursuant to

Table 1 Detection modes used for the estimation of transition parameters in three-state and five-state Markov chain models

Detection modes	Information encoded for estimating the parameters of natural history	Number	Contribution to estimating parameters	Likelihood function (see Appendix)
Three-state model				
1. Prevalent screen				
(1) Normal	True negative +false negative cases	31245	Preclincial incidence rate, MST, sensitivity and specificity	Equation (1)
(2) Prevalent cancer	PCDP breast cancer detected at first screen	130		Equation (2)
2. Later screen				
(1) Normal	True negative case +false negative cases (staying in the PCDP) at subsequent screen	52874	Preclincial incidence rate, MST, sensitivity and specificity	Equation (5)
(2) Incident cancer	PCDP Breast cancers detected at subsequent screen	159		Equation (6)
3. Interval cancer	Clinical breast cancer (newly diagnosed cases and false negative cases surfacing to clinical phase at time t) developed between screens	129 ^a	Preclincial incidence rate, MST, sensitivity	Equation (7)
4. Refuser cancer				
Never come to screen	Clinical breast cancers arising from non-participant developed after follow-up time t_2	20	Preclincial incidence rate and MST	Equation (13)
5. Post-screening cancer (after last invitation)	Breast cancers occurring after last invitation with follow-up until the end of study	247		Equation (13)
Five-state Markov model				
1. Prevalent Screen				
(1) Normal	True negative +false negative cases	31245		Equation (14)
(2) Prevalent cancer				
Localized tumor	Localized PCDP breast cancer detected at first screen	75	Preclincial incidence rate and stage-specific MST, sensitivity	Equation (15)
Non-localized tumor	Non-localized PCDP breast cancer detected at first screen	24		Equation (16)
2. Later Screen				
(1) Normal	True negative +false negative cases (staying in the PCDP) at subsequent screen	52874		Equation (20)
(2) Incident cancer				
Localized tumor	Localized PCDP breast cancer detected at subsequent screen	94	Preclincial incidence rate and stage-specific MST, sensitivity	Equation (21)
Non-localized tumor	Non-localized PCDP breast cancer detected at subsequent screen	46		Equation (22)
3. Interval Cancer				
Localized tumor	Clinical breast cancer (newly diagnosed cases and false negative cases surfacing to clinical phase at time t)	61	Preclincial incidence rate and stage-specific MST, sensitivity	Equation (23)
Non-localized tumor	by the classification of whether to have the spread of regional lymph node	57		Equation (24)

Table 1 continued

Detection modes	Information encoded for estimating the parameters of natural history	Number	Contribution to estimating parameters	Likelihood function (see Appendix)
4. Refuser cancer				
Never come to screen				
Localized tumor	Clinical breast cancers arising from non-participant by the classification of whether to have the spread of regional lymph node	7	Preclinical incidence rate and stage-specific MST	Equation (25)
Non-localized tumor		12		Equation (26)
5. Post-screening cancer				
Localized tumor	Clinical breast cancers occurring after last invitation with follow-up time until the end of study by the classification of whether to have the spread of regional lymph node	146	Preclinical incidence rate and stage-specific MST	
Non-localized tumor		79		

t_1 , the time period since the last negative screening; t_2 , the time period from the first invitation to the end of observation

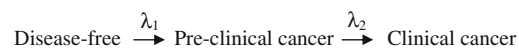
^a These included 22, 34, 44, and 29 cancers surfacing to clinical phase by 0–6 months, 6–12 months, 12–18 months, and 18–24 months

the criteria of Finnish Cancer Registry by assessing whether breast tumor had regional lymph node spread or distant metastases. Information on each detection mode is also presented in a similar manner like the three-state model.

Statistical analysis for estimation of natural history parameters

We used two multi-state Markov models to depict the natural history of breast cancer as shown in Fig. 1. A three-state model illustrated the progression of breast cancer from disease free to the preclinical screen-detectable phase to the clinical phase. Information provided from each detection mode mentioned above for estimating the parameters of disease natural history is presented and delineated in Table 1. The contribution of each detection mode to relevant parameters is also presented in Table 1, including the pre-clinical incidence rate, the MST, sensitivity, and specificity. In the five-stage model, stage of the disease using the definition of advanced breast cancer mentioned above was incorporated. In Fig. 1, λ_1 , λ_2 , λ_3 , and λ_4 are the parameters, the transition rates of moving from one state of the Markov chain to another in an instantaneous period of time, to reflect the intensity of progression from one state to another state. The inverse of the transition rate $1/\lambda$ is the MST. The methodology for the three-state model and five-state model for breast cancer screening follows Chen et al.'s method [9, 10]. The details of model specification and the probabilities for different

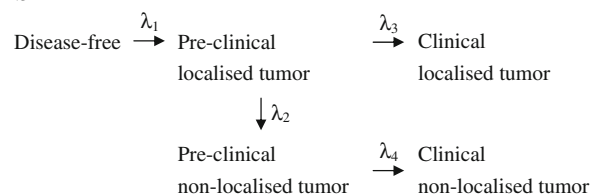
a



λ_1 : transition rate from disease-free to pre-clinical cancer (preclinical incidence rate)

λ_2 : transition rate from pre-clinical cancer to clinical cancer

b



λ_1 : transition rate from disease-free to pre-clinical localized tumor

λ_2 : transition rate from pre-clinical localized tumor to pre-clinical non-localized tumor

λ_3 : transition rate from pre-clinical localized tumor to clinical localized tumor

λ_4 : transition rate from pre-clinical non-localized tumor to clinical non-localized tumor

Non-localized tumor: the tumor metastasized to the lymph node

Fig. 1 Multi-state Markov model for breast cancer. **a** Three-state Markov model. **b** Five-state Markov model

detection modes with annotation for the three-state model are given in “[Appendix](#)”. The details of the corresponding equations for the five-state model are given with supplemental files online (http://homepage.ntu.edu.tw/~ntucbsc/tony_e.htm). The 1-year and 2-year transition probabilities for each possible transition were also computed following

the Chen et al. method [2, 3]. Further, to consider self-selection bias, the likelihood function based on the refuser group was also taken into account. It was assumed that the attendance of women was independent of the progression of breast cancer. In the three-state model, the transition parameters were estimated separately in women aged 50–54 and in women aged 55–59. For the five-state model, we assumed different progression rates from normal to preclinical localized cancer in 1988–1991, 1992–1996, and 1997–2000. In this model, the screening period was parameterized as a covariate to assess how it affects the progression rate from normal to preclinical localized phase. The SAS package with PROC IML command was used to estimate the maximum likelihood estimate (MLE) of all the parameters and their 95% confidence intervals.

Computer simulation

According to the estimated transition rates and sensitivity in the five-state model, the computer simulation was performed to evaluate the effect of different inter-screening intervals on the rate of advanced cancer. Four hypothetical groups of women, each has a total number of 33,375 women, were invited to breast cancer screening for a 12-year period. Four screening regimes were assigned to each group, annually, biennially, triennially, and only once at the end of the screening period (control group). By the application of the Chen et al. methods [2, 3] together with transition parameters estimated from the likelihood based on data from Pirkanmaa center of Finland, effectiveness of reducing advanced cancer under each screening regime would be predicted. The impacts of inter-screening interval, sensitivity, and attendance rate on the reduction in the rate of advanced breast cancer could be assessed.

Results

Table 1 shows the number of breast cancer by different detection modes, encoded information for estimating the parameters of disease natural history. The service screening program organized by Pirkanmaa Center found 130 prevalent screen-detected breast cancers, 159 screen-detected breast cancers at subsequent screens, 129 interval cancers (including 22, 34, 44, and 29 cases occurring 0–6 months, 6–12 months, 12–18 months, and 18–24 months by time since last negative screens) during 1988–2000. For non-participants, we found 20 refuser cases. We ascertained 247 PSC after last invitations with the follow-up time until 2000.

Table 2 shows the estimated transition parameters by the two Markov models. In three-state Markov model, annual preclinical incidence rate was 0.0025 (95% CI = 0.0022–0.0028) for women aged 50–59 years.

Annual transition rate from the PCDP to the clinical phase was 0.4956 per year, which yields 2.02 years of the MST. The sensitivity for detecting the preclinical breast cancer was 84.83% and the specificity was 99.97%. Age-specific MSTs and sensitivity estimates are also presented with 1.92 years and 83.75% for women aged 50–54 years and with 2.34 years and 89.48% for women aged 55–59 years.

In five-state Markov model, annual transition rate from the PCDP to the clinical phase was 0.2897 per year and 1.2230 per year for localized breast tumor and non-localized breast tumor, respectively. The estimated transition rate from preclinical localized phase to preclinical non-localized phase was 0.3371 per year. This suggests that the localized PCDP cancers are more likely to progress to the non-localized PCDP than surface to the clinical localized cancer given the sensitivity. The estimate of the MST was 2.04 years for localized tumor and 0.82 year for non-localized tumor. Assuming the sensitivity of preclinical non-localized tumor 100% the estimate of sensitivity for detecting preclinical localized breast tumor was 68.21%. Assuming variation in incidence (non-constant rate model), the annual incidence rate of the PCDP localized cancer was higher in the period 1988–1996 and was lower in the period 1997–2000.

Table 3 shows the 1-year and 2-year transition probabilities for different state transitions estimated from the transition parameters in different models. In three-state model, almost 39% of the PCDP developed into the clinical phase within 1 year and 63% entered clinical phase within 2 years. In five-state model, the difference of transition probabilities from the pre-clinical phase to the clinical phase was quite large between localized tumor and non-localized tumor. Only 33% of pre-clinical localized tumors were estimated to progress to clinical localized cancer or non-localized cancers compared with 71% of pre-clinical non-localized cancer within 1 year. The corresponding figures were 60% for localized tumor and 91% for non-localized tumor within 2 years.

Figure 2 shows the validation between the two models by comparing the two predicted cumulative incidence rate of non-localized tumor by biennial screening program with the empirical data from Pirkanmaa biennial screening program. The model which parameterized screening period as a covariate had a better fit compared to the model with the constant rate of preclinical localized cancer.

Figure 3 represents the cumulative incidence of non-localized tumor by different inter-screening intervals according to the results of computer simulation. The cumulative incidence of non-localized tumor of control group is about 120 per 10,000 during 12 years. Compared to the control group, annual, biennial, and triennial screening programs can reduce 53, 37, and 28% of non-localized tumor, respectively.

Table 4 shows the simulation results of the relative risk to develop non-localized tumor of different screening

Table 2 Estimated parameters for progression rate and the sensitivity in three-state Markov model and five-state Markov model

Parameters	Estimates	95% CI
Three-state model ^a		
50–59		
Normal → preclinical cancer (λ_1)	0.0025	(0.0022, 0.0028)
Preclinical cancer → clinical cancer (λ_2)	0.4956	(0.3816, 0.6097)
Mean sojourn time ($1/\lambda_2$)	2.02	(1.64, 2.62)
Sensitivity	84.83%	(74.88%, 94.79%)
Specificity	99.97%	(99.89%, 100%)
50–54		
Normal → preclinical cancer (λ_1)	0.0025	(0.0022, 0.0027)
Preclinical cancer → clinical cancer (λ_2)	0.5207	(0.4057, 0.6356)
Mean sojourn time ($1/\lambda_2$)	1.92	(1.57, 2.46)
Sensitivity	83.75%	(71.26%, 96.23%)
55–59		
Normal → preclinical cancer (λ_1)	0.0025	(0.0021, 0.0029)
Preclinical cancer → clinical cancer (λ_2)	0.4269	(0.3131, 0.5408)
Mean sojourn time ($1/\lambda_2$)	2.34	(1.85–3.19)
Sensitivity	89.48%	(76.56%, 100%)
Five-state model		
50–59		
Normal → preclinical N(–) (λ_1)	0.0025	(0.0023, 0.0027)
Preclinical N(–) → preclinical N(+) (λ_2)	0.3371	(0.2549, 0.4192)
Preclinical N(–) → clinical N(–) (λ_3)	0.2897	(0.2186, 0.3609)
Mean sojourn time $\left(\frac{1}{\frac{(\lambda_2 + \lambda_3)^+}{\lambda_2} + \lambda_4} \right)$	2.04	
Preclinical N(+) → clinical N(+) (λ_4)	1.2230	(0.9259, 1.5201)
Mean sojourn time ($1/\lambda_4$)	0.82	(0.66, 1.08)
Sensitivity of preclinical N(–) cancer	68.21%	(54.63%, 81.79%)
^b Period as a covariate for λ_1		
Normal → preclinical N(–) (λ_1)		
Period 1988–1991	0.0026	(0.0023, 0.0028)
Period 1992–1996	0.0026	(0.0022, 0.0031)
Period 1997–2000	0.0020	(0.0015, 0.0028)
Preclinical N(–) → preclinical N(+) (λ_2)	0.3298	(0.2488, 0.4109)
Preclinical N(–) → clinical N(–) (λ_3)	0.2828	(0.2126, 0.3531)
Preclinical N(+) → clinical N(+) (λ_4)	1.2052	(0.9054, 1.505)
Sensitivity of preclinical N(–) cancer	67.56%	(54.04%, 81.07%)

Goodness-of-fit for three-state model $X^2 = 2.69$, d.f. = 4, P -value = 0.61

Goodness-of-fit for five-state model $X^2 = 12.12$, d.f. = 7, P -value = 0.10

Goodness-of-fit for five-state model (piecewise method) $X^2 = 37.88$, d.f. = 29, P -value = 0.13

^a The estimation was independently performed for three age groups

^b Baseline period: 1988–1991

regime by different screening sensitivity of localized tumor. It shows that the biennial screening regime could reduce large proportion of non-localized tumor even though the sensitivity of localized tumor is not high. The reduction ranged from 33 to 46% when the sensitivity was changed from 60 to 90%. Table 5 shows the impact of attendance rate on the effectiveness of reducing advanced breast cancer. It can be seen that the effectiveness of intensive screening with poor attendance rate is the same as that of infrequent screening with high attendance rate. For example, the reduction in rate of advanced breast cancer

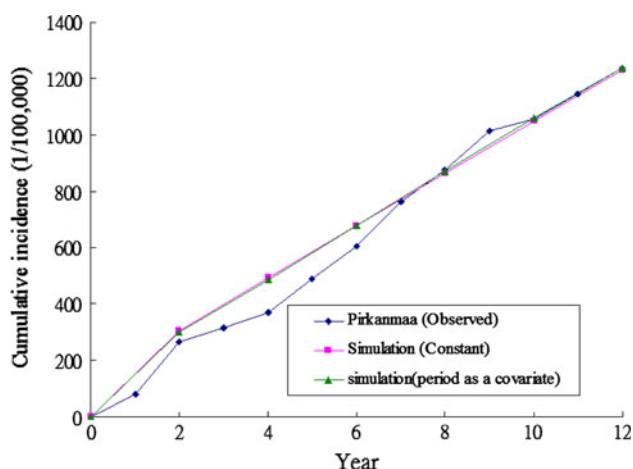
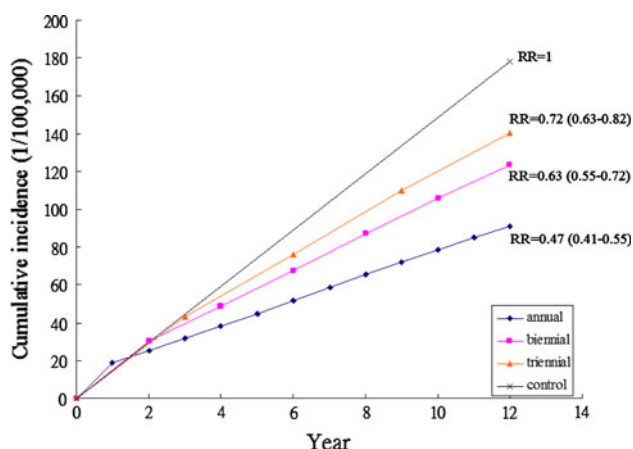
for annual screening with 60% attendance rate was close to that for biennial screening with 90% attendance rate.

Discussion

By the application of two multi-state Markov processes to the data from the routine two-yearly breast cancer screening regime in Finland, we estimated the pre-clinical incidence rate and the MST staying at the PCDP of breast cancer with and without the classification of stage, taking

Table 3 Estimated transition probabilities from three-state model and five-state model

Initial state	Final state	Transition probability	
		1 year	2 years
Three-state model			
Pre-clinical cancer	Pre-clinical cancer	0.6092	0.3711
	Clinical cancer	0.3908	0.6289
Five-state model			
Pre-clinical localized cancer	Pre-clinical localized cancer	0.5343	0.2855
	Pre-clinical non-localized cancer	0.1357	0.1124
	Clinical localized cancer	0.2152	0.3302
	Clinical non-localized cancer	0.1148	0.2719
Pre-clinical non-localized cancer	Pre-clinical non-localized cancer	0.2943	0.0866
	Clinical non-localized cancer	0.7057	0.9134

**Fig. 2** Cumulative incidence of non-localized breast cancer of observed data from Pirkanmaa center and of simulation results**Fig. 3** Cumulative incidence of non-localized tumor by different screening regimes

into account measurement errors (sensitivity and specificity) and self-selection bias based on data from non-participants. The MST of preclinical detectable phase for women aged 50–59 years was estimated to be 2.02 years in current study after adjustment for 15% false negative cases missed

at screen. Compared to previous studies, the estimation using the data from Swedish Two-County trial was 3.3 years for the same age group [5]. Norwegian Breast Cancer Screening Programme study estimated the MST to be 4 years or below after excluding intraductal carcinomas, adjusting for extra background incidence and correcting for possible bias due to opportunistic screening between screening rounds by using non-linear least-square regression approach for the estimation of a three-state Markov chain model [6]. The data used in the Norwegian study was only from the prevalent screening with a higher proportion of slow-growing tumors in screen-detected cancers compared to the subsequent incidence rounds. Therefore, the estimate of the MST is longer than that from the studies including data on incident screen.

Some potential factors could contribute to the short MST estimated in the current study. First, the follow-up time of our screened cohort, particularly aged 55–59 years may not be sufficient longer for identifying slow-growing (long sojourn time) breast tumor with the potential of surfacing to clinical phase due to symptom and signs. In contrast, the estimated MST on women aged 50–59 years with the data from the two-county study was based on at least 20 years of follow-up when the corresponding figure was estimated. Our cohort for women aged 50–59 years have 12-year follow-up. Hence, a woman aged 59 years invited to screen in the year 1988 can be only followed over time until 71 years. The shorter follow-up period precluded the identification of slow-growing tumor until breast cancers occur in the old age even data on PSC have been collected and incorporated into the model. The shorter follow-up together with the cessation of screening after 60 years of age, leading to the truncation of slow-growing breast tumor, may result in a lower MST estimate, particularly women aged 55–59 years. This should be validated in future with a further long-term follow-up study.

Second, as sensitivity and MST are negatively correlated the higher the sensitivity the shorter the sojourn time. Using

Table 4 Relative risk of non-localized breast cancer of different screening regime by screening sensitivity

	Sensitivity of localized tumor (%)	Control group	Screening annually	Screening biennially	Screening triennially
	68.2 ^a	1	0.47 (0.41, 0.55)	0.63 (0.55, 0.72)	0.72 (0.63, 0.82)
	60	1	0.51 (0.44, 0.59)	0.67 (0.59, 0.76)	0.75 (0.66, 0.85)
	80	1	0.42 (0.36, 0.49)	0.58 (0.50, 0.66)	0.67 (0.59, 0.77)
^a Estimate from five-state Markov model	90	1	0.38 (0.33, 0.45)	0.54 (0.47, 0.62)	0.64 (0.56, 0.73)

Table 5 Relative risk of non-localized breast cancer by attendance rate with 68.2^a sensitivity for localized breast cancer

	Attendance rate (%)	Control group	Screening annually	Screening biennially	Screening triennially
	100	1	0.44 (0.38, 0.51)	0.61 (0.53, 0.69)	0.70 (0.62, 0.80)
	90	1	0.49 (0.43, 0.57)	0.65 (0.57, 0.74)	0.73 (0.64, 0.83)
	60	1	0.66 (0.58, 0.75)	0.76 (0.67, 0.86)	0.82 (0.73, 0.93)
^a Estimate from five-state Markov model	30	1	0.83 (0.74, 0.94)	0.88 (0.78, 0.99)	0.91 (0.81, 1.03)

the Markov model together with detection modes (screen-detected plus interval cancers) to separate the role of test sensitivity (measurement error) accounting for interval cancer from that of sojourn time (biological property), we reckon the main reason for short sojourn time is, to a larger extent, due to rapid progression from the PCDP to clinical phase, which, in turn, leads to higher proportion of interval cancer (30.9%), higher than that in the Swedish two-county trial (25.8%) even though the inter-screening interval was shorter in Finnish service screening [11, 12], and the sensitivity was only slightly lower. This supports the postulate of rapid tumor progression from the PCDP to clinical phase leading to the occurrence of interval cancers. The biological and organized factors accounting for this remain unclear and need to be investigated. The increase in breast density due to the use of hormone replacement therapy and improved access to diagnostic services were proposed to explain the increased interval cancer from 1991 to 1999 in Finland [11] or in other European countries. The shortened time period from the symptoms appeared to the access of the medical services also contributed to the faster progression rate from the preclinical phase to the clinical phase. However, this should be interpreted with great caution before being validated by long-term follow-up empirical data. In spite of lower MST, we think such an underestimation of MST would not affect the evaluation of relative effectiveness of breast cancer screening regarding the effect of inter-screening interval, sensitivity, and attendance rate because the relative rather than absolute benefit was of interest in this study.

When allowing different progression rates for localized and non-localized tumors, the MST was estimated to be 2.2 years for preclinical localized tumor and 0.7 year for preclinical non-localized tumor, respectively, after adjustment for 32% false negative localized breast tumor missed at screen (68% sensitivity for localized breast cancer). This

result quite corresponds to the biological viewpoint that the progression rate for non-localized tumor is much faster than localized tumor. Note that the estimated MST for localized breast cancer also takes into account the latent progression from the localized PCDP cancer to non-localized PCDP cancer. The finding that this latent transition rate doubled the transition rate from the PCDP to the clinical phase for localized breast cancer implies the benefit of breast cancer screening with mammography is to arrest this latent progression. Compared to the 1-year transition probabilities derived from five-state model in Swedish Two-County Trial study [2], the present results show higher probabilities to enter the clinical phase from the preclinical phase within 1 year due to the faster transition rates observed in current study. The transition probability from preclinical phase to clinical phase with non-localized cancer is approximately six times for non-localized tumor compared to localized tumor within 1 year for both the Swedish study and this study. In addition, the current study demonstrated that screening biennially could make 37% reduction of non-localized cancer from computer simulation. This result was consistent with the predicted effect that screening regularly could reduce large proportion of node positive tumors by using the data from the Swedish Two-County Trial [3].

Unlike previous studies using either detection method or incidence method to estimate the sensitivity of breast cancer screening [11, 13]. This is the first study to report the sensitivity of breast cancer screening in Finland using Markov model approach. The total sensitivity was estimated to be 85%, and the sensitivity for localized tumor was around 68%. When estimating sensitivity using either detection method or incidence method, the interval cancer is regard as the cancer which was overlooked at previous screening. This will result in underestimation of sensitivity

due to some interval cancer arise between screening rounds in reality. Therefore, it is reasonable that the estimated sensitivity in current study is higher than the previous estimation by using maximum likelihood approach based on the development of Markov process.

One limitation in this study is that as mentioned earlier a shorter follow-up period may lead to underestimation of MST, particularly for women aged 55–59 years. Although this may not affect the evaluation of relative effectiveness of screening program, the absolute value of estimated parameters on natural history such as MST estimates would be interpreted with great caution by program evaluator and health policy-maker when they are compared with that from previous studies as different follow-up times and screening policies may affect the estimation of parameters of natural history.

In conclusion, we demonstrated how to estimate natural history parameters and sensitivity and specificity using a 12-year follow-up of two-yearly routine mammographic screening regime implemented in Pirkanmaa, Finland. We applied these parameters to assess how changing the inter-screening interval and sensitivity affect the rate of advanced cancer. However, these results should be validated with a further follow-up data.

Acknowledgment This research work was supported by the FiDi-Pro Research Project of Tampere School of Public Health Granted from Academy of Finland. It was also supported by the National Science Council of Taiwan (NSC 96-2628-B-002-096-Mr3).

Appendix

Three-state Model

Let the stochastic process of disease natural history of breast cancer denoted by a random variable $X(t)$, the outcome of which is defined by a state space $W = \{1, 2, 3\}$, where states 1, 2, and 3 stand for free of breast cancer, the PCDP, and clinical cancer, respectively. In this three-state model, the transition probabilities from previous state (i) to current state (j) during time t can be represented by Table 6.

The application of this transition matrix to data on breast cancer screening has been described in full elsewhere [10]. The detailed likelihood functions for estimating the natural history parameters are decomposed by round of screens and detection modes.

Prevalent Screen

Suppose women invited to first screen (prevalent screen) at age m , the probabilities of having negative screening result (P_{s1_1}) and positive screen-detected results (P_{s1_2}) using transition probabilities and sensitivity (sen) and specificity (spe) are written as follows.

$$\begin{aligned}
 &P_{s1_1}(\text{Probability of having negative screening result at first screen}) \\
 &= \text{Probability of being true negative} + \text{probability of being false negative} \\
 &= \frac{\left(\text{Probability of free of cancer at age of entry } m \times \text{specificity} \right. \\
 &\quad \left. + \text{probability of preclinical cancer at age of entry } m \times \text{probability of being missed} \right)}{\left(\text{Probability of free of cancer} + \text{probability of preclinical cancer at age of entry } m \right)} \\
 &= \frac{P_{11}(m) \times \text{spe} + P_{12}(m) \times (1 - \text{sen})}{P_{11}(m) + P_{12}(m)} \quad (1)
 \end{aligned}$$

$$\begin{aligned}
 &P_{s1_2}(\text{Probability of having positive screen results at first screen}) \\
 &= \text{Probability of being false positive} + \text{probability of being true positive} \\
 &= \frac{\left(\text{Probability of free of cancer at age of entry } m \times \text{probability of being false positive} \right. \\
 &\quad \left. + \text{probability of preclinical cancer at age of entry } m \times \text{sensitivity} \right)}{\left(\text{Probability of free of cancer} + \text{probability of preclinical cancer at age of entry } m \right)} \\
 &= \frac{P_{11}(m) \times (1 - \text{spe}) + P_{12}(m) \times \text{sen}}{P_{11}(m) + P_{12}(m)} \quad (2)
 \end{aligned}$$

Table 6 Transition probabilities from the previous state to the current state during time t ($P_{ij}(t)$)

Previous state (i)	Current state (j)		
	Normal ($j = 1$)	Preclinical cancer ($j = 2$)	Clinical cancer ($j = 3$)
Normal ($i = 1$)	$P_{11}(t)^*$	$P_{12}(t)$	$P_{13}(t)$
Preclinical cancer ($i = 2$)	0	$P_{22}(t)$	$P_{23}(t)$
Clinical cancer ($i = 3$)	0	0	1

* $P_{11}(t)$ denotes the transition probability from normal state to normal state during time t

$P_{12}(t)$, $P_{13}(t)$, $P_{22}(t)$ and $P_{23}(t)$ denote the transition probability for other transition processes in the same rationale

Those women who have negative screening results at first screen are composed of those who are actually disease free (true negative) or misclassified (false negative), whereas those women who have positive screening results at first screen consist of who are actually disease (true positive) or misclassified (false positive).

asymptomatic women missed at first screen with the respective proportions denoted by $tn(1)$ (true negative at first screen) and $fn(1)$ (false negative at first screen), respectively. The formulae of these two probabilities are expressed as follows.

$tn(1)$ (Probability of being true negative at first screen)

$$\begin{aligned}
 &= \frac{\text{Probability of free of cancer at age of entry } m \times \text{specificity}}{\left(\text{Probability of free of cancer at age of entry } m \times \text{specificity} \right.} \\
 &\quad \left. + \text{probability of preclinical cancer at age of entry } m \times \text{probability of being missed} \right) \quad (3) \\
 &= \frac{P_{11}(m) \times \text{spe}}{P_{11}(m) \times \text{spe} + P_{12}(m) \times (1 - \text{sen})}
 \end{aligned}$$

$fn(1)$ (Probability of being false negative at first screen)

$$\begin{aligned}
 &= \frac{\text{Probability of preclinical cancer at age of entry } m \times \text{probability of being missed}}{\left(\text{Probability of free of cancer at age of entry } m \times \text{specificity} \right.} \\
 &\quad \left. + \text{probability of preclinical cancer at age of entry } m \times \text{probability of being missed} \right) \quad (4) \\
 &= \frac{P_{12}(m) \times (1 - \text{sen})}{P_{11}(m) \times \text{spe} + P_{12}(m) \times (1 - \text{sen})}
 \end{aligned}$$

Incident Screen

Owing to false negative cases missed at prevalent screen, the underlying population for the second screen after the prevalent screen consist of disease free (true negative) and

Assume the false negative cases missed at first screen can be detected at the second screen, the probabilities of screen-negative women (P_{s2_1}) and screen-detected cases (P_{s2_2}) at the second screen and the probabilities of interval cancer (P_{s2_1}) diagnosed before second screen and refuser

cases (P_{s2_R}) at second screen but diagnosed at time t after first screen are expressed as follows.

approximation to the instantaneous time (dt) used in the derivative of the probability density function. For example,

$$\begin{aligned}
 &P_{s2_1}(\text{Probability of having negative screening result at second round}) \\
 &= \left(\begin{array}{l} \text{Probability of being true negative at first round} \times \text{probability of staying normal during time } t \\ \times \text{specificity} \end{array} \right) \\
 &+ \left(\begin{array}{l} \text{probability of being true negative at first round} \times \text{transition probability from normal to preclinical during time } t \\ \times \text{probability of being missed} \end{array} \right) \\
 &= \text{tn}(1) \times P_{11}(t) \times \text{spe} + \text{tn}(1) \times P_{12}(t) \times (1 - \text{sen})
 \end{aligned} \tag{5}$$

$$\begin{aligned}
 &P_{s2_2}(\text{Probability of having positive result at second round}) \\
 &= \left(\begin{array}{l} \text{Probability of being true negative at first round} \times \text{probability of staying normal during time } t \\ \times \text{probability of being false positive} \end{array} \right) \\
 &+ \left(\begin{array}{l} \text{probability of being true negative at first round} \times \text{transition probability from normal to preclinical during time } t \\ \times \text{sensitivity} \end{array} \right) \\
 &+ (\text{probability of being false negative at first round} \times \text{probability of staying preclinical state during time } t) \\
 &= \text{tn}(1) \times P_{11}(t) \times (1 - \text{spe}) + \text{tn}(1) \times P_{12}(t) \times \text{sen} + \text{fn}(1) \times P_{22}(t)
 \end{aligned} \tag{6}$$

$$\begin{aligned}
 &P_{s2_1}(\text{Probability of interval cancer between first and second round}) \\
 &= \left(\begin{array}{l} \text{Probability of being true negative at first round} \\ \times \text{probability of surfacing to clinical state during the short time period } \Delta t \text{ before the end of time } t \end{array} \right) \\
 &+ \left(\begin{array}{l} \text{probability of being false negative at first round} \\ \times \text{probability of surfacing to clinical state during the short time period } \Delta t \text{ before the end of time } t \end{array} \right) \\
 &= \text{tn}(1) \times [P_{11}(t - \Delta t) \times P_{13}(\Delta t) + P_{12}(t - \Delta t) \times P_{23}(\Delta t)] + \text{fn}(1) \times [P_{22}(t - \Delta t) \times P_{23}(\Delta t)]
 \end{aligned} \tag{7}$$

$$\begin{aligned}
 &P_{s2_R}(\text{Probability of refuser cases after attending first round}) \\
 &= \left(\begin{array}{l} \text{Probability of being true negative at first round} \\ \times \text{probability of surfacing to clinical state during the short time period } \Delta t \text{ before the end of time } t \end{array} \right) \\
 &+ \left(\begin{array}{l} \text{probability of being false negative at first round} \\ \times \text{probability of surfacing to clinical state during the short time period } \Delta t \text{ before the end of time } t \end{array} \right) \\
 &= \text{tn}(1) \times [P_{11}(t - \Delta t) \times P_{13}(\Delta t) + P_{12}(t - \Delta t) \times P_{23}(\Delta t)] + \text{fn}(1) \times [P_{22}(t - \Delta t) \times P_{23}(\Delta t)]
 \end{aligned} \tag{8}$$

The first components of the Eqs. 5–7 delineate the progression of true negative subjects after prevalent screen and the second components give delineate the progression of false negative cases missed at prevalent screen. Note that, in the above formulae, t is the time interval between the prevalent and the second screen, inter-screening interval, or time after first screen for the refuser, and Δt , say one month, which is an

the first component of the bracket on the right side of the Eq. 7, $P_{11}(t - \Delta t) \times P_{13}(\Delta t) + P_{12}(t - \Delta t) \times P_{23}(\Delta t)$, which is called compound probability, is an approximation to $dP_{13}(t)/dt$. The merit of using compound probability has been described in Duffy et al. study [14] and Kay [15] as it can accommodate the rapid and slow progression of breast cancer.

Similar to the prevalent screen, the underlying population for the third screen after the second screen round is also composed of two categories, disease free and asymptomatic women missed at second screen, with the respective probabilities of $tn(2)$ and $fn(2)$ with formulae as follows.

screen round can be derived in a similar manner by the incorporation of $tn(2)$ and $fn(2)$ into the likelihood function.

$$\begin{aligned}
 & tn(2) (\text{Probability of being true negative at second round}) \\
 &= \frac{\left(\begin{array}{l} \text{Probability of being true negative at first round} \\ \times \text{probability of staying normal during time } t \text{ between first and second round} \times \text{specificity} \end{array} \right)}{\left[\begin{array}{l} \left(\begin{array}{l} \text{Probability of being true negative at first round} \\ \times \text{probability of staying normal during time } t \text{ between first and second round} \times \text{specificity} \end{array} \right) \\ + \left(\begin{array}{l} \text{Probability of being true negative at first round} \\ \times \text{probability of transition from normal to preclinical cancer during time } t \text{ between first and second round} \\ \times \text{probability of being missed} \end{array} \right) \end{array} \right]} \\
 &= \frac{tn(1) \times P_{11}(t) \times spe}{tn(1) \times P_{11}(t) \times spe + tn(1) \times P_{12}(t) \times (1 - sen)} \quad (9)
 \end{aligned}$$

$$\begin{aligned}
 & fn(2) (\text{Probability of being false negative at second screen}) \\
 &= \frac{\left(\begin{array}{l} \text{Probability of true negative at first round} \\ \times \text{probability of transition from normal to preclinical cancer during time } t \text{ between first and second round} \\ \times \text{probability of being missed} \end{array} \right)}{\left[\begin{array}{l} \left(\begin{array}{l} \text{Probability of being true negative at first round} \\ \times \text{probability of staying normal during time } t \text{ between first and second round} \times \text{specificity} \end{array} \right) \\ + \left(\begin{array}{l} \text{Probability of true negative at first round} \\ \times \text{probability of transition from normal to preclinical cancer during time } t \text{ between first and second round} \\ \times \text{probability of being missed} \end{array} \right) \end{array} \right]} \\
 &= \frac{tn(1) \times P_{12}(t) \times (1 - sen)}{tn(1) \times P_{11}(t) \times spe + tn(1) \times P_{12}(t) \times (1 - sen)} \quad (10)
 \end{aligned}$$

The above formulae can be simplified as

$$tn(2) = \frac{P_{11}(t) \times spe}{P_{11}(t) \times spe + P_{12}(t) \times (1 - sen)} \quad (11)$$

$$fn(2) = \frac{P_{12}(t) \times (1 - sen)}{P_{11}(t) \times spe + P_{12}(t) \times (1 - sen)}. \quad (12)$$

The likelihood function for screen-detected cases, screen-negative women, and interval cancer at the third

The Refuser Group

The probability of developing breast cancer for those who never come to screen (P_{NA}) is expressed as follows. In the following formula, m represents the age of the first invitation to the screening program, t_2 represents the time period from the first invitation to the year of the diagnosis of breast cancer.

P_{NA} (Probability of developing breast cancer for those who never come to screen)

$$\begin{aligned}
 & \left[\begin{aligned} & \left(\begin{aligned} & \text{Probability of free of cancer at age of first invitation } m \\ & \times \text{probability of surfacing to clinical state during the short time period } \Delta t \text{ before the end of time } t_2 \end{aligned} \right) \\ & + \left(\begin{aligned} & \text{probability of preclinical cancer at age of first invitation } m \\ & \times \text{probability of surfacing to clinical state during the short time period } \Delta t \text{ before the end of time } t_2 \end{aligned} \right) \end{aligned} \right] \\
 & = \frac{(\text{Probability of free of cancer} + \text{probability of preclinical cancer at age of entry } m)}{[P_{11}(m) + P_{12}(m)]} \\
 & = \frac{[P_{11}(m) \times P_{11}(t_2 - \Delta t) \times P_{13}(\Delta t) + P_{11}(m) \times P_{12}(t_2 - \Delta t) \times P_{23}(\Delta t) + P_{12}(m) \times P_{22}(t_2 - \Delta t) \times P_{23}(\Delta t)]}{[P_{11}(m) + P_{12}(m)]} \quad (13)
 \end{aligned}$$

Table 7 Transition probability from previous state to current state during time t ($P_{ij}(t)$)

Previous state (i)	Current state (j)				
	Normal ($j = 1$)	Preclinical localized cancer ($j = 2$)	Preclinical non-localized cancer ($j = 3$)	Clinical localized cancer ($j = 4$)	Clinical non-localized cancer ($j = 5$)
Normal ($i = 1$)	$P_{11}(t)^*$	$P_{12}(t)$	$P_{13}(t)$	$P_{14}(t)$	$P_{15}(t)$
Preclinical localized cancer ($i = 2$)	0	$P_{22}(t)$	$P_{23}(t)$	$P_{24}(t)$	$P_{25}(t)$
Preclinical non-localized cancer ($i = 3$)	0	0	$P_{33}(t)$	$P_{34}(t)$	$P_{35}(t)$
Clinical localized cancer ($i = 4$)	0	0	0	$P_{44}(t)$	$P_{45}(t)$
Clinical non-localized cancer ($i = 5$)	0	0	0	0	1

* $P_{11}(t)$ denotes the transition probability from normal state to normal state during time t .

$P_{12}(t)$, $P_{13}(t)$, $P_{14}(t)$, $P_{15}(t)$, $P_{22}(t)$, $P_{23}(t)$, $P_{24}(t)$, $P_{25}(t)$, $P_{33}(t)$, $P_{34}(t)$, $P_{35}(t)$, $P_{44}(t)$ and $P_{45}(t)$ denote the transition probability for other transition processes in the same rationale

The Eq. 13 is also applied to PSC (see Table 1).

Five-state model

The likelihood functions using for the five-state model are developed in the same way as the three-state model. In the five-state model, the state space is changed as $W = \{1, 2, 3, 4, 5\}$, where states 1, 2, 3, 4, and 5 denote free of breast cancer, preclinical localized cancer, preclinical non-localized cancer, clinical localized cancer, and clinical non-localized cancer, respectively. The transition probabilities from state i to state j during time t can be represented by Table 7.

As in three-state model, the likelihood functions for the estimation of parameters are also decomposed by rounds of screen and detection modes. The details of likelihood functions are given with supplemental files and available online (http://homepage.ntu.edu.tw/~ntucbsc/tony_e.htm).

References

1. Tabar L, Fagerberg G, Chen HH, Duffy SW, Smart CR, Gad A, Smith RA (1995) Efficacy of breast cancer screening by age. *Cancer* 75:2507–2517
2. Chen HH, Duffy SW, Tabar L, Day NE (1997) Markov chain models for progression of breast cancer. Part I: tumour attributes and the preclinical screen-detectable phase. *J Epidemiol Biostat* 2:9–23
3. Chen HH, Duffy SW, Tabar L, Day NE (1997) Markov chain models for progression of breast cancer. Part II: prediction of outcomes for different screening regimes. *J Epidemiol Biostat* 2:25–35
4. Lai MS, Yen MF, Kuo HS, Koong SL, Chen THH, Duffy SW (1998) Efficacy of breast-cancer screening for female relatives of breast-cancer-index cases: Taiwan multicentre cancer screening (TAMCAS). *Int J Cancer* 78:21–26
5. Tabar L, Fagerberg G, Chen HH, Duffy SW, Smart CR, Gad A, Smith RA (1995) Efficacy of breast cancer screening by age. *Cancer* 1995:2507–2517
6. Weedon-Fekjær H, Vatten LJ, Aalen OO, Lindqvist B, Tretli S (2005) Estimating mean sojourn time and screening test sensitivity in breast cancer mammography screening: new results. *J Med Screen* 12:172–178
7. Sarkeala T, Anttila A, Forsman H, Luostarinen T, Saarenmaa I, Hakama M (2004) Process indicators from ten centres in the Finnish breast cancer screening programme from 1991 to 2000. *Eur J Cancer* 40:2116–2125. doi:10.1016/j.ejca.2004.06.017
8. Wu JC, Anttila A, Yen AM, Hakama M, Saarenmaa I, Sarkeala T, Malila N, Auvinen A, Chiu SY, Chen THH (2009) Evaluation of breast cancer service screening programme with a Bayesian approach: mortality analysis in a Finnish region. *Breast Cancer Res Treat* (in press). doi:10.1007/s10549-009-0604-x

9. Chen HH, Duffy SW, Tabar L (1996) A markov chain method to estimate the tumour progression rate from preclinical to clinical phase, sensitivity and positive predictive value for mammography in breast cancer screening. *The Stat* 45:307–317
10. Chen THH, Kuo HS, Yen MF, Lai MS, Tabar L, Duffy SW (2000) Estimation of sojourn time in chronic disease screening without data on interval cases. *Biometrics* 56:167–172
11. Sarkeala T, Hakama M, Saarenmaa I, Hakulinen T, Forsman H, Anttila A (2006) Episode sensitivity in association with process indicators in the Finnish breast cancer screening program. *Int J Cancer* 118:174–179. doi:[10.1002/ijc.21310](https://doi.org/10.1002/ijc.21310)
12. Duffy SW, Day NE, Tabar L, Chen HH, Smith TC (1997) Markov models of breast tumor progression: some age-specific results. *J Natl Cancer Inst Monogr* 22:93–97
13. Anttila A, Koskela J, Hakama M (2002) Programme sensitivity and effectiveness of mammography service screening in Helsinki, Finland. *J Med Screen* 9:153–158
14. Duffy SW, Chen HH, Tabar L, Day NE (1995) Estimation of mean sojourn time in breast cancer screening using a Markov chain model of both entry to and exit from the preclinical detectable phase. *Statist Med* 14:1531–1543
15. Kay R (1986) A Markov model for analyzing cancer markers and disease states in survival studies. *Biometrics* 42:855–865