Article

SMMR
STATISTICAL METHODS IN MEDICAL RESEARCH

Statistical Methods in Medical Research
0(0) 1–21

DOI: 10.1177/0962280216682284
journals.sagepub.com/home/smm
SAGE

Bayesian negative-binomial-family-based multistate Markov model for the evaluation of periodic population-based cancer screening considering incomplete information and measurement errors

Chen-Yang Hsu,[1] Ming-Fang Yen,[2] Anssi Auvinen,[3] Yueh-Hsia Chiu[4] and Hsiu-Hsi Chen[1]

Abstract

Population-based cancer screening is often asked but hardly addressed by a question: "How many rounds of screening are required before identifying a cancer of interest staying in the pre-clinical detectable phase (PCDP)?" and also a similar one related to the number of screens required for stopping screening for the low risk group. It can be answered by using longitudinal follow-up data on repeated rounds of screen, namely periodic screen, but such kind of data are rather complicated and fraught with intractable statistical properties including correlated multistate outcomes, unobserved and incomplete (censoring or truncation) information, and imperfect measurements. We therefore developed a negative-binomial-family-based discrete-time stochastic process, taking sensitivity and specificity into account, to accommodate these thorny issues. The estimation of parameters was implemented with Bayesian Markov Chain Monte Carlo method. We demonstrated how to apply this proposed negative-binomial-family-based model to the empirical data similar to the Finnish breast cancer screening program.

Keywords

Multistate model, geometric model, negative binomial model, disease natural history, cancer screening

1 Introduction

Population-based organized service screening program is characterized by sort of follow-up data involving repeated observations with regular and predetermined interval.[1–3] Such panel data offer an opportunity to elucidate the progression of disease and associated factors. While population-based screening for cancer such as breast and colon and rectum has increasingly gained attention,[4–7] a series of subsidiary issues have been raised such as inter-screening interval and age to begin with screen.[8–13] From the practical aspect of screening, both are related to the costs incurred in the number of screening rounds required before detecting a cancer given a routine screening policy. Hence, these issues are related to the planning of an organized screening program as well as the allocation of medical resources.

Although these arguments have been addressed in a continuous-time Markov model characterized by the rate of being susceptible to pre-clinical detectable phase (PCDP) and the mean sojourn time with[3,9,11–15] and without[15,16] adjustment for measurement errors, a practical question that is often asked but has been rarely addressed is "How many screening rounds are required for detecting a cancer of interest staying in the pre-clinical (asymptomatic)

[1]Division of Biostatistics, College of Public Health, National Taiwan University, Taipei, Taiwan
[2]School of Oral Hygiene, College of Oral Medicine, Taipei Medical University, Taipei, Taiwan
[3]Tampere School of Health Science, University of Tampere, Tampere, Finland
[4]Department of Health Care Management, College of Management, Chang Gung University, Tao-Yuan, Taiwan

Corresponding author:
Hsiu-Hsi Chen, Division of Biostatistics, College of Public Health, National Taiwan University, Room 533, No. 17, Hsu-Chow Road, Taipei 100, Taiwan.
Email: chenlin@ntu.edu.tw

detectable phase (PCDP)?''. The similar question is also asked like ''Can a subject be categorized as very low risk for stopping screening after several rounds of screening with negative results?'' These questions motivate us to develop the alternative models in contrast to the pervious well-established continuous-time Markov models by estimating the average rounds of screens denoted by S_N before detecting the cancer given a predetermined inter-screening interval in the light of geometric distribution at individual level and negative binomial distribution at population level. Given various scenarios of inter-screening interval and age to begin with screen, it is straightforward to estimate S_N given the two variables. However, the number of screening rounds before detecting a cancer is influenced by many factors such as the natural course of disease progression, the performance of the screening tool (sensitivity and specificity), and the characteristics of the target population such as age.[12–14,17,18] In addition to temporal natural history of breast cancer progression, the estimation of sensitivity would further complicate the estimation of S_N as the higher the sensitivity the fewer the rounds of screen are required.

Suppose the number of screening rounds needed to detect disease (denoted by n) in a regular service screening program follows the geometric distribution and the corresponding probability mass function is expressed by $(1-p)^{n-1}\times\ p$ given the probability (denoted by p) of developing asymptomatic disease (staying in the PCDP) that can be detected by a specific screening tool. Note that the complementary probability $(1-p)$ represents screenees free of disease. The expected value of S_N can be derived in a canonical way. However, the direct application of the geometric model or related series is not straightforward like two-state outcome data as the empirical data on population-based cancer screening are characterized by correlated and multistate outcome of disease status, left-, right-, and interval-censoring and truncation due to the restriction of study period and logistics of follow-up, and imperfect measurements of screening tool. To accommodate these thorny issues while answering the question and dealing with the random variable related to geometric and negative binomial distribution, we developed a discrete-time stochastic process rather than a continuous-time stochastic process. Besides, the former can also dispense with the requirement of exact time to occurrence and diagnosis of different states of disease progressions of cancer. All of these motivate us to envisage a generalized stochastic process, focusing on negative-binomial-family-based stochastic model in this article, to deal with these complicated statistical properties. From the practical aspect, we demonstrated how the proposed negative-binomial-family-based stochastic model can be applied to the empirical data similar to the Finnish breast cancer screening program for women aged 50–59 years.

The rest of this article is organized as follows. Section 2 provides the rationale for how and why the empirical data similar to the Finnish breast cancer screening program can be used to demonstrate the application of the proposed negative-binomial-family-based stochastic model. Section 3 gives the details of model specification beginning with the notion of the backbone of generalized linear stochastic model (GLISM) for embracing various types of regression-based stochastic processes in section 3.1, then focusing on discrete-time and discrete-state Markov process with negative-binomial-family-based underpinning in section 3.2, giving the description of statistical properties of screening data in section 3.3, and developing the key formula for accommodating incomplete information in section 3.4 and measurement errors in sections 3.5 from which those were derived including the marginal probabilities in section 3.6 and logistic and negative-binomial-family-based link function for assessing the effect of covariates allowing for over-dispersion in sections 3.7, the formulation of likelihood function in section 4, and the number of screening required for decision-making in section 5. Section 6 delineates Bayesian Markov Chain Monte Carlo method for estimation of parameters. Section 7 shows the estimated results by fitting a series of negative-binomial-family-based models. Section 8 gives the discussion of clinical applications and implications using the proposed negative-binomial-family-based stochastic process.

2 Empirical data

The Finnish nationwide biennial mammographic screening program was implemented and targeted to women aged 50–59 years since 1988.[19] The biennial inter-screening interval yielded multi-state outcomes identified in each discrete-time interval (including free of cancer, screen-detected and clinically-detected cancer (consisting of interval cancer and cancers from non-participants), namely detection modes defined in the theory of screening). More importantly, as age range of screening was between 50 and 59 years, the follow-up time was restricted, and the accuracy of mammography was not perfect, these unique characteristics provide us an opportunity to deal with left (the entry of age from 50 years) and right truncation (the termination of screening at 60 years of age), censoring, and measurement errors mentioned above. In addition to these inherent statistical properties, it was a good empirical data for demonstrating the application of our proposed negative-binomial-family-based stochastic model as the process indicators of the Finnish breast cancer screening program (such as compliance rate, recall

Table 1. The observed frequencies of invasive breast cancer by detection mode and number of screening round at Pirkanmaa center, Finland 1988–2000.

	Attenders				Non-attenders	
Round	Free of breast cancer	Screen-detected breast cancer	Interval cancer	Breast cancer diagnosed after follow-up	Free of breast cancer	Breast cancer diagnosed after follow-up
1	29,215	128	–	328	3,604	56
2	24,483	49	46	304	2,939	28
3	17,828	65	29	222	1,960	34
4	11,438	32	27	125	1,291	14
5	6,841	15	18	63	691	7

rate, and detection rate) are complete and comprehensive and the assessment of the efficacy and the performance of mammographic examination showed an optimal performance of the program.[2,3,19] More importantly, information collected from this screening program was kept in the form of individual data rather than aggregate one. Data used for the following analysis are of women aged 50–59 years invited by Pirkanmaa screening center but excluding women who had been diagnosed by breast cancer due to the presence of clinical symptoms and signs, which are often defined as clinically detected breast cancer before their first invitation. This raises the first statistical issue on the truncation of unobservable data (left truncation).

Regarding the screening data, the program collected complete recorded information on screening history, screen-detected cancer at each round of screen, and the interval cancer obtained by linking mass screening registry data with nationwide cancer registry. Details on the design, process indicators, and evaluation of the efficacy of the screening program have been published in full elsewhere.[2,3,19] Table 1 shows the frequencies of invasive breast cancer from 1988 to 2000 at each screening round by detection mode for women who attended regular screening until the end of follow-up and non-participants. For attenders, frequencies of screen-detected breast cancer and interval cancer (cancer diagnosed between two screens) due to the presence of clinical symptoms and signs are listed by each round of screen. While women attended more than five times of screens, they were treated as censored cases because they were observed as the status free of breast cancer although they may be missed at screen, which would be handled by the consideration of measurement errors in the following section 3.5.

Less than five times of screening round would be expected for some women who were free of breast cancer because of the following reasons. First, women who were invited to first screen when older than 51 years and were free of cancer before 60 years. Second, there were a fraction of women who had entered the screening program after 1992 but the follow-up ended in 2000. Third, some participants were invited but attended the repeated screen with irregular interval. There were also some women moving to other areas or dying before 60. Table 1 also lists the frequencies of women who refused to attend screen and identified as breast cancer or censoring cases during the follow-up times. All of the statistical issues on censoring and truncation will be handled and delineated in the following section 3.4.

3 Model specification

3.1 GLISM

A stochastic process for depicting the state of a subject at time t is denoted by a random variable, say X_t, for which its possible outcomes lie within a state space denoted by Ω. Following the concept of generalized linear model (GLM), the effect of relevant covariates on the transition between states can be modeled by using an appropriate link function. The expected value for the transition from state g at time t_{m-1} to state h at time t_m is denoted as $E(X_{t_m} = h | X_{t_{m-1}} = g)$, denoted by $\mu(g \to h$ between t_{m-1} and $t_m)$ with shorthand notation represented by $\mu_{gh}(t_m)$. The expected value, $\mu_{igh}(t_m)$, is linked through appropriate link function denoted by $g(\mu_{igh}(t_m))$ with the relevant covariates x_{ikgh} (the k_{th} covariate) given individual i. For individual i with the outcome of transition from g to h given time t_m, the generalized equation to model the effect of covariates is expressed as follows

$$g\left(\mu_{igh}(t_m)\right) = \sum_{k=1}^{p} \beta_{kgh} x_{ikgh}(t_m) + \sum_{k=p+1}^{q} \beta_{kgh} x_{ikgh} \tag{1}$$

incorporating p time-dependent covariates and q time-constant covariates in the model. Here, $\mu_{igh}(t_m)$ is a vector of quantities related to transition probabilities or rates depending on the type of stochastic processes with the combination of time and state following Cox and Miller[20] proposition. As a matter of fact, this is an extension of univariate to multivariate GLM with Markov process underpinning. Theoretically, equation (1) can be used to accommodate different distributions of multistate outcome with various types of link function. In analogy to the framework of GLM that can be flexibly applied to various types of random components (such as discrete and also continuous-dependent variable), the multistate outcomes defined within the state space in the proposed stochastic process can also be classified into a variety of multivariate random components, such as discrete-state and discrete-time (Markov chain), discrete-state and continuous time (Markov process), continuous sate and discrete-time (simple random walk model), continuous state and continuous time (Brownian motion model). The details of each type of stochastic process refer to Cox and Miller[20] and Chiang.[21] While these stochastic processes have been well established and applied in different areas, the unifying framework for using the concept of GLM to link multistate outcomes of various types with covariates of interest has been barely addressed in literatures. The idea is that we regard the multistate events defined in the state space of stochastic processes as the multivariate outcome and deal with the residual variation after the classification of states with the incorporation of measured covariates of interest. This notion is embodied by the statistical framework called a GLISM.

3.2 Negative-binomial-family-based discrete-state and discrete-time stochastic model

Although the proposed GLISM model can be applied to different fields, we here only lay emphasis on the application of discrete-time and discrete-state stochastic process here due to the limited space and the practical questions raised by population-based cancer screening data as mentioned above.

The notations and basic discrete-time and discrete-state stochastic process applied to the following population-based periodic cancer screening data are defined as follows. In a screening program, state space can be classified into free of breast cancer (state 0), detected by screening before surfacing to clinical phase (CP) (preclinical detectable phase, PCDP (state 1)), and breast cancer developing symptom and signs and surfacing to CP (state 2). Thus, the state space can be written as $\Omega = \{0, 1, 2\}$. An attendee is eligible for periodically attending the screening program until she is detected as a case of breast cancer at screen or turning into clinical case between screens. Thus, the evolution of the state beginning from state 0, through state 1, and absorbed to state 2 can be delineated by using a stochastic process.

Let $X_{(t_m, t_{m-1})}$ (where m is the mth screen) denote the state of a subject detected at t_m $(m = 1, \ldots, r)$ or between rounds of screen (i.e. t_{m-1} and t_m) with the maximum round of screen equal to r. The joint probability, $\Pr(X_{t_1}, X_{t_2}, \ldots, X_{t_{m-1}}, X_{t_m}, \ldots, X_{t_r})$, given an individual sample path can be decomposed into a series of conditional probabilities and the initial probability by using the Markov property shown as follows

$$\Pr(X_{t_1}, X_{t_2}, \ldots, X_{t_{m-1}}, X_{t_m}, \ldots, X_{t_r}) = \Pr(X_{t_r} | X_{t_{r-1}}) \ldots \Pr(X_{t_m} | X_{t_{m-1}}) \ldots \Pr(X_{t_1}) \tag{2}$$

The transition probabilities that generate the sequence of the stochastic process can thus be expressed by

$$\mathbf{P} = \begin{array}{c} \\ 0 \\ 1 \\ \vdots \\ s \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 1 & \cdots & s \end{array} \\ \begin{bmatrix} P_{00} & P_{01} & \cdots & P_{0s} \\ P_{10} & P_{11} & \cdots & P_{1s} \\ \vdots & \vdots & \ddots & \vdots \\ P_{s0} & P_{s1} & \cdots & P_{ss} \end{bmatrix} \end{array} \tag{3}$$

where $P_{gh} = \Pr(X_{t_m} = h | X_{t_{m-1}} = g)$.

To delineate population-based periodic screening data with the transition probabilities contained in equation (3) under the context of GLISM, the multinomial distribution was therefore assigned as the random component. Let a random variable, $X_{igh}(t_m) \sim multinomial\ [P_{ig0}, P_{ig1}, \ldots, P_{igs}]$ given an individual i denote the number of transitions starting from g $(g = 0, \ldots, s)$ to h $(h = 0, \ldots, s)$ following the matrix of transition

probabilities as above, equation (1) using a log link function can be expressed as follows

$$\log\left(\frac{P_{igh}}{P_{ig0}}\right) = \sum_{k=1}^{p} \beta_{kgh} x_{ikgh}(t_m) + \sum_{k=p+1}^{q} \beta_{kgh} x_{ikgh} \tag{4}$$

We then linked the multinomial model with the geometric model by the application of first passage time to give the quantity of the required number for first detecting the event of interest, namely the first transition from state g in initial time to state h captured by a random variable T_{gh}. We then developed generalized negative binomial model and its power-based family series models applied to periodical screening data later.

Following the transition probabilities from equation (2) together with T_{gh}, we have

$$\Pr\left(T_{gh} = m\right) = \Pr\left(X_{t_m} = h; X_{t_n} \neq h, n = 1, \ldots, t_{m-1} | X_0 = g\right) \tag{5}$$

This is the first passage time defined in the conventional Markov process. Note that g and h in the breast cancer screening data lie within $\Omega = \{0,1,2\}$ defined as above. Let P_{01} denote the probability of being detected as screen-detected cancer at each screening round with the transition from state 0 to 1 and P_{02} denote the probability of surfacing to interval cancer by time since last negative screen with transition from state 0 to 2. The probability of being free of cancer at each screening round is $P_{00} = (1 - P_{01} - P_{02})$. The detailed algorithm using spectral decomposition of deriving the expected value of first passage time given the three-state Markov chain is delineated in Appendix 1. The random variable, X_{i0h}, $h \in \{0, 1, 2\}$, representing the transition to these possible destinations for normal subject i is captured by multinomial distribution, $\sim multi[P_{i00}, P_{i01}, P_{i02}]$. Let T_{01} be the number of screening rounds needed to detect the preclinical cancer and T_{02} be that of developing interval cancer. In our three state Markov chain, the variable of interest is thus the first transition from normal to PCDP (T_{01}) and to CP (T_{02}). The relationship between the multinomial parameters given an attendee starting from normal state ($g = 0$) is derived as follows

$$\begin{aligned} f_{01}^{(m)} &= \Pr(T_{01} = m) = \Pr\left(X_{t_m} = 1; X_{t_n} \neq 1, n = 0, \ldots, t_{m-1} | X_0 = 0\right) \\ &= P_{00}^{m-1} P_{01} \end{aligned} \tag{6}$$

and

$$\begin{aligned} f_{02}^{(m)} &= \Pr(T_{02} = m) = \Pr\left(X_{t_m} = 2; X_{t_n} \neq 2, n = 0, \ldots, t_{m-1} | X_0 = 0\right) \\ &= P_{00}^{m-1} P_{02} \end{aligned} \tag{7}$$

assuming the screening procedure with 100% accuracy which will be relaxed later by the introduction of measurement errors parameters.

Following the notation mentioned above, the random component of the specified transition with the parameters of the transition probabilities encoded in equation (3) can be expressed as the expected value defined in equations (6) and (7) to describe the observed panel data on a total of r rounds of screen. However, in order to be explicit to relate the transition probabilities to T_{gh}, we still retain the transition probabilities that are regarded as the expected values of multivariate outcome. Therefore, the probability of the number of screening rounds needed to detect preclinical cancer equals to z ($1 \leq z \leq r$) is thus

$$\Pr(T_{01} = z) = P_{00}^{z-1} \times P_{01} \tag{8}$$

and that of developing interval caner is

$$\Pr(T_{02} = z) = P_{00}^{z-1} \times P_{02} \tag{9}$$

3.3 Property of screening data

Suppose we have five ($r = 5$) rounds of screen, Figure 1 gives an illustration of empirical screening data assuming 100% sensitivity and specificity (this assumption will be relaxed later). Subject 2 was detected as breast cancer in

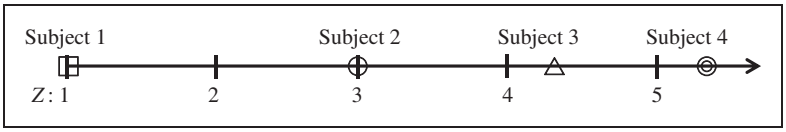

Figure 1. An illustration of hypothetical cases from population-based periodical screening data.
Subject 1 (□): detected as PCDP at first round of screening (left truncation).
Subject 2 (○) : detected as PCDP at third round of screening (subsequent screen-detected cancer).
Subject 3 (△): surfacing to CP between fourth and fifth round of screening (interval cancer).
Subject 4 (◎): not detected as PCDP and also CP till the end of screening program (right censoring).

the PCDP at the third round. The state transition was from 0 to 1 between second and third rounds of screen. The likelihood function contributed from this subject follows equation (8) with the GLISM with z equal to 3. Subject 3 was diagnosed as interval cancer (breast cancer surfacing to the CP. The state transition was from 0 to 2 between 4th and 5th round of screen. The likelihood function contributed from this subject follows equation (9) with z equal to 5.

3.4 Probabilities for right censoring and left truncation

There are two unique characteristics of empirical screening data as indicated before that require the modification of the two transition probabilities, including right censored cases (subject 4) and left truncation (subject 1).

3.4.1 Right censoring

Subjects attending screening program may remain in the state of free of breast cancer (state 0) till the end of observation resulting in the censored observation (see Table 1). To take censored case (subject 4, not detected as screen-detected and interval cancers) into account, the probability of the number of screening rounds to be detected cancer larger than r can be expressed by the following formula

$$\Pr(T_{01} > r, T_{02} > r) = P_{00}^{r} \times P_{01} \times [1 + P_{00} + P_{00}^{2} + \cdots] + P_{00}^{r} \times P_{02} \times [1 + P_{00} + P_{00}^{2} + \cdots] = P_{00}^{r} \tag{10}$$

3.4.2 Left truncation

Women who had developed cancer before the first screen were not eligible for attending screening program. Therefore, the probability of observing a subject already surfacing to the CP in the first screen round (prevalent screen, $Pr(T_{02} = 1)$) should be zero (truncated) for the observed empirical data. Initial distribution for such a Markov process would be changed from multinomial distribution to binomial distribution with $Bin(n, P^*)$ with P^* written as follows

$$P^* = \Pr(T_{01} = 1) = P_{01}/(1 - P_{02}) \tag{11}$$

which is the likelihood contribution for subject 1 in Figure 1.

The likelihood contribution for each type of subject being detected as PCDP of CP at different screening round could be similarly derived using equations (8) to (11) based on the generalized geometric model with Markov process underpinning. According to the hypothetical data shown in Table 2, we developed a geometric model to describe the panel data derived from organized service screening with regular interval. The theoretical probability distribution for the hypothetical panel data in the geometric form is listed in Table 2.

3.5 Incorporating measurement errors

Since the accuracy of screening program is often not 100% in reality, we also considered the measurement errors in our geometric model. Note that when the accuracy of screening test is not 100%, the empirical data in Table 1 should not correspond to the geometric distribution because the number of screening rounds of becoming screen-detected cancer and the number of screening round of becoming interval cancer are related to the screening results

Table 2. The geometric likelihood function for the hypothetical screening data.

	Conditional probability[a]	
Number of screen rounds	Screen-detected cancer (PCDP)	Interval cancer (CP)
1	$P_{01}/(1-P_{02})$	0
2	$P_{00}P_{01}/(1-P_{02})$	$P_{00}P_{02}/(1-P_{02})$
3	$P_{00}^2P_{01}/(1-P_{02})$	$P_{00}^2P_{02}/(1-P_{02})$
4	$P_{00}^3P_{01}/(1-P_{02})$	$P_{00}^3P_{02}/(1-P_{02})$
⋮	⋮	⋮
r	$P_{00}^{r-1}P_{01}/(1-P_{02})$	$P_{00}^{r-1}P_{02}/(1-P_{02})$
$>r$	$P_{00}^{r}/(1-P_{02})$	

[a]Conditional on the zero observation of interval cancer by the first screening round.

Table 3. The probabilities of observing a subject as preclinical cancer (O_{01}) and clinical one (O_{02}) at zth screening round conditional on the subject developing preclinical cancer (T_{01}) and clinical cancer (T_{02}) at nth screening round.

	$\Pr(O_{01}=z \mid T_{01}=n)$	$Pr(O_{02}=z \mid T_{01}=n)$	$Pr(O_{01}=z \mid T_{02}=n)$	$Pr(O_{02}=z \mid T_{02}=n)$
$z<n$	$(1-\alpha)^{z-1}\alpha$	0	$(1-\alpha)^{z-1}\alpha$	0
$z=n$	$(1-\alpha)^{n-1}(1-\beta)$	0	0	$(1-\alpha)^{n-1}$
$n+1 \le z \le r$	$(1-\alpha)^{n-1}\beta^{z-n}(1-P_{12})^{z-n}(1-\beta)$	$(1-\alpha)^{n-1}\beta^{z-n}(1-P_{12})^{z-n-1}\,P_{12}$	0	0
$\Pr(O_{01}>r,\, O_{02}>r \mid T_{01}=n)$	$(1-\alpha)^{n-1}\beta^{r-n+1}(1-P_{12})^{r-n}$		–	
$\Pr(O_{01}>r,\, O_{02}>r \mid T_{02}=n)$	–		0	

of previous screening rounds and are influenced by the sensitivity and specificity of screening program. However, the true number of screening rounds to develop cancer, either preclinical or clinical, is a geometric distribution.

Following notation defined above, let T_{01} be the random variable of the true number of screening rounds to develop preclinical cancer and T_{02} be that of the true number of screening rounds to develop clinical cancer. Denote the random variable of observed number of screening rounds to develop preclinical cancer by O_{01} and that of the observed number of screening rounds to develop clinical cancer by O_{02}. There are two types of measurement errors related to the quality of population-based screening tool rather than the accuracy of diagnostic tool:

(1) the attendee is detected as having breast cancer (screen positive), while the underlying status is free of breast cancer (false positive); and
(2) the attendee is detected as free of breast cancer (screen negative), while the underlying status is having disease but still in the preclinical phase (false negative).

Let α be the probability of false positive of the screening result when the time dimension of measurement error is taken into account. β be the probability of false negative of the screening result. Accordingly, $(1-\alpha)$ is the specificity and $(1-\beta)$ is the sensitivity of the screening program. Since there are measurement error involved in the screening process, the probabilities of having such an observation given the underlying status are expressed in terms of transition probabilities, α and β. The probabilities of observing O_{01} and O_{02} given T_{01} and T_{02} are summarized in Table 3 and elaborated as follows.

3.5.1 Probability of detecting the subject with breast cancer at sth screening round given underlying disease progression to breast cancer at nth round (T_{01}=n)

The relationship between the true number of screening round and the observed number of screening round for the subject detected as a cancer case can be established as follows. When $T_{01}=n$ $(1 \le n \le r)$, namely a woman has already entered into the preclinical phase of breast cancer at nth screening round, the probability of observing breast cancer of her at different screening rounds can be written as $\Pr(O_{01}=z|T_{01}=n)$ when identified as a screen-detected cancer at zth screening round and $\Pr(O_{02}=z|T_{01}=n)$ found as interval cancer by

zth screening round. $\Pr(O_{01} > r, O_{02} > r|T_{01} = n)$ denotes the probability of observed number of screening round to develop cancer longer than r.

For $n = z$, the observed status of the subject is the same as the true disease status and the probability of having such an observation given the true disease state is thus

$$\Pr(O_{01} = z|T_{01} = n) = (1-\alpha)^{n-1}(1-\beta), \quad z = n \tag{12}$$

For $z < n$, the screening process would correctly categorize attendee for $z-1$ rounds but yield a false positive result at the zth round and thus the conditional probability is

$$\Pr(O_{01} = z|T_{01} = n) = (1-\alpha)^{z-1}\alpha, \quad z < n. \tag{13}$$

For $n+1 \le z \le r$, the screening process would correctly classify the subject free of disease for the first $n-1$ rounds but fail to detect the preclinical disease following $z-n$ screening rounds until the zth screening round. Although the subject has entered the preclinical phase at the nth screening round she may be missed by the screening process, and would not surface to clinical disease during the following $z-n$ rounds of screening. The conditional probability is thus written as

$$\Pr(O_{01} = z|T_{01} = n) = (1-\alpha)^{n-1}\beta^{z-n}(1-P_{12})^{z-n}(1-\beta), \quad n+1 \le z \le r \tag{14}$$

where P_{12} is the probability of developing clinical cancer at each round when a woman has already entered the preclinical phase of breast cancer.

The conditional probability $\Pr(O_{02} = z|T_{01} = n)$ is zero for $z < n$ and $z = n$ due to the progressive assumption that the development of clinical disease must pass through the preclinical phase.

For $n+1 \le z \le r$, the conditional probability $\Pr(O_{02} = z|T_{01} = n)$ can be derived as follows

$$\Pr(O_{02} = z|T_{01} = n) = (1-\alpha)^{n-1}\beta^{z-n}(1-P_{12})^{z-n-1}P_{12}, \quad n+1 \le z \le r \tag{15}$$

representing the screening process of true negative for $n-1$ rounds followed by false negative for $z-n$ rounds after the subjects has already entered into the preclinical phase of breast cancer but without surfacing to CP until the zth round.

The conditional probability of having censored observations can be derived as follows

$$\Pr(O_{01} > r, O_{02} > r|T_{01} = n) = (1-\alpha)^{n-1}\beta^{r-n+1}(1-P_{12})^{r-n} \tag{16}$$

Representing screening process of true negative for $n-1$ rounds followed by false negative results while she has already entered into the preclinical disease of breast cancer till the end of follow-up for $r-n+1$ rounds without surfacing to clinical disease for remaining $r-n$ rounds.

3.5.2 *Probability of detecting the subject with breast cancer at zth screening round given underlying disease progression to clinical breast cancer at nth round ($T_{02} = n$)*

The probabilities of O_{01} and O_{02} conditional on T_{02} are derived in a similar manner. The conditional probability $\Pr(O_{01} = z|T_{02} = n)$ equals zero for $z = n$ and $n+1 \le z \le r$ due to progressive assumption as indicated above.

The conditional probability for $z < n$ can be written as follows

$$Pr(O_{01} = z|T_{02} = n) = (1-\alpha)^{z-1}\alpha, \quad z < n \tag{17}$$

due to the true negative for $z-1$ rounds but detecting the subject as cancer case at zth round while she has not surfaced to clinical disease yet.

The zero probabilities of $\Pr(O_{01} > r, O_{02} > r\,|T_{02} = n)$ and $\Pr(O_{02} = s|T_{02} = n)$ when $z < n$ and $n+1 \le z \le r$ result from the same argument of biological plausibility.

For $z = n$, the conditional probability can be derived as follows

$$Pr(O_{02} = z|T_{02} = n) = (1-\alpha)^{n-1}, \quad z = n \tag{18}$$

reflecting the true negative for the $n-1$ rounds before the subject surfacing to clinical disease.

3.6 Marginal probability of being observed incorporating measurement errors

According to the conditional probabilities (12) to (18) established above to link the true number of screening rounds of developing cancer and the observed number of screening rounds of developing cancer, we can derive the probability of observed number of screening round to develop preclinical cancer ($\Pr(O_{01}=z)$, $1 \le z \le r$), the probability of observed number of screening round to develop clinical cancer ($\Pr(O_{02}=z)$, $1 \le z \le r$), and the probability of observed number of screening round to develop cancer larger than r ($\Pr(O_{01}>r, O_{02}>r)$) given that $\Pr(T_{01}=z)$ and $\Pr(T_{02}=z)$ are still based on geometric distribution.

$$\begin{aligned}\Pr(O_{01}=z) = &\sum_{n=1}^{r}[\Pr(T_{01}=n)\times\Pr(O_{01}=z|T_{01}=n)] + \sum_{n=1}^{r}[\Pr(T_{02}=n)\times\Pr(O_{01}=z|T_{02}=n)] \\ &+ \Pr(T_{01}>r, T_{02}>r)\times\Pr(O_{01}=z|T_{01}>r, T_{02}>r)\end{aligned} \tag{19}$$

$$\begin{aligned}\Pr(O_{02}=z) = &\sum_{n=1}^{r}[\Pr(T_{01}=n)\times\Pr(O_{02}=z|T_{01}=n)] + \sum_{n=1}^{r}[\Pr(T_{02}=n)\times\Pr(O_{02}=z|T_{02}=n)] \\ &+ \Pr(T_{01}>r, T_{02}>r)\times\Pr(O_{02}=z|T_{01}>r, T_{02}>r)\end{aligned} \tag{20}$$

and

$$\begin{aligned}\Pr(O_{01}>r, O_{02}>r) = &\sum_{n=1}^{r}[\Pr(T_{01}=n)\times\Pr(O_{01}>r, O_{02}>r|T_{01}=n)] \\ &+ \sum_{n=1}^{r}[\Pr(T_{02}=n)\times\Pr(O_{01}>r, O_{02}>r|T_{02}=n)] \\ &+ \Pr(T_{01}>r, T_{02}>r)\times\Pr(O_{01}>r, O_{02}>r|T_{01}>r, T_{02}>r)\end{aligned} \tag{21}$$

The characteristics of screening data are retained and the binomial initial distribution as in equation (11) was adopted in the form of $P^{*}=\Pr(O_{01}=1)=P_{01}/(1-P_{02})$.

3.7 Logistic regression model and negative binomial regression models

3.7.1 Geometric regression model

According to previous studies, the probability of disease progression and sensitivity may vary between age groups[3,14,18,22] and may differ between prevalent and subsequent screen.[23] The proposed model can be extended to incorporate relevant covariates based on the framework of GLISM expressed in equation (1). We thus incorporated the effect of covariate into the model by using a traditional logistic link to assess the effect of age groups of the attendants on two transition probabilities, one from free of cancer to preclinical cancer (P_{01}) and the other from preclinical caner to clinical cancer (P_{12}) as follows

$$\begin{aligned}P_{i,01}(X_{i,age}, X_i) &= \frac{\exp(\alpha_{01}+\beta_{01}X_{i,age}+\beta_{02}X_i)}{1+\exp(\alpha_{01}+\beta_{01}X_{i,age}+\beta_{02}X_i)} \\ P_{i,12}(X_{i,age}) &= \frac{\exp(\alpha_{11}+\beta_{11}X_{i,age})}{1+\exp(\alpha_{11}+\beta_{11}X_{i,age})}\end{aligned} \tag{22}$$

where $X_{i,age}$ indicates whether an attendee was 55 years and older when entering the program and X_i indicates the subsequent screens ($X_i=1$) or prevalent one ($X_i=0$).The logistic form used for the incorporation of the effect of age and calendar years into sensitivity is expressed by

$$Sensitivity_i(age_i, year_i) = (1-\beta_i(age_i, year_i)) = \frac{\exp(\alpha_{20}+\beta_{21}age_i+\beta_{22}year_i)}{1+\exp(\alpha_{20}+\beta_{21}age_i+\beta_{22}year_i)} \tag{23}$$

where β_{21} and β_{22} are the two regression coefficients corresponding to the covariate of the increment of age of the attendant and that of calendar year, respectively, on the sensitivity of screening program given individual i. We used the age of 55 years and the calendar year of 1994 as the reference groups and treated the two covariates as interval scale.

3.7.2 Negative binomial regression models

Modeling the observed sequence of stochastic process based on the geometric distribution may suffer from the over-dispersion, as studied extensively in previous works on count models.[24] This may be a particular concern when applying the proposed generalized model to data on cancer screening since we did not consider all factors accounting for the occurrence of PCDP and CP such as genetic factors and inherited histological factors of neoplastic tissues. Also, the linear and additive relationship between predictors and the transition probabilities may not hold. In addition, unobserved factors and correlations such as variation between areas associated with the occurrence of PCDP and CP may also lead to over-dispersion. The proposed approach can be extended to negative binomial model to cope with these concerns. As the geometric model is a special case of negative binomial model, the progression of disease mentioned above can be easily extended under the concept of with GLM.[24] We first extended the geometric model of equation (22) to the negative binomial regression with the parameterization of NB-2 by including the dispersion parameter r_1 and r_2 as follows

$$\begin{aligned} P_{i,01}(X_{i,age}, X_i) &= \frac{\gamma_1 \mu_{i,01}}{1 + \gamma_1 \mu_{i,01}} \\ P_{i,12}(X_{i,age}) &= \frac{\gamma_2 \mu_{i,12}}{1 + \gamma_2 \mu_{i,12}} \end{aligned} \tag{24}$$

where $\mu_{i,01} = \exp(\eta_{i,01})$ and $\mu_{i,12} = \exp(\eta_{i,12})$ which are in turn determined by $\eta_{i,01} = \alpha_{01} + \beta_{01} X_{i,age} + \beta_{02} X_i$ and $\eta_{i,12} = \alpha_{11} + \beta_{11} X_{i,age}$, respectively. The dispersion parameters, γ_1 and γ_2, of negative binomial distribution account for unobserved heterogeneity of transition probabilities between individuals.

We further extended the negative binomial regression expressed in equation (24) to the generalized negative binomial model (NB-P model) to allow for more flexibility of the variance function related to the power-based family series as follows

$$\begin{aligned} P_{i,01}(X_{i,age}, X_i) &= \frac{\mu_{i,01}}{\mu_{i,01} + \gamma_1 \mu_{i,01}^{Q_1}} \\ P_{i,12}(X_{i,age}) &= \frac{\mu_{i,12}}{\mu_{i,12} + \gamma_2 \mu_{i,12}^{Q_2}} \end{aligned} \tag{25}$$

In addition to the dispersion parameters, γ_1 and γ_2, the power parameters, Q_1 and Q_2, can account for the extra heterogeneity of transition probabilities based on the NB-P model.

4 Likelihood function

According to the geometric probability distribution mentioned above, the total likelihood of the screening data assuming 100% accuracy can be expressed as follows

$$\Pr(n_{z_1}, n_{z_2}, n_{r+1} | P_{01}, P_{12}) \propto [\Pr(T_{01} = z_1)]^{n_{z_1}} [\Pr(T_{02} = z_2)]^{n_{z_2}} [\Pr(T_{01} > r, T_{02} > r)]^{n_{r+1}} \tag{26}$$

representing a multivariate geometric distribution derived from multinomial distribution based on formulas (8) to (11). The total likelihood function is thus the product of the number of observed frequencies of PCDP and CP of z_1 and z_2 rounds ranged from 1,2,...r as follows

$$L = \prod_{z_1=1}^{r} \left\{ [\Pr(T_{01} = z_1)]^{n_{z_1}} \right\} \times \prod_{z_2=2}^{r} \left\{ [\Pr(T_{02} = z_2)]^{n_{z_2}} \right\} \times [\Pr(T_{01} > r, T_{02} > r)]^{n_{r+1}} \tag{27}$$

After incorporating measurement errors using marginal probabilities of observed screening rounds from equations (19) to (21), the total likelihood function for the observed screening data is thus expressed by the following form

$$L = \prod_{z_1=1}^{r} \left\{ [\Pr(O_{01} = z_1)]^{n_{z_1}} \right\} \times \prod_{z_2=2}^{r} \left\{ [\Pr(O_{02} = z_2)]^{n_{z_2}} \right\} \times [\Pr(O_{01} > r, O_{02} > r)]^{n_{r+1}} \tag{28}$$

Equations (27) and (28) were used to estimate the parameters contained in P dominating the realization of number of screening rounds to detect an attendee as PCDP or CP denoted by T_{01}, T_{02} and O_{01}, O_{02} with and without considering measurement errors for the proposed geometric model. Using the logistic link and the generalized equation of equation (3), P_{01} is determined by $\alpha_{01}, \beta_{01}, \text{and} \beta_{02}$, and P_{02} further determined by α_{11}, β_{11} according to equations (22) and (23). When we extended the model to the negative binomial family using NB-2 parameterization, the dispersion parameters, γ_1 and γ_2, together with α_{01}, β_{01}, β_{02} and α_{11}, β_{11} determine the transition probabilities P_{01} and P_{12}, respectively, as specified by equation (24). Power parameters, Q_1 and Q_2, were incorporated to increase the flexibility of transition probabilities when the NB-2 model was extended to NB-P model as specified by equation (25).

5 Number of screening rounds for decision-making

5.1 Expected number of screening rounds for detecting a cancer

According to the estimated probabilities of disease progression, the expected number of screening round required for detecting a cancer through screening activity can be computed by the following formula

$$\sum_{z=1}^{r} \frac{\Pr(T_{01} = z | T_{02} > 1) \times z}{\sum_{x=1}^{r} \Pr(T_{01} = x | T_{02} > 1)} \tag{29}$$

Formula (29) can be extended to derive the expected number of screening round required for detecting a cancer making allowance for measurement errors as follows

$$\sum_{z=1}^{r} \frac{\Pr(O_{01} = z | O_{02} > 1) \times z}{\sum_{x=1}^{r} \Pr(O_{01} = x | O_{02} > 1)} \tag{30}$$

5.2 Number of screening rounds required for stopping screening

The estimated parameters can be also applied to calculating number of screening rounds to stopping screening after a series of negative screening results, which is related to the risk of developing breast cancer for those attending multiple screening rounds. The risk can be calculated on the basis of the proposed geometric model displayed in Table 2 and the estimated results in Table 5 taking into account measurement errors. For subjects attending $r-1$ rounds of breast cancer screening with negative results, the probabilities of being detected as breast cancer case at the state of PCDP and CP at rth round are $P_{00}^{r-1} P_{01}$ and $P_{00}^{r-1} P_{02}$, respectively. Using first-step analysis and Markov property, to determine the number of subsequent screening rounds after prevalent screen (first screen) required for stopping screening, is equivalent to finding the chance of detecting cancer during screens and between or after screen given the number of subsequent screens to catch up the expected cumulative risk of the underlying age-specific population denoted by p_s which is often available from cancer registry or estimated from the previous study.[3] Following the concept of first-step analysis, the number of screening rounds required for stopping screening is thus derived by

$$1 + \frac{p_s}{P_{01} + P_{02}} \tag{31}$$

where 1 presents first screen and the second component represents the expected number of incident (subsequent) screens.

6 Bayesian Markov chain Monte Carlo method

Estimation of parameters with the application of the proposed model to empirical data would also involve the parameters of high dimension. Also, the incorporation of parameters of measurement errors into the model rendered the estimation procedure intractable. We therefore considered the use of Bayesian Markov chain Monte Carlo method as it takes advantages of deriving posterior marginal distribution by integrating out the rest of parameters in the model with a sampling algorithm such as the Gibbs sampler. We thus applied the full Bayesian approach in conjunction with Markov chain Monte Carlo method to data on breast cancer screening to derive the marginal posterior distributions of the parameters of interest. The Bayesian method underpinning the

Table 4. Estimated results of transition probabilities applying geometric model, NB-2 model, and NB-P model without incorporating measurement error.

	Geometric model		NB-2 model		NB-P model	
Parameter	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI
Transition probability						
P_{01}	0.004	(0.003, 0.005)	0.004	(0.003, 0.005)	0.004	(0.003, 0.005)
P_{12}	0.657	(0.520, 0.810)	0.661	(0.515, 0.833)	0.646	(0.504, 0.792)
β_{02}(Subsequent round)	−0.35	(−0.57, −0.12)	−0.35	(−0.58, −0.12)	−0.43	(−0.75, −0.14)
OR, subsequent round	0.71	(0.56, 0.88)	0.70	(0.56, 0.88)	0.65	(0.47, 0.87)
Dispersion parameter						
γ_1	–	–	0.707	(0.230, 1.641)	0.388	(0.140, 1.161)
γ_2	–	–	1.124	(0.331, 3.112)	2.886	(0.659, 6.819)
Power parameter						
Q_1	–	–	–	–	0.159	(0.007, 0.365)
Q_2	–	–	–	–	1.822	(0.070, 4.688)

framework of GLM to link the effect of relevant covariates in conjunction with the constructed likelihood function expressed in equations (27) and (28) entails the estimation of parameters based on the collected empirical data. The full conditional posterior distribution can be constructed by using the proposed likelihood function and the specified prior distributions. Non-informative priors with uniform distributions were assigned. The uniform distributions elicited for the prior distributions are summarized in Appendix Table S-1.

A Gibbs sampler was used to derive samples of a stationary posterior distribution by which inferences on the parameters of interest could be made. The initial values for parameters were assigned with reference to the results of maximum likelihood estimates (MLE). To minimize the correlation between iterations a thinning interval of 20 with a burn-in interval of 100,000 followed by 100,000 iterations were used. Such an MCMC simulation yielded a total of 5000 updated posterior samples. An illustration of trace plots of simulations for the parameters of NB-P model is given in Appendix Figure S-1.

The principle of the comparisons between models with and without the incorporation of measurement errors as well as the inclusion of covariates on transition probabilities and measurement errors is based on the value of deviance information criterion (DIC).

7 Results

7.1 Models with and without considering measurement errors

The estimated results based on geometric model, negative binomial model (NB-2), and the power-based negative binomial model (NB-P) without considering measurement error are listed in Table 4. The estimated probability of detecting PCDP during a screening round (P_{01}) and that of surfacing to CP from the PCDP (P_{12}) were similar for three models with the point estimate around 0.004 and 0.65, respectively, with similar precision on estimates. Compared with prevalent round of screen, subsequent rounds had lower chance of detecting PCDP with the order of OR estimated as 0.71 (95% credible interval (CI) 0.56–0.88) when applying the geometric model. The corresponding results based on the NB-2 model and NB-P model revealed the same direction but slightly lower estimates. For NB-2 mode, the dispersion parameters, γ_1 and γ_2, were estimated as 0.71 (95% CI: 0.23–1.64) and 1.12 (95% CI: 0.33–3.11), respectively, indicating that the geometric model was acceptable for the data. The estimates of dispersion parameters with the application of the NB-P model were 0.39 (95% CI: 0.14–1.16) and 2.89 (95% CI: 0.66–6.82), indicating lacking of statistical significance. The power parameters, Q_1 and Q_2, corresponding to the transition from free of breast cancer to PCDP and PCDP to CP, were estimated as 0.16 (95% CI: 0.007–0.365) and 1.82 (0.07–4.69), respectively.

The estimated results of the proposed models with the incorporation of measurement errors including the probability of being false negative (β) and false positive (α) expressed by sensitivity ($1-\beta$) and specificity ($1-\alpha$) are listed in Table 5. The estimated results for specificity were uniformly high (99.9%) in three models. Estimated results for sensitivity were around 83% ranging from 82.2% to 83.9%. The chance of detecting PCDP for subsequent screens was still lower than that for prevalent screen but the estimated ORs ranging from 0.56

Table 5. Estimated results of transition probabilities applying geometric model, NB-2 model, and NB-P model incorporating measurement error.

	Geometric model		NB-2 model		NB-P model	
Parameter	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI
Transition probability						
P_{01}	0.004	(0.003, 0.005)	0.004	(0.002, 0.006)	0.004	(0.003, 0.006)
P_{12}	0.656	(0.472, 0.839)	0.720	(0.413, 0.967)	0.612	(0.331, 0.895)
β_{02}(Subsequent round)	−0.48	(−0.83, −0.17)	−0.51	(−0.88, −0.19)	−0.58	(−0.95, −0.21)
OR, subsequent round	0.62	(0.44, 0.85)	0.60	(0.42, 0.83)	0.56	(0.39, 0.81)
Measurement error						
$1-\alpha$ (Specificity)	99.9	(99.8, 100)	99.9	(99.8, 100)	99.9	(99.8, 100)
$1-\beta$ (Sensitivity)	83.9	(62.1, 95.3)	83.0	(60.8, 95.3)	82.2	(55.3, 94.9)
Dispersion parameter						
γ_1	–	–	0.669	(0.156, 1.456)	0.357	(0.140, 1.055)
γ_2	–	–	2.181	(0.140, 6.298)	2.886	(0.484, 6.988)
Power parameter						
Q_1	–	–	–	–	0.165	(0.008, 0.376)
Q_2	–	–	–	–	1.814	(0.076, 4.668)

to 0.62 were lower compared with the corresponding figures without considering measurement errors (see Table 4). The estimated results of dispersion parameters, γ_1 and γ_2, for NB-2 and NB-P were similar to those without considering measurement errors and all of them were lacking of statistical significance.

7.2 Effects of covariates on disease progression and sensitivity

Using the proposed GLISM, we constructed a model allowing for the heterogeneity of disease progression among the subpopulations of different age groups and screening rounds with the consideration of the effect of advancing age and calendar year on sensitivity simultaneously. Table 6 shows the model incorporating both the heterogeneity on transition probabilities as well as the effect of covariates on sensitivity using the proposed GLISM. Compared with the young age group, attendees aged 55 years and older had a borderline lower risk of the occurrence of PCDP with the estimated OR ranging from 0.96 to 1 depending on different models, although the 95% CIs covered one. The old age group also had lower risk of progressing from PCDP to CP with the order of OR estimated as 0.67 in geometric model. The corresponding results using NB-2 and NB-P model were 0.73 (95% CI: 0.14–5.38) and 0.98 (95% CI: 0.15–6.42), respectively. Considering the effect of covariates on sensitivity, the advancing age and calendar year on sensitivity also showed positive associations with the order of ORs estimated around 1.0 to 1.03.

7.3 Model comparison

The comparisons between models based on DIC values are listed in Table 7. Considering models without incorporating measurement errors, a decrease in DIC in NB-2 model (975469.7) was 5.1 compared with the geometric model (975474.8), reflecting the improvement with the inclusion of dispersion parameters (γ_1 and γ_2). Although the NB-P model provides more flexible in the change of transition probabilities, there was lacking of significant difference of DIC values between NB-P (975467.0) and NB-2 models (975469.7). The incorporation of measurement errors resulted in modest improvement in three proposed models. After allowing for measurement errors including false positive and false negative in the model, the incorporation of covariates on transition probabilities and measurement errors did not further reduce the DIC values.

7.4 Expected number of screening rounds required for detecting a breast cancer

Applying the proposed GLSM based on geometric distribution and negative binomial distribution to data on breast cancer screening, the average number of screening rounds for an attendee to being detected as PCDP

Table 6. Incorporating the effect of covariates on both transition probabilities and sensitivity based on geometric model, NB-2 model, and NB-P model.

	Geometric model	NB-2 model			NBP model	
Parameter	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI
Transition probability						
P_{01}	0.004	(0.003, 0.005)	0.004	(0.002, 0.005)	0.004	(0.002, 0.005)
P_{12}	0.694	(0.520, 0.813)	0.794	(0.481, 0.957)	0.697	(0.442, 0.918)
β_{02}(Subsequent round)	−0.47	(−0.82, −0.16)	−0.51	(−0.87, −0.15)	−0.58	(−0.97, −0.22)
β_{01}(age group, P_{01})	0.00	(−0.41, 0.39)	−0.04	(−0.45, 0.35)	−0.04	(−0.44, 0.39)
β_{11}(age group, P_{12})	−0.40	(−1.85, 1.77)	−0.31	(−2.00, 1.68)	−0.02	(−1.90, 1.86)
OR, subsequent round	0.63	(0.44, 0.86)	0.60	(0.42, 0.86)	0.56	(0.38, 0.81)
OR, age group (P_{01})	1.00	(0.66, 1.48)	0.96	(0.64, 1.42)	0.96	(0.64, 1.48)
OR, age group (P_{12})	0.67	(0.16, 5.86)	0.73	(0.14, 5.38)	0.98	(0.15, 6.42)
Measurement error						
$1-\alpha$(Specificity)	99.9	(99.8, 100)	99.9	(99.8, 100)	99.9	(99.8, 100)
$1-\beta$(Sensitivity)	88.2	(68.9, 95.3)	87.7	(67.9, 95.3)	88.1	(69.4, 95.3)
β_{21}(age)	0.00	(−0.36, 0.29)	0.01	(−0.37, 0.32)	0.01	(−0.35, 0.30)
β_{22}(year)	0.03	(−0.27, 0.31)	0.02	(−0.29, 0.30)	0.02	(−0.27, 0.29)
OR, age	1.00	(0.70, 1.34)	1.01	(0.69, 1.38)	1.01	(0.70, 1.35)
OR, year	1.03	(0.76, 1.36)	1.02	(0.75, 1.35)	1.02	(0.76, 1.34)
Dispersion parameter						
γ_1	–	–	0.600	(0.180, 1.482)	0.345	(0.139, 1.056)
γ_2	–	–	2.821	(0.331, 6.984)	3.252	(0.591, 7.094)
Power parameter						
Q_1	–	–	–	–	0.170	(0.009, 0.369)
Q_2	–	–	–	–	1.274	(0.045, 3.911)

Table 7. DIC value of models.

Model	Geometric model	NB-2 model	NB-P model
Models without incorporating covariates			
Without considering measurement error	975474.8	975469.7	975467.0
Incorporating measurement error	975469.6	975466.7	975462.4
Models incorporating covariates			
Effect of age on transition probabilities	975472.1	975470.4	975468.1
Effects of age and year on sensitivity	975474.2	975472.7	975464.9
Effect of covariates on transition probabilities and sensitivity	975476.0	975475.4	975471.7

(underlying disease progression) and that of being detected by screening program considering measurement errors (observed number of screening) given a total of five rounds of screening can be derived using equations (29) and (30), respectively. The average number of screening round required for detecting PCDP was around 2.77. Since the screening tool was not perfect, the average number of screening round required for detecting PCDP ranged from 2.79 to 2.81.

After taking the effect of age and calendar year into account, the number of screening rounds required to detect an attendee as PCDP was from 2.81 to 2.83 for women younger than 55 years and from 2.76 to 2.79 for the old age group with the sensitivity ranging from 87% to 88% (see Table 6). Based on the results of NB-P model and assuming 80% sensitivity, the average number of screening rounds required for detecting an asymptomatic case increased from 2.79 to 2.81 for old age group. When the sensitivity is improved to 95%, the

average number of screens is reduced to 2.78 based on the NB-P model. Detailed results are listed in Appendix Table S-2.

7.5 Number of screening rounds required for stopping screening

Following equation (31), suppose a screening program is offered for women from 50 years until 60 years of age. The expected cumulative risk, p_s, was 0.025 according to Finnish cancer registry or the previous study.[3] The probability of detecting breast cancer during screening rounds and that of developing interval cancer taking measurement errors into account can be calculated with equation (31) on the basis of parameters derived from Table 5. The number of screening rounds required for stopping screening was around eight rounds, three additional screens after five rounds of routine screens with two years of inter-screening interval until 70 years of age.

8 Discussion

There are numerous similar previous works using the Markov process (discrete-state and continuous time) applied to population-based screening data [3,11,14,15,17,25,26] but very few studies have enlarged on how the GLM model with multivariate outcome can be adapted to such a kind of Markov chain. The periodic population-based screening is one type of classical example. Here, we extend the GLM to GLSM, particularly with emphasis on negative binomial GLM, to study the number of screens required before detecting the screen-detected case but not surfacing as interval cancer (multistate outcome) and number of screening rounds required for stopping screening for low risk subjects. The advantage of using the negative binomial model is to render the question like "how many rounds of screen are needed to detect an asymptomatic breast cancer?" "how many rounds of screen are needed to stop screening after a series of negative screen results?" become explicit. Under the generalized model, it is very straightforward to model the effect of relevant covariates on the average rounds of screen required through link function in order to answer a specific question of whether and how this estimate varied with age, for example. Hence, different age groups may require different rounds of screens before detecting an asymptomatic breast cancer. In our case, a 57-year-old woman (old age group) may require 2.79 rounds of screen before detecting asymptomatic breast cancer but not surfacing to symptomatic breast cancer based on the estimates of NB-P model (Table S-2).

The incorporation of measurement errors renders such a stochastic process powerful for predicting the previous questions with different performances of screening tools as listed in Appendix Table S-2. The proposed GLISM also provides an approach to model the effect of inter-screen interval with the consideration of sensitivity given different levels of risk based on the number of interval cancer as elaborated in Appendix 2 and Table S-3.

It should be noted that sensitivity and specificity here are not like the accuracy of diagnosis of disease as used in clinical science. Although the accuracy of non-disease for positive results of mammography has been ruled out by further diagnostic workup false positive rate here is not for the accuracy of diagnosis but is related to the estimate of positive predictive value (PPV) of mammography screening in the domain of quality assurance of population-based screening. Similarly, the sensitivity here is also not for the accuracy of diagnosis but is used for assessing the quality of population-based screening program that is related to negative screenees (too many to be confirmed with diagnostic workup immediately after screen) but missed at screen and discovered later on due to the presence of symptoms and signs (called interval cancers). The links of false negative and false positive rate and PPV with population-based cancer screening have been elaborated in the previous studies.[13] It should be also noted that as both measurement errors are related to population and also time to follow-up, depending on mean sojourn time, they could not be directly observed but could be possibly estimated by using a sophisticated statistical method with a discrete-time stochastic process as proposed here.

When the time dimension of measurement error is taken into account in the scenario of population-based screening, it is therefore better to estimate the joint effect of sensitivity and specificity of screening tool as proposed here because both are related to the mean sojourn time, the average duration of the PCDP, equivalent to our estimate of P_{12} as demonstrated in the previous study.[17] We could estimate it separately by using a two-stage procedure based on the sensitivity, the mean sojourn time (derived from P_{12}), and preclinical incidence rate (derived from P_{01}) against the observed prevalent case at first screen[13] estimated from the proposed model without considering false positive rate. However, doing so may lose chance of estimating the joint effect of

both estimates. To demonstrate this, we compared the proposed model with and without considering false positive rate and found those false positive cases without being captured by specificity would lead to a slightly higher estimate of the probability of developing breast cancer (P_{01}, 0.004 vs. 0.005), which, in turn, also affects the transition rate of P_{12} (0.694 vs. 0.596) and also sensitivity (88.2% versus 91.8%) although the majority of the effects of covariates still remained robust. The DIC value for the model with false positive rate was lower than that without considering it by 4.1, which was statistically significant ($p < 0.05$). From the viewpoint of statistical technique, the proposed model here may be more sophisticated but flexible as the model without considering the false positive rate parameter can be easily accommodated by specifying the parameters of false positive rate (α) as zero corresponding to a specificity of 100%.

In addition to the estimation of joint effect of sensitivity and specificity, the proposed model with the incorporation of false positive rate may also capture overdiagnosis, an extreme case of false positive case, in population-based cancer screening although it is not a perfect model for separating those from false positive cases. It can be seen that the probability of developing breast increased from 0.004 to 0.005 if the model without considering false positive rate (assuming a zero false positive rate). This may be accounted for by the shift of overdiagnosis from the part of false positive cases to the incidence rate of breast cancer when the joint effect of sensitivity and specificity could not be modeled. However, the proposed model is still not specific to exactly estimate the proportion of overdiagnosis. It should be also noted that whether to precisely estimate overdiagnosis may also affect the estimate of number of screening rounds required for stopping screening. The topic on overdiagnosis is therefore very important but is beyond the scope of this study and will become an ongoing subject.

In addition to using geometric distribution, we also extended the proposed GLSM to the negative binomial distribution (NB-2) and its power-based parameterization (NB-P). Geometric distribution is a special case of negative binomial distribution and can be derived by specifying the dispersion parameter (γ_1 and γ_2) as one. The incorporation of dispersion parameter enables the model to account for excessive variation arising from correlation and heterogeneity between observations. The estimated extent of dispersion was higher for the transition from PCDP to CP (γ_2: 2.18 (95% CI: 0.14–6.30)) than the occurrence of PCDP (γ_1: 0.67 (95% CI: 0.16–1.46)), although the estimated results were not deviated from one with statistical significance (NB-2 model incorporating measurement errors, Table 5).

Another interest in the application of stochastic process is to elucidate the role of relevant factors on different stages of disease progression. The proposed GLSM is flexible for incorporating factors of interest as covariates into the model through appropriate link function for both transition probabilities corresponding to the movements between stages. In our application to breast cancer screening, the probability for detecting PCDP is of great interest and was significantly lower for subsequent rounds compared with prevalent round (OR: 0.63 (95% CI: 0.44–0.86), the geometric model in Table 6). In addition to assessing the effect of age group on the stages of disease transition, the incorporation of age group as a covariate also enables the model to cope with the heterogeneities of the occurrence and disease progression of breast cancer shown in previous studies[12–14,18] by allowing for the transition probabilities varying with age groups.

In conclusion, we developed negative-binomial-family-based stochastic process under the backbone of GLISM extending from the framework of GLM to deal with the required number of screening round before detecting multistate outcomes of cancer and number of screening required to stop screening using the empirical data on periodical population-based cancer screening. The proposed GLISM is flexible in several aspects, including accommodating the nature of multistate disease progression to various types of distributions, assessing the effect of covariates on different transitions of disease through appropriate link functions, and making allowance for measurement errors. We applied the proposed method to an empirical data similar to Finnish breast cancer screening program with unique statistical characteristics to first estimate the parameters of disease natural history and measurement errors and then applied these parameters to answer the expected number of screens for detecting a cancer and for stopping screening. Our current work not only makes contribution to the practical aspect of cancer screening in determining inter-screening interval and age to begin with screen given limited resources but also proposes a prototype of generalized stochastic model for accommodating different types of stochastic processes, which can be further developed in the future.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

This study was supported by the Ministry of Science and Technology (grant number MOST 103-2118-M-002-005-MY3; MOST-105-2811-M-002-216).

References

1. Chen HH, Chiu YH, Luh DL, et al. Community-based multiple screening model: design, implementation, and analysis of 42,387 participants. *Cancer* 2004; **100**: 1734–1743.
2. Wu JC, Anttila A, Yen AMF, et al. Evaluation of breast cancer service screening programme with a Bayesian approach: mortality analysis in a Finnish region. *Breast Cancer Res Treat* 2010; **121**: 671–678.
3. Wu JC, Hakama M, Anttila A, et al. 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. *Breast Cancer Res Treat* 2010; **122**: 553–566.
4. Hewitson P, Glasziou P, Watson E, et al. Cochrane systematic review of colorectal cancer screening using the fecal occult blood test (hemoccult): an update. *Am J Gastroenterol* 2008; **103**: 1541–1549.
5. Tabár L, Vitak B, Chen HH, et al. Swedish two-county trial: impact of mammographic screening on breast cancer mortality during 3 decades. *Radiology* 2011; **260**: 658–663.
6. Independent UK Panel on Breast Cancer Screening. The benefits and harms of breast cancer screening: an independent review. *Lancet* 2012; **380**: 1778–1786.
7. Brenner H, Stock C and Hoffmeister M. Effect of screening sigmoidoscopy and screening colonoscopy on colorectal cancer incidence and mortality: systematic review and meta-analysis of randomised controlled trials and observational studies. *BMJ* 2014; **348**: g2467.
8. Tabár L, Faberberg G, Day NE, et al. What is the optimum interval between mammographic screening examinations? An analysis based on the latest results of the Swedish two-county breast cancer screening trial. *Br J Cancer* 1987; **55**: 547–551.
9. Chen HH, Duffy SW and Day NE. Markov chain models for progression of breast cancer. Part I: tumour attributes and the preclinical screen-detectable phase. *J Epidemio Biostat* 1997; **2**: 9–23.
10. Chen HH, Duffy SW, Tabár L, et al. Markov chain models for progression of breast cancer. Part II: prediction of outcomes for different screening regimes. *J Epidemio Biostat* 1997; **2**: 23–35.
11. Chen HH, Yen AM-F and Tabár L. A stochastic model for calibrating the survival benefit of screen-detected cancers. *J Am Stat Assoc* 2012; **107**: 1339–1359.
12. Chen HH, Tabar L, Fagerberg G, et al. Effect of breast cancer screening after age 65. *J Med Screen* 1995; **2**: 10–14.
13. Duffy SW, Chen HH, Tabar L, et al. Sojourn time, sensitivity and positive predictive value of mammography screening for breast cancer in women aged 40-49. *Int J Epidemiol* 1996; **25**: 1139–1145.
14. Duffy SW, Day NE, Tabár L, et al. Markov models of breast tumor progression: some age-specific results. *J Natl Cancer Inst Monogr* 1997; **22**: 93–97.
15. Chen HH, Kuo HS, Yen MF, et al. Estimation of sojourn time in chronic disease screening without data on interval cases. *Biometrics* 2000; **56**: 167–172.
16. Hsieh HJ, Chen TH and Chang SH. Assessing chronic disease progression using non-homogeneous exponential regression Markov models: an illustration using a selective breast cancer screening in Taiwan. *Stat Med* 2002; **21**: 3369–3382.
17. Chen HH, Duffy SW and Tabár L. 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. *J R Stat Soc Series D* 1996; **45**: 307–317.
18. Chiu SY, Duffy S, Yen AM, et al. Effect of baseline breast density on breast cancer incidence, stage, mortality, and screening parameters: 25-year follow-up of a Swedish mammographic screening. *Cancer Epidemiol Biomarkers Prev* 2010; **19**: 1219–1228.
19. Sarkeala T, Anttila A, Forsman H, et al. Process indicators from ten centres in the Finnish breast cancer screening programme from 1991 to 2000. *Eur J Cancer* 2004; **40**: 2116–2125.
20. Cox DR and Miller HD. *The Theory of Stochastic Processes*. London: Chapman and Hall/CRC Press, 1977.
21. Chiang CL. *Introduction to Stochastic Process and their Applications, Ch 5*. New York: Jhon Wiley & Sons, 1980.
22. Tabár L, Fagerberg G, Chen HH, et al. Efficacy of breast cancer screening by age. New results from the Swedish Two-County Trial. *Cancer* 1995; **75**: 2507–2517.
23. Walter SD and Day NE. Estimation of the duration of a pre-clinical disease state using screening data. *Am J Epidemiol* 1983; **118**: 865–886.
24. Hilbe JM. *Negative Binomial Regression*. 2nd ed. Ch 10. Cambridge: Cambridge University Press, 2011.
25. Chen HH, Duffy SW and Tabár L. A mover-stayer mixture of Markov chain models for the assessment of dedifferentiation and tumour progression in breast cancer. *J Appl Stat* 1997; **24**: 265–278.
26. Jackson CH, Jit M, Sharples LD, et al. Calibration of complex models through Bayesian evidence synthesis: a demonstration and tutorial. *Med Decis Making* 2015; **35**: 148–161.

Appendix 1. The derivation of *n*th step transition probability of three-Markov model

The transition matrix for a three-state Markov chain following the expression in equation (3) in the main text can be specified as

$$\mathbf{P} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{bmatrix} P_{00} & P_{01} & P_{02} \\ P_{10} & P_{11} & P_{12} \\ P_{20} & P_{21} & P_{22} \end{bmatrix} \end{array} \tag{32}$$

where $P_{00}=1-P_{01}-P_{02}$, $P_{11}=1-P_{12}-P_{10}$, and $P_{22}=1-P_{20}-P_{21}$. Also note the zero probabilities of returning from PCDP to normal (P_{10}) and from CP to normal (P_{20}) and PCDP (P_{21}) due to the progressive disease assumption of breast cancer. Suppose that $\mathbf{P}$ has distinct eigenvalues, λ_0, λ_1, λ_2, which can be written as

$$\begin{aligned} \lambda_0 &= 1 \\ \lambda_1 &= 1 - \frac{s + \sqrt{s^2 - 4d}}{2} \\ \lambda_2 &= 1 - \frac{s - \sqrt{s^2 - 4d}}{2} \end{aligned} \tag{33}$$

where

$$\begin{aligned} s &= P_{01} + P_{02} + P_{10} + P_{12} + P_{20} + P_{21} \\ d &= (P_{01}P_{12} + P_{01}P_{20} + P_{01}P_{21} + P_{02}P_{10} + P_{02}P_{12} + P_{02}P_{21} + P_{10}P_{20} + P_{10}P_{21} + P_{12}P_{20}). \end{aligned}$$

The corresponding eigenvectors are thus

$$\begin{aligned} \mathbf{q}_0^T &= [1, 1, 1] \\ \mathbf{q}_1^T &= \left[\frac{g_1 + r_1\lambda_1}{k}, \frac{-g_0 - r_0\lambda_1}{k}, 1\right] \\ \mathbf{q}_2^T &= \left[\frac{g_1 + r_1\lambda_2}{k}, \frac{-g_0 - r_0\lambda_2}{k}, 1\right] \end{aligned} \tag{34}$$

where

$$\begin{aligned} k &= P_{01}P_{20}^2 - P_{10}P_{21}^2 + P_{20}P_{21}(P_{01} + P_{02} - P_{10} - P_{12}) \\ r_0 &= P_{10}P_{20} + P_{10}P_{21} + P_{12}P_{20} \\ r_1 &= P_{01}P_{20} + P_{01}P_{21} + P_{02}P_{21} \\ g_0 &= P_{10}(P_{02} + P_{20} + 2P_{21} - 1) + P_{10}P_{21}(P_{21} - 1) + P_{12}P_{20}(P_{01} + P_{02} + P_{20} + P_{21} - 1), \\ g_1 &= P_{01}(P_{12} + 2P_{20} - 1) + P_{02}(P_{10} + P_{12} + P_{20} - 1) \end{aligned}$$

The transition probability matrix $\mathbf{P}$ can be decomposed into

$$\mathbf{Q}\begin{bmatrix} \lambda_0 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{bmatrix}\mathbf{Q}^{-1} \tag{35}$$

derived from the result of standard algebra of matrix. The stationary distribution of the three state Markov chain is thus

$$\begin{aligned} \pi &= \left[\frac{r_0}{R}, \frac{r_1}{R}, \frac{r_2}{R}\right] \\ r_2 &= P_{01}P_{12} + P_{02}P_{10} + P_{02}P_{12} \\ R &= r_0 + r_1 + r_2 \end{aligned} \tag{36}$$

Since the random variables T_{01} and T_{02} follow a geometric type distribution with parameters P_{01} and P_{02}, the expected values can be derived as follows

$$\begin{aligned} E(T_{01}) &= \sum_{n=1}^{\infty} n f_{01}^{(n)} = \sum_{n=1}^{\infty} n P_{00}^{n-1} P_{01} = P_{01} \sum_{n=1}^{\infty} n P_{00}^{n-1} = P_{01}\left[\sum_{n=1}^{\infty} P_{00}^{n-1} + \sum_{n=2}^{\infty} P_{00}^{n-1} + \cdots\right] \\ &= P_{01}\left[\frac{1}{1-P_{00}} + \frac{P_{00}}{1-P_{00}} + \frac{P_{00}^2}{1-P_{00}} \cdots\right] = \frac{P_{01}}{(1-P_{00})^2} \end{aligned} \tag{37}$$

and

$$E(T_{02}) = \sum_{n=1}^{\infty} n f_{02}^{(n)} = \frac{P_{02}}{(1-P_{00})^2} \tag{38}$$

The expected number of time units (C) for detecting subjects with cancer including those detected during screen rounds and surfacing to clinical phase as interval cancer is thus

$$E(C) = E(T_{01}) + E(T_{02}). \tag{39}$$

In the illustration of the natural history of breast cancer, the transition probability matrix is

$$\mathbf{P} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{bmatrix} 0.993 & 0.004 & 0.003 \\ 0 & 0.343 & 0.657 \\ 0 & 0 & 1 \end{bmatrix} \end{array} \tag{40}$$

according to the estimated results of the geometric model in Table 4. The probabilities of the reversion of disease status are zero ($P_{10} = 0$, $P_{20} = 0$, and $P_{21} = 0$) due to the progressive nature. The values of r_0 and r_1 are thus zero and $r_2 = 1$, which gives the stationary distribution of

$$\pi = [0, 0, 1]$$

The expected values, E(T_{01}) and E(T_{02}), given the transition probability matrix are thus 81.6 and 61.2, respectively, and E(C) is calculated as 142.8. Considering a screening program with the inter-screening interval of two years provided to a cohort of 33375 women,[3] the average person-year for observing a breast cancer case is thus 233.72 (33375/142.8). Using the average age of 55 years for the attendees, the average person-year for an attendee to be observed as a breast cancer case is thus 4.25 (233.72/55), corresponding to 2.12 screen round of the program.

Appendix 2. Optimal inter-screen interval for population with different risks

Given the estimated probability of disease progression, namely P_{01} and P_{12}, corresponding to the risk of developing breast cancer at incipient stage (PCDP) and that of progress to CP, the optimum inter-screen interval can be determined with the reference to the number of interval cancers. As an illustration, considering a base-case of applying two-year inter-screen interval to the average risk population of 50, 000 subjects with P_{01} and P_{12} of 0.004 and 0.656, respectively, according to the results derived from the geometric model listed in Table 4. During the follow-up period, 299 interval cancers (CP) and 643 screen-detected cancers (PCDP) would be identified by five rounds of screen including one prevalent screen and four subsequent screens with the inter-screen interval of two years. Suppose the program is applied to the young age group in which the probability of disease progression (P_{12}) was estimated as 0.69 derived from the results of Table 6 and formula (22) in the main text, the expected numbers of interval cancer cases were estimated as 253 and 321 for the inter-screen intervals of 1.5 and 2 years, respectively. On the other hand, if the program with a two-year interval between screening rounds was applied to the old age group ($P_{12} = 0.6$), the expected number of interval cases was estimated as 274. Populations with different risk profiles have different number of interval cancer and thus require different inter-screen intervals. Table S-3 lists the expected number of interval cancer and PCDP for screening program with different inter-screening intervals.

Table S-1. Prior distribution of parameters.

	Parameter	Prior
Transition probability		
Free of cancer → PCDP	Baseline (α_{01})	Uniform(−6,−4)
	Dispersion parameter (log(γ_1))	Uniform(−2,2)
	Power parameter (Q_1)	Uniform(0,5)
PCDP → CP	Baseline (α_{11})	Uniform(0,1.5)
	Dispersion parameter (log(γ_2))	Uniform(−2,2)
	Power parameter (Q_2)	Uniform(0,5)
Effect of covariate		
Free of cancer → PCDP	Age (β_{01})	Uniform(−0.5,0.5)
	Subsequent screening (β_{02})	Uniform(−1,0)
PCDP → CP	Age (β_{11})	Uniform(−2,2)
Measurement error		
Baseline	Specificity (α_{10})	Uniform(4,12)
	Sensitivity(α_{20})	Uniform(0,3)
Effect of covariate (Sensitivity)	Age (β_{21})	Uniform(−1,0.5)
	Year (β_{22})	Uniform(−1.5,0.5)

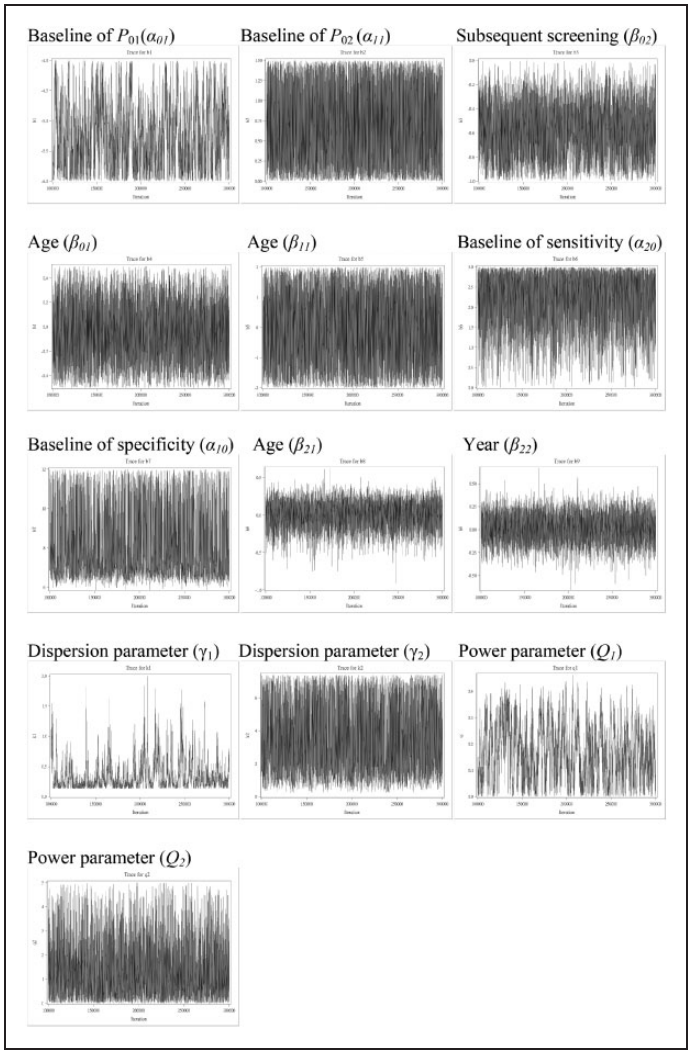

Figure S-1. Trace plots for the NB-P model considering measurement errors and the effects of covariates.

Table S-2. Number of screen rounds required for detecting PCDP based on geometric model, NB-2 model, and NB-P model.

	Geometric model		NB-2 model		NB-P model	
Model	Estimate	95% CI	Estimate	95% CI	Estimate	95% CI
Incorporation measurement error						
Underlying disease progression	2.77	(2.58, 2.92)	2.75	(2.56, 2.92)	2.77	(2.61, 2.91)
Observed number of screening	2.79	(2.62, 2.94)	2.78	(2.60, 2.95)	2.81	(2.65, 2.94)
Incorporation the effects of covariates on transition probability and sensitivity						
Underlying disease progression						
Young age group	2.77	(2.60, 2.94)	2.76	(2.56, 2.93)	2.77	(2.61, 2.92)
Old age group	2.77	(2.60, 2.94)	2.76	(2.56, 2.93)	2.77	(2.62, 2.93)
Observed number of screening						
Young age group	2.83	(2.60, 3.07)	2.81	(2.55, 3.11)	2.83	(2.60, 3.07)
Old age group	2.78	(2.51, 3.05)	2.76	(2.43, 3.05)	2.79	(2.51, 3.05)

Table S-3. Number of interval cancer (CP) and PCDP by different inter-screen intervals and risks profiles of population with size of 50,000 during 8 years of follow-up.

	Inter screen interval (year)			
Detection mode	1	1.5	2	4
Young age group ($P_{12} = 0.69$)				
PCDP	784	722	655	493
Interval cancer (CP)	191	253	321	483
Old age group ($P_{12} = 0.60$)				
PCDP	772	722	643	521
Interval cancer (CP)	145	195	274	396
Average risk population ($P_{12} = 0.66$)				
PCDP	772	715	643	498
Interval cancer (CP)	170	227	299	444