IISE Transactions on Healthcare Systems Engineering

ISSN: 2472-5579 (Print) 2472-5587 (Online) Journal homepage: http://www.tandfonline.com/loi/uhse21

Analyzing overdiagnosis risk in cancer screening: A case of screening mammography for breast cancer

Mahboubeh Madadi, Mohammadhossein Heydari, Shengfan Zhang, Edward Pohl, Chase Rainwater & Donna L. Williams

To cite this article: Mahboubeh Madadi, Mohammadhossein Heydari, Shengfan Zhang, Edward Pohl, Chase Rainwater & Donna L. Williams (2018) Analyzing overdiagnosis risk in cancer screening: A case of screening mammography for breast cancer, IISE Transactions on Healthcare Systems Engineering, 8:1, 2-20, DOI: 10.1080/24725579.2017.1396512

To link to this article: https://doi.org/10.1080/24725579.2017.1396512

Accepted author version posted online: 30 Oct 2017.
Published online: 14 Dec 2017.

IISE TRANSACTIONS ON HEALTHCARE SYSTEMS ENGINEERING
2018, VOL. 8, NO. 1, 2–20
https://doi.org/10.1080/24725579.2017.1396512

Analyzing overdiagnosis risk in cancer screening: A case of screening mammography for breast cancer

Mahboubeh Madadi[a], Mohammadhossein Heydari[b], Shengfan Zhang[b], Edward Pohl[b], Chase Rainwater[b], and Donna L. Williams[c]

[a]Department of Industrial Engineering, Louisiana Tech University, Ruston, LA, USA; [b]Department of Industrial Engineering, University of Arkansas, Fayetteville, AR, USA; [c]School of Public Health, Louisiana State University Health Sciences Center, New Orleans, LA, USA

ABSTRACT

Overdiagnosis is defined as the diagnosis of an asymptotic cancer that would not have presented clinically in a patient's lifetime in the absence of screening. Quantifying overdiagnosis is difficult, since it is impossible to distinguish between a cancer that would cause symptoms in the patient lifetime and the ones that would not. In this study, a mathematical framework is developed to estimate the lifetime overdiagnosis and cancer mortality risks associated with cancer screening policies. We also develop an optimization model to extract screening policies with minimum overdiagnosis and lifetime breast cancer mortality risk. The proposed optimization model is highly nonlinear with complex structure. Therefore, we linearize the optimization model by introducing new decision variables and restructuring the equations to solve it optimally. We utilize existing data on breast cancer for average-risk women and evaluated mammography screening policies in terms of their associated lifetime overdiagnosis and breast cancer mortality risk. Optimal policies with minimum overdiagnosis and mortality risks are derived. The optimal policies outperform the existing in-practice policies by recommending more frequent screenings at younger ages, as the cancer is more aggressive and the remaining life expectancy is higher for younger patients.

ARTICLE HISTORY
Received October 2016
Accepted September 2017

KEYWORDS
Overdiagnosis; cancer screening; nonlinear optimization; linearization

1. Introduction

Since the advent of cancer screening technologies, a substantial reduction in cancer mortality has been observed, and organized screening programs have led to a shift from late-stage diagnosis to early-stage detection. Early detection of cancer enables a broader range of treatment options, less intensive chemotherapy with fewer side-effects, as well as higher survival rates. However, early detection of cancer through screening also leads to detection of cancers (mostly cancers in early stages) that are not life-threatening and would not cause any problem in the patients' lifetimes. Therefore, there is a trade-off between mortality reduction and risk of overdiagnosis due to early detection in any screening program.

Overdiagnosis is defined as the diagnosis of a cancer through screening that would not have presented clinically in a woman's lifetime in the absence of screening (Welch and Black, 2010). Overdiagnosis, although known as the major harm associated with cancer screening, has not been investigated thoroughly. In this study, we address the overdiagnosis issue in preventive healthcare since it can adversely affect people's lives and cause physical and psychosocial harms by unnecessary labeling patients with a lifelong diagnosis and unneeded treatments and surveillance (Carter et al., 2015). Overdiagnosis also causes economic harm by unnecessarily contributing to the rising cost of healthcare (Black, 2000; Welch and Black, 2010). Note that overdiagnosis is different from false positive. Overdiagnosis occurs when a disease is diagnosed correctly, but the diagnosed cancer will not cause harm, suggesting that the treatment for the disease is not needed. False positive is an initial test result that suggests the presence of a disease, but later is proven not to be present (with additional testing). There are two possible explanations for overdiagnosis: (1) the cancer never progresses (or, in fact, regresses); or (2) the cancer progresses slowly enough that the patient dies from a competing cause before the cancer becomes symptomatic (Black, 2000). In other words, overdiagnosis occurs when "very slow" growing cancers (more precisely, at a slow enough pace that individuals die from something else before the cancer ever causes symptoms) are detected. The second explanation incorporates the interaction of three factors: the tumor size at detection, its growth rate, and the competing risks of mortality for the patient. Thus, even a rapidly growing cancer may still represent overdiagnosis if detected when it is very small or in a patient with limited life expectancy (Jørgensen and Gøtzsche, 2009). Figure 1 shows the case when overdiagnosis occurs. As it is not possible to distinguish between lethal and harmless cancers, all detected cancers are treated, and overdiagnosis and overtreatment are therefore inevitable.

Quantifying overdiagnosis, however, is challenging because it is impossible, at the time of diagnosis, to distinguish between an overdiagnosed cancer and one that will become clinical later. There are various studies that quantify overdiagnosis resulting from cancer screenings. The magnitude of overdiagnosis esti-

CONTACT Mahboubeh Madadi madadi@latech.edu Industrial Engineering, Louisiana Tech University, Ruston, LA 71272-0046, USA.
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/uhse.

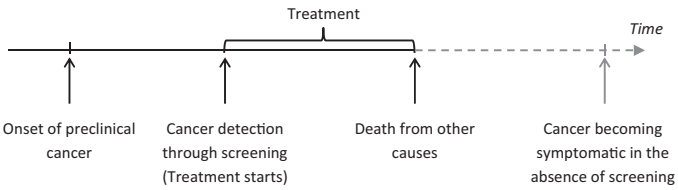

Figure 1. Representation of overdiagnosis in cancer screening.

mates varies widely from one study to another. For example, according to Welch and Black (2010), 25% of mammographically detected breast cancers are overdiagnosed. However, based on Jørgensen and Gøtzsche (2009), one in three breast cancers detected in publicly organized mammography screening programs is overdiagnosed. In an earlier study in 2008, Duffy et al. (2008) estimated the breast cancer overdiagnosis risk to be 39%. However, in a later study (Duffy and Parmar, 2013) after a prolonged follow-up of a screening program in England and Wales, Duffy and Parmar reported an overdiagnosis risk of 7–8% for a biennial screening schedule. They found this estimate to be more plausible than their own previous estimate of 39%.

Most of the studies quantifying cancer overdiagnosis are cohort studies or randomized controlled trial (RCT) follow-up studies which estimate the overdiagnosis risk based on the excess incidence in a screened population compared to an unscreened reference population during the screening period (Welch and Black, 2010; Jørgensen and Gøtzsche, 2009; Duffy et al., 2008; Duffy and Parmar, 2013). However, the downside of these studies is the substantial time and resource requirements due to the need to follow-up observations over a long period of time in order to get a reliable overdiagnosis estimate. In addition, a desirable RCT or observational study must compare screened and unscreened patients with the same underlying risk factors and representing the same historical period and region, from the onset of screening until death (which can be over 30 years). Therefore, many of these studies are subject to selection bias and confounding. Another approach in estimating overdiagnosis is through mathematical modeling or simulation. There are a few relevant studies in the literature that propose a mathematical framework/simulation for estimating overdiagnosis risk. Davidov and Zelen (2004) developed a mathematical model to estimate the probability of overdiagnosis for cancer screening. They applied their model to hypothetical early detection programs for prostate cancer. Gunsoy et al. (2014) developed a Markov simulation model for the evaluation of mammography screening policies in a cohort of British women born in 1935–40. They evaluated nine different screening strategies to quantify the impact of the screening frequency (annual and triennial), starting and ending ages on breast cancer mortality reduction and overdiagnosis. Seigneurin et al. (2011) developed a stochastic simulation model and an approximate Bayesian computation approach to quantify overdiagnosis in French women aged 50–69 years old.

In this article, we propose a mathematical framework to estimate a patient's lifetime overdiagnosis and cancer mortality risks for different screening policies. We apply our estimation model to breast cancer data and evaluate a wide range of policies for an average-risk woman, given that she will develop breast cancer in her lifetime. Compared to previous studies, our estimation model provides a more flexible framework for modeling different uncertainty sources such as cancer sojourn time (i.e., the time interval between the onset of a detectable preclinical cancer and the point when the cancer progresses to the clinical stage, causing symptoms). Unlike previous studies with Markovian assumptions (Gunsoy et al., 2012, 2014), in the proposed approach the distribution of sojourn time is not limited to exponential and can take any plausible form. Note that the assumption of exponential sojourn time has some serious limitations: (1) an exponential distribution has a mode at zero that implies an instant transition from preclinical cancer to clinical cancer; (2) an exponential distribution has a fast decaying tail, which does not adequately account for slow growing tumors; and (3) the memoryless property of the exponential distribution implies that the sojourn time and remaining sojourn time upon cancer detection through screening have the same distribution, which does not capture the real characteristics of cancer growth. In our model, however, the sojourn time can take any acceptable form; e.g., lognormal distribution that has been identified in the literature (Peer et al., 1993).

Although overdiagnosis and overtreatment are inevitable in any screening program, tailoring screening strategies can control the probability of overdiagnosis, and thus decrease the negative effects associated with overdiagnosis and overtreatment. Recently, there have been several studies on the design and optimization of screening and surveillance policies for different types of cancers with respect to different health outcomes, such as remaining life expectancy, mortality risks, false positives, etc. For example, Maillart et al. (2008) evaluated a broad range of screening mammography policies and generated a set of efficient policies, measured by a lifetime breast cancer mortality risk metric and expected number of mammograms. Madadi et al. (2015) developed a partially observed Markov chain to evaluate a broad set of static and dynamic policies in terms of breast cancer mortality risks and total quality-adjusted life years (QALYs). Erenay et al. (2014) identified optimal individualized colonoscopy screening and surveillance policies with the objective of maximizing QALYs for colorectal cancer care. Harms due to cancer screening have also been addressed in previous studies. For example, Ayer et al. (2012) reported the expected number of false positives, and expected number of mammograms (to quantify the disutilities of screening, such as pain, distress, etc.) for different screening mammography policies. However, none of these studies considered overdiagnosis, the most salient risk of screening, in their analyses. In this article, we propose an optimization model to simultaneously minimize a patient's lifetime risk of breast cancer overdiagnosis and breast cancer mortality risk. We use overdiagnosis and mortality risks as the objective functions in our analysis, since overdiagnosis and breast cancer mortality reductions are known to be the most important harm and benefit of screening, respectively.

The remainder of this article is organized as follows. In Section 2, the proposed models for quantifying cancer overdiagnosis and mortality risk, as well as the optimization model, are presented. Restructuring and linearization of the proposed model are also presented in this section. Section 3 presents the data sources and parameter estimates for our numerical studies. In Section 4, numerical studies and results for breast cancer mammography screening are presented. Moreover, optimal policies with minimum overdiagnosis risk and breast cancer

mortality risk are obtained. Finally, Section 5 summarizes the findings and suggests future research opportunities. Note that, throughout the article, we assume that patients comply with the screening policies perfectly.

2. Mathematical formulation

In this section, we first present our mathematical framework for estimating lifetime cancer overdiagnosis and mortality risks for a woman, given that she develops breast cancer in her lifetime. Then, the optimization model developed based on the overdiagnosis and mortality risks formulations is presented in Section 2.2.

2.1. Estimation model

In the proposed estimation model, we incorporate uncertainty in the onset of detectable preclinical cancer, variation in the cancer sojourn time, and competing causes of death. The goal is to develop a mathematical framework to quantify cancer overdiagnosis and mortality risks associated with screening strategies. The following is the list of notations used in the problem formulation.

Notation

m Number of screening examinations in a prescribed screening policy.

τ Screening schedule $\tau = \{\tau_0, \tau_1, ..., \tau_{m+1}\}$ where τ_i's. $(i = 1 \dots m)$ are the decision variables representing the age at which a mammogram should be prescribed. Note that τ_0 and τ_{m+1} are the fixed beginning and ending points of the patient's follow-up period and therefore are not decision variables. No screening is scheduled at ages τ_0 and τ_{m+1}.

T Random variable representing the patient's age at the onset of the detectable preclinical cancer.

S Random variable representing the cancer sojourn time. Note that random variables S and T are dependent because, depending on the age that cancer onsets, the growth rate and therefore the cancer sojourn time are different.

S_j Random variable representing the remaining cancer sojourn time (also known as forward recurrence time) measured from age τ_j.

Ξ Patient health state space, Ξ ={CF,SC,CC}, where CF, SC, and CC represent a cancer-free individual, a patient with screen-detected breast cancer, and a patient with clinical (symptomatic) breast cancer, respectively.

R_u^ξ Random variable representing the remaining life years of a patient in health state $\xi \in \Xi$ at age u. Note that, for the purpose of brevity, we denote the remaining life years at age τ_j by R_j^ξ.

$f_T(t)$ Probability density function of cancer onset time T.

$g_{S|T}(s)$ Conditional probability density function of the cancer sojourn time, given that the cancer preclinical onset is at age T.

$g^j_{S|T}(s,t)$ Conditional probability density function of the forward recurrence time measured from age τ_j, given that the cancer preclinical onset is at age T.

$h_u^\xi(r)$ Conditional probability density function of R_u^ξ, the remaining life years of an individual in health state ξ, given that the patient has survived to age u.

$H_u^\xi(r)$ Conditional cumulative distribution function of R_u^ξ, the remaining life years of an individual in health state ξ, given that the patient has survived to age u.

α_j Sensitivity of mammography (probability of detecting a cancer when it is present) at age τ_j.

Ω_τ Lifetime overdiagnosis risk associated with screening schedule τ.

Θ_τ Lifetime breast cancer mortality risk associated with screening schedule τ.

$\Theta_{\tau,1}$ Lifetime breast cancer mortality risk associated with screening schedule τ when the cancer is diagnosed through screening.

$\Theta_{\tau,2}$ Lifetime breast cancer mortality risk associated with screening schedule τ when the cancer becomes symptomatic.

$\beta(t, \tau_{j-1}, \tau_j)$ Probability that a patient survives to age τ_j, given that she has survived to τ_{j-1} when the cancer onset is at age t.

$\gamma(t, \tau_{j-1}, \tau_j)$ Probability that a cancer with onset at age t has not become symptomatic up to age τ_j, given that it was not symptomatic at age τ_{j-1}.

Note that random variable T is dependent on the conditional remaining life years of a patient (R_j^{CF}, R_j^{SC}, and R_j^{CC}), since knowing the cancer onset time gives us information about patient's age and the expected remaining life years. However, S is independent from the patient's conditional life years, since knowing the sojourn time does not provide information on how long the patient lives. Remaining sojourn time can extend to after patient death. The case that the remaining sojourn time exceeds the patient's remaining life years is when overdiagnosis happens. In addition, the three random variables, R_j^{CF}, R_j^{SC}, and R_j^{CC}, are conditionally dependent, since the three health states CF, SC, and CC are mutually exclusive, and knowing the patient's state and its associated life years provides information about life years associated with the other two states. For example, if we know that a patient with clinical cancer (in state CC) lives for 15 years and then dies from breast cancer ($R_j^{CC} = 15$), it means that the patient will not die from other causes in the next 15 years, so $R_j^{CF} > 15$.

2.1.1. Lifetime overdiagnosis risk

In the case of overdiagnosis, the cancer is detected through a screening examination and treatment starts upon cancer detection. However, if the breast cancer was not detected through screening, the patient would have died from a competing cause before the breast cancer grows to the clinical stage. In other

words, the cancer grows slowly enough that the patient would die from a cause other than breast cancer before the cancer became symptomatic.

Assume the preclinical cancer onset $T = t$ occurs in interval $[\tau_{i-1}, \tau_i)$; i.e., $t \in [\tau_{i-1}, \tau_i)$, $i = 1, 2, \ldots, m+1$. Let D_{ij} be the event that the cancer is diagnosed at the j^{th} screening (at age τ_j, $j = i, i+1, \ldots, m$), given that $T = t \in [\tau_{i-1}, \tau_i)$. Therefore, the conditional probability of overdiagnosis ($\omega_{ij}(t)$) is the probability that the patient's remaining cancer sojourn time (S_j) is greater than the patient's remaining life years in the absence of cancer (R_j^{CF}) at the time of cancer detection (τ_j); i.e.,

$$\omega_{ij}(t) = \Pr(S_j > R_j^{CF} | T = t, D_{ij}) = \int_0^{\infty} g_{S|T}^{j}(s,t) \cdot H_j^{CF}(s) ds. \tag{1}$$

Note that the remaining life years of a cancer-free patient are independent of the cancer onset and remaining sojourn time. The upper bound of the remaining sojourn time is infinity ($+\infty$), implying that it is possible that the cancer would never advance to the clinical stage or show symptoms. In addition, the conditional probability density function of the cancer sojourn time ($g_{S|T}(s)$) and the conditional probability density function of the forward recurrence time (the remaining sojourn time, $g_{S|T}^{j}(s,t)$ at age τ_j) are related through the following equation.

$$\begin{aligned} g_{S|T}^{j}(s,t) &= \Pr(S = \tau_j - T + s | S > \tau_j - T, T = t) \\ &= \frac{g_{S|T}(\tau_j - t + s)}{\overline{G}_{S|T}(\tau_j - t)} \\ &= \frac{g_{S|T}(u - t)}{\overline{G}_{S|T}(\tau_j - t)}, \quad s > 0, \tau_j < u < \tau_{j+1}, \end{aligned} \tag{2}$$

where $u = \tau_j + s$ is the age at which the cancer becomes symptomatic and $\overline{G}_{S|T}(.)$ is the survival function of the cancer sojourn time when the cancer onset is at $T = t$.

In addition, diagnosing the cancer at age τ_j through screening implies that (1) the cancer has not become symptomatic yet; and (2) the patient has survived (has not died from other causes) to age τ_j. Let R^{CF} be the patient's life years when the patient dies from other causes of death. The probability that the patient does not die from other causes before age τ_j, given that she has survived to age τ_{j-1}, is given in Equation (3); and the probability that cancer with onset $t \in [\tau_{i-1}, \tau_i)$ has not developed to a clinical stage at age τ_j yet (i.e., the cancer sojourn time is greater than $\tau_j - t$, Figure 2), given that the cancer had not developed to a clinical stage up to age τ_{j-1}, is given in Equation (4).

$$\beta(t, \tau_{j-1}, \tau_j) = \frac{\Pr(R^{CF} > \tau_j)}{\Pr(R^{CF} > \tau_{j-1})}. \tag{3}$$

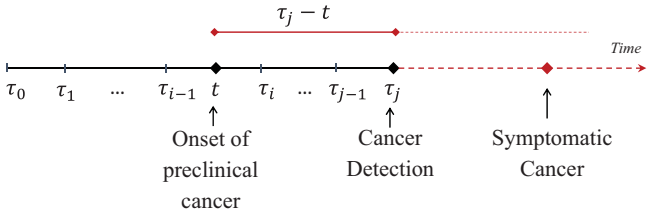

Figure 2. Representation of a hypothesized screening policy, cancer onset and detection.

$$\gamma(t, \tau_{j-1}, \tau_j) = \frac{\Pr(S > \tau_j - T | T = t)}{\Pr(S > \tau_{j-1} - T | T = t)} = \frac{\overline{G}_{S|T}(\tau_j - t)}{\overline{G}_{S|T}(\tau_{j-1} - t)}. \tag{4}$$

Note in the case that $j = i$, $\beta(t, \tau_{j-1}, \tau_j)$ and $\gamma(t, \tau_{j-1}, \tau_j)$ reduce to $\frac{\Pr(R^{CF} > \tau_j)}{\Pr(R^{CF} > t)}$ and $Pr(S > \tau_i - t)$, respectively. In addition, $\beta(t, \tau_{j-1}, \tau_j)$ is a function of t only when $j = i$.

Diagnosing a cancer with onset in the i^{th} screening interval at the j^{th} screening ($j > i$) implies that the cancer was missed in all previous screenings ($i^{th}, (i+1)^{th}, \ldots, (j-1)^{th}$) and then detected at the j^{th} examination. Moreover, if the cancer is detected at the first screening scheduled after its onset ($i = j$), the probability of cancer detection is equal to the sensitivity of the screening test at age τ_j (i.e., α_j). Therefore,

$$\Pr(D_{ij}) = \begin{cases} \alpha_j, & j = i, \\ \alpha_j \prod_{l=i}^{j-1} (1 - \alpha_l), & j > i. \end{cases} \tag{5}$$

In addition, assume that D_i is the event that the cancer is diagnosed through screening, given that its preclinical onset is in the i^{th} screening interval. Then, $D_i = \cup_{j=i}^{m} D_{ij}$, and since D_{ij}'s are mutually exclusive events, the probability that a cancer with preclinical onset in the i^{th} interval is diagnosed through screening is

$$\Pr(D_i) = \sum_{j=i}^{m} \Pr(D_{ij}). \tag{6}$$

Hence, the probability of overdiagnosis, given that the cancer onset is at $t \in [\tau_{i-1}, \tau_i)$, is

$$\omega_i(t) = \sum_{j=i}^{m} \beta(t, \tau_{j-1}, \tau_j) \gamma(t, \tau_{j-1}, \tau_j) \frac{\Pr(D_{ij})}{\Pr(D_i)} \omega_{ij}(t), \tag{7}$$

which calculates the probability that the patient survives and does not develop any symptoms up to the screening age τ_j, given that she had survived and not developed symptoms until τ_{j-1}, the cancer gets detected at the j^{th} screening and, eventually, the patient dies from a competing cause before the cancer grows to the symptomatic size.

Since cancer onsets at different intervals are mutually exclusive, the probability of overdiagnosis for screening schedule $\tau = \{\tau_0, \tau_1, \tau_2, \ldots \tau_m, \tau_{m+1}\}$ using the law of total probability is

$$\begin{aligned} \Omega_\tau = \sum_{i=1}^{m} \int_{\tau_{i-1}}^{\tau_i} f_T(t) \omega_i(t) dt &= \sum_{i=1}^{m} \sum_{j=i}^{m} \frac{Pr(D_{ij})}{Pr(D_i)} \\ &\times \int_{\tau_{i-1}}^{\tau_i} f_T(t) \beta(t, \tau_{j-1}, \tau_j) \gamma(t, \tau_{j-1}, \tau_j) \\ &\times \int_0^{\infty} g_{S|T}^{j}(s,t) \cdot H_j^{CF}(s) ds dt. \end{aligned} \tag{8}$$

2.1.2. Lifetime cancer mortality risk

To calculate the lifetime cancer mortality risk, two cases need to be considered: (1) when the cancer is diagnosed through a screening examination; and (2) when the cancer progresses to a clinical stage and becomes symptomatic in the interval between two prescribed screening tests. These two distinct cases are considered because symptomatic cancers, as opposed to screening detected cancers, are more advanced and may lead to higher cancer mortality risk (Vecchiato et al., 2010).

Suppose that the cancer preclinical onset is in the i^{th} prescribed screening interval; i.e., $T = t \in [\tau_{i-1}, \tau_i)$. The lifetime cancer mortality risks for these two cases are discussed in the following. Consider the first case when the cancer is diagnosed at age τ_j (the j^{th} screening test, $j \geq i$). In this case, the probability that the patient dies from breast cancer is

$$\nu_1^{\tau_j} = \Pr(R_j^{SC} < R_j^{CF}) = \int_0^{\infty} h_j^{CF}(r) H_j^{SC}(r) dr. \tag{9}$$

Given that the cancer is detected at the j^{th} screening test, the patient should survive to age τ_j. The probability that the patient does not die from other causes prior to age τ_j, given that she has survived to age τ_{j-1}, is provided in Equation (3). In addition, detecting the cancer at the j^{th} screening test implies that the cancer sojourn time is greater than $\tau_j - t$. This is also suggesting that the cancer does not develop any symptoms up to age τ_{j-1}. The associated probability of this event, as presented in Equation (4), is $\gamma(t, \tau_{j-1}, \tau_j)$. In addition, the probability that the cancer is diagnosed at the j^{th} screening is presented in Equation (5). Therefore, using the law of total probability, the unconditional probability that a screen-detected cancer with onset $t \in [\tau_{i-1}, \tau_i)$ causes death is

$$\sum_{j=i}^{m} \beta(t, \tau_{j-1}, \tau_j) \gamma(t, \tau_{j-1}, \tau_j) \Pr(D_{ij}) \nu_1^{\tau_j}, \tag{10}$$

which incorporates the patient's survival probability up to age τ_j, and the probabilities that her cancer has not progressed to the symptomatic stage up to screening age τ_j, the cancer is detected through screening and the patient eventually dies from breast cancer.

Since the onsets of cancer at different intervals are mutually exclusive, using the law of total probability, the lifetime breast cancer mortality risk for the first case is

$$\Theta_{\tau,1} = \sum_{i=1}^{m} \sum_{j=i}^{m} \Pr(D_{ij}) \nu_1^{\tau_j} \times \int_{\tau_{i-1}}^{\tau_i} f_T(t) \beta(t, \tau_{j-1}, \tau_j) \gamma(t, \tau_{j-1}, \tau_j) dt. \tag{11}$$

The second case considers the situation in which the cancer becomes symptomatic. Assume that the cancer becomes symptomatic at age $u \in (\tau_{j-1}, \tau_j]$, $j = i+1, \ldots, m+1$. In this case, the conditional breast cancer lifetime mortality risk at age u is

$$\nu_2(u) = \Pr(R_u^{CC} < R_u^{CF}) = \int_0^{\infty} h_u^{CF}(r) H_u^{CC}(r) dr. \tag{12}$$

The fact that the cancer gets symptomatic at age $u \in (\tau_{j-1}, \tau_j)$ suggests that it was not detected through screening tests up to age τ_{j-1}, which occurs with probability $\prod_{l=i}^{j-1}(1 - \alpha_l)$. In addition, the probability that the cancer becomes symptomatic at age $u \in (\tau_{j-1}, \tau_j]$ is $g_{S|T}^{j-1}(u - \tau_{j-1}, t)$ based on Equation (2). Moreover, similar to the first case, the probability that the patient does not die from a competing cause, given that she has survived to age τ_{j-1} (when she received a false negative screening result), is $\beta(t, \tau_{j-1}, u)$ based on Equation (3). Therefore, the associated cancer mortality risk when the cancer onset is in the interval $[\tau_{i-1}, \tau_i)$, and the cancer becomes symptomatic at age u in the interval (τ_{j-1}, τ_j), $j = i+1, \ldots, m+1$ is

$$\prod_{l=i}^{j-1}(1 - \alpha_l) \int_{\tau_{j-1}}^{\tau_j} g_{S|T}^{j-1}(u - \tau_{j-1}, t) \beta(t, \tau_{j-1}, u) \nu_2(u) du, \tag{13}$$

which calculates the probability that the cancer does not get detected through previous screenings and the probability that the individual dies from breast cancer when the cancer becomes symptomatic in the interval (τ_{j-1}, τ_j). Summing over all screening intervals after the cancer onset interval and unconditioning on the cancer onset, the cancer mortality risk is

$$\psi_1 = \sum_{i=1}^{m+1} \int_{\tau_{i-1}}^{\tau_i} f_T(t) \sum_{j=i+1}^{m+1} \prod_{l=i}^{j-1}(1 - \alpha_l) \times \int_{\tau_{j-1}}^{\tau_j} g_{S|T}^{j-1}(u - \tau_{j-1}, t) \beta(t, \tau_{j-1}, u) \nu_2(u) du dt. \tag{14}$$

In addition, it is possible that the cancer becomes symptomatic before the first scheduled screening, in the interval (t, τ_i). In such a case, the following is the associated cancer mortality risk, which has a similar logic as Equation (14).

$$\psi_2 = \sum_{i=1}^{m+1} \int_{\tau_{i-1}}^{\tau_i} f_T(t) \int_t^{\tau_i} g_{S|T}(u - t) \beta(t, \tau_{i-1}, u) \nu_2(u) du dt. \tag{15}$$

Equation (15) calculates the probability that the cancer onsets at time t and it becomes symptomatic at age u, the patient does not die from a competing cause in the interval (τ_{i-1}, u) (i.e., $\beta(t, \tau_{i-1}, u)$), and she eventually dies from breast cancer ($\nu_2(u)$).

Therefore, the lifetime cancer mortality risk of screening schedule τ for a symptomatic cancer is

$$\Theta_{\tau,2} = \psi_1 + \psi_2. \tag{16}$$

Considering both cases of cancer detection, the lifetime cancer mortality risk for screening schedule τ is

$$\Theta_{\tau} = \Theta_{\tau,1} + \Theta_{\tau,2}. \tag{17}$$

2.2. Optimization model

In this section, we develop a bi-objective model in which we simultaneously minimize cancer overdiagnosis and mortality risk. We propose a mixed integer linear model to minimize a linear function of overdiagnosis and breast cancer mortality risks associated with screening policies. For this purpose, we first fix the value of m, number of screenings, and solve the problem for the fixed value of m. Then, across all values of m, we choose the policy that provides the minimum objective function value. The optimization model for a fixed value of m is given in Model (18). In the optimization model, for the purpose of ease of practical implementation, we limit the number of possible interval length switchings in screening policies. Therefore, the number of times that a policy can switch from one screening interval length to another is limited ($\leq N$). Moreover, the time intervals between two consecutive screenings are multiples of a predefined value δ. In addition, since there is medical evidence suggesting that no screening should take place for older women (Schonberg et al., 2009), we define an upper limit for the latest age at which a patient can undergo a mammogram (T_s). Note that T_s is the end of the decision horizon and is different from τ_{m+1}, which is the maximum age expectancy of an individual. In other words,

τ_{m+1} is the end of patient's follow-up period while T_s is the end of the decision horizon. In this section, we also assume that α_j is not age-dependent ($\alpha_j = \alpha, \forall j$), since the problem becomes very complex and extremely difficult to solve optimally when α is a function of a patient's age. The proposed optimization model can then be written as

$$\min \quad (\Omega_\tau, \Theta_\tau) \tag{18a}$$

$$\text{s.t.} \quad \tau_j < \tau_{j+1}, \qquad \forall j \in \{1, \ldots, m\}, \tag{18b}$$

$$\sum_{j=1}^{m-1} I_j \le N, \tag{18c}$$

$$I_j \le M\left|(\tau_{j+1} - \tau_j) - (\tau_j - \tau_{j-1})\right|, \quad \forall j \in \{1, \ldots, m-1\}, \tag{18d}$$

$$\left|(\tau_{j+1} - \tau_j) - (\tau_j - \tau_{j-1})\right| \le MI_j, \quad \forall j \in \{1, \ldots, m-1\}, \tag{18e}$$

$$\tau_j \in \{\tau_0 + \delta, \tau_0 + 2\delta, \ldots, T_s\}, \quad \forall j \in \{1, \ldots, m\}, \tag{18f}$$

$$I_j \in \{0, 1\}, \quad \forall j \in \{1, \ldots, m-1\}, \tag{18g}$$

where Ω_τ and Θ_τ are formulated in Equations (8) and (17) and represent the cancer overdiagnosis and lifetime mortality risks associated with screening policy τ, respectively. Constraint (18b) determines the order of screening decisions. Constraint (18c) ensures that the number of times that a policy switches from one screening interval length to another is limited to N. To count the number of times switching intervals occur, we define binary variable I_j, $j = 1, \ldots, m-1$, which is equal to 0 if two consecutive screening intervals have the same length (i.e., $\tau_{j+1} - \tau_j = \tau_j - \tau_{j-1}$), and 1 otherwise. If two consecutive screening intervals have the same length (i.e., $\tau_{j+1} - \tau_j = \tau_j - \tau_{j-1}$), then Constraints (18d) and (18e) are equivalent to $I_j \le 0$ and $0 \le MI_j$. Since I_j is a binary variable, this results in $I_j = 0$. Otherwise, if two consecutive screening intervals have different lengths (i.e., $(\tau_{j+1} - \tau_j) - (\tau_j - \tau_{j-1}) = c \ne 0$), then Constraints (18d) and (18e) are equal to $I_j \le Mc$ and $c \le MI_j$, which results in $I_j = 1$. Therefore, Equations (18d) and (18e) calculate the value of binary variable I_j, $j = 1, \ldots, m-1$. Moreover, M in Equations (18d) and (18e) represents a large number. Because $(\tau_{j+1} - \tau_j) - (\tau_j - \tau_{j-1}) \le \tau_{m+1} - \tau_0$, we define $M \equiv \tau_{m+1} - \tau_0$ in Equations (18d) and (18e), where, as defined in Section 2.1, τ_0 and τ_{m+1} are fixed parameters denoting the beginning and end points of the follow-up period. Additionally, Constraint (18f) determines the ages from which the screening tests should be prescribed, in which δ represents the minimum interval length between two consecutive screening tests.

To solve Model (18), we use a Pareto multi-objective method to convert the multi-objective model to a single-objective model, as presented in the following.

$$\min \quad a_\Omega \Omega_\tau + a_\Theta \Theta_\tau, \tag{19}$$
$$\text{Constraints (18b)–(18g).}$$

where a_Ω and a_Θ are the decision maker's preference weights for the overdiagnosis and cancer morality risks, respectively. Model (19) is nonlinear/nonconvex and the objective function is not monotone with respect to variables τ_j, $j = 1, \ldots, m$, suggesting that Model (19) is difficult to solve directly; therefore, we develop an equivalent mixed integer linear model which can be solved optimally. The following is the list of sets, parameters, and decision variables introduced in the linearized model.

Sets

- Γ_i For $i \in \{1, \ldots, m\}$, the set of all possible times that screening i can occur
- $\overline{\Gamma}_j$ For $j \in \{1, \ldots, m\}$, the set of possible times n such that cancer onset occurs in interval $[n - \delta, n)$, prior to screening j
- Δ_j For $j \in \{1, \ldots, m+1\}$, the set of all possible times for two consecutive screening tests (τ_{j-1}, τ_j)
- $\Lambda_{i,n}$ For $i \in \{1, \ldots, m\}$ and $n \in \Gamma_i$, the set of all possible times for two consecutive screening tests (τ_{i-1}, τ_i) such that the screening time τ_{i-1} is smaller and screening time τ_i is equal to or greater than time n
- Υ_j For $j \in \{1, \ldots, m\}$, the set of possible times (n, τ_{j-1}, τ_j) for two consecutive screening tests τ_{j-1} and τ_j, when the cancer onset time is $n \in \{\tau_0 + \delta, \ldots, \tau_j\}$
- Φ_j For $j \in \{2, \ldots, m\}$, the set of possible times (n, τ_{j-1}, τ_j) for two consecutive screening tests τ_{j-1} and τ_j, when the cancer onset time is $n \in \{\tau_0 + \delta, \ldots, \tau_{j-1}\}$

Parameters

- T_s Maximum screening age
- $Q_n^{k,\ell}$ The overdiagnosis probability given that cancer onset is in $[n - \delta, n)$, cancer is diagnosed at age ℓ and previous screening is scheduled at age k
- $R_n^{k,\ell}$ Cancer mortality risk for the first case (screen-detected cancer), given that cancer onset is in $[n - \delta, n)$, cancer is diagnosed at age ℓ and previous screening is scheduled at age k
- $U_n^{k,\ell}$ Cancer mortality risk for the second case (symptomatic cancer), given that cancer onset is in $[n - \delta, n)$ and cancer becomes symptomatic in screening interval (k, ℓ) where $n < k$
- $V^{k,\ell}$ Cancer mortality risk for the second case (symptomatic cancer), given that cancer becomes symptomatic before the first scheduled screening after its onset (i.e., cancer onset occurs after age k and cancer becomes symptomatic before age ℓ)

Decision variables

- $x_i^{k,\ell}$ $\begin{cases} 1 \text{ if } \tau_{i-1} = k \text{ and } \tau_i = \ell, \\ 0 \text{ otherwise,} \end{cases}$ $(\forall i \in \{1, \ldots, m\},\ (k, \ell) \in \Delta_i)$
- y_i^n $\begin{cases} 1 \text{ if } n \in \{\tau_{i-1} + \delta, \ldots, \tau_i\}, \\ 0 \text{ otherwise,} \end{cases}$ $(\forall i \in \{1, \ldots, m\},\ n \in \{\tau_0 + \delta, \ldots, T_s\})$
- $z_{j,n}^{\text{OV}}$ For $j \in \{1, \ldots, m\},\ n \in \{\tau_0 + \delta, \ldots, T_s\}$, the probability of detecting a cancer at the j^{th} screening, given that the cancer onset is in $[n - \delta, n)$ and the cancer would eventually be detected through screening
- $z_{j,n}^{\text{MF}}$ For $j \in \{1, \ldots, m\},\ n \in \{\tau_0 + \delta, \ldots, T_s\}$, the probability of detecting a cancer at the j^{th} screening, given that the cancer onset is in $[n - \delta, n)$

$z_{j,n}^{\mathrm{MS}}$ For $j \in \{2, \ldots, m\}$, $n \in \{\tau_0 + \delta, \ldots, T_s\}$, the probability that a cancer with an onset $[n-\delta, n)$ is missed in all screenings which occurred before screening j

z_n^{MB} For $n \in \{\tau_0 + \delta, \ldots, T_s\}$, the probability that a cancer with onset $[n-\delta, n)$ is missed in all screenings

Using the definition of the sets, we now state the mathematical representation of the sets.

$$\Gamma_i = \{\tau_0 + i\delta, \ldots, T_s - (m-i)\delta\}, \quad \forall i \in \{1, \ldots, m\}, \tag{20}$$

$$\overline{\Gamma}_j = \{\tau_0 + \delta, \ldots, T_s - (m-j)\delta\}, \quad \forall j \in \{1, \ldots, m\}, \tag{21}$$

$$\Delta_1 = \{(k, \ell) \mid k = \tau_0, \ell \in \{k + \delta, \ldots, T_s - (m-1)\delta\}\}, \tag{22}$$

$$\Delta_j = \{(k, \ell) \mid k \in \{\tau_0 + (j-1)\delta, \ldots, T_s - (m-j+1)\delta\}, \ \ell \in \{k + \delta, \ldots, T_s - (m-j)\delta\}\}, \quad \forall j \in \{2, \ldots, m\}, \tag{23}$$

$$\Delta_{m+1} = \{(k, \ell) \mid k \in \{\tau_0 + m\delta, \ldots, T_s\}, \ell = \tau_{m+1}\}, \tag{24}$$

$$\Lambda_{1,n} = \{(k, \ell) \mid k = \tau_0, \ell \in \{n, \ldots, T_s - (m-1)\delta\}\}, \quad \forall n \in \Gamma_1, \tag{25}$$

$$\Lambda_{i,n} = \{(k, \ell) \mid k \in \{\tau_0 + (i-1)\delta, \ldots, n - \delta\}, \ \ell \in \{n, \ldots, T_s - (m-i)\delta\}\}, \quad \forall i \in \{2, \ldots, m\}, \ n \in \Gamma_i, \tag{26}$$

$$\Upsilon_1 = \{(n, k, \ell) \mid k = \tau_0, \ell \in \{\tau_0 + \delta, \ldots, T_s - \ (m-1)\delta\}, n \in \{\tau_0 + \delta, \ldots, \ell\}\}, \tag{27}$$

$$\Upsilon_j = \{(n, k, \ell) \mid k \in \{\tau_0 + (j-1)\delta, \ldots, T_s - (m-j+1)\delta\}, \ell \in \{k + \delta, \ldots, T_s - (m-j)\delta\}, n \in \{\tau_0 + \delta, \ldots, \ell\}\}, \quad \forall j \in \{2, \ldots, m\}, \tag{28}$$

$$\Phi_j = \{(n, k, \ell) \mid k \in \{\tau_0 + (j-1)\delta, \ldots, T_s - (m-j+1)\delta\}, \ell \in \{k + \delta, \ldots, T_s - (m-j)\delta\}, n \in \{\tau_0 + \delta, \ldots, k\}\}, \quad \forall j \in \{2, \ldots, m\}. \tag{29}$$

According to Constraint (18f), for screening $i \in \{1, \ldots, m\}$, the screening times are greater than τ_0 in (20). For screening $j \in \{1, \ldots, m\}$, $\overline{\Gamma}_j$ defines the set of ages n such that the cancer onset starts in age interval $[n-\delta, n)$. We define the set of possible values of consecutive screening tests (τ_0, τ_1) in (22) and (τ_{j-1}, τ_j) in (23), in which τ_0 is a fixed parameter. Variables $x_j^{k,\ell}$ are defined to replace variables τ_j, $j = 1, \ldots, m$, in the nonlinear model and represent a screening policy in the equivalent linearized model. More specifically, variable $x_i^{k,\ell}$ is equal to 1 if the screening time for screenings $i-1$ and i are equal to k and ℓ, respectively. To ensure that variables $x_j^{k,\ell}$ represent a unique screening policy, the following constraints should be imposed in the model.

$$\sum_{(k,\ell)\in\Delta_j} x_j^{k,\ell} = 1, \quad \forall j \in \{1, \ldots, m\}, \tag{30}$$

$$\sum_{(k,n)\in\Delta_j} x_j^{k,n} = \sum_{(n,\ell)\in\Delta_{j+1}} x_{j+1}^{n,\ell}, \quad \forall j \in \{1, \ldots, m\}, \ n \in \Gamma_j, \tag{31}$$

where (30) ensures that exactly one testing time is assigned to each screening and (31) implies that the ending point of the $j-1^{\text{th}}$ screening interval is the beginning point of the j^{th} screening interval.

Based on the definition, variable y_i^n is equal to 1 if $n \in \{\tau_{i-1} + \delta, \ldots, \tau_i\}$ is a number between screening times τ_{i-1} and τ_i, and $\Lambda_{i,n}$ defines the set of all possible times for two consecutive screening tests (τ_{i-1}, τ_i) such that τ_{i-1} is smaller than n, and τ_i is equal to or greater than n. Therefore, we have

$$y_i^n = \sum_{(k,\ell)\in\Lambda_{i,n}} x_i^{k,\ell}, \quad \forall i \in \{1, \ldots, m\}, \ n \in \Gamma_i, \tag{32}$$

$$y_i^n = 0, \forall i \in \{1, \ldots, m\}, \ n \in \{\tau_0 + \delta, \ldots, T_s\} \setminus \Gamma_i. \tag{33}$$

Because screening i cannot occur in $\{\tau_0 + \delta, \ldots, T_s\} \setminus \Gamma_i$, Equation (33) is valid. According to the definition of variables $z_{j,n}^{\upsilon}$, $\upsilon \in \{\mathrm{OV}, \mathrm{MF}, \mathrm{MS}\}$ and z_n^{MB}, we have

$$z_{j,n}^{\mathrm{OV}} = \sum_{i=1}^{j} \frac{\Pr(D_{ij})}{\Pr(D_i)} y_i^n, \forall j \in \{1, \ldots, m\}, n \in \{\tau_0 + \delta, \ldots, T_s\}, \tag{34}$$

$$z_{j,n}^{\mathrm{MF}} = \sum_{i=1}^{j} \Pr(D_{ij}) y_i^n, \quad \forall j \in \{1, \ldots, m\}, \ n \in \{\tau_0 + \delta, \ldots, T_s\}, \tag{35}$$

$$z_{j,n}^{\mathrm{MS}} = \sum_{i=1}^{j-1} (1-\alpha)^{j-i} y_i^n, \quad \forall j \in \{2, \ldots, m\}, n \in \{\tau_0 + \delta, \ldots, T_s\}, \tag{36}$$

$$z_n^{\mathrm{MB}} = \sum_{i=1}^{m} (1-\alpha)^{m-i} y_i^n, \quad \forall n \in \{\tau_0 + \delta, \ldots, T_s\}. \tag{37}$$

The term $\Pr(D_{ij})/\Pr(D_i) y_i^n$ defines the probability of detecting cancer in the j^{th} screening, given that cancer onset occurs in time $[n-\delta, n) \in [\tau_{i-1}, \tau_i]$ and the cancer is eventually detected. Because cancer onsets at different intervals are mutually exclusive, Equation (34) calculates $z_{j,n}^{\mathrm{OV}}$. Similarly, Equations (35), (36) and (37) calculate values of $z_{j,n}^{\mathrm{MF}}$, $z_{j,n}^{\mathrm{MS}}$ and z_n^{MB}, respectively.

Note that parameters $Q_n^{k,\ell}$, $R_n^{k,\ell}$ and $U_n^{k,\ell}$ do not carry out index j, because they are only dependent on the cancer onset and the screening times for the current and previous screenings. Define parameters $Q_n^{k,\ell}$, $R_n^{k,\ell}$, $U_n^{k,\ell}$, and $V^{k,\ell}$ as follows:

$$Q_n^{k,\ell} = \int_{n-\delta}^{n} f_T(t)\beta(t,k,\ell)\gamma(t,k,\ell) \int_0^{\infty} g_{S|T}^{\ell}(s,t).H_{\ell}^{CF}(s)dsdt, \quad \forall k \in \{\tau_0, \ldots, T_s - \delta\}, \ \ell \in \{k + \delta, \ldots, T_s\}, n \in \{\tau_0 + \delta, \ldots, \ell\}. \tag{38}$$

$$R_n^{k,\ell} = \nu_1^{\tau_j} \int_{n-\delta}^{n} f_T(t)\beta(t,k,\ell)\gamma(t,k,\ell)dt, \tag{39}$$

$$\forall k \in \{\tau_0, \ldots, T_s - \delta\},\ \ell \in \{k+\delta, \ldots, T_s\}, n \in \{\tau_0 + \delta, \ldots, \ell\}.$$

$$U_n^{k,\ell} = \int_{n-\delta}^{n} \int_{k}^{\ell} g_{S|T}^{k}(u-k,t)\beta(t,k,u)\nu_2(u)dudt, \tag{40}$$

$$\forall k \in \{\tau_0 + \delta, \ldots, T_s - \delta\},\ \ell \in \{k+\delta, \ldots, T_s\}, n \in \{\tau_0 + \delta, \ldots, k\}.$$

$$U_n^{\ell,\tau_{m+1}} = \int_{n-\delta}^{n} \int_{\ell}^{\tau_{m+1}} g_{S|T}^{\ell}(u-\ell,t)\beta(t,u,\ell)\nu_2(u)dudt, \tag{41}$$

$$\forall \ell \in \{\tau_0 + \delta, \ldots, T_s\}, n \in \{\tau_0 + \delta, \ldots, \ell\}.$$

$$V^{k,\ell} = \int_{k}^{\ell} f_T(t) \int_{t}^{\ell} g_{S|T}^{\ell}(u-t)\beta(t,k,u)\nu_2(u)dudt,$$

$$\forall k \in \{\tau_0, \ldots, T_s\},\ \ell \in \{\{k+\delta, \ldots, T_s\} \cup \{\tau_{m+1}\}\}. \tag{42}$$

In the following, we rewrite the cancer overdiagnosis and mortality risks derived in Sections 2.1.1 and 2.1.2 in terms of the new decision variables defined in this section. For the overdiagnosis risk, using Equations (30) and (38), the probability of overdiagnosis for the j^{th} screening when the cancer onset is in $[n-\delta, n)$ can be written as

$$\int_{n-\delta}^{n} f_T(t)\beta(t,\tau_{j-1},\tau_j)\gamma(t,\tau_{j-1},\tau_j) \times \int_{0}^{\infty} g_{S|T}^{j}(s,t).H_j^{CF}(s)dsdt = \sum_{(k,\ell)\in\Delta_j} Q_n^{k,\ell} x_j^{k,\ell},$$

$$\forall j \in \{1, \ldots, m\},\ n \in \overline{\Gamma}_j. \tag{43}$$

The following shows the cancer overdiagnosis risk Ω_τ, the mortality risk for the first case $\Theta_{\tau,1}$, and the elements of the mortality risk for the second case (ψ_1 and ψ_2) in terms of the new decision variables and parameters. For more details, please refer to Appendices A to C.

$$\Omega_\tau = \sum_{j=1}^{m} \sum_{(n,k,\ell)\in\Upsilon_j} Q_n^{k,\ell} x_j^{k,\ell} z_{j,n}^{\text{OV}}. \tag{44}$$

$$\Theta_{\tau,1} = \sum_{j=1}^{m} \sum_{(n,k,\ell)\in\Upsilon_j} R_n^{k,\ell} x_j^{k,\ell} z_{j,n}^{\text{MF}}. \tag{45}$$

$$\psi_1 = \sum_{j=2}^{m} \sum_{(n,k,\ell)\in\Phi_j} U_n^{k,\ell} x_j^{k,\ell} z_{j,n}^{\text{MS}} + \sum_{(n,k,\ell)\in\Upsilon_m} U_n^{\ell,\tau_{m+1}} x_m^{k,\ell} z_n^{\text{MB}}. \tag{46}$$

$$\psi_2 = \sum_{j=1}^{m} \sum_{(k,\ell)\in\Delta_j} V^{k,\ell} x_j^{k,\ell} + \sum_{(k,\ell)\in\Delta_m} V^{\ell,\tau_{m+1}} x_m^{k,\ell}. \tag{47}$$

Therefore, Model (19) can be rewritten in terms of the new decision variables as

$$\min a_\Omega \Bigg(\sum_{j=1}^{m} \sum_{(n,k,\ell)\in\Upsilon_j} Q_n^{k,\ell} x_j^{k,\ell} z_{j,n}^{\text{OV}} \Bigg)$$

$$+ a_\Theta \Bigg(\sum_{j=1}^{m} \sum_{(n,k,\ell)\in\Upsilon_j} R_n^{k,\ell} x_j^{k,\ell} z_{j,n}^{\text{MF}} + \sum_{j=2}^{m} \sum_{(n,k,\ell)\in\Phi_j} U_n^{k,\ell} x_j^{k,\ell} z_{j,n}^{\text{MS}} + \sum_{(n,k,\ell)\in\Upsilon_m} U_n^{\ell,\tau_{m+1}} x_m^{k,\ell} z_n^{\text{MB}} + \sum_{j=1}^{m} \sum_{(k,\ell)\in\Delta_j} V^{k,\ell} x_j^{k,\ell} + \sum_{(k,\ell)\in\Delta_m} V^{\ell,\tau_{m+1}} x_m^{k,\ell} \Bigg), \tag{48a}$$

$$\text{s.t. } x_j^{k,\ell} \in \{0,1\}, \quad \forall j \in \{1, \ldots, m\},\ (k,\ell) \in \Delta_j, \tag{48b}$$

$$z_{j,n}^{\upsilon} \geq 0, \quad \forall j \in \{1, \ldots, m\},\ n \in \overline{\Gamma}_j,\ \upsilon \in \{\text{OV}, \text{MF}\}, \tag{48c}$$

$$z_{j,n}^{\text{MS}} \geq 0, \quad \forall j \in \{2, \ldots, m\},\ n \in \overline{\Gamma}_{j-1}, \tag{48d}$$

$$z_n^{\text{MB}} \geq 0, \quad \forall n \in \overline{\Gamma}_m, \tag{48e}$$

Constraints (18c)–(18g) and Constraints (30)–(37).

Model (48) is still nonlinear. We now linearize Constraints (18c) to (18g) as follows:

$$\sum_{(k,\ell)\in\Delta_{j+1}} x_{j+1}^{k,\ell}(\ell - k) - \sum_{(k,\ell)\in\Delta_j} x_j^{k,\ell}(\ell - k) = e_j^+ - e_j^-, \quad \forall j \in \{1, \ldots, m-1\}, \tag{49}$$

$$e_j^+ \leq M d_j^+, \quad \forall j \in \{1, \ldots, m-1\}, \tag{50}$$

$$e_j^- \leq M d_j^-, \quad \forall j \in \{1, \ldots, m-1\}, \tag{51}$$

$$d_j^+ + d_j^- = 1, \quad \forall j \in \{1, \ldots, m-1\}, \tag{52}$$

$$(e_j^+ + e_j^-) \leq M I_j, \quad \forall j \in \{1, \ldots, m-1\}, \tag{53}$$

$$I_j \leq M(e_j^+ + e_j^-), \quad \forall j \in \{1, \ldots, m-1\}, \tag{54}$$

$$\sum_{j=1}^{m-1} I_j \leq N, \tag{55}$$

$$d_j^+, d_j^-, I_j \in \{0,1\}, \quad \forall j \in \{1, \ldots, m-1\}, \tag{56}$$

$$e_j^+, e_j^- \geq 0, \quad \forall j \in \{1, \ldots, m-1\}, \tag{57}$$

where $\sum_{(k,\ell)\in\Delta_j} x_j^{k,\ell}(\ell - k)$ in Equation (49) is equal to $\tau_j - \tau_{j-1}$. The absolute value of $|(\tau_{j+1} - \tau_j) - (\tau_j - \tau_{j-1})| = e_j^+ + e_j^-$ is calculated in Equations (49) to (52). Since e_j^+, e_j^- and $(e_j^+ + e_j^-)$ are less than or equal to $|(\tau_{j+1} - \tau_j) - (\tau_j - \tau_{j-1})| \leq \tau_{m+1} - \tau_0$, we define $M \equiv \tau_{m+1} - \tau_0$, in which, as defined in Section 2.1, τ_0 and τ_{m+1} are fixed parameters denoting the beginning and end points of the follow-up horizon. We also use variables $w_{j,n}^{\upsilon,k,\ell} \equiv x_j^{k,\ell} z_{j,n}^{\upsilon}$ ($\upsilon \in \{\text{OV}, \text{MF}, \text{MS}\}$) to linearize $x_j^{k,\ell} z_{j,n}^{\upsilon}$ and $w_n^{\text{MB},k,\ell} \equiv x_m^{k,\ell} z_n^{\text{MB}}$ to linearize $x_m^{k,\ell} z_n^{\text{MB}}$ in Objective (48a). Noting that $z_{j,n}^{\upsilon} \leq 1$, $\upsilon \in \{\text{OV}, \text{MF}, \text{MS}\}$, and $z_n^{\text{MB}} \leq 1$, the mixed integer linear model is given by

$$\min\ a_{\Omega}\left(\sum_{j=1}^{m}\sum_{(n,k,\ell)\in\Upsilon_j} Q_n^{k,\ell} w_{j,n}^{\mathrm{OV},k,\ell}\right) + a_{\Theta}\Bigg(\sum_{j=1}^{m}\sum_{(n,k,\ell)\in\Upsilon_j} R_n^{k,\ell} w_{j,n}^{\mathrm{MF},k,\ell} + \sum_{j=2}^{m}\sum_{(n,k,\ell)\in\Phi_j} U_n^{k,\ell} w_{j,n}^{\mathrm{MS},k,\ell}$$
$$+ \sum_{(n,k,\ell)\in\Upsilon_m} U_n^{\ell,\tau_{m+1}} w_n^{\mathrm{MB},k,\ell} + \sum_{j=1}^{m}\sum_{(k,\ell)\in\Delta_j} V^{k,\ell} x_j^{k,\ell} + \sum_{(k,\ell)\in\Delta_m} V^{\ell,\tau_{m+1}} x_m^{k,\ell}\Bigg), \tag{58a}$$

$$\text{s.t. } w_{j,n}^{\upsilon,k,\ell} \le x_j^{k,\ell}, \quad \forall j\in\{1,\dots,m\},\ (n,k,\ell)\in\Upsilon_j,\ \upsilon\in\{\mathrm{OV},\mathrm{MF}\}, \tag{58b}$$

$$w_{j,n}^{\upsilon,k,\ell} \le z_{j,n}^{\upsilon}, \quad \forall j\in\{1,\dots,m\},\ (n,k,\ell)\in\Upsilon_j,\ \upsilon\in\{\mathrm{OV},\mathrm{MF}\}, \tag{58c}$$

$$w_{j,n}^{\upsilon,k,\ell} \ge z_{j,n}^{\upsilon} - (1-x_j^{k,\ell}), \quad \forall j\in\{1,\dots,m\},\ (n,k,\ell)\in\Upsilon_j,\ \upsilon\in\{\mathrm{OV},\mathrm{MF}\}, \tag{58d}$$

$$w_{j,n}^{\upsilon,k,\ell} \ge 0, \quad \forall j\in\{1,\dots,m\},\ (n,k,\ell)\in\Upsilon_j,\ \upsilon\in\{\mathrm{OV},\mathrm{MF}\}, \tag{58e}$$

$$w_{j,n}^{\mathrm{MS},k,\ell} \le x_j^{k,\ell}, \quad \forall j\in\{2,\dots,m\},\ (n,k,\ell)\in\Phi_j, \tag{58f}$$

$$w_{j,n}^{\mathrm{MS},k,\ell} \le z_{j,n}^{\mathrm{MS}}, \quad \forall j\in\{2,\dots,m\},\ (n,k,\ell)\in\Phi_j, \tag{58g}$$

$$w_{j,n}^{\mathrm{MS},k,\ell} \ge z_{j,n}^{\mathrm{MS}} - (1-x_j^{k,\ell}), \quad \forall j\in\{2,\dots,m\},\ (n,k,\ell)\in\Phi_j, \tag{58h}$$

$$w_{j,n}^{\mathrm{MS},k,\ell} \ge 0, \quad \forall j\in\{2,\dots,m\},\ (n,k,\ell)\in\Phi_j, \tag{58i}$$

$$w_n^{\mathrm{MB},k,\ell} \le x_m^{k,\ell}, \quad \forall (n,k,\ell)\in\Upsilon_m, \tag{58j}$$

$$w_n^{\mathrm{MB},k,\ell} \le z_n^{\mathrm{MB}}, \quad \forall (n,k,\ell)\in\Upsilon_m, \tag{58k}$$

$$w_n^{\mathrm{MB},k,\ell} \ge z_n^{\mathrm{MB}} - (1-x_m^{k,\ell}), \quad \forall (n,k,\ell)\in\Upsilon_m, \tag{58l}$$

$$w_n^{\mathrm{MB},k,\ell} \ge 0, \quad \forall (n,k,\ell)\in\Upsilon_m, \tag{58m}$$

Constraints (30)–(37), Constraints (48b)–(48e), and Constraints (49)–(57).

3. Model input

The probability distribution of breast cancer onset is the most challenging distribution to estimate in this study, since cancer onset cannot be observed or measured directly. Parmigiani and Skates (2001) developed a de-convolution method to estimate the age of disease onset distribution based on the natural history of the disease, the disease incidence rate, competing causes of death, etc. They then used the singular value decomposition method to solve their developed model numerically. The breast cancer preclinical onset age distribution is calculated using the exact calculation method of Parmigiani and Skate.

There are various studies in the literature estimating the distribution of breast cancer sojourn time. Some of these studies used an exponential specification in estimating the sojourn time distributions (Duffy et al., 1997; Shen and Zelen, 2001, 2005; Tabár et al., 2000). However, the assumption of an exponential duration in the preclinical stage has some limitations. The first concern is the implausible assumption of mode at zero, which corresponds to an instant transition from preclinical to clinical stage, and the fast decaying tail, which does not adequately account for slow–growing tumors (Parmigiani and Skates, 2001). The second limitation is due to the memoryless property of the exponential distribution, which implies that the sojourn time and remaining sojourn time upon cancer detection through screening have the same distribution. In other words, the hazard function of sojourn time is constant, implying that, as time passes, the instantaneous probability that the cancer develops to a clinical stage is constant. However, it would be more plausible that the instantaneous probability of developing to a clinical stage be a non-monotonic function of time. A second distribution proposed for modeling cancer sojourn time is the lognormal distribution. Spratt et al. (1986) postulated a lognormal distribution for sojourn time based on the growth patterns of breast tumors. In this study, we examine both exponential and lognormal sojourn time distributions. For the exponential case, the age-specific (age groups of 40–49, 50–59, 60+) rate parameters are adopted from the mean sojourn time provided in Duffy et al. (1997), Shen and Zelen (2001, 2005), and Tabár et al. (2000). For the case with lognormal sojourn time, the age-specific distribution parameters are estimated using the median and upper 95% quantile matching Peer et al. (1993). Note that the mean sojourn times for the exponential distributions are very close to the mean sojourn times of the lognormal distributions for all the age groups considered. The estimated mean sojourn time for the three age groups of 40–49, 50–59, and 60+ are respectively 3.2, 4.7, and 5.2 for the exponential case, and 2.9, 4.5, and 5.9 for the lognormal case. The estimated hazard functions of lognormal breast cancer sojourn time for the three age groups (40–49, 50–59, 60+) are presented in Fig. 3. As the results show, the hazard functions for this case are increasing at first and start decreasing slowly a few years after the cancer onset. This suggests that if no symptom appears by a certain amount of time after the cancer onset, then the instantaneous probability that the cancer actually becomes symptomatic starts to decrease (because the tumor stops growing or it regresses).

The probability distribution of the remaining life years for a cancer-free individual is extracted from the 2008 US life table for females (Centers for Disease Control and Prevention, 2012) and the Surveillance, Epidemiology and End Results (SEER) data (National Concer Institute, 2015). SEER data are used to exclude the probability of death from breast cancer in the life table and adjust the probabilities for women without breast

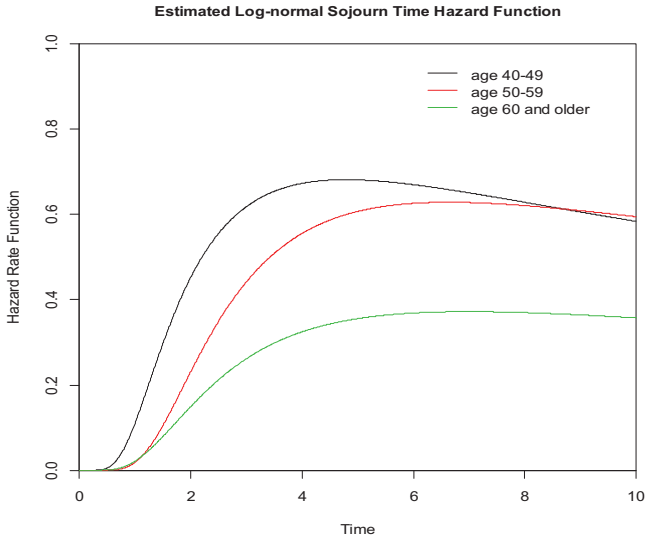

Figure 3. Estimated lognormal breast cancer sojourn time for different age groups.

cancer. We estimate the probability density of remaining life years for the two cases of cancer detection (screen-detected and symptomatic breast cancer) using the data available in Schairer et al. (2004) and Zhang (2011). They reported the survival probabilities of patients for four different breast cancer stages: in situ, localized, regional and distant breast cancer. We adopt the stage distributions of screen-detected and symptomatic breast cancer from Bleyer and Welch (2012) and Plevritis et al. (2007), respectively, and adjust the probability density of remaining life years for the two cases. For example, let q_j^{ξ} be the probability that a patient is in state ξ and is diagnosed with a cancer in stage j (where $j \in J = \{\text{in situ, localized, regional, distant}\}$), and $p_{i,j}^{\xi}$ be the probability that an individual of age i who is in state ξ and cancer stage of j dies in the age interval $(i, i+1]$. Therefore, the probability that an individual of age i and in state ξ dies in the age interval $(i, i+1]$ is

$$p_i^{\xi} = \sum_{j \in J} q_j^{\xi} p_{i,j}^{\xi}. \tag{59}$$

The mortality probabilities in CDC (n.d.), Schairer et al. (2004) and Zhang (2011) are presented in tabular format, and therefore we used a piecewise-constant density function to model the probability density functions. Let p_i^{ξ} be the probability that an individual of age i in health state ξ dies in the age interval $(i, i+1]$. Therefore,

$$\Pr(R^{\xi} \le k) = 1 - \prod_{i \le k} (1 - p_i^{\xi}), \tag{60}$$

and

$$h_j^{\xi}(r) = \frac{\Pr(R^{\xi} \le k+1) - \Pr(R^{\xi} \le k)}{\Pr(R^{\xi} \ge j)}, \quad \forall\, k < r \le k+1,\ k > j. \tag{61}$$

Lastly, age–specific mammography screening sensitivities are extracted from Kerlikowske et al. (2000). Table 1 presents the summary of data sources for the different inputs (parameters and distributions) incorporated in the model.

4. Results

Disease onset and conditional remaining life year functions do not have a general closed form and are only available in tabular format. This makes analytical calculation of integration challenging, especially when the sojourn time follows a lognormal distribution. Therefore, in this study, we exploit the Monte Carlo integration method to calculate the integrations in the proposed model. Monte Carlo integration methods are sampling methods to calculate complicated integrations, based on the central limit theorem and the law of large numbers. Please refer to Robert and Casella (2013) for more details.

To validate the models, the probabilities of developing breast cancer in the next 10 years and breast cancer mortality risks calculated by the model are compared with those reported by the American Cancer Society (ACS). We compared the age-adjusted incidence rates in the model with the rates reported by the ACS (n.d.). Table 2 presents the 10-year incidence likelihood for different age groups based on our model (calculated based on estimated f_T) and the ACS report. Note that the ACS reports the incidence likelihood in the next 10 years for the general population. However, in our model we focus on the population with breast cancer; i.e., the probability of cancer onset in a 10-year time window for a patient, given that she develops breast cancer in her lifetime. Therefore, the probabilities reported by the ACS

Table 1. Data sources for the model input estimations.

Description	Notation/Parameter	Reference
Distribution of preclinical detectable breast cancer onset	$f(.)$	Parmigiani and Skates (2001)
Preclinical sojourn time distribution (exponential case) parameters (mean of S)	λ	Shen and Zelen (2001, 2005), and Tabár et al. (2000)
Preclinical sojourn time distribution (lognormal case) parameters (mean and standard deviation of S)	μ, σ	Peer et al. (1993)
Probability density of remaining life years of a cancer-free individual	h_u^{CF}	The US life table for females Centers for Disease Control and Prevention (2012), and SEER National Concer Institute (2012)
Age– and stage–specific probability distribution of death from breast cancer	h_u^{SC}, h_u^{CC}	Schairer et al. (2004) and Zhang (2011)
Stage distribution of breast cancers	q_j^{ξ}	Bleyer and Welch (2012) (screen-detected cancers), Plevritis et al. (2007) (clinically detected cancers)
Age–specific mammography sensitivity	α_j	Kerlikowske *et al.* (2000)

Table 2. Comparison of age-adjusted incidence rate of the model and the ACS data.

	The probability of developing BC in the next 10 years			
	Population with BC		General Population	
Age	ACS	Model	ACS	Model
40	0.110	0.109	0.014	0.013
50	0.184	0.182	0.023	0.023
60	0.280	0.278	0.035	0.034
70	0.312	0.310	0.039	0.038

(fifth column) are normalized by the lifetime chance of developing breast cancer (1 in 8) to calculate the associated probabilities for the population with breast cancer (second column). As the results show, the estimated onsets are very close to the data provided by the ACS. In addition, we compared our estimated breast cancer mortality risk with the the survival rates reported by the ACS (2015). Based on the ACS report, the mortality rate (survival rate) for women diagnosed with breast cancer after 15 years is 22% (78%), which is comparable with the lifetime mortality risks calculated in our model. For more details, please refer to Section 4.1.

The results presented in this section have two parts. In the first part, we evaluate different screening policies, in terms of the breast cancer overdiagnosis and mortality risks. The second part presents the optimal screening policies obtained.

4.1. Policy evaluation

In this section, we present the cancer overdiagnosis and lifetime mortality risks of different policies. We evaluate different static as well as dynamic screening policies. We consider two types of dynamic policies with one and two switching ages. In dynamic policies with one switching age (one-switch policies), a patient starts with one screening interval and then switches to another screening interval at the switching age. Similarly, for policies with two switching ages (two-switch policies), a patient switches the screening frequency at two points in her life.

The two most commonly referred mammography screening policies are the American Cancer Society (ACS) and the US Preventive Services Task Force (USPSTF) policies. Previously, the ACS recommended annual screening mammography, beginning at age 40. Recently, the ACS has changed the breast cancer screening guidelines. In the new ACS policy, women are recommended to receive annual mammograms between ages 45 to 54 and switch to getting mammograms every two years afterward. Based on the new ACS policy, the screening should continue as long as a woman is in good health and is expected to live 10 more years or longer. According to the USPSTF guideline, screening mammograms should be done every two years between age 50 and 74 for women at average risk of breast cancer.

Table 3 presents the policies evaluated in this study. Since the new ACS policy does not specify an age to stop screening, we consider three variations of the new ACS policy with stopping ages of 80, 90, and 100. Including the USPSTF policy and the three variations of the new ACS policy, in total we evaluate 244 policies. We use vectors of length three, five, and seven to represent static, one-switch policies, and two-switch policies, respectively. Static policies are presented by (a_s, i_1, a_e), in which a_s, and a_e are the starting and ending age, respectively, and i_1 is the screening interval length. Policies with two screening intervals (one-switch policies) are represented by $(a_s, i_1, a_1, i_2, a_e)$, where a_s and a_e are defined the same as above; a_1 is the age that the patient switches from one interval to another; and i_1 and i_2 represent the first and second screening intervals, respectively. Similarly, $(a_s, i_1, a_1, i_2, a_2, i_3, a_e)$ represents a policy with three screening intervals (two-switch policies) where a_e, i_1, a_2, i_2, a_e are defined, same as above and a_2, and i_3 represent the second switching age and the length of third screening interval, respectively. For example, policy (40,1,50,2,60,3,80) recommends that women receive annual screening tests between age 40 and 50, then switch to biennial screenings up to age 60, and continue screenings every three years up to age 80. Note that the old ACS policy, the USPSTF policy and the three variations of the new ACS policies are represented as (40,1,100), (50,2,74), (45,1,54,2,80), (45,1,54,2,90), and (45,1,54,2,100), respectively.

Figure 4 presents the breast cancer overdiagnosis and mortality risks of the policies considered in this article along with the efficient frontier policies for the case with (a) exponential and (b) lognormal sojourn time distributions. Efficient frontier policies are policies for which overdiagnosis (mortality) risk cannot be improved in value without degrading the mortality (overdiagnosis) risk.

Note that the policies, associated overdiagnosis risks are slightly different for the two cases of exponential and lognormal sojourn time with overdiagnosis risk of the lognormal case being slightly higher. Although slightly different, the overdiagnosis risks for the two distributions are close. In addition, the mortality risk for the lognormal case is smaller than the corresponding mortality risk for the exponential case. This is due to the memoryless property of exponential case and the nonmonotone behavior of lognormal sojourn time hazard function. In the lognormal sojourn time case, as discussed earlier, if a cancer is not symptomatic for an amount of time after its preclinical onset, the probability of instantaneous symptomatic cancer decreases with time, which implies that the probability that the cancer does not grow to a symptomatic size is higher, resulting in a higher overdiagnosis risk and a lower mortality risk. For example, suppose that the cancer preclinical onset is at age 50 and cancer sojourn time follows a lognormal distribution. In such case, if the cancer is not detected (either through screening or symptoms) within the next four years, the probability that it grows to symptomatic size afterwards decreases. However, if the sojourn time is exponential and the cancer is not detected up to age a, the probability that cancer becomes symptomatic at any

Table 3. Screening policies considered in the numerical analysis: a_s, and a_e are the starting and stopping ages; i_1, i_2, and i_3 are the screening interval lengths; and a_1 and a_2 are the interval length switching ages.

Policy	a_s	i_1	a_1	i_2	a_2	i_3	a_e
Static	40,50	1,2,3	—	—	—	—	80,90,100
Dynamic (one–switch policies)	40,50	1,2,3	50,60,70	1,2,3	—	—	80,90,100
Dynamic (two–switch policies)	40,50	1,2,3	50,60,70	1,2,3	60,70,80	1,2,3	80,90,100

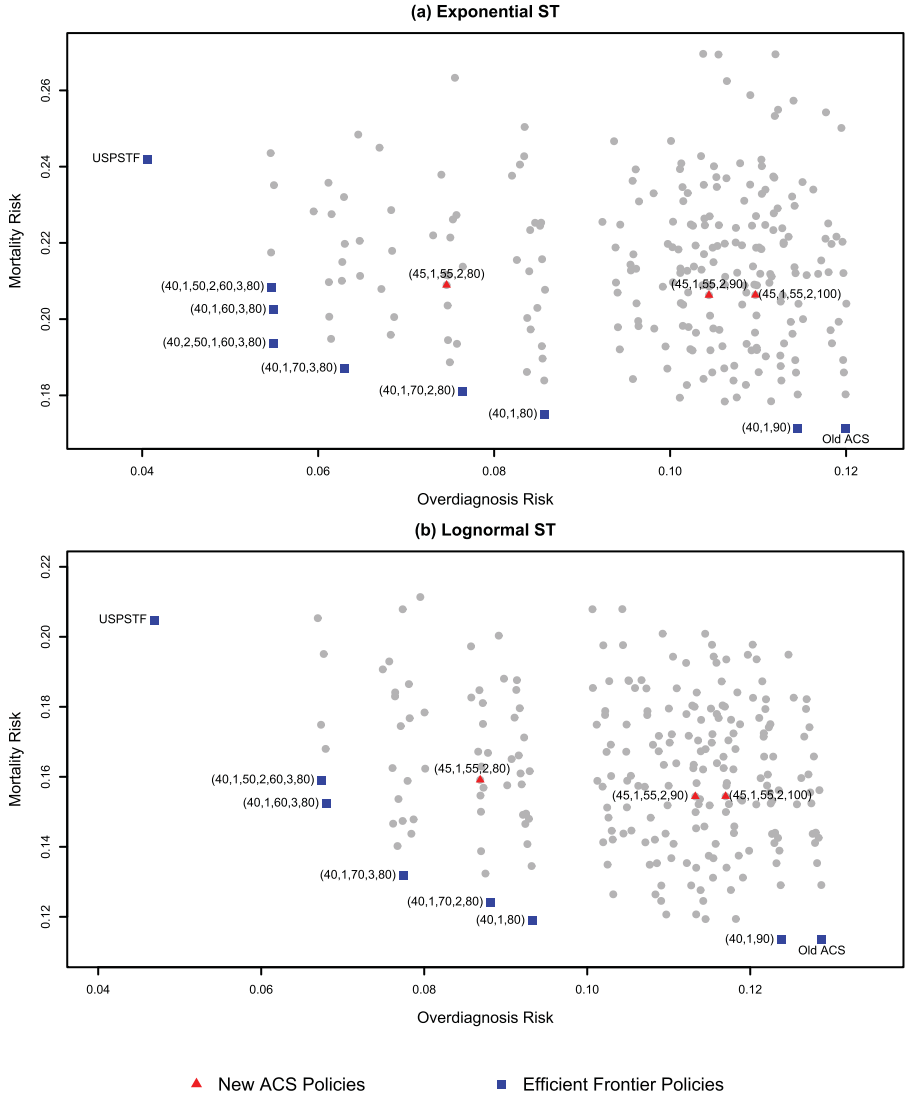

Figure 4. Breast cancer overdiagnosis and mortality risks of the screening policies in Table 3.

age afterwards is constant, regardless of the time that has passed since the cancer onset. Therefore, the overdiagnosis and mortality risks in lognormal sojourn time case are higher and lower, respectively.

The overdiagnosis risks for both distributions are comparable with the reported values in the literature (Duffy and Parmar, 2013; De Gelder et al., 2011; Gunsoy et al., 2014). The overdiagnosis risk estimates for biennial screening schedules are in line with the values reported in the literature: Duffy and Parmar (2013) estimate 7–8%, and De Gelder's estimates 7.2%, De Gelder et al. (2011). In addition, our result is in line with Gunsoy's estimate of 5.6% for triennial screening strategies (Gunsoy et al., 2014). The breast cancer mortality risks are also comparable with the ACS report (2015). Based on the ACS report, the 10-year and 15-year survival rate of breast cancer patients are 83% and 78%, respectively, suggesting about 17% and 22% risk of mortality in 10 and 15 years after detection, which are in line with our results. Note that our results present the risks for the population of individuals with breast cancer (given that the patient develops breast cancer in her lifetime).

The results in Fig. 4 show that the old ACS policy and the USPSTF policy are among efficient frontier policies for both exponential and lognormal sojourn time cases. In fact, the USPSTF policy and the old ACS policy have the lowest overdiagnosis and breast cancer mortality risk among the evaluated policies, respectively. Most of the efficient frontier policies (except for the old ACS and (40,1,90)) recommend that women stop screening at or before age 80. This is because, after age 80, the probability of death from competing causes significantly increases, which results in high overdiagnosis risk. In addition, most efficient policies (except for the USPSTF) recommend women that start screenings at age 40. This maintains both the overdiagnosis and mortality risks at a lower level. In addition, in terms of the distribution of screening tests in the efficient frontier policies, more frequent tests are recommended in the age interval 40 to 60 when breast cancer onset is more likely.

4.2. Optimization results

In this section, we present the optimal policies extracted for two different decision horizons of age of 40 to 80, and 45 to

80. The starting age of 40 is considered ($\tau_0 = 39$), since it is the earliest age among the recommended starting ages for routine mammography screening and the likelihood of developing breast cancer prior to age 40 is very low (1 in 1899 (American Cancer Society, 2015)). We also consider possible starting age of 45, as it is recommended by the new ACS policy, updated in late 2015. We also assume that the latest age a patient can undergo screening (T_s) is 80 since, based on a previous study, screening women over the age of 80 would cause more harms than benefits (Schonberg et al., 2009). In addition, a patient's life expectancy drops drastically after this age and therefore the probability of overdiagnosis increases. Note that in our analysis $\tau_{m+1} = 100$, since the patient is followed up to age 100 to calculate her breast cancer mortality risk. In our optimization analysis, the mammography sensitivity is assumed to be independent of age, and its value is calculated as the weighted average of age-dependent sensitivity values used in Section 4.1. We also assume the minimum interval between two subsequent mammograms should be more than one year ($\delta = 1$ year) and the number of times a policy can switch between screening intervals is limited to two (i.e., N=2). We vary the overdiagnosis risk weight (a_Ω) and mortality risk weight (a_Θ) from 0 to 1 with 0.1 increments ($a_\Omega + a_\Theta = 1$). Therefore, the decision maker can choose the optimal policy based on his/her preference on the overdiagnosis and mortality risk weights.

A server with an Intel core i32 with 3.1 GHz and 768 GB RAM and CPLEX 12.4 is used to solve the mixed integer programming models to optimality. Note that the computational time for each individual problem with a fixed number of screenings m is no more than five hours.

4.2.1. Numerical results

The optimal policies along with the associated overdiagnosis (Ω^*_τ) and breast cancer mortality risks (Θ^*_τ) are presented in Tables 4 through 7. Note that, in Tables 4 through 7, a shaded cell represents a recommended mammography screening while an empty cell represents "no screening." The results are presented for different combinations of a_Ω and a_Θ in range 0.1 to 0.9. Note that for the case when $a_\Omega = 1$, the optimal policy is "no screening," which results in the overdiagnosis risk of zero and mortality risks of 0.33, and 0.34 for the exponential and lognormal case, respectively. In addition, in the case of $a_\Theta = 1$, the optimal policy is screening every year over the decision horizon, which results in the minimal breast cancer mortality risks.

Tables 4 and 5 present the optimal policies for the exponential sojourn time case for the two decision horizons of 40–80 and 45–80, respectively. Note that the solving time of Model (58) depends on τ_0 and τ_{m+1}. In practice, we are interested to study screenings where $\tau_0 \geq 39$ and $\tau_{m+1} \leq 100$. In the following, we show that Cplex is able to solve Model (58) efficiently for $\tau_0 = 39$ and $\tau_{m+1} = 100$, so we are able to solve the problem efficiently for the practical values of cancer screening. The optimal policies have similar structures for both decision horizons and also across different risk weights. The optimal policies recommend more frequent mammograms at younger ages and no mammogram screenings as the patients get older. All screening policies recommend starting annual screenings at age 40, and the stopping age depends on the two risks associated weights. As a_Ω

Table 4. Optimal policies: Exponential sojourn time case between age 40 to 80.

a_Ω	a_Θ	Policy (40 41 42 43 44 45 46 47 48 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80)	m^*	Ω^*_τ	Θ^*_τ
0.9	0.1		15	0.0054	0.2901
0.8	0.2		20	0.0106	0.2570
0.7	0.3		24	0.0197	0.2288
0.6	0.4		27	0.0267	0.2153
0.5	0.5		30	0.0349	0.2052
0.4	0.6		33	0.0446	0.1976
0.3	0.7		38	0.0606	0.1895
0.2	0.8		41	0.0652	0.1875
0.1	0.9		41	0.0652	0.1875

Table 5. Optimal policies: Exponential sojourn time case between age 45 to 80.

		Policy																																											
a_Ω	a_Θ	40	41	42	43	44	45	46	47	48	48	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	m^*	Ω^*_τ	Θ^*_τ
0.9	0.1						■	■	■	■	■	■	■	■	■	■																											10	0.0052	0.3042
0.8	0.2						■	■	■	■	■	■	■	■	■	■	■	■	■	■	■																						15	0.0104	0.2709
0.7	0.3						■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■																	20	0.0195	0.2428
0.6	0.4						■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■														23	0.0265	0.2293
0.5	0.5						■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■											26	0.0346	0.2192
0.4	0.6						■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■									28	0.0410	0.2139
0.3	0.7						■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■			34	0.0603	0.2034
0.2	0.8						■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	36	0.06503	0.2015
0.1	0.9						■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	36	0.06503	0.2015

Table 6. Optimal policies: Lognormal sojourn time case between age 40 to 80.

		Policy																																											
a_Ω	a_Θ	40	41	42	43	44	45	46	47	48	48	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	m^*	Ω^*_τ	Θ^*_τ
0.9	0.1	■	■	■	■	■	■	■	■	■	■	■		■		■		■		■		■																					16	0.0084	0.2233
0.8	0.2	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■		■		■		■		■																22	0.0166	0.1719
0.7	0.3	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■		■		■		■													26	0.0207	0.1596
0.6	0.4	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■													29	0.0233	0.1539
0.5	0.5	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■		■										31	0.0297	0.1464
0.4	0.6	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■		■		■								32	0.0495	0.1373
0.3	0.7	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■		■		■		■		■	37	0.0598	0.1264
0.2	0.8	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■		■		■	39	0.0604	0.1261
0.1	0.9	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■		■	40	0.0610	0.1260

Table 7. Optimal policies: Lognormal sojourn time case between age 45 to 80.

a_Ω	a_Θ	Policy (40 41 42 43 44 45 46 47 48 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80)	m^*	Ω^*_τ	Θ^*_τ
0.9	0.1		11	0.0082	0.2456
0.8	0.2		18	0.0164	0.1942
0.7	0.3		22	0.0204	0.1819
0.6	0.4		25	0.0231	0.1763
0.5	0.5		26	0.0295	0.1688
0.4	0.6		28	0.0402	0.1597
0.3	0.7		32	0.0589	0.1450
0.2	0.8		34	0.0602	0.1440
0.1	0.9		35	0.0608	0.1432

increases, fewer screenings (m) are recommended, and screenings are scheduled at younger ages to avoid overdiagnosis. This is intuitively correct, since the probability of death from a competing cause and, as a result, the likelihood of overdiagnosis is lower at younger ages. However, as a_Θ increases, more screening tests are prescribed to detect the cancer through screening and prevent symptomatic cancer that is associated with higher breast cancer mortality risk. In addition, as the results show, the overdiagnosis and breast cancer mortality risks increase and decrease, respectively, as the number of screenings and a_Θ increases.

Tables 6 and 7 present the results for the decision horizon of 40 to 80 and 45 to 80, respectively, where the sojourn time follows a lognormal distribution. As the results suggest, the optimal policies, structures are very similar to the exponential case, with more frequent screenings at the beginning and more spread-out mammograms at older ages. In addition, the optimal numbers of screenings m^* for the lognormal case are very similar to those in the exponential case across different weight combinations (e.g., $m^* = 30$ for the exponential case vs. $m^* = 31$ for the lognormal case when $a_\Omega = a_\Theta = 0.5$ for the decision horizon of 40 to 80.) However, in the lognormal case, the stopping screening age tends to be larger, specially when $a_\Theta > 0.5$ ($\tau_m = 77$ vs. $\tau_m = 80$ for exponential and lognormal, respectively, when $a_\Theta = 0.7$). This translates to less frequent (e.g., biennial) screenings toward the end of the decision horizon. This happens due to the characteristics of exponential and lognormal distributions. As discussed earlier, exponential distribution mode at zero implies an instant transition from preclinical to clinical stage, which is not realistic. This implies that when the sojourn time is exponentially distributed, the probability that the cancer transitions from its preclinical onset to clinical stage (become symptomatic) within one year is 19% on average. This value is 3% for the case of lognormal distribution. Note that these probabilities are calculated as the weighted average of the probabilities of transitioning to the clinical stage within one year of cancer onset for different age groups, with the weights being the proportion of the US population in the associated age groups. In addition, the memoryless property of exponential distribution implies a constant probability of instantaneous symptomatic cancer. Therefore, the probability that a cancer becomes symptomatic remains the same throughout the patient's life. However, in the lognormal case, as discussed earlier, the hazard functions are increasing at first but start decreasing after some time, implying that if a cancer is not symptomatic after a while, its probability of becoming symptomatic decreases. These properties imply that a patient should undergo mammogram screenings more frequently in order to catch cancer through screening to decrease cancer mortality risk when sojourn time is exponentially distributed. However, in the case of lognormal sojourn time, since the probability of cancer becoming symptomatic decreases after some time, the screenings are more spread out.

5. Conclusion

Preventive health services with advanced technologies, although known to detect diseases in early stages when patients are more

likely to be successfully treated, have ignited a debate on overdiagnosis. Ideally, screening interventions aim to detect diseases that will ultimately cause harm, and the purpose of screening interventions is to advance the detection time, when the disease is in its early stages and is more likely to be treated. However, there is always the risk of overdiagnosis and overtreatment when detecting a disease in its early stages. Overdiagnosis of a disease is defined as the diagnosis of an asymptomatic disease having no signs or symptoms, which would have never become symptomatic during an individual's remaining lifetime.

In this study, we derive the equations for the probability of lifetime breast cancer overdiagnosis and mortality risks. Although applied to breast cancer, the proposed model can be generalized to calculate risks for other types of cancer. We evaluate the lifetime overdiagnosis and mortality risks associated with in-practice and alternative mammography screening policies, including the old and new ACS and the USPSTF policies. In addition, we derive optimal screening policies with minimum linear combination of overdiagnosis and mortality risk. The initial optimization model is nonlinear and very complex to solve optimally. Therefore, we restructure the initial model by introducing new decision variables and linearizing the model. The optimization results imply that more frequent screenings should be performed at younger ages and few to no mammogram test-s at older ages.

There are some limitations and future research directions for this study. First, there is no universal agreement on the sojourn time distribution. More clinical and statistical research on finding the sojourn time would help quantify the overdiagnosis risk more accurately. As limitation on the available data sources can cause some ambiguity and inconsistency in the parameter estimates, robust optimization can be applied in the future to ensure that the optimal policies are robust against this limitation. In addition, in this study, the optimal policies are derived at the population level. A future direction would be to optimize screening policies at the individual level. A multi-stage stochastic programming model or Markov decision process could be developed that depends on the status of the individual and the data received at each screening. In the current model, we assume that mammography sensitivity is fixed in contrast with the fact that the sensitivity increases as women age. A possible future direction would be to consider age-dependent mammography sensitivity. Implementing this, however, requires a major restructure of the optimization model presented here. Note that the dynamics of sensitivity have already been incorporated in the estimation models but not in the optimization model. Another possible direction would be to investigate the impact of overdiagnosis on a patient's quality of life (e.g., QALYs) by quantifying the physical, psychosocial and economic harms that overdiagnosis causes. In addition, overdiagnosis risk is a function of the stage at which the cancer is detected. In this study, we do not incorporate different breast cancer stages in the overdiagnosis calculation. A very interesting future direction would be to investigate overdiagnosis risk as a function of the stage of a detected cancer. Moreover, in this study, we assume perfect adherence to screening guidelines, which results in overestimation of overdiagnosis risk. A possible future direction would be incorporation of uncertainty in patient's adherence since, in reality, patients do not comply with cancer screening policies completely.

References

American Cancer Society (2015) Breast Cancer Facts & Figures 2015-2016. https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/breast-cancer-facts-and-figures/breast-cancer-facts-and-figures-2015-2016.pdf (accessed June 2017).

Ayer, T., Alagoz, O., and Stout, N. K. (2012) A POMDP approach to personalize mammography screening decisions. *Operations Research*, **60**(5), 1019–1034.

Black, W. C. (2000) Overdiagnosis: An underrecognized cause of confusion and harm in cancer screening. *Journal of the National Cancer Institute*, **92**(16), 1280–1282.

Bleyer, A. and Welch, H. G. (2012) Effect of three decades of screening mammography on breast-cancer incidence. *New England Journal of Medicine*, **367**(21), 1998–2005.

Carter, J. L., Coletti, R. J., and Harris, R. P. (2015) Quantifying and monitoring overdiagnosis in cancer screening: A systematic review of methods. *BMJ*, **350**(5), g7773–g7773.

Centers for Disease Control and Prevention (2012) United States Life Tables, 2008. http://www.cdc.gov/nchs/data/nvsr/nvsr61/nvsr61_03.pdf (accessed June 2016).

Davidov, O. and Zelen, M. (2004) Overdiagnosis in early detection programs. *Biostatistics*, **5**(4), 603–613.

De Gelder, R., Heijnsdijk, E. A., Van Ravesteyn, N. T., Fracheboud, J., Draisma, G., and De Koning, H. J. (2011) Interpreting overdiagnosis estimates in population-based mammography screening. *Epidemiologic Reviews*, **27**(33), 111–121.

Duffy, S. W., Day, N. E., Tabár, L., Chen, H. H., and Smith, T. C. (1997) Markov models of breast tumor progression: some age-specific results. *Journal of the National Cancer Institute. Monographs*, **22**, 93–97.

Duffy, S. W., Lynge, E., Jonsson, H., Ayyaz, S., and Olsen, A. H. (2008) Complexities in the estimation of overdiagnosis in breast cancer screening. *British Journal of Cancer*, **99**(7), 1176–1178.

Duffy, S. W. and Parmar, D. (2013) Overdiagnosis in breast cancer screening: the importance of length of observation period and lead time. *Breast Cancer Research*, **15**(3), 1–9.

Erenay, F. S., Alagoz, O., and Said, A. (2014) Optimizing colonoscopy screening for colorectal cancer prevention and surveillance. *Manufacturing & Service Operations Management*, **16**(3), 381–400.

Gunsoy, N. B., Garcia-Closas, M., and Moss, S. M. (2012) Modelling the overdiagnosis of breast cancer due to mammography screening in women aged 40 to 49 in the united kingdom. *Breast Cancer Research*, **14**(6), 1.

Gunsoy, N. B., Garcia-Closas, M., and Moss, S. M. (2014) Estimating breast cancer mortality reduction and overdiagnosis due to screening for different strategies in the united kingdom. *British Journal of Cancer*, **110**, 2412–2419.

Jørgensen, K. J. and Gøtzsche, P. C. (2009) Overdiagnosis in publicly organised mammography screening programmes: systematic review of incidence trends. *BMJ*, **339**, b2587.

Kerlikowske, K., Carney, P. A., Geller, B., Mandelson, M. T., Taplin, S. H., Malvin, K., Ernster, V., Urban, N., Cutter, G., Rosenberg, R., and Ballard-Barbash, R. (2000) Performance of screening mammography among women with and without a first-degree relative with breast cancer. *Annals of Internal Medicine*, **133**(11), 855–863.

Madadi, M., Zhang, S., and Henderson, L. M. (2015) Evaluation of breast cancer mammography screening policies considering adherence behavior. *European Journal of Operational Research*, **247**(2), 630–640.

Maillart, L. M., Ivy, J. S., Ransom, S., and Diehl, K. (2008) Assessing dynamic breast cancer screening policies. *Operations Research*, **56**(6), 1411–1427.

National Cancer Institute (2012) Surveillance, epidemiology and end results. http://seer.cancer.gov/archive/csr/1975_2009_pops09/results_merged/topic_lifetime_risk.pdf (accessed December 2015).

Parmigiani, G. and Skates, S. (2001) Estimating distribution of age of the onset of detectable asymptomatic cancer. *Mathematical and Computer Modelling*, **33**(12), 1347–1360.

Peer, P. G. M., Van Dijck, J. A. A. M., Hendriks, J. H. C. L., Holland, R., and Verbeek, A. L. M. (1993) Age-dependent growth rate of primary breast cancer. *Cancer*, **71**(11), 3547–3551.

Plevritis, S. K., Salzman, P., Sigal, B. M., and Glynn, P. W. (2007) A natural history model of stage progression applied to breast cancer. *Statistics in Medicine*, **26**(3), 581–595.

Robert, C. P. and Casella, G. (2013) *Monte Carlo Statistical Methods*. Springer Science & Business Media, New York, NY, United States.

Schairer, C., Mink, P. J., Carroll, L., and Devesa, S. S. (2004) Probabilities of death from breast cancer and other causes among female breast cancer patients. *Journal of the National Cancer Institute*, **96**(17), 1311–1321.

Schonberg, M. A., Silliman, R. A., and Marcantonio, E. R. (2009) Weighing the benefits and burdens of mammography screening among women age 80 years or older. *Journal of Clinical Oncology*, **27**(11), 1774–1780.

Seigneurin, A., François, O., Labarère, J., Oudeville, P., Monlong, J., and Colonna, M. (2011) Overdiagnosis from non-progressive cancer detected by screening mammography: Stochastic simulation study with calibration to population based registry data. *BMJ*, **343**, d7017.

Shen, Y. and Zelen, M. (2001) Screening sensitivity and sojourn time from breast cancer early detection clinical trials: Mammograms and physical examinations. *Journal of Clinical Oncology*, **19**(15), 3490–3499.

Shen, Y. and Zelen, M. (2005) Robust modeling in screening studies: Estimation of sensitivity and preclinical sojourn time distribution. *Biostatistics*, **6**(4), 604–614.

Spratt, J. S., Greenberg, R. A., and Heuser, L. S. (1986) Geometry, growth rates, and duration of cancer and carcinoma in situ of the breast before detection by screening. *Cancer Research*, **46**(2), 970–974.

Tabár, L., Vitak, B., Chen, H. H., Duffy, S. W., Yen, M. F., Chiang, C. F., Krusemo, U. B., Tot, T., and Smith, R. A. (2000) The swedish two-county trial twenty years later: Updated mortality results and new insights from long-term follow-up. *Radiologic Clinics of North America*, **38**(4), 625–651.

Vecchiato, F., Strabbioli, M., Zorzi, M., Baracco, S., and Ciatto, S. (2010) Interval cancers in breast cancer screening: Comparison of stage and biological characteristics with screen-detected cancers or incident cancers in the absence of screening. *Tumori*, **96**(2), 198–201.

Welch, H. G. and Black, W. C. (2010) Overdiagnosis in cancer. *Journal of the National Cancer Institute*, **102**(9), 605–613.

Zhang, S. (2011) *Modeling the complexity of breast cancer under conditions of uncertainty*. North Carolina State University, Raleigh, NC, United States.

Appendix A: Proof of Equation (44)

$$\Omega_\tau = \sum_{j=1}^{m}\sum_{i=1}^{j}\frac{\Pr(D_{ij})}{\Pr(D_i)}\int_{\tau_{i-1}}^{\tau_i} f_T(t)\beta(t,\tau_{j-1},\tau_j)\gamma(t,\tau_{j-1},\tau_j)\int_0^\infty g^j_{S|T}(s,t).H^{CF}_j(s)dsdt, \tag{62}$$

$$= \sum_{j=1}^{m}\sum_{i=1}^{j}\frac{\Pr(D_{ij})}{\Pr(D_i)}\sum_{n\in\overline{\Gamma}_i}\left(\int_{n-\delta}^{n} f_T(t)\beta(t,\tau_{j-1},\tau_j)\gamma(t,\tau_{j-1},\tau_j)\int_0^\infty g^j_{S|T}(s,t).H^{CF}_j(s)dsdt\right)y_i^n, \tag{63}$$

$$= \sum_{j=1}^{m}\sum_{i=1}^{j}\frac{\Pr(D_{ij})}{\Pr(D_i)}\sum_{q\in\{\tau_0+\delta,\ldots,T_s\}}\left(\int_{q-\delta}^{q} f_T(t)\beta(t,\tau_{j-1},\tau_j)\gamma(t,\tau_{j-1},\tau_j)\int_0^\infty g^j_{S|T}(s,t).H^{CF}_j(s)dsdt\right)y_i^q, \tag{64}$$

$$= \sum_{j=1}^{m}\sum_{q\in\{\tau_0+\delta,\ldots,T_s\}} z^{\mathrm{OV}}_{j,q}\left(\int_{q-\delta}^{q} f_T(t)\beta(t,\tau_{j-1},\tau_j)\gamma(t,\tau_{j-1},\tau_j)\int_0^\infty g^j_{S|T}(s).H^{CF}_j(s)dsdt\right), \tag{65}$$

$$= \sum_{j=1}^{m}\sum_{q\in\overline{\Gamma}_j} z^{\mathrm{OV}}_{j,q}\left(\int_{q-\delta}^{q} f_T(t)\beta(t,\tau j-1,\tau_j)\gamma(t,\tau_{j-1},\tau_j)\int_0^\infty g^j_{S|T}(s,t).H^{CF}_j(s)dsdt\right), \tag{66}$$

$$= \sum_{j=1}^{m}\sum_{q\in\overline{\Gamma}_j} z^{\mathrm{OV}}_{j,q}\sum_{(k,\ell)\in\Delta_j} Q_q^{k,\ell}x_j^{k,\ell} = \sum_{j=1}^{m}\sum_{(q,k,\ell)\in\Upsilon_j} Q_q^{k,\ell}x_j^{k,\ell}z^{\mathrm{OV}}_{j,q}, \tag{67}$$

where we change the order of summations over i and j in Equation (62). Moreover, Equation (63) is valid because of the definition of variable y_i^n; Equations (64) and (65) follow from (33) and the definition of $z^{\mathrm{OV}}_{j,q}$, respectively. Equation (66) follows from (33) and the fact that $\Gamma_j \subseteq \overline{\Gamma}_j$, the first and second equalities of (67), follow from Equation (43), and the definition of set Υ_j, respectively.

Appendix B: Proof of Equation (45)

$$\Theta_{\tau,1} = \sum_{j=1}^{m}\sum_{i=1}^{j}\Pr(D_{ij})\nu_1^{\tau_j}\int_{\tau_{i-1}}^{\tau_i} f_T(t)\beta(t,\tau_{j-1},\tau_j)\gamma(t,\tau_{j-1},\tau_j)dt, \tag{68}$$

$$= \sum_{j=1}^{m}\sum_{i=1}^{j}\Pr(D_{ij})\sum_{n\in\overline{\Gamma}_i}\left(\nu_1^{\tau_j}\int_{n-\delta}^{n} f_T(t)\beta(t,\tau_{j-1},\tau_j)\gamma(t,\tau_{j-1},\tau_j)dt\right)y_i^n, \tag{69}$$

$$= \sum_{j=1}^{m}\sum_{i=1}^{j}\Pr(D_{ij})\sum_{q\in\{\tau_0+\delta,\ldots,T_s\}}\left(\nu_1^{\tau_j}\int_{q-\delta}^{q} f_T(t)\beta(t,\tau_{j-1},\tau_j)\gamma(t,\tau_{j-1},\tau_j)dt\right)y_i^q, \tag{70}$$

$$= \sum_{j=1}^{m} \sum_{q\in\{\tau_0+\delta,\ldots,T_s\}} z_{j,q}^{\mathrm{MF}} \left(\nu_1^{\tau_j} \int_{q-\delta}^{q} f_T(t)\beta(t,\tau_{j-1},\tau_j)\gamma(t,\tau_{j-1},\tau_j)dt \right), \tag{71}$$

$$= \sum_{j=1}^{m} \sum_{q\in\overline{\Gamma}_j} z_{j,q}^{\mathrm{MF}} \left(\nu_1^{\tau_j} \int_{q-\delta}^{q} f_T(t)\beta(t,\tau_{j-1},\tau_j)\gamma(t,\tau_{j-1},\tau_j)dt \right), \tag{72}$$

$$= \sum_{j=1}^{m} \sum_{q\in\overline{\Gamma}_j} z_{j,q}^{\mathrm{MF}} \sum_{(k,\ell)\in\Delta_j} R_q^{k,\ell} x_j^{k,\ell} = \sum_{j=1}^{m} \sum_{(q,k,\ell)\in\Upsilon_j} R_q^{k,\ell} x_j^{k,\ell} z_{j,q}^{\mathrm{MF}}, \tag{73}$$

where the orders of summations over i and j are changed in Equation (68); Equation (69) follows from the definition of variable y_i^n; Equation (70) holds because of (33); Equation (71) follows from the definition of $z_{j,q}^{\mathrm{MF}}$; Equation (72) holds because of (33) and $\Gamma_j \subseteq \overline{\Gamma}_j$; the second equality of (73) results from the definition of Υ_j; and the first equality of (73) holds because

$$\nu_1^{\tau_j} \int_{q-\delta}^{q} f_T(t)\beta(t,\tau_{j-1},\tau_j)\gamma(t,\tau_{j-1},\tau_j)dt = \sum_{(k,\ell)\in\Delta_j} R_q^{k,\ell} x_j^{k,\ell}, \ \forall j \in \{1,\ldots,m\}, \ q \in \overline{\Gamma}_j, \tag{74}$$

where Equation (74) follows from (30).

Appendix C: Proof of Equation (46) and (47)

$$\psi_1 = \sum_{j=2}^{m+1} \sum_{i=1}^{j} (1-\alpha)^{j-i} \int_{\tau_{i-1}}^{\tau_i} \int_{\tau_{j-1}}^{\tau_j} g_{S|T}^{j-1}(u-\tau_{j-1},t)\beta(t,\tau_{j-1},u)\nu_2(u)dudt, \tag{75}$$

$$= \sum_{j=2}^{m+1} \sum_{i=1}^{j} (1-\alpha)^{j-i} \sum_{n\in\overline{\Gamma}_i} \left(\int_{n-\delta}^{n} \int_{\tau_{j-1}}^{\tau_j} g_{S|T}^{j-1}(u-\tau_{j-1},t)\beta(t,\tau_{j-1},u)\nu_2(u)dudt \right) y_i^n, \tag{76}$$

$$= \sum_{j=2}^{m+1} \sum_{i=1}^{j} (1-\alpha)^{j-i} \sum_{q\in\{\tau_0+\delta,\ldots,T_s\}} \left(\int_{q-\delta}^{q} \int_{\tau_{j-1}}^{\tau_j} g_{S|T}^{j-1}(u-\tau_{j-1},t)\beta(t,\tau_{j-1},u)\nu_2(u)dudt \right) y_i^q, \tag{77}$$

$$\begin{aligned} &= \sum_{j=2}^{m} z_{j,q}^{\mathrm{MS}} \sum_{q\in\{\tau_0+\delta,\ldots,T_s\}} \left(\int_{q-\delta}^{q} \int_{\tau_{j-1}}^{\tau_j} g_{S|T}^{j-1}(u-\tau_{j-1},t)\beta(t,\tau_{j-1},u)\nu_2(u)dudt \right) \\ &\quad + \sum_{q\in\{\tau_0+\delta,\ldots,T_s\}} z_q^{\mathrm{MB}} \left(\int_{q-\delta}^{q} \int_{\tau_m}^{\tau_{m+1}} g_{S|T}^{m}(u-\tau_m,t)\beta(t,\tau_m,u)\nu_2(u)dudt \right), \end{aligned} \tag{78}$$

$$\begin{aligned} &= \sum_{j=2}^{m} z_{j,q}^{\mathrm{MS}} \sum_{q\in\overline{\Gamma}_{j-1}} \left(\int_{q-\delta}^{q} \int_{\tau_{j-1}}^{\tau_j} g_{S|T}^{j-1}(u-\tau_{j-1},t)\beta(t,\tau_{j-1},u)\nu_2(u)dudt \right) \\ &\quad + \sum_{q\in\overline{\Gamma}_m} z_q^{\mathrm{MB}} \left(\int_{q-\delta}^{q} \int_{\tau_m}^{\tau_{m+1}} g_{S|T}^{m}(u-\tau_m,t)\beta(t,\tau_m,u)\nu_2(u)dudt \right), \end{aligned} \tag{79}$$

$$= \sum_{j=2}^{m} z_{j,q}^{\mathrm{MS}} \sum_{q\in\overline{\Gamma}_{j-1}} \sum_{(k,\ell)\in\Delta_j} U_q^{k,\ell} x_j^{k,\ell} + \sum_{q\in\overline{\Gamma}_m} z_q^{\mathrm{MB}} \sum_{(k,\ell)\in\Delta_m} U_q^{\ell,\tau_m} x_j^{k,\ell}, \tag{80}$$

$$= \sum_{j=2}^{m} \sum_{(q,k,\ell)\in\Phi_j} U_q^{k,\ell} x_j^{k,\ell} z_{j,q}^{\mathrm{MS}} + \sum_{(q,k,\ell)\in\Upsilon_m} U_q^{\ell,\tau_{m+1}} x_m^{k,\ell} z_q^{\mathrm{MB}}, \tag{81}$$

where we change the order of summations over i and j in Equation (75); Equation (76) results from the definition of variable y_i^n; Equation (77) results from (33); Equation (78) follows from the definition of $z_{j,q}^{\mathrm{MS}}$ and z_q^{MB}. Equation (79) is correct because of (33) and $\Gamma_j \subseteq \overline{\Gamma}_j$; and Equation (80) is valid because

$$\int_{q-\delta}^{q}\int_{\tau_{j-1}}^{\tau_j} g_{S|T}^{j-1}(u-\tau_{j-1},t)\beta(t,\tau_{j-1},u)\nu_2(u)dudt = \sum_{(k,\ell)\in\Delta_j} U_q^{k,\ell}x_j^{k,\ell}, \qquad \forall j\in\{2,\dots,m\},\ q\in\overline{\Gamma}_{j-1}, \tag{82}$$

$$\int_{q-\delta}^{q}\int_{\tau_m}^{\tau_{m+1}} g_{S|T}^{m}(u-\tau_m,t)\beta(t,\tau_m,u)\nu_2(u)dudt = \sum_{(k,\ell)\in\Delta_m} U_q^{\ell,\tau_{m+1}}x_m^{k,\ell}, \qquad \forall q\in\overline{\Gamma}_m, \tag{83}$$

where Equations (82) and (83) follow from (30).

For ψ_2 we have

$$\psi_2 = \sum_{j=1}^{m}\sum_{(k,\ell)\in\Delta_j} V^{k,\ell}x_j^{k,\ell} + \sum_{(k,\ell)\in\Delta_m} V^{\ell,\tau_{m+1}}x_m^{k,\ell}, \tag{84}$$

which holds because of the definition of parameter $V^{k,\ell}$ and because

$$\int_{\tau_{i-1}}^{\tau_i} f_T(t)\int_{t}^{\tau_i} g_{S|T}(u-t)\beta(t,\tau_{i-1},u)\nu_2(u)dudt = \sum_{(k,\ell)\in\Delta_i} V^{k,\ell}x_i^{k,\ell},\ \forall i\in\{1,\dots,m\}, \tag{85}$$

$$\int_{\tau_m}^{\tau_{m+1}} f_T(t)\int_{t}^{\tau_{m+1}} g_{S|T}^{m+1}(u-t)\beta(t,\tau_m,u)\nu_2(u)dudt = \sum_{(k,\ell)\in\Delta_m} V^{\ell,\tau_{m+1}}x_m^{k,\ell}, \tag{86}$$

where Equations (85) and (86) follow from (30).