Published in final edited form as:
Stat Biosci. 2011 December ; 3(2): 169–186. doi:10.1007/s12561-011-9032-7.

Risk-specific optimal cancer screening schedules: an application to breast cancer early detection

Charlotte Hsieh Ahern,
Division of Biostatistics, Dan L. Duncan Cancer Center at Baylor College of Medicine, One Baylor Plaza, BCM600, Houston, TX 77030, USA

Yi Cheng, and
Department of Mathematical Sciences, Indiana University South Bend, 1700 Mishawaka Avenue, South Bend, IN 46634, USA

Yu Shen
Department of Biostatistics, UT-M.D. Anderson Cancer Center, PO Box 301402, Unit 411, Houston, TX 77230, USA

Charlotte Hsieh Ahern: chahern@bcm.edu; Yi Cheng: ycheng@iusb.edu; Yu Shen: yshen@mdanderson.org

Abstract

The optimal schedules for breast cancer screening in terms of examination frequency and ages at examination are of practical interest. A decision-theoretic approach is explored to search for optimal cancer screening programs which should achieve maximum survival benefit while balancing the associated cost to the health care system. We propose a class of utility functions that account for costs associated with screening examinations and value of survival benefit under a non-stable disease model. We consider two different optimization criteria: optimize the number of screening examinations with equal screening intervals between exams but without a pre-fixed total cost; and optimize the ages at which screening should be given for a fixed total cost. We show that an optimal solution exists under each of the two frameworks. The proposed methods may consider women at different levels of risk for breast cancer so that the optimal screening strategies will be tailored according to a woman's risk of developing the disease. Results of a numerical study are presented and the proposed models are illustrated with various data inputs. We also use the data inputs from the Health Insurance Plan of New York (HIP) and Canadian National Breast Screening Study (CNBSS) to illustrate the proposed models and to compare the utility values between the optimal schedules and the actual schedules in the HIP and CNBSS trials. Here, the utility is defined as the difference in cure rates between cases found at screening examinations and cases found between screening examinations while accounting for the cost of examinations, under a given screening schedule.

Keywords

Cancer screening; Non-stable disease model; Optimal exam times; Screening sensitivity; Sojourn time; Utility function

1 Introduction

Screening for solid tumors may lead to high cure rates when early detection of the disease is combined with effective treatments. While screening is broadly utilized for detecting breast, colon, cervical, and other cancers, the optimal strategy in terms of examination frequency and ages at screening remains unclear; the guidelines for cancer screening vary even among major cancer societies in the United States and United Kingdom. For example, the National

Cancer Institute recommends screening mammography for breast cancer every 1 to 2 years for women ages 40 and older [9], the American Cancer Society recommends annual mammography beginning at age 40 [1], and the United Kingdom provides screening mammography every three years for women ages 50 to 70 [10]. The U.S. Preventive Services Task Force recently released an updated recommendation statement which recommends biennial screening mammography beginning at age 50 [19]. International debates on screening policies in terms of ideal screening intervals and impacts are ongoing.

Cost-effectiveness analyses via simulations have largely been used to assess the trade-off between the costs and benefits of a finite set of different screening programs [3,5,7,8,14,17]. In contrast to the broad attention paid to modeling the screening process and evaluating the corresponding benefit and cost using simulations, less attention has been paid to the optimization of screening schedules using mathematical models. While a simulation model may assess the benefits and costs within a finite set of screening programs and identify a cost-effective program in the set, a mathematical model has the advantage of searching for a global optimum given a utility function, where optimal strategy refers to the optimal ages at examination or the optimal number of examinations within a continuous screening horizon. Limited literature is available that mathematically explores the optimization of examination schedules in the setting of early detection [6,11,18,20].

Tsodikov and Yakovlev (1991) addressed the problem of optimal screening with non-periodic examinations by minimizing both diagnosis delay time and cost due to screening under a fixed cost. Alternatively, Lee and Zelen (1998) considered using a predetermined threshold value, such as probability of an individual being in the preclinical state, to determine when examinations should be given, or using schedule sensitivity to determine the effectiveness of different screening programs. Zelen (1993) used a weighted utility function, which is linear in the probabilities of screening detection and symptom-based detection, to investigate optimal scheduling for early detection programs under the stable disease model, in which transition from the healthy state into the preclinical state is independent of age. Similarly, Parmigiani (1993) took a decision theoretic approach to account for the expected number of examinations and losses due to mortality and other negative factors. Although his formulation is fairly general, optimal schedules are derived only for two special cases: when there is only one exam, or with perfect exam sensitivity. Though their utility functions are different, both Zelen and Parmigiani independently concluded that optimal equal interval examinations can be derived only under restrictive assumptions such as perfect sensitivity of the examination under the stable disease model. Although the assumption of a stable disease model significantly simplifies the formulations, it can be unrealistic for a potentially long screening horizon, which is defined to be the time interval between the first and last exams. Moreover, equal interval screening examinations may have practical advantages in their convenience and feasibility of implementation at the population level of screening.

Early detection programs often involve huge costs and a long-term commitment from society. Therefore, it is important to evaluate the magnitude of survival benefit and to understand the trade-off between costs of examination and gains in survival. We generalize the work of Zelen (1993) from a stable disease model to the more general non-stable disease model, and from a fixed cost (i.e. number of examinations) to a utility function that includes a non-fixed cost/number of exams. In particular, we investigate optimal scheduling programs by taking into consideration age-dependent preclinical incidence of disease as well as the costs of screening exams and the dollar value of benefit. The number of examinations may or may not be fixed for a given screening horizon. We will explore two distinct frameworks for the utility function given the estimated age-dependent preclinical incidence, screening sensitivity, and sojourn time distribution. Because an equal screening interval has implementation advantages at the population level, one of our focuses is to search for the

optimal equal interval within a specified screening horizon under the nonstable disease model with a screening sensitivity that may not be perfect. We will further derive the optimal ages for examinations within a specified screening horizon given the number of examinations (or a fixed budget) for which the optimal screening program may not have equal screening intervals.

The remainder of the paper is organized as follows. In Section 2 we describe the models and basic notations and prove the existence of the optimal solution under the specified utility functions. In Section 3 we illustrate the methods using numerical studies and compare the utility values of the optimal solution with actual breast cancer screening trials. We provide a concluding discussion in Section 4.

2 Models and Methods

2.1 Preliminaries

Assume that screening examinations are scheduled at ages t_i, $i = 0, \ldots, n$ within a screening horizon of t_0 to T, where $0 < t_0 < t_1 < \cdots < t_n \leq T$. Let β be the sensitivity of the screening test, $0 < \beta \leq 1$. For an individual in the screening population, there are three mutually exclusive possible states: healthy state (H), preclinical state (P), and clinically diagnosed state (C). The utility function is a numerical function that sensibly reflects the effects under consideration, such as screening sensitivity, preclinical disease incidence, and the preclinical sojourn time distribution. Under an assumed stable disease model, Zelen (1993) proposed a utility function that represents a weighted difference between the probability of screening detection ($Pr(SD)=\sum_{i=0}^{n} D_i(\beta)$) and the probability of clinical detection ($Pr(CD)=\sum_{i=1}^{n} I_i(\beta)$) for a screening program with $n + 1$ examinations:

$$U_{n+1}=A\sum_{i=0}^{n} D_i(\beta) - B\sum_{i=1}^{n} I_i(\beta), \tag{2.1}$$

where the weights A and B are non-negative, $A > B$, and

$D_i(\beta) = P$(disease is detected preclinically by the screening examination at t_i),

$I_i(\beta) = P$(disease is clinically diagnosed in $(t_{i-1},t_i]$).

The weights may, for example, represent the probability of a cure when disease is found through a screening examination or a clinical case. The idea is to maximize the difference in the proportion of cures between the two groups with the objective of maximizing the probability of cure in the screening program. In the utility function defined in (2.1), the first term is $Pr(cure\ for\ SD) = Pr(cure|SD)Pr(SD)$, where $Pr(cure|SD) = A$ and $Pr(SD)$ is a function of screening sensitivity and preclinical sojourn time distribution for a given screening program. Similarly, $Pr(cure\ for\ CD) = Pr(cure|CD)Pr(CD)$. Thus, the focus is maximizing the utility function defined as a weighted difference between the two probabilities, with the cure rates for the two types of diagnosis as weights. In his approach, Zelen considered a fixed screening horizon for a fixed number of examinations at times t_0, $t_1, \ldots, t_n$. The optimal examination times, $t_1, \ldots, t_{n-1}$, where $t_0 < t_1 < t_2 < \ldots < t_{n-1} < t_n = T$, are determined by maximizing the utility function (2.1) under these assumptions.

The following notations are helpful in facilitating the computation of the utility because the corresponding terms are building blocks. Let $w(t)$ be the probability of transition from the healthy state to the preclinical state, $P(t)$ be the prevalence of disease at time t, $q(t)$ be the probability density function of the preclinical sojourn time, and $Q(t)$ be the survival function

for $q(t)$. The rate at which an individual who entered the preclinical state in $(t_{i-1},t_i]$ will enter the clinically diagnosed state at time t, denoted as $P_i(t_i)q_i(t)$, can be expressed as

$$P_i(t_i)q_i(t)=\int_{t_{i-1}}^{t_i} w(x)q(t_i - x+t)dx,$$

where

$$P_i(t_i)=\int_{t_{i-1}}^{t_i} w(x)Q(t_i - x)dx.$$

Given the sensitivity β and a screening horizon $[0,T]$ for a fixed number of examinations at times $0 = t_0,t_1,\ldots,t_n = T$, Zelen derived the following general expressions for the probability that the disease is detected at the ith screening exam ($D_i(\beta)$) and the probability that an individual is incident at age t after i exams ($I_i(\beta,t)$). Mathematically,

$$D_0(\beta)=\beta P(t_0) \tag{2.2}$$

$$D_i(\beta)=\beta\left\{\sum_{j=0}^{i-1}(1-\beta)^{i-j}P_j(t_j)Q_j(t_i - t_j)+\int_{t_{i-1}}^{t_i} w(x)Q(t_i - x)dx\right\} \tag{2.3}$$

for $i = 1,\ldots,n$, where expression (2.3) is derived based on the partition that an individual in the preclinical state detected by the screening test at t_i may enter the state in one and only one of the following time intervals: $(t_{j-1},t_j]$ for $j = 1, \ldots, i$. A similar partition is useful in deriving the probability of being a clinical case in the ith interval. The cumulative incidence of clinical (interval) cases in the ith interval is:

$$I_i(\beta)=\int_{t_{i-1}}^{t_i} I_i(\beta,t)dt, \tag{2.4}$$

where

$$I_i(\beta,t)=\sum_{j=0}^{i-1}(1-\beta)^{i-j}P_j(t_j)q_j(t - t_j)+\int_{t_{i-1}}^{t} w(x)q(t - x)dx. \tag{2.5}$$

In contrast to Zelen's framework, we aim at more general settings for the nonstable disease model by allowing the transition rate to be age-dependent and presenting two types of utility functions. In both frameworks, $w(x)$, the age-dependent transition rate from the healthy state to the preclinical state at age x, is assumed to be a continuous function in x. Although settings in earlier works showed that choosing equal intervals may not be optimal unless the exam sensitivity under the stable disease model is 1 [11,20], optimal equal intervals may exist under the more general nonstable disease model. Recall that the screening horizon for our proposed frameworks is $[t_0, T]$ for $0 < t_0 < T$, where t_0 is the age at the initial screening

test. We present the utility structure and the verifications of the existence of an optimal solution for each of the two scenarios in Section 2.2 and Section 2.3, respectively.

2.2 Framework 1: Equal Intervals

Assuming equal intervals, we present a utility function in which the number of exams n, or equivalently the interval length between exams, denoted as Δ, in a screening schedule is unknown. When exams are equally spaced, the n (or Δ) that maximizes the utility function will define the optimal schedule.

Using the expressions for $D_i(\beta)$ and $I_i(\beta)$ given by (2.2) through (2.5), we define the utility function as follows:

$$U(n)=C_1\left(A\sum_{i=0}^{n}D_i(\beta)-B\sum_{i=1}^{n}I_i(\beta)\right)-C_2(n+1), \tag{2.6}$$

where the first term in the difference is the value of the screening benefit and the second term is the cost of screening for a screening program with $n + 1$ exams. If preclinical disease has been identified at t_i, or an incident has occurred in $(t_i,t_{i+1}]$, the diagnosed case will be removed from the screening program thereafter. Here the benefit is defined as the difference in cure rates between preclinical cases found at screening examinations compared to clinical cases, where A is the probability of a cure when disease is found through screening, and B is the probability of a cure for a clinical case. Assuming that the overall cure rate is higher for screening-detected patients than that for non-screening detected patients ($A > B$) and the cure rates are constant, we want to maximize the difference in the proportion of cures between the two groups with the objective of maximizing the probability of cure in the screening program. The use of cure rates as weights avoids potential complications due to length bias and lead-time bias, which may result from using alternative weights such as prolonged survival time. C_1 is the average dollar value of survival benefit due to one life cured, and C_2 is the average cost of a screening exam. With a fixed screening horizon of ages $[t_0, T]$, we assume that $n + 1$ examinations at times $t_0, t_1, \ldots, t_n$, $0 < t_0 < t_1 < \ldots < t_n = T$, are equally spaced with a common difference $\Delta = t_i - t_{i-1}$. Thus we have the following expressions:

$$\Delta=\frac{T-t_0}{n},$$
$$t_i=t_0+i\Delta=t_0+\frac{i}{n}(T-t_0)$$

for $i = 1, \ldots, n$. Using these expressions in the utility function defined by (2.6), we can find the n^* that maximizes $U(n)$ and therefore solve for the optimal Δ^*.

2.2.1 The Existence of an Optimal Solution and its Asymptotic Expression—

We provide a sketch of the proof for the existence of an optimal solution under some mild assumptions. Similar to prior work in the literature [4,16,20], the sojourn time in the preclinical state is assumed to follow an exponential distribution with mean $\mu = 1/\lambda$. Because breast cancer incidence increases with age and the incidence is near zero for women under a certain age (e.g., 18), we then assume that $w(t)$ has a piecewise linear relationship with age of the form $k_1t - k_0$, where k_1 and k_0 are constants, and $w(t) = 0$ for $t < k_0/k_1$. One difficulty in proving the existence of an optimal n is that the utility function (2.6) involves the summations of n, which is the variable to be optimized. The proof thus takes multiple steps that require the closed forms of the summations by using properties of the geometric series in the expression of the utility. Below we provide the key ideas of the proof. A detailed presentation of the proof can be found in Online Resource 1.

Let $D_n=\sum_{i=0}^{n} D_i(\beta)$ and $I_n=\sum_{i=1}^{n} I_i(\beta)$. Using Taylor expansion and properties of the geometric series, we obtain the following approximate expressions for D_n and I_n:

$$D_n \doteq a_0+a_1\frac{1}{n}+O\left(\frac{1}{n^2}\right), \quad I_n \doteq b_1\frac{1}{n}+O\left(\frac{1}{n^2}\right), \tag{2.7}$$

where

$$a_0=k_1\mu\left\{\mu(1-\beta)\exp(-\lambda(t_0-\frac{k_0}{k_1}))+(t_0-\mu)(1-\beta)+\frac{\lambda(T^2-t_0^2)}{2}\right\}-k_0\mu\left\{(1-\beta)+\lambda(T-t_0)\right\}, \tag{2.8}$$

$$a_1=k_1\left\{-\mu\exp(-\lambda(t_0-\frac{k_0}{k_1}))\frac{(1-\beta)}{\beta}-\frac{(T+t_0-\mu)(1-\beta)}{\beta}-\frac{\lambda(T+t_0)(T-t_0)(2-\beta)}{4\beta}\right\}(T-t_0)+k_0\left\{\frac{2(1-\beta)}{\beta}+\frac{(T-t_0)\lambda(2-\beta)}{2\beta}\right\}(T-t_0), \tag{2.9}$$

and

$$b_1=(T-t_0)\left\{(k_1(t_0-\mu)-k_0)\left(\frac{\lambda(T-t_0)(2-\beta)}{2\beta}+\frac{(1-\beta)}{\beta}\right)+k_1(T-t_0)\left(\frac{\lambda(T-t_0)(2-\beta)}{4\beta}+\frac{2-\beta}{2\beta}\right)+\mu k_1\exp(-\lambda(t_0-\frac{k_0}{k_1}))($$

(2.10)

By inserting the approximations of D_n and I_n in (2.7) into the utility function (2.6), we have

$$U(n)=C_1A\left(a_0+a_1\frac{1}{n}\right)-C_1B\left(b_1\frac{1}{n}\right)-C_2(n+1)+O\left(\frac{1}{n^2}\right). \tag{2.11}$$

Based on the above equation, we derive the conditions for the existence of a unique optimal n^*, which may be a continuous positive value, such that $U(n)$ achieves its maximum at $n = n^*$. Let $x = 1/n$, and let h denote the approximation of $U(n)$ by ignoring the last term in (2.11), where

$$h(x)=C_1A(a_0+a_1x)-C_1B(b_1x)-C_2\left(\frac{1}{x}+1\right).$$

Note that these approximations will be more accurate when n is relatively large. The first two derivatives of h can be easily computed as follows:

$$h'(x)=C_1Aa_1 - C_1Bb_1+C_2\frac{1}{x^2}$$

and

$$h''(x)=\frac{-2C_2}{x^3},$$

where $h''(x) < 0$ because C_2 and x are always positive. Setting $h'(x) = 0$, any critical point of $h(x)$, denoted as x_c, must satisfy

$$x_c^2=\frac{C_2}{C_1(Bb_1 - Aa_1)}. \tag{2.12}$$

Given the condition $k_0/k_1 < (t_0 + T - 2\mu)/2$, it is sufficient to ensure $b_1 > 0$ in (2.10). For a_1 in (2.9), $a_1 < 0$ if the sufficient condition $k_0/k_1 < (t_0 + T - \mu)/2$ holds. It is then clear that $k_0/k_1 < (t_0 + T - 2\mu)/2$ is a sufficient condition for both $a_1 < 0$ and $b_1 > 0$ to hold (see Online Resource 1 for detailed verifications of the sufficient conditions). For realistic scenarios pertaining to the natural history and incidence of breast cancer, the aforementioned condition can be easily satisfied when there is a long screening horizon relative to the mean of the sojourn time and the low incidence of the disease. Therefore, the right-hand side of (2.12) is always positive. Thus, the unique maximum of h is

$$x_c=\sqrt{\frac{C_2}{C_1(Bb_1 - Aa_1)}}$$

according to the second derivative test since $h''(x) < 0$ for any $x > 0$.

The asymptotic optimal Δ^* can be directly obtained from the expression of the asymptotic optimal n^*:

$$n^*=\sqrt{\frac{C_1(Bb_1 - Aa_1)}{C_2}} \text{ and } \Delta^*=(T - t_0)\sqrt{\frac{C_2}{C_1(Bb_1 - Aa_1)}}.$$

In order to have $n^* \geq 1$, we require

$$\frac{C_1(Bb_1 - Aa_1)}{C_2} \geq 1.$$

Because the benefit of saving a life (C_1 converted to a dollar value) is much larger than the dollar value for one screening examination, C_2 (i.e., $C_1 \gg C_2$), this condition is satisfied in

general. Otherwise, any regular screening examination would not be a realistic option at the population level.

2.3 Framework 2: Fixed Budget

This framework applies to the setting of a limited healthcare budget for a screening program. Under a fixed budget, or equivalently a fixed number of examinations, we present a utility function where, given the starting age t_0, the ages of examination $\{t_1, \ldots, t_n\}$ are to be determined to maximize the utility function. This optimal schedule is solved for a prespecified number of exams and screening horizon. The utility function for a screening program with $n + 1$ exams at times $t_0, t_1, \ldots, t_n$ is defined as follows:

$$U(t_1, \ldots, t_n) = C_1 \left(A \sum_{i=0}^{n} D_i(\beta, t_0, \ldots, t_i) - B \sum_{i=1}^{n+1} I_i(\beta, t_0, \ldots, t_i) \right) - C_2(n+1), \tag{2.13}$$

where the expressions $D_i(\beta) = D_i(\beta, t_0, \ldots, t_i)$ and $I_i(\beta) = I_i(\beta, t_0, \ldots, t_i)$ are given by (2.3), (2.4), and (2.5), with the first exam at age t_0. Because n in (2.13) is fixed, the last term reflects the total cost of screening examinations and plays no role in determining optimal screening ages, $t_0, \cdots, t_n$. Under the stable disease model, the above utility function reduces to Zelen's utility function.

We fix the age of first examination (t_0) and the maximum age of examination (T), for example $t_0 = 40$ and $T = 79$, according to current screening guidelines. We are thus searching for the optimal schedule of $n + 1$ examinations at times $t_0, t_1, \ldots, t_n$, which may or may not be equally spaced at interval lengths of $\Delta_1, \Delta_2, \ldots, \Delta_n$, respectively, and where $\Delta_i = t_i - t_{i-1}$ for $i = 1, \ldots, n$. The proof that an optimal *solution* exists under this framework is given below.

2.3.1 The Existence of an Optimal Solution with a Fixed Budget—We can prove that $U(t_1, \ldots, t_n)$ achieves a maximum at $\{t_1^*, \ldots, t_n^*\}$ by using the Extreme Value Theorem [13]. Because the examination ages are bounded by the starting and ending ages t_0 and T, the set $\{(t_1, \ldots, t_n): t_0 \le t_1 \le t_2 \le \ldots \le t_n \le T\}$ is compact. It is clear that the utility function U is a composition of continuous functions on the compact set $\{(t_1, \ldots, t_n): t_0 \le t_1 \le t_2 \le \ldots \le t_n \le T\}$. Therefore, the Extreme Value Theorem implies that a solution of $\{t_1, \ldots, t_n\}$ exists to maximize $U(t_1, \ldots, t_n)$.

3 Numerical Studies

We apply the theoretical results developed in Section 2 to search for the optimal screening program for breast cancer under the two distinct frameworks. We also give examples using sensitivity and preclinical sojourn time estimates from the Health Insurance Plan (HIP) trial and Canadian National Breast Screening Study (CNBSS) to illustrate the proposed models and compare the utility values between the optimal schedules and the actual schedules in the two trials.

3.1 Illustration of Framework 1

To illustrate the results for finding the optimal number of equally spaced exams within a fixed screening horizon, we specify input parameter values for the utility function $U(n)$ defined in (2.6). The input values listed in Table 1 are based on practical estimators from previous breast cancer screening studies. The cure rates $A = 0.8$ and $B = 0.52$ were estimated based on a long-term follow-up study of survival in breast cancer patients [12]. The cure rate for each cohort is a weighted average of cure rates given patients disease stage at diagnosis,

and the weights represented by the percentage of early (and late) stage disease serve the purpose of distinguishing between screen detected cases and clinically detected cases. We consider a range of reasonable values for the mean sojourn time in the preclinical state and the sensitivity of mammography examination [15,16]. The ratio of the cost of examination to value of life saved (C_1/C_2) is varied at \$50, \$100, or \$200, where we fix the exam cost at \$150 and choose the value of life saved due to a cure using the commonly accepted threshold that states that a year of life is valued at \$50,000. For example, if the average expected gain in survival due to a cure for screening versus clinical detection in a screened cohort made up of healthy women and women with disease is 2, 4, or 7 months, then the corresponding value of life saved is \$7,500, \$15,000, or \$30,000 based on a \$50,000 per year threshold value.

The values k_0 and k_1 in the transition (to the preclinical state) rate function $w(x)$ – also called the preclinical incidence function – are chosen to approximately reflect an age-dependent incidence function of a population with an average risk for breast cancer ($k_1 = 0.0001$ and $k_0 = 0.0025$) and an increased risk for breast cancer ($k_1 = 0.00015$ and $k_0 = 0.0030$). The values for an average-risk population were chosen to reflect a rough linear approximation of incidence estimates from the North American Association of Central Cancer Registries [2]. The corresponding lifetime risks for these estimates are approximately 15% and 27% for the average-risk and increased-risk populations, respectively. According to the assumptions for Framework 1 in Table 1, the Δ (and thus n) that maximizes the utility can be solved using the expressions in Section 2.2.1, where the optimal number of examinations $n^* = (T - t_0)/\Delta^*$ is rounded down to the nearest integer (Table 2). The results for $C_1/C_2 = 50$ and $C_1/C_2 = 100$ are presented. A visualization of the utility function versus n is shown in Figure 1 for the case $C_1/C_2 = 100$, $\beta = 0.8$, and $\mu = 3$ for the two risk scenarios and illustrates the downward concavity of the function.

Under the framework with a constant screening interval, the optimal schedule for an average-risk population with a mean sojourn time of $\mu = 2$ and $C_1/C_2 = 100$ in Table 2 includes exams in approximately 1.5 to 2 year intervals when the sensitivity varies from 0.7 to 0.9. The optimal screening interval of two years is consistent with the most recent recommendation from the U.S. Preventive Services Task Force for women at average risk for breast cancer. As expected, as the mean of sojourn times (μ) increases, the length of the optimal screening interval Δ^* also increases, i.e., the examinations may be spaced further apart. Also, as the screening sensitivity (β) increases, the interval between the exams Δ^* increases. Better exam performance allows for an increased reliability of the diagnosis and thus less frequent examinations. The results also showed that the optimal time duration between examinations is shorter when the disease incidence is higher in the higher-risk group. This implies that a higher-risk group requires more frequent examinations to achieve the maximum utility. To validate this conclusion against a potentially lower cure rate associated with high-risk disease, we used lower cure rates of A=0.70 and B=0.40 in a sensitivity analysis and found that the same relative results hold. We also considered the situations where screening and early detection may not always offer benefit (e.g. prostate cancer) by using a low cure rate for screening detection of A=0.60, so that the difference between A and B=0.52 is small. We found that the optimal intervals between exams are longer when the benefit of screening is smaller.

Depending on the ratio of costs C_1/C_2, the optimal n and Δ vary. As the ratio increases, the optimal Δ becomes smaller. This reflects that when the cost for screening examinations is much smaller relative to the value of life to be gained, there is an incentive to have more frequent or intensive screening examinations, which leads to shorter screening intervals.

3.2 Illustration of Framework 2

With a fixed number of screening exams, we illustrate the finding of an optimal set of exam ages $\{t_1, \ldots, t_n = T\}$ within a screening horizon $[t_0, T]$ given the data inputs in Table 1 and using the utility function $U(t_1, \ldots, t_n)$ defined in (2.13). We search for the optimal screening ages while setting the starting age at 40 and ending age at 60 or 80 under various combinations of mean sojourn times and screening sensitivities. We consider two scenarios for the given number of exams, 5 and 10 (n = 4,9), for illustration, using the same two incidence profiles $w(x)$ as in the previous illustration. Note that under this framework, the optimal exam times do not depend on the exam costs or life value, which is pre-fixed. The results listed in Table 3 are obtained by solving the set of exam ages $t_1, \ldots, t_n$ that maximize the utility defined by (2.13).

We made the following general observations. First, an increase in mean sojourn time in the preclinical state leads to an increase in intervals between exams, except in the first interval. An increase in screening sensitivity has a similar small impact on the optimal screening ages. Note that breast cancer incidence increases with age, which results in a much larger interval between the first and second examinations than for the subsequent intervals, and the intervals between exams consistently decrease with age. It is also expected that the intervals between exams are shorter for shorter screening horizons and longer for longer screening horizons when the total number of screening exams is fixed. Only the results of the average risk scenario are presented in Table 3, however we found it interesting to note from our results that across the two different risk scenarios the screening interval is slightly smaller under the increased-risk scenario for the first interval only; i.e., the second examination is given at an earlier age than in the average-risk scenario, whereas the subsequent screening intervals are close to those under the average-risk scenario. In essence, optimal screening ages are very similar under the two different risk scenarios when the total number of the screening exams and the screening horizon are prefixed. Although we did not present the results for $t_0 = 50$, the same observations carry over.

As suggested by a reviewer, we conducted a sensitivity analysis which assumes a two-stage disease model where the sensitivity is constant in each stage. This takes into account the potential increase in sensitivity towards the end of the preclinical sojourn time, when the tumor size is likely to have increased. To assess the impact, we incorporated a higher sensitivity at the time of the last two exams. Results showed that increasing the sensitivity in the second stage (at the last two exams) results in minor changes in the optimal schedule.

3.3 Examples

To further illustrate the proposed methods, we apply the two frameworks to assess the utility/optimality of the screening interval of two actual breast cancer screening trials. Given estimates of the mean preclinical sojourn time and exam sensitivity from the HIP and CNBSS studies [15], we compute optimal screening intervals (ages) of examination and compare the corresponding utility values between the computed optimal schedule and the actual HIP and CNBSS screening schedule for each framework. We use the same parameters as those used in the numerical illustrations for an average-risk cohort: $A = 0.80$, $B = 0.52$, $C_1/C_2 = 100$, $k_1 = 0.0001$ and $k_0 = 0.0025$.

3.3.1 Framework 1—Under Framework 1, we use estimates of the mean sojourn times and exam sensitivities of the HIP and CNBSS trials from Shen and Zelen (2001). If the screening horizon is between ages $t_0 = 40$ and $T = 80$, the estimated optimal constant interval between exams for the HIP estimates of $\mu = 2.5$ years and $\beta = 0.70$ is $\Delta^* = 1.727$ years, which is equivalent to the optimal number of exams $n^* = 24$. In the HIP study, participants were given screening examinations annually. If we compare the utility values

using the calculated optimal time interval $\Delta^* = 1.727$ and the time interval used in the study, $\Delta = 1.0$, our results show that the computed Δ^* is superior to the annual exams given in the study since $U(\Delta^* = 1.727) > U(\Delta = 1)$, where $U(\Delta^* = 1.727) = -5406.6$ and $U(\Delta = 1.0) = -79063.0$.

Similarly, we consider ages of entry $t_0 = 40$ or $t_0 = 50$ in the CNBSS trial with $T = 80$. For age of entry of 40 years, the estimates of $\mu = 1.9$, $\beta = 0.9$ yield an optimal interval between exams of $\Delta^* = 1.886$ with $n^* = 22$. If the age of entry is 50 years, the optimal interval between exams is $\Delta^* = 2.053$ with $n^* = 15$ using the estimates of $\mu = 3.1$, $\beta = 0.8$. Again, $U(\Delta^*) > U(\Delta = 1)$ in both cases, which supports the conclusion above.

3.3.2 Framework 2—We also assess the HIP and CNBSS screening programs under the second framework. Since the actual studies each allowed for a maximum of four screening exams with varying ages of entry, we present the results under the different screening horizons listed in Table 4, for a pre-fixed number of exams $n + 1 = 4$. The computed optimal examination ages are not equally spaced as they were in the studies; the time intervals instead decrease as age increases within the screening horizon. Similar to the comparison between the average-risk and high-risk scenarios for this framework in Section 3.2, the second exam after t_0 is given earlier as t_0 increases since incidence increases with age. Note that given the estimated inputs in the HIP study, the optimal intervals are almost constant at close to three years for the screening horizon of 60 to 69, given a total of four examinations. However, for the younger screening cohort [40,49), the length of the first screening interval is one year longer than that of the last interval. We compare the utility values using the computed optimal exam times to the utility values assuming annual examinations beginning at age t_0. Clearly, the utility for the unequally spaced examinations is superior to the utility of using annual exams in both the HIP and CNBSS trials.

4 Discussion

Decision theoretic approaches are powerful tools used to search for optimal cancer screening schedules while taking into account the trade-off between the value of benefit and cost of screening in the utility function. We have shown with mathematical proofs that optimal solutions for screening programs exist for each of the two proposed frameworks under the non-stable disease model. Compared to the existing theoretical approaches in the literature, our proposed models address more general and practical problems by balancing the trade-off for monetary costs of a screening program under the non-stable disease model, with the stable disease scenario as a special case. The optimal schedule obtained by maximizing the corresponding utility function will depend on the screening sensitivity, preclinical sojourn time distribution, age-dependent disease incidence, cost of each screening examination, and the value of screening benefit, which are all incorporated into our models.

Under the first framework, the derived asymptotic algebraic expressions ensure the existence of the optimal solution and also provide the solution for the optimal number of examinations under the assumption of equal intervals. For the purpose of policy recommendation at the population level, the implementation of equal screening intervals may be more feasible than the implementation of unequally spaced exams. This framework accounts for the costs of screening and the dollar value of life saved through screening detection versus clinical detection. Under the second proposed framework, the optimal ages of examination can be found given a prespecified number of exams (equivalent to a fixed cost) within a given screening horizon. With the current healthcare system, there are often limited budgets for cancer prevention and early detection programs. For a system with such a constraint, our model can estimate the best ages for receiving a finite number of screening examinations to achieve the most benefit within a given age horizon.

In contrast to theoretical approaches such as the ones described in this article, simulation-based approaches can be used to identify a cost-effective screening schedule from a finite set of pre-determined schedules, as opposed to a global optimum. Although it is not impossible, it could be extremely computationally intensive to perform a grid search for the global optimum in simulation studies, especially with multi-dimensional parameters. Simulation-based studies have more flexibility but are limited by their computational intensity. They require the generation of subject-level data including the natural history and individual disease progression for each of a large number of subjects. Characteristics of each subject over time are simulated using assumed distributions and parameters such as disease incidence, preclinical sojourn time, and age of onset. As an example, simulation-based studies [7] were used as support to the U.S. Preventive Services Task Force's recently updated screening recommendation. In the studies, 20 pre-identified screening strategies were compared using outcomes such as number of mammograms, life-years gained, and overdiagnosis. The results of this study showed that of the 20 screening strategies that were considered, biennial mammography beginning from age 50 achieved the most benefit with less harm. While this result is compelling, the simulation-based model is limited to comparing the 20 pre-identified strategies. In contrast, our mathematical model has the ability to identify less conventional screening schedules (e.g. non-constant intervals between exams) that may in fact be more beneficial with less risk/cost, while using less time and computer resources and similarly accounting for the key data inputs such as exam sensitivity and preclinical sojourn time.

One limitation of the use of mathematical models is the reduced flexibility of individualized parameter assumptions. In order to facilitate a set of expressions that are easy to work with, we incorporate population-level assumptions for key parameters such as disease incidence, exam sensitivity, and exponential sojourn time distribution. For instance, the assumption that the sojourn time in the preclinical state is exponentially distributed may not necessarily be the best fit, but it is widely used to model the sojourn time and seems to be a reasonable choice that leads to meaningful results in breast cancer screening studies. Alternative models for the preclinical incidence and distributions for the sojourn time where the mode is in the interior of the support, such as the lognormal or Weibull distributions, may be further explored in future work. Presently, incorporating alternative sojourn time distributions such as those stated above pose algebraic difficulties in multiple integration. However the reader should be aware that potential changes in outcomes may result from using an alternative distribution which may reflect a more realistic scenario of tumor growth.

Moreover, we have not explicitly accounted for the cost of false-positive exams in the utility functions or included a component for competing risks for death. However, the average cost due to false-positive screening examinations may be incorporated into the cost of screening examinations given the information for the rate of false-positive tests for a screening modality. Additionally, while we are aware that exam sensitivity may also be age- and tumor size-dependent, the inclusion of such a dependency in sensitivity is theoretically challenging to prove the existence of the optimal solution and we defer this to future research. In the meantime, a sensitivity analysis was performed to assess the impact of a constant sensitivity within each stage of a two-stage disease model, where the sensitivity is increased in the second stage. We found that a non-constant sensitivity has a minor impact on the results.

We acknowledge that both simulation-based and theoretical approaches have their pros and cons. A method combining the two complementary approaches may allow more efficient investigation of cost-effective screening policies in future studies. Although the approaches for the two studies are different, our study results are reasonably consistent with results from the simulation-based studies by Mandelblatt et al. (2009) [7]; both studies found that for a

fixed screening interval, screening every two years achieves comparable or increased benefit compared to annual screening while balancing the cost/harm for women at average risk of breast cancer.

This work was intended to contribute to the identification of optimal screening policies by proposing a set of utility functions that integrate different components, including both survival benefit and costs associated with screening examinations, under a nonstable disease model. Using the proposed model structures, we can search for different optimal schedules for patients at different levels of risk for disease. The proposed methods may guide health policy makers by shedding some light on alternative, less conventional optimal strategies for cost-effective screening programs. Although our methods and models are motivated by and applied to breast cancer early detection research, they can be used to evaluate any cancer or chronic disease for which screening is contemplated.

5 Electronic Supplementary Materials

Online Resource 1, referenced in Section 2, is available as electronic supplementary material.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was completed in part while the second author was visiting the Department of Biostatistics at the University of Texas, M.D. Anderson Cancer Center. The authors would like to thank Dr. Marvin Zelen for helpful discussions. This work was supported by a National Institutes of Health grant CA0769466.

References

1. American Cancer Society. Detailed guide: breast cancer. Can breast cancer be found early?. 2006 [Accessed Dec 2008]. http://www.cancer.org/
2. American Cancer Society. Breast cancer facts and figures 2009–2010. 2010 [Accessed Aug 2010]. http://www.cancer.org/Research/CancerFactsFigures/BreastCancerFactsFigures/
3. Boer R, Plevritis S, Clarke L. Diversity of model approaches for breast cancer screening: a review of model assumptions by The Cancer Intervention and Surveillance *Network* (CISNET) Breast Cancer Groups. Stat Methods Med Res. 2004; 13:525–538. [PubMed: 15587437]
4. Day NE, Walter SD. Simplified models of screening for chronic disease: estimation procedures from mass screening programmes. Biometrics. 1984; 40:1–14. [PubMed: 6733223]
5. Knudsen AB, McMahon PM, Gazelle GS. Use of modeling to evaluate the cost-effectiveness of cancer screening programs. J Clin Oncol. 2007; 25:203–208. [PubMed: 17210941]
6. Lee SJ, Zelen M. Scheduling periodic examinations for the early detection of disease: applications to breast cancer. J Am Stat Assoc. 1998; 93:1271–1281.
7. Mandelblatt JS, Cronin KA, Bailey S, Berry DA, deKonig HJ, Draisma G, Huang H, Lee SJ, Munsell M, Plevritis SK, Ravdin P, Schechter CB, Sigal B, Stoto MA, Stout NK, van Ravesteyn NT, Venier J, Zelen M, Feuer EJ. Effects of mammography screening under different screening schedules: model estimates of potential benefits and harms. Ann Intern Med. 2009; 151:738–747. [PubMed: 19920274]
8. Mandelblatt JS, Schechter CB, Yabroff KR, Lawrence W, Dignam J, Muennig P, Chavez Y, Cullen J, Fahs M. Benefits and costs of interventions to improve breast cancer outcomes in African American women. J Clin Oncol. 2004; 22:2554–2566. [PubMed: 15173213]
9. National Cancer Institute. Breast cancer: screening and testing. 2007 [Accessed Dec 2008]. http://www.cancer.gov/cancertopics/screening/breast/

10. NHS Cancer Screening Programmes. NHS breast screening programme. 2008 [Accessed Dec 2008]. http://cancerscreening.org.uk/breastscreen/
11. Parmigiani G. On optimal screening ages. J Am Stat Assoc. 1993; 88:622–628.
12. Rosen PR, Groshen S, Saigo PE, Kinne DW, Hellman S. A long-term follow-up study of survival in stage I (T1N0M0) and stage II (T1N1M0) breast carcinoma. J Clin Oncol. 1989; 7:355–366. [PubMed: 2918331]
13. Rudin, W. Principles of mathematical analysis. 3. McGraw-Hill Science; New York: 1976.
14. Shen Y, Parmigiani G. A model-based comparison of breast cancer screening strategies: mammograms and clinical breast examination. Cancer Epidemiol Biomarkers Prev. 2005; 14:529–532. [PubMed: 15734983]
15. Shen Y, Zelen M. Screening sensitivity and sojourn time from breast cancer early detection trials: mammograms and physical examinations. J Clin Oncol. 2001; 9:3490–3499. [PubMed: 11481355]
16. Shen Y, Zelen M. Parametric estimation procedures for screening programmes: stable and nonstable disease models for multimodality case finding. Biometrika. 1999; 86:503–515.
17. Stout NK, Rosenberg MA, Trentham-Dietz A, Smith MA, Robinson SM, Fryback DG. Retrospective cost-effectiveness analysis of screening mammography. J Natl Cancer Inst. 2006; 98:774–782. [PubMed: 16757702]
18. Tsodicov AD, Yakovlev AY. On the optimal policies of breast cancer screening. Math Biosci. 1991; 107:21–45. [PubMed: 1806106]
19. U.S. Preventive Services Task Force. Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med. 2009; 151:716–726. [PubMed: 19920272]
20. Zelen M. Optimal scheduling of examinations for the early detection of disease. Biometrika. 1993; 80:279–293.

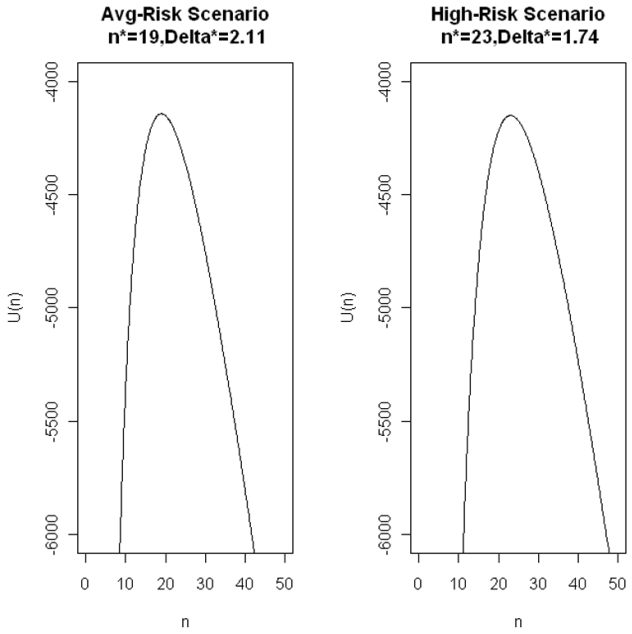

Fig. 1.

Plots of utility function against n when $\frac{c_1}{c_2}$=100, β = 0.8, μ = 3, under the average-risk (k1 = 0.0001, k_0 = 0.0025) and high-risk (k_1 = 0.00015, k_0 = 0.0030) scenarios.

Table 1

Parameter assumptions under Frameworks 1 and 2.

Parameter	Value
Framework 1	
Starting age of screening	$t_0 = 40$
Ending age of screening	$T = 80$
Cure rate (screening-detected)	$A = 0.8$
Cure rate (clinically detected)	$B = 0.52$
Sojourn time distribution	*Expo*(*mu;*), *mu;* = {2,4}
Exam sensitivity	$\beta = \{0.7,0.8,0.9\}$
Value of life saved by cure	$C_1 = \{\$7, 500, \$15, 000, \$30, 000\}$
Exam cost	$C_2 = \$150$
Preclinical Incidence Function*	$\{k_0, k_1\} = \{0.0001, 0.0025\}, \{0.00015, 0.0030\}$
Framework 2	
Starting age of screening	$t_0 = \{40, 50\}$
Maximum age of screening	$T = \{60,80\}$
Cure rate (screening-detected)	$A = 0.8$
Cure rate (clinically detected)	$B = 0.52$
Sojourn time distribution	*Expo*(*mu;*), *mu;* = {2,4}
Exam sensitivity	$\beta = \{0.7,0.9\}$
Number of Exams	$n + 1 = \{5,10\}$
Preclinical Incidence Function*	$\{k_0, k_1\} = \{0.0001, 0.0025\}, \{0.00015, 0.0030\}$

*$w(t) = k_1 t - k_0$

Table 2

Framework 1: Table of optimal n and Delta under different scenarios for $\frac{C_1}{C_2}=50$, or equivalently C_2 = \$150 and C_1 = \$7500, or the value of about 2 months; and $\frac{C_1}{C_2}=100$, or equivalently C_2 = \$150 and C_1 = \$15000, or the value of about 4 months

		Increased-Risk $k_1 = 0.00015$ $k_0 = 0.0030$	Average-Risk $k_1 = 0.0001$ $k_0 = 0.0025$
		$\frac{C_1}{C_2}=50$	
$\mu = 2$	$\beta = 0.7$	$n^* = 23$, $\Delta^* = 1.807$	$n^* = 19$, $\Delta^* = 2.194$
	$\beta = 0.8$	$n^* = 20$, $\Delta^* = 2.021$	$n^* = 17$, $\Delta^* = 2.454$
	$\beta = 0.9$	$n^* = 18$, $\Delta^* = 2.252$	$n^* = 15$, $\Delta^* = 2.735$
$\mu = 4$	$\beta = 0.7$	$n^* = 16$, $\Delta^* = 2.513$	$n^* = 14$, $\Delta^* = 3.050$
	$\beta = 0.8$	$n^* = 15$, $\Delta^* = 2.823$	$n^* = 12$, $\Delta^* = 3.427$
	$\beta = 0.9$	$n^* = 13$, $\Delta^* = 3.163$	$n^* = 11$, $\Delta^* = 3.841$
		$\frac{C_1}{C_2}=100$	
$\mu = 2$	$\beta = 0.7$	$n^* = 32$, $\Delta^* = 1.278$	$n^* = 26$, $\Delta^* = 1.552$
	$\beta = 0.8$	$n^* = 28$, $\Delta^* = 1.429$	$n^* = 24$, $\Delta^* = 1.735$
	$\beta = 0.9$	$n^* = 26$, $\Delta^* = 1.592$	$n^* = 21$, $\Delta^* = 1.934$
$\mu = 4$	$\beta = 0.7$	$n^* = 23$, $\Delta^* = 1.777$	$n^* = 19$, $\Delta^* = 2.157$
	$\beta = 0.8$	$n^* = 21$, $\Delta^* = 1.996$	$n^* = 17$, $\Delta^* = 2.423$
	$\beta = 0.9$	$n^* = 18$, $\Delta^* = 2.237$	$n^* = 15$, $\Delta^* = 2.716$

Table 3

Framework 2: Table of optimal Δ_i and t_i, rounded to the nearest tenth, under different scenarios for $T = \{60,80\}$, $t_0 = 40$, $\beta = \{0.7, 0.9\}$, $\mu = \{2, 4\}$, $n = \{4, 9\}$, $A = 0.8$, and $B = 0.52$ for the average-risk scenario ($k_1 = 0.0001$, $k_0 = 0.0025$).

			$\beta = 0.7$		$\beta = 0.9$	
			$\mu = 2$	$\mu = 4$	$\mu = 2$	$\mu = 4$
$T = 60$	$n = 4$	t_1, Δ_1	48.2,8.2	47.3,7.3	47.5,7.5	46.8,6.8
		t_2, Δ_2	52.4,4.2	52.1,4.8	52.0,4.4	51.8,5.0
		t_3, Δ_3	55.8,3.4	56.1,4.0	55.6,3.6	56.1,4.3
		t_4, Δ_4	58.7,2.9	59.7,3.6	58.8,3.2	60.0,3.8
	$n = 9$	t_1, Δ_1	43.6,3.6	43.1,3.1	43.2,3.2	42.9,2.9
		t_2, Δ_2	46.4,2.8	45.8,2.8	46.0,2.7	45.6,2.7
		t_3, Δ_3	48.8,2.4	48.3,2.5	48.4,2.4	48.0,2.4
		t_4, Δ_4	51.0,2.2	50.6,2.3	50.6,2.2	50.3,2.3
		t_5, Δ_5	53.0,2.0	52.7,2.1	52.7,2.1	52.4,2.1
		t_6, Δ_6	54.9,1.9	54.7,2.0	54.7,2.0	54.5,2.0
		t_7, Δ_7	56.6,1.8	56.7,1.9	56.5,1.9	56.4,1.9
		t_8, Δ_8	58.3,1.7	58.5,1.8	58.3,1.8	58.3,1.9
		t_9, Δ_9	59.9,1.6	60.0,1.5	60.0,1.7	60.0,1.7
$T = 80$	$n = 4$	t_1, Δ_1	64.5,24.5	59.6,19.6	63.4,23.4	58.2,18.2
		t_2, Δ_2	69.8,5.3	67.1,7.4	69.1,5.7	66.2,8.0
		t_3, Δ_3	74.1,4.3	73.0,5.9	73.8,4.7	72.6,6.5
		t_4, Δ_4	77.9,3.8	78.1,5.1	77.9,4.1	78.3,5.7
	$n = 9$	t_1, Δ_1	54.2,14.2	49.7,9.7	52.2,12.2	48.3,8.3
		t_2, Δ_2	59.0,4.7	55.4,5.6	57.3,5.0	53.9,5.7
		t_3, Δ_3	62.8,3.8	60.0,4.6	61.3,4.1	58.7,4.7
		t_4, Δ_4	66.1,3.3	64.0,4.0	64.9,3.6	62.9,4.2
		t_5, Δ_5	69.1,3.0	67.7,3.6	68.2,3.3	66.8,3.9
		t_6, Δ_6	71.8,2.8	71.0,3.4	71.2,3.0	70.3,3.6

	$\beta = 0.7$		$\beta = 0.9$	
	$\mu = 2$	$\mu = 4$	$\mu = 2$	$\mu = 4$
t_7, Δ_7	74.4,2.6	74.2,3.2	74.0,2.8	73.7,3.4
t_8, Δ_8	76.8,2.4	77.2,3.0	76.7,2.7	76.9,3.2
t_9, Δ_9	79.2,2.3	80.0,2.8	79.3,2.6	80.0,3.1

Table 4

Example of second framework for 4 exams using corresponding sensitivity and preclinical sojourn time estimates from the HIP and CNBSS studies.

	Screening Horizon $[t_0, T)$	Optimal ages of exam $t_1^*, t_2^*, t_3^*, t_4^*$	Optimal intervals between exams $\Delta_1^*, \Delta_2^*, \Delta_3^*, \Delta_4^*$	$U(t_1^*, t_2^*, t_3^*, t_4^*)$	$U(t_1,t_2,t_3,t_4)^a$
HIP	[40,49)	40.0,43.4,46.1,48.6	3.4,2.8,2.4	−500.6	−605.6
	[50,59)	50.0,53.0,55.8,58.4	3.0,2.8,2.6	−439.2	−583.9
	[60,69)	60.0,62.9,65.7,68.3	2.9,2.8,2.6	−377.1	−562.1
CNBSS	[40,49)	40.0,43.3,46.0,48.4	3.3,2.7,2.4	−487.4	−611.8
	[50,59)	50.0,53.2,56.2,58.9	3.2,3.0,2.8	−295.4	−498.0

[a] $t_i = t_0 + i$