The Influence of Disease Risk on the Optimal Time Interval between Screens for the Early Detection of Cancer: A Mathematical Approach

James F. O'Mahony, PhD, Joost van Rosmalen, PhD, Nino A. Mushkudiani, PhD, Frans-Willem Goudsmit, MSc, Marinus J. C. Eijkemans, PhD, Eveline A. M. Heijnsdijk, PhD, Ewout W. Steyerberg, PhD, J. Dik F. Habbema, PhD

The intervals between screens for the early detection of diseases such as breast and colon cancer suggested by screening guidelines are typically based on the average population risk of disease. With the emergence of ever more biomarkers for cancer risk prediction and the development of personalized medicine, there is a need for risk-specific screening intervals. The interval between successive screens should be shorter with increasing cancer risk. A risk-dependent optimal interval is ideally derived from a cost-effectiveness analysis using a validated simulation model. However, this is time-consuming and costly. We propose a simplified mathematical approach for the exploratory analysis of the implications of risk level on optimal screening interval. We develop a mathematical model of the optimal screening interval for breast cancer screening. We verified the results by programming the simplified model in the MISCAN-Breast microsimulation model and comparing the results. We validated the results by comparing them with the results of a full, published MISCAN-Breast cost-effectiveness model for a number of different risk levels. The results of both the verification and validation were satisfactory. We conclude that the mathematical approach can indicate the impact of disease risk on the optimal screening interval. ***Key words:*** *breast cancer; mammography; oncology; cost-effectiveness analysis; simulation methods.* ***(Med Decis Making XXXX;XX:XX–XX)***

Screening for preclinical disease can improve health outcomes by enabling early treatment of cancer and its precursor lesions. Screening has been proven to improve health outcomes in a number of cancers, including breast cancer.[1–4] Increasing the screening frequency can yield greater health gains, but these are traded off against the greater costs and risks of more intensive screening, such as false positives, overdiagnosis, and repeated x-ray exposure.

Randomized controlled trials have been crucial in establishing the effectiveness of breast, lung, and

Received 30 July 2012 from Department of Public Health, Erasmus MC–University Medical Centre Rotterdam, Rotterdam, the Netherlands (JFO'M, JvR, FWG, EAMH, EWS, JDFH); Department of Health Policy and Management, Trinity College Dublin, Dublin, Ireland (JFO'M); Department of Biostatistics, Erasmus MC–University Medical Centre Rotterdam, Rotterdam, the Netherlands (JvR); Statistics Netherlands, Den Haag, the Netherlands (NAM); and Department of Biostatistics, UMC–University Medical Centre Utrecht, Utrecht, the Netherlands (MJCE). J. F. O'Mahony was supported by a PhD fellowship provided by the Health Research Board of Ireland. The Health Research Board of Ireland is a publicly funded body that supports research into population health. E. W. Steyerberg and J. D. F. Habbema received grand support from ZonMw, The Netherlands Organisation for Health Research and Development. ZonMw is a publicly funded body that supports research in health and health care in the Netherlands. Revision accepted for publication 22 February 2014.

DOI: 10.1177/0272989X14528380

Supplementary material for this article is available on the *Medical Decision Making* Web site at http://mdm.sagepub.com/supplemental.

Address correspondence to James F. O'Mahony, Department of Health Policy and Management, Trinity College Dublin, 2-4 Foster Place, Dublin, D2, Ireland; e-mail: omahonj1@tcd.ie.

colon cancer screening. These trials assessed various screening age ranges and screening frequencies. However, the trialed screening strategies are not necessarily optimal. It would be impossible to conduct controlled trials to identify optimal strategies due to the number of possible alternatives, time constraints, and ethical considerations.[5] Therefore, simulation models of screening based on published trial results have been developed to estimate the effectiveness and cost-effectiveness of alternative screening strategies. These cost-effectiveness analyses (CEAs) typically only estimate the optimal age range and interval for screening average-risk populations. Thus, strictly speaking, the optimal policies suggested by such CEAs only apply to average-risk individuals.

Those at higher risk of disease have greater potential to gain from more intensive screening than those at average risk. This is reflected in screening guidelines for those with a family history of breast or colorectal cancer, which recommend screening with a higher frequency and earlier starting age than for the general population.[6,7] However, since CEA estimates are usually unavailable for high-risk groups, screening guidelines for such groups are typically based on expert opinion.

The question of optimal screening intensity is of increased significance with improved knowledge of biomarkers of individual disease risk. These risk markers include genetic and biochemical markers but also family history and behavioral risk factors such as smoking, diet, and physical activity. Growing knowledge of prognostic markers will lead to risk groups becoming increasingly differentiated, thus expanding the scope for risk-specific screening strategies.

A full cost-effectiveness analysis for each level of disease risk is time-consuming and costly and requires the input of experienced modelers. There is a need for a quick first approximation of what an elevated risk implies for the screening interval. The purpose of this paper is therefore to present an easy-to-apply tool for a first estimation of the optimal screening interval for high-risk groups. This will provide an informed basis for adapting screening guidelines for high-risk groups where specific CEA estimates are not yet available.

Previous Literature

Most current screening models are simulation models, using cohort or microsimulation approaches. Simulation is useful as it can easily handle complex screening models with many parameters and health states. Knudsen and others[8] and Ramsey and others[9] have provided useful overviews of simulation models and their application to cancer screening. The MISCAN microsimulation model used in this study is an example of such a simulation model.[10–12]

The cost-effectiveness of screening can also be estimated using mathematical models. These models offer the advantages that they can be analytically solved for optimal solutions and that these solutions can then be rapidly applied in other applications. However, requiring that the solution to a model be mathematically tractable may limit the model's level of realism. Alagoz and others[13] reviewed cancer screening models, including a number of the mathematical models with applications to various cancers. Ivy[14] provided a detailed review of a number of models applied to breast cancer screening, giving particular consideration to models considering test scheduling and the optimization of screening. Below we give an overview of some of the most relevant mathematical models regarding screening programs.

Early applications of mathematical models in cancer screening were to estimate the preclinical sojourn time of disease and the sensitivity of screening tests.[15–17] The utility of these models is that they permit the estimation of parameters that are typically not observed using observable screening data.

Other models have attempted not just to describe disease development and screening performance but to consider how screening intervals might be adjusted to maximize the mortality reduction or life-years saved by screening.[18] An example of this is the model by Weiss and Lincoln,[19] which formally demonstrates why regular periodic screening is more effective than irregular screening. Another early example of a mathematical model of optimal screening intervals is that of Kirch and Klein.[20] They presented a relatively simple model of screening, showing that the optimal screening interval varies proportionally to the square root of the incidence rate. This relationship is derived under the simplifying assumptions of perfect screening sensitivity and a constant duration of the preclinical period during which disease is screen-detectable.

Zelen[21] presented a more detailed model of screening that allows for both imperfect sensitivity and a random duration of the preclinical period. The model uses a utility function for the detection of disease weighted by the probability of cure for screen-detected and clinically detected cancer. The optimal screening intervals are found as those maximizing the utility function subject to a constraint on the number

of screens, which serves as a proxy for a budget constraint. Zelen illustrated the model using the example of breast cancer screening under the assumption of constant disease incidence.

Lee and Zelen[22] extended the model presented by Zelen[21] to consider the possibility that disease incidence varies with age. The measure of screening efficacy used in the study is schedule sensitivity, which is the proportion of total disease cases detected by the screening program over the period of its implementation. The optimal strategy is found using what is described as a threshold method, whereby each screen is scheduled following the initial screen so that the probability of an individual having preclinical disease at each screen is equal to that at the first screen. Lee and Zelen[22] also considered how screening intervals and age ranges might vary in high-incidence groups. They found that the optimal screening schedule for a high-risk group starts at a younger age and has shorter intervals relative to that of the average-risk population.

Baker[5] presented a mathematical model of breast cancer screening based on tumor growth rates, from which the probabilities of detection or death are derived. The model finds optimal policies that minimize an objective function of the costs of disease, which includes the costs of screening and life lost. It does not consider how the optimal screening interval changes with disease incidence.

Parmigiani[23] presented a more general model in which screening sensitivity can vary with age and duration of the preclinical state. The model also describes an objective function containing the costs of screening and the value of health gained. Illustrative applications of this model for breast cancer screening have been given by Parmigiani[24] and by Parmigiani and Kamlet.[25] Neither study considered how screening intensities should vary between risk subgroups.

Eddy[26] presented a mathematical model of cancer screening with applications to a number of different cancers. The model is solved numerically to estimate the costs and effects of alternative screening strategies. Sensitivity analyses accompanying the main results show how the total and incremental effects of screening strategies differ between risk subgroups in cases including breast and cervical cancer.[26,27] However, Eddy did not explicitly consider incremental cost-effectiveness ratios or how their variation between risk subgroups might be used to optimize screening.

Finally, we note the recent application of partially observable Markov models as another technique to estimate how to optimize screening effectiveness. Maillart and others[28] used a partially observable Markov model to investigate how the breast cancer lifetime mortality risk can be minimized by switching to a longer screening interval to accommodate a lengthening sojourn time with age. Similarly, Ayer and others[29] used a partially observable Markov decision process (POMDP) to find dynamic screening intervals that maximize the expected quality-adjusted life-years (QALYs) by periodically adjusting the screening interval to account for the estimated disease risk based on factors such as age and screening history. However, while both models consider the optimization of the health effects of screening, neither considers the optimization of cost-effectiveness.

This overview shows that there are already a number of detailed mathematical models of screening effectiveness. However, almost none of these studies have been conducted in the context of conventional CEA, whereby the ratio of incremental costs to incremental health effects is compared with a cost-effectiveness threshold to identify the optimal screening interval given the threshold willingness-to-pay. In addition, the motivation for investigating optimal screening intervals in most studies is to understand how the interval might vary with age rather than between risk subgroups.

Contribution

The purpose of this paper is to present a simple mathematical model to estimate the relationship between disease incidence and the optimal cost-effective screening interval, while accounting for imperfect screening sensitivity and specificity and a random duration of the preclinical disease state.

Our model of the optimal screening interval is more sophisticated than that of Kirch and Klein.[20] While our model is less detailed than that of Parmigiani,[23] it is easier to apply and more accessible in its exposition. It differs from the models by Lee and Zelen,[22] Baker,[5] and others that do not consider interval optimization in terms of conventional CEA.

Unlike other studies in the literature, our model has the benefit of being verified and validated against a published microsimulation CEA model, namely the MISCAN-Breast model. The verification entails checking that the mathematical model and MISCAN yield the same results when the same simplified assumptions used in the mathematical model are programmed in MISCAN. The validation entails checking whether the mathematical model approximates

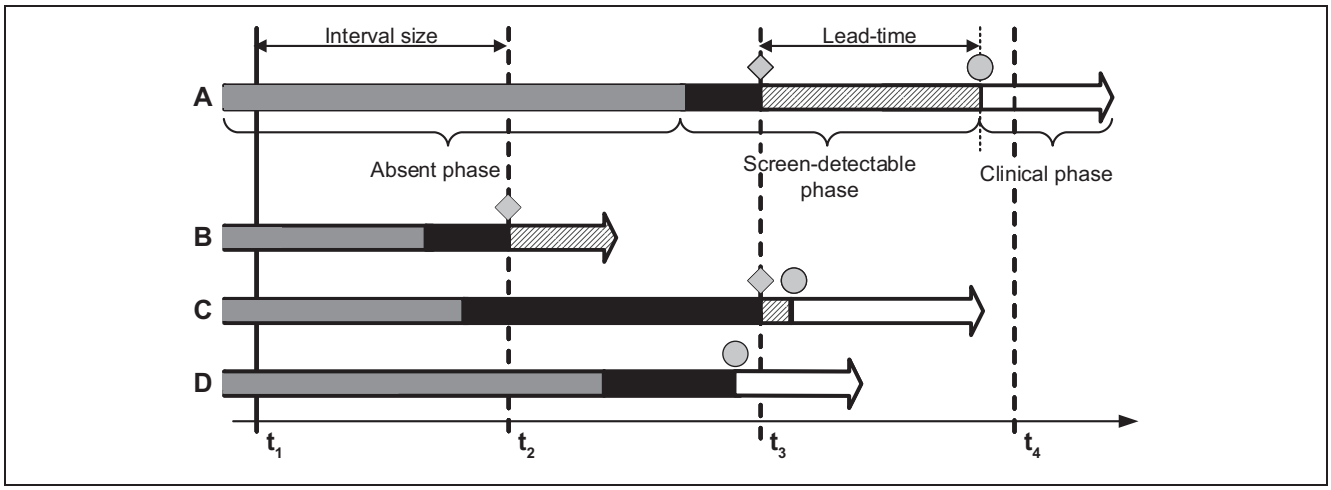

Figure 1 Natural history of disease for 4 screened subjects, developing from no disease to a screen-detectable phase and on to the clinical phase. Screening occurs at t_i*, screen detection at the diamond, clinical diagnosis at the circle, and death at the arrowhead. The lead time is shown between screen-detection and clinical diagnosis. Subject A is detected at* t_3*; B is detected at* t_2 *but dies from another cause and would have never entered the clinical phase; C is missed at* t_2 *due to imperfect test sensitivity but is detected at* t_3*; D is not screen-detected, due to a short preclinical phase.*

the change in screening intervals with disease incidence as calculated by the full MISCAN microsimulation model: that is, whether our relatively simple mathematical model adequately captures the underlying relationship between disease risk and the optimal interval. For the validation, we use the MISCAN model as applied in a recently published CEA of actual screening policy guidelines.[30]

METHODS

The Mathematical Model

We present a mathematical model of the cost-effectiveness of screening. The model is based on a natural history of cancer consisting of 3 phases (Figure 1).[31] These phases are motivated by the possible detection of the tumor or its precursors by a screening test,[32–34] with the example of breast cancer screening in mind.

The 3 stages in the model are 1) the ''no-disease'' stage, in which a subject does not have the disease; 2) the ''screen-detectable'' phase, during which the patient has preclinical disease that can be detected by the screening test; and 3) the ''clinical'' phase, by when the disease has been diagnosed due to symptoms. The benefits of a screening program are realized by detecting and treating subjects in the screen-detectable phase, thereby postponing mortality in some subjects. To maintain a simple mathematical model, we did not explicitly consider death from other causes during the natural history process.

The optimal screening interval depends on several parameters. Table 1 presents an overview of all parameters used in the mathematical model. Table 2 presents the assumptions made in the full mathematical model. The most important parameters can be described as follows.

1. The incidence rate (r) describes the annual probability that a healthy individual will enter the preclinical phase. Assuming that only healthy individuals may enter the preclinical phase and that the hazard is constant over time leads to an exponentially distributed time to the start of the preclinical phase.
2. The mean duration ($\bar{D}$) of the screen-detectable phase (in years) reflects the rate of disease progression.
3. The screening sensitivity (s_n) represents the probability that a screening test will correctly identify an individual with preclinical disease as having cancer.
4. The screening specificity (s_p) represents the probability that a screening test will correctly identify an individual without disease as disease-free.
5. The treatment effect (T) of a screen-detected case is the postponement of mortality due to early detection. The maximum treatment effect is achieved when a disease is detected as soon as it enters the preclinical phase. The mean treatment effect ($\bar{T}$) occurs when disease is screen-detected at the midpoint of the screen-detectable phase. $\bar{T}$ and T are expressed as QALYs gained.

Table 1 Functions and Parameters Used in the Mathematical Model

Function or Parameter	Description
r	The annual incidence rate of preclinical cancer among patients in the absent phase
$\bar{D}$	The mean duration of the screen-detectable phase
D	The realized duration of the screen-detectable phase
p	The number of screenings during the screen-detectable phase
I	The interval (in years) between consecutive screenings
t_i	The time between the ith screen and the start of the screen-detectable phase
$\bar{T}$	The mean treatment effect (i.e., the number of QALYs gained when a cancer is screen-detected at the midpoint of the screen-detectable phase)
$T(D,I,i,t_1)$	The realized treatment effect (i.e., the number of QALYs gained per screen-detected cancer detected at the ith screen)
$T_p(D,I,t_1)$	The total expected treatment effect (i.e., the number of QALYs gained by detecting a cancer during the screen-detectable phase as a function of t_1)
$T_E(D,I)$	The total expected treatment effect (i.e., the number of QALYs for a cancer detected during the screen-detectable phase)
$E(D,I)$	The annual expected health effects (number of QALYs gained) as a function of the interval and the duration of the screen-detectable phase
C	The total costs of the screening program
c_0	The fixed cost of the screening program
c_t	The cost of the screening test
c_d	The cost of diagnostics
c	The total variable costs per additional screening test
s_n	The screening sensitivity (i.e., the probability that a screen of an individual with preclinical disease leads to a cancer diagnosis)
s_p	The screening specificity (i.e., the probability that a screen of an individual without disease leads to a correct diagnosis)
h	The QALY decrement of a false-positive screen result due to unnecessary diagnostics
H	The expected QALY decrement due to false-positive screen per test
W	The cost-effectiveness threshold, in euros per QALY gained

Note: QALY, quality-adjusted life-year.

6. The annual costs (C) of a screening program include the fixed organizational costs of the program (c_0) and the variable costs (c) of a screening test. The fixed costs are sufficiently small that they can be disregarded. The variable cost (c) is given by

$$c = c_t + (1 - s_p)c_d, \tag{1}$$

where c_t is the cost of the test and c_d is the cost of diagnostic follow-up to a positive screen result. We make the simplifying assumptions that the cost of diagnostics only applies to false positives and that the proportion of screening tests that are false positive is $1 - s_p$. We also assume that c_t, c_d, and s_p are constant, which implies that c is constant.

We develop a model that allows for a treatment effect that declines with the time spent in the screen-detectable phase, imperfect screening sensitivity and specificity, and an exponential distribution of the duration of the screen-detectable phase.

Cost-Effectiveness Analysis

We considered the optimal screening interval from a CEA perspective. CEA is the standard framework for assessing costs and benefits of health care interventions.[35] Alternative screening intensities can be compared in terms of their incremental cost-effectiveness ratios (ICERs). The ICER is the ratio of the additional costs incurred (C) to the additional health effects gained (E) when a screening program is compared with less intensive screening programs or no screening at all:

$$\text{ICER} = \frac{\partial C}{\partial E}. \tag{2}$$

The ICER is typically compared with a cost-effectiveness threshold, which represents the marginal willingness-to-pay per QALY gained. We assume the existence of a threshold value given by W. A screening program is then acceptable if its

Table 2 List of Assumptions of the Full Mathematical Model

The disease follows a 3-stage process, in which individuals start in the absent phase and then proceed to the preclinical phase and the clinical phase. Individuals in the absent phase and in the preclinical phase are screened; screen-detection can only occur in the preclinical phase.
The incidence rate (annual probability of entering the preclinical phase) is constant over age.
The time of the start of the preclinical phase is exponentially distributed but can be approximated by a uniform distribution because the total incidence over the screening interval is small.
Individuals cannot develop disease more than once.
The preclinical sojourn time is exponentially distributed, with a constant mean duration.
Test sensitivity and specificity are constant.
Test results are independent between screenings.
The rate of false positives applies to all women screened irrespective of true disease state.
All screening moments are evenly spaced throughout the screening program.
The timing of the screenings is independent of the time of the start of the preclinical phase.
The beneficial effects of early disease detection decline linearly through the preclinical phase from their maximum at the start of the preclinical phase to zero at the moment of clinical detection. These beneficial effects do not depend on age or other factors.
The cost per screen does not vary.
Each screen imposes a fixed cost and harm due to false positives.
The false-positive rate applies to all women screened irrespective of true disease state.
Screening imposes no health burden other than that due to false positives.
Treatment costs are constant and do not depend on the mode of detection.
It is cost-effective to screen individuals at least once during their lifetime.

ICER is not greater than the threshold W (i.e., ICER $\leq$ W). Both costs and effects depend on I. Then from a cost-effectiveness perspective, it must hold that

$$\text{ICER} = \frac{\partial C}{\partial E} = \frac{\partial C}{\partial I} \Big/ \frac{\partial E}{\partial I} \leq W. \tag{3}$$

Assuming that it is cost-effective to screen individuals at least once during their lifetime, the optimal screening interval is the interval with an ICER equal to W. If the ICER is larger than W for every value of I, then screening is not cost-effective at all.

The costs per individual per year, for a certain interval length, are the costs per screening divided by the interval length (in years):

$$C = \frac{c}{I}. \tag{4}$$

The derivative of Equation 4 needed to calculate the ICER is

$$\frac{\partial C}{\partial I} = -\frac{c}{I^2}. \tag{5}$$

This represents the change in costs per change in interval length. Equation 5 will be used throughout this paper.

The assumptions of the mathematical model are presented sequentially. In our model, the duration of the preclinical phase, D, is a random variable. We first calculate the expected treatment effect conditional on the distribution of the preclinical phase for a given value of D. Once these mathematical results have been obtained, we then integrate out the value of D given its probability distribution. The results of the mathematical model are then compared with those from the simplified MISCAN model to verify the results and with those from the full MISCAN model to validate the mathematical model's findings.

We first assume that the treatment effect of a screen-detected cancer declines linearly during the screen-detectable phase. This assumption is based on the fact that the probability that screen-detection prevents mortality for a cancer that would otherwise be lethal depends on the amount of lead time. If the lead time is 0, so that the time of screen-detection and the time of clinical detection coincide, no treatment effects should be expected. The duration of the screen-detectable phase is assumed constant and the screening sensitivity is assumed to be perfect. The treatment effect of screening at the beginning of the screen-detectable phase is $2\bar{T}$ and declines linearly to 0 at the end of the phase, leading to an average effect over the screen-detectable phase of $\bar{T}$. Consequently, the treatment effect of a screen-detected cancer is

$$\frac{D-t}{D} 2\bar{T}, \tag{6}$$

where t is the time between the start of the screen-detectable phase and the moment of screen detection.

Imperfect Screening Test Sensitivity and Specificity

We then add the assumption of imperfect screening sensitivity and specificity. The screening sensitivity and specificity are assumed constant. We also assume that the probability of screen detection is

independent between screens. The probability that a preclinical cancer is detected at the *i*th screening in the screen-detectable phase given that it was not detected at the $i-1$ previous screenings is

$$s_n(1-s_n)^{i-1}. \tag{7}$$

Imperfect specificity means there will be false-positive test results, which will lead to a loss of utility. The expected utility decrement per screening is

$$H = (1 - s_p)h, \tag{8}$$

where the utility decrement of a false positive is given by h.

The expected benefit of the *i*th screening can be expressed as the product of the probability that the *i*th test is positive (i.e., Equation 7) and the expected treatment effect of detection conditional on the test being positive, which is the treatment effect if a cancer is found as defined in Equation 6.

Combining Equations 6 and 7 yields the expected treatment effects for the *i*th screening during a screen-detectable phase:

$$T(D, t_i) = \frac{D - t_i}{D} 2\bar{T}s_n(1-s_n)^{i-1}, \tag{9}$$

where t_i is the time between the start of the screen-detectable phase and the *i*th screening. Because the interval between the screenings is constant, the *i*th screening occurs at time

$$t_i = t_1 + (i-1)I. \tag{10}$$

We first consider the case where $I \leq D$. In that case, D/I will have a value between two positive integers k and $k+1$, such that $kI \leq D < (k+1)I$, and the number of screens during the screen-detectable phase will be either k or $k+1$. In Equation 11, the function T given in Equation 9 is subject to a change of variables, based on the expression for t_i given in Equation 10, which yields

$$T(D, I, i, t_1) = \frac{D - t_1 - (i-1)I}{D} 2\bar{T}s_n(1-s_n)^{i-1} \text{ for } I \leq D. \tag{11}$$

Let p denote the number of screenings during the screen-detectable phase. The total expected treatment effects obtained during the preclinical phase, given values of D, I, and t_1 (i.e., the time between the start of screen-detectable phase and the first screening during the screen-detectable phase), are given by

$$T_p(D, I, t_1) = \sum_{i=1}^{p} T(D, I, i, t_1) \text{ for } I \leq D. \tag{12}$$

We assume that the timing of the screenings is independent of the timing of the start of the screen-detectable phase. Because the timing of the start of the screen-detectable phase is exponentially distributed, the distribution of t_1 is a truncated exponential with density function $f(t_1) = \frac{re^{-rt_1}}{1-e^{-rI}}$ between 0 and I. The total incidence during the screen interval is rI. When this is small (e.g., <5%), the distribution of t_1 can be approximated using the uniform distribution with density $f(t_1) = \frac{1}{I}$ between 0 and I. We use this approximation in the remainder of this paper. Note that when applying the model for other, more common diseases, one should check whether this approximation is reasonable. The expected treatment effects during the screen-detectable phase, that is, the expectation of $T_p(D,I,t_1)$ with respect to t_1, are thus given by

$$T_E(D, I) = \int_0^I f(t_1)T_P(D, I, t_1)dt_1 = \frac{1}{I}\int_0^I \sum_{i=1}^{p} T_P(D, I, i, t_1)dt_1 = \frac{1}{I}\int_0^I \sum_{i=1}^{p} \frac{D - t_1 - (i-1)I}{D} 2\bar{T}s_n(1-s_n)^{i-1}dt_1 \text{ for } I \leq D. \tag{13}$$

To calculate the expected treatment effects per year, we multiply the total expected treatment effects during the screen-detectable phase with the probability that the screen-detectable phase starts in a given year, which yields r times $T_E(D,I)$. The annual expected health effects, as a function of D and I, are the expected treatment effects per year minus the expected cost QALY decrement of false positives, which yields

$$E(D, I) = rT_E(D, I) - \frac{H}{I} = \frac{r}{I}\int_0^I \sum_{i=1}^{p} \frac{D - t_1 - (i-1)I}{D} 2\bar{T}s_n(1-s_n)^{i-1}dt_1 - \frac{H}{I} = \frac{2\bar{T}r}{I}\left[\int_0^{D-kI} \sum_{i=1}^{k+1} \frac{D-t_1-(i-1)I}{D} s_n(1-s_n)^{i-1}dt_1 + \int_{D-kI}^{I} \sum_{i=1}^{k} \frac{D-t_1-(i-1)I}{D} s_n(1-s_n)^{i-1}dt_1\right] - \frac{H}{I}, \text{ for } I \leq D, \tag{14}$$

where the first integral relates to when $p = k + 1$ screening intervals fall within the screen-detectable phase and the second integral relates to when $p = k$ screenings fall within the screen-detectable phase. The derivation of Equation 15 from Equation 14 uses the approximation $k \approx D/I$ and is shown in Appendix I (available online).

$$E(D,I) = \frac{\overline{T}r}{I}\left[\frac{I^2}{s_n D}\left(1-(1-s_n)^{\frac{D}{T}}\right)(s_n-2)+2I\right]-\frac{H}{I} \text{ for } I \le D. \quad (15)$$

In the case in which $I > D$, there is at most 1 screening during the screen-detectable phase. We thus calculate the expected health effects as

$$E(D,I) = r\int_0^I f(t_1)T(t_1)dt_1 - \frac{H}{I} = \frac{r}{I}\int_0^I T(t_1)dt_1 - \frac{H}{I} =$$
$$\frac{r}{I}\int_0^D T(t_1)dt_1 + \frac{r}{I}\int_D^I T(t_1)dt_1 - \frac{H}{I} =$$
$$\frac{r}{I}\int_0^D T(t_1)dt_1 - \frac{H}{I} = \frac{r}{I}\int_0^D \frac{D-t_1}{D}2\bar{T}s_n(1-s_n)^{i-1}dt_1 - \frac{H}{I} =$$
$$2\bar{T}\frac{r}{I}\int_0^D \frac{D-t_1}{D}s_n dt_1 - \frac{H}{I} = \frac{\bar{T}rs_nD-H}{I} \text{ for } I>D. \quad (16)$$

Exponentially Distributed Preclinical Phase

Now consider D as an exponentially distributed random variable with the following probability density function $f(D) = \frac{1}{\bar{D}}e^{-D/\bar{D}}$. Other distributions could be used for the preclinical phase, such as the Weibull distribution. The annual expected health effects of screening for $I \le D$ and $I > D$ from Equations 15 and 16 are now combined with the exponential distribution of the screen-detectable phase, which shows the annual expected health effects as a function of the screening interval to be

$$E(I) = \int_0^\infty f(D)E(D,I)dD =$$
$$\int_0^I \frac{e^{-D/\bar{D}}}{\bar{D}}\left[\frac{\bar{T}r}{I}s_nD-\frac{H}{I}\right]dD +$$
$$\int_I^\infty \frac{e^{-D/\bar{D}}}{\bar{D}}\left[\frac{\bar{T}r}{I}\left[\frac{I^2}{s_nD}\left(1-(1-s_n)^{\frac{D}{T}}\right)(s_n-2)+2I\right]-\frac{H}{I}\right]dD =$$
$$\int_0^I \frac{e^{-D/\bar{D}}}{\bar{D}}\frac{\bar{T}r}{I}s_nDdD +$$
$$\int_I^\infty \frac{e^{-D/\bar{D}}}{\bar{D}}\frac{\bar{T}r}{I}\left[\frac{I^2}{s_nD}\left(1-(1-s_n)^{\frac{D}{T}}\right)(s_n-2)+2I\right]dD$$
$$-\frac{H}{I}\int_0^\infty \frac{e^{-D/\bar{D}}}{\bar{D}}dD =$$
$$\frac{\bar{T}r}{I}\left[\int_0^I \frac{e^{-D/\bar{D}}}{\bar{D}}s_nDdD +\right.$$
$$\left.\int_I^\infty \frac{e^{-D/\bar{D}}}{\bar{D}}\frac{I^2}{s_nD}\left(1-(1-s_n)^{\frac{D}{T}}\right)(s_n-2)+2I\right]dD\Bigg]-\frac{H}{I}, \quad (17)$$

where the term within the brackets on the penultimate line refers to screening intervals longer than the duration of the screen-detectable phase, and the following term within the brackets on the final line refers to intervals shorter than the screen-detectable phase.

We can form an expression for the ICER by differentiating Equation 17 with respect to I and combining it with Equations 3 and 5. Mathematica 4.0 (Wolfram, Champaign, IL) was used to differentiate Equation 17, and the resulting expression for the ICER is reported in Appendix II (available online). Using Mathematica, we evaluated the ICER for a range of screening intervals using the parameter described above. Using Mathematica we solve for values of I such that ICER = W. The model is used to solve for such values of I for a range of incidence rates for which the cost-effectiveness criterion of ICER = W can be met.

Verification and Validation in a Breast Cancer Example

The models are applied to breast cancer screening. Most of the parameter values in the model are specified to correspond to a published CEA of mammography screening in Switzerland.[30] We consider women between 50 and 70 years old, for whom the incidence rate for breast cancer is assumed to be 0.00225 per woman per year, as assumed in the Swiss study.[30] We assume a mean duration of the screen-detectable phase of 3 years, in line with estimates for Canadian women over the age of 50.[36] We assume a test sensitivity and specificity of 66% and 97%, respectively, in line with estimates from a study of women aged between 50 and 69 used to inform the Swiss CEA.[37]

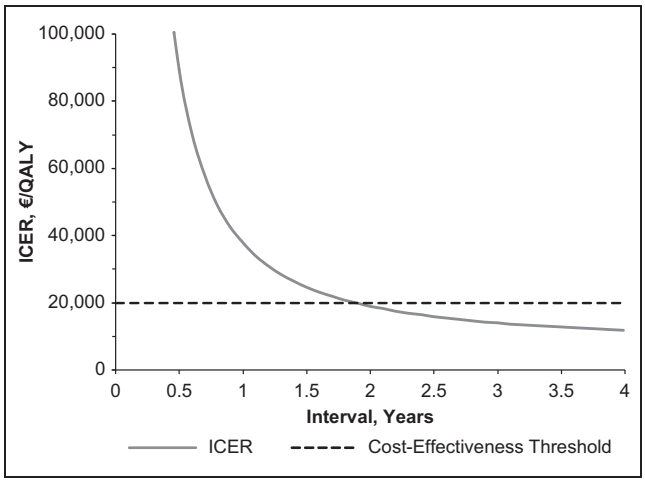

Figure 2 Variation of the incremental cost-effectiveness ratio (ICER) with the screening interval with a cost-effectiveness threshold of €20,000 per quality-adjusted life-year (QALY).

The costs per screening test and costs of diagnostics following a false positive are assumed to be €138 and €650, respectively, also in line with the costs used in the Swiss CEA. The threshold, W, is set at €20,000 per QALY.[38] The benefit of detecting a cancer in the screen-detectable phase varies with the length of the screening interval in the models. However, we assume a mean treatment effect, $\bar{T}$, of 4.2 QALYs per screen-detected cancer, to match the approximate QALY gain per screen-detected cancer in the Swiss CEA under the assumption of biennial screening. Finally, we also assume a QALY penalty of 0.001 for each false positive, again in line with assumptions used in the Swiss CEA.

We verified our mathematical model against the MISCAN-Breast model. The MISCAN model was simplified significantly to match the structural assumptions of the mathematical model and used the same parameter values. Using this simplified version of MISCAN, we estimated the cost and effects of screening intervals ranging from 18 to 0.2 years over an 18-year screening program. We then computed the ICERs of the efficient screening intervals and found the optimal screening interval with reference to the cost-effectiveness threshold. Linear interpolation was used to approximate the optimal interval in cases where the threshold fell between the ICERs of 2 adjacent strategies. This process was repeated for a range of incidence rates to derive the relationship between incidence and the optimal interval for comparison with the mathematical model.

We validated the mathematical model against a full version of MISCAN-Breast used in a recently published CEA of screening in Switzerland.[30] This version of MISCAN did not have any of the structural simplifications used in the verification exercise. The methods used in the validation for identifying the relationship between the incidence rate and the optimal screening interval were the same as those described for the verification exercise, except that we assessed 22 different screening intervals ranging from 9 to 0.2 years in the fixed screening age range of 50 to 69. The model is as described by de Gelder and others,[30] except that we simulated a single birth cohort of previously unscreened women rather than a population for the sake of simplicity.

We conducted a sensitivity analysis to investigate the effect of different cost-effectiveness thresholds on the model validation. We also conducted a sensitivity analysis to assess how the relationship between incidence rates and relative optimal screening interval varies with discount rates of 0%, 3%, and 5% applied to costs and effects.

RESULTS

The derived function for the ICER was evaluated over a range of interval lengths, resulting in the gray line in Figure 2. In this case, the intersection of the ICER curve and the threshold indicates an optimal interval of approximately 1.9 years. The incidence rate was then varied to generate new ICER curves, which were combined with a threshold of €20,000 per QALY, to determine the relationship between incidence, r, and the optimal interval shown with the gray line in Figure 3B.

Verification

The results of the verification analysis are presented in Figure 3. Figure 3A shows the ICER estimates at different incidence rates from the simplified MISCAN model and the mathematical model. The MISCAN estimates are shown as discrete points because MISCAN estimates the ICER of discrete changes to the screening interval, while the mathematical model estimates the effect of a continuous change to screening intervals.

Figure 3A shows the ICERs from the MISCAN and mathematical models to be very close. The ICERs from the MISCAN model are consistently marginally lower than those from the mathematical model. This is to be expected, as the continuous ICER estimates of the mathematical model necessarily result in higher ICERs than the ICER estimates based on discrete

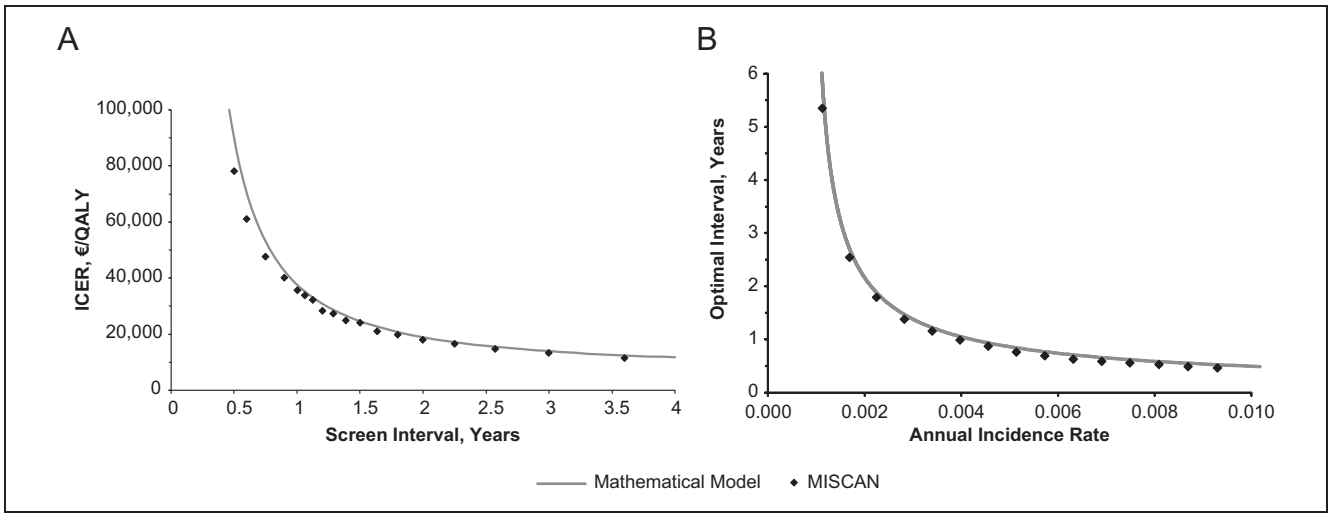

Figure 3 (A) Verification against a simplified version of the MISCAN model showing the variation of the incremental cost-effectiveness ratio (ICER) with the screening interval. (B) Verification against a simplified version of the MISCAN model showing the variation of the optimal screening interval with the annual incidence rate at a cost-effectiveness threshold of €20,000 per quality-adjusted life-year (QALY).

changes to the screening interval made by the MISCAN model. The relationship between the incidence rate and the optimal interval is shown in Figure 3B for the simplified MISCAN and the mathematical models. Again, the results indicate a close match between the 2 models; the small difference between the estimates occurs because the MISCAN model simulates discrete changes to the interval, while the interval changes continuously in the mathematical model.

Validation

The optimal screening intervals estimated by the mathematical model did not exactly match those derived from the complete MISCAN model. However, the purpose of our mathematical model is to show the relative change in optimal screening intensity with a change in disease incidence rather than to estimate the optimal level of screening intensity at any one point. Consequently, we expressed the optimal screening interval at each incidence rate as an index relative to the optimal interval at the average-risk incidence rate. This relationship is depicted in Figure 4, which shows a close agreement between the models on the relationship between the incidence rate and the relative screening intensity, although there is some deviation at incidence rates below the baseline average risk.

The results of a sensitivity analysis for different levels of the cost-effectiveness threshold are shown in Appendix III (available online). The results show that the match between models was maintained at other threshold levels.

Illustrative Application

A practical application of the model can demonstrate its benefit in tailoring screening strategies to risk subgroups. Consider the example of the current breast screening program in the UK, in which women aged between 50 and 70 receive a mammogram every 3 years. Assuming this program is optimal at a given willingness-to-pay threshold, we consider how the screening intervals should be adjusted for women at elevated risk.

Using the Cuzick-Tyrer IBIS risk calculator to estimate the risk of breast cancer given a woman's characteristics and family history,[39,40] we consider a 50-year-old, postmenopausal, parous woman whose age at the birth of her first child was 22 and whose mother had breast cancer at age 72. The risk calculator suggests that this woman's relative risk (RR) of developing breast cancer is approximately 1.5 compared with the average population risk annual incidence rate. This woman's annual probability of developing cancer can be approximated by multiplying the average incidence by the RR. Entering an annual incidence 1.5 times the average in our mathematical model results in an optimal interval 0.65 times that of the average-risk population, which would be closely approximated by an interval of 2 years. Alternatively, if we assume the same woman's 3 sisters had breast cancer at age 42, the risk

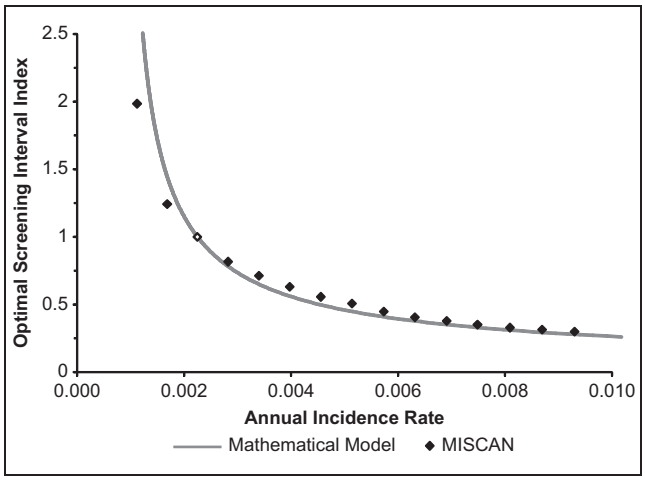

Figure 4 Validation against a complete version of the MISCAN model showing the relative change in the optimal screening interval with the annual incidence rate at a cost-effectiveness threshold of €20,000 per quality-adjusted life-year (QALY) and the optimal screening interval indexed to 1 at the baseline annual incidence rate of 0.00225.

calculator suggests an RR of approximately 3, yielding an optimal screening interval 0.36 times that of the average-risk population, which could be pragmatically be approximated by a 1-year interval.

We can estimate the health benefits of tailoring the interval to the incidence rate using the full MISCAN model used in the model validation. Reducing the screening interval for women with an RR of 1.5 from 3 to 2 years results in a discounted total of 18 additional QALYs gained per 1000 women screened. This represents a 21% increase over what would be achieved with an interval of 3 years. Similarly, the lifetime disease-specific mortality would be reduced from 4.2% to 4.0%, while the lifetime probability of an advanced cancer represented by the incidence of T2+ tumors would be reduced from 5.7% to 5.3%. Alternatively, moving from a 3-year interval to annual screening for women with an RR of 3 would save an additional 93 QALYs per 1000 women screened, representing a 53% gain. The lifetime disease-specific mortality would be reduced from 8.5% to 7.3%, and the lifetime incidence of T2+ tumors would be reduced from 11.5% to 9.6%.

DISCUSSION

Using a mathematical model, we estimated the relationship between the incidence rate and the length of the optimal screening interval. Our results show that a mathematical model can provide a clear insight into how disease risk influences the optimal screening interval. Furthermore, it shows how a small number of parameters can be used to estimate the change in the optimal screening interval for individual risk profiles.

Reality is more complex than represented in our model. For example, breast cancer incidence increases with age, whereas our mathematical model assumes it to be constant. This would suggest that the optimal screening interval would become shorter with age. However, the potential benefit of early detection decreases with age, due to increased competing risk of other-cause mortality,[41] potentially partially offsetting a reduction in the optimal interval with age. In principle, our mathematical model could be applied to different age groups separately, using age-specific disease incidence and expected health benefits, allowing it to specify age-specific adjustments to the optimal interval. However, since most cancer screening programs use screening intervals of equal length across all ages, providing age-specific adjustments to intervals would be of limited use.

One of the simplifying assumptions used in our mathematical model was the inclusion of no treatment costs. This assumption can be justified on the grounds that the differences in treatment costs between screen-detected and clinically detected cancers are not large relative to the costs of screening. We checked the validity of this assumption by conducting a sensitivity analysis to determine whether the relative change in the optimal screening interval with disease incidence was different when treatment costs are dropped from the full MISCAN model. We found no meaningful difference between the relative change in the optimal screening interval with disease incidence between the full MISCAN model with and without treatment costs (results not shown).

Like previous mathematical models, the model presented here uses undiscounted costs and effects. In practice, resource allocation is made on the basis of discounted outcomes. Discounting could be included in the model, but at the cost of additional complexity. We conducted a sensitivity analysis to assess how the optimal screening interval varies with disease incidence using the MISCAN model when discounting is applied (results not shown). Although the length of the optimal interval changed with discounting, the relative relationship between incidence and the optimal interval index did not vary substantially.

Our model complements current research considering how screening can be optimized to disease risk. However, our model differs from much of the work described in the introduction, in that rather

than optimizing effectiveness it explicitly optimizes cost-effectiveness for a given threshold. Similarly, our model differs from the recent work of Maillart and others[28] and Ayer and others,[29] using the POMDP approach to consider how fully personalized, dynamic screening schedules might be specified and updated for each individual. By contrast, we use a parsimonious mathematical model that is analytically solved to find how the frequency of a population screening program should be adjusted for specific risk groups.

The advantage of the mathematical model is that it is a fast and convenient method for understanding the influence of relative disease risk on the optimal screening interval. It requires few input parameters and yields estimates quickly. As such, it offers a useful means for interpreting the impact of biomarkers as potential risk predictors on the optimal interval.[42] In particular, it offers a simple way of assessing the significance of a large number of biomarkers with a cumulative effect on the optimal screening interval, without the need for separate analyses of each biomarker separately.

CONCLUSION

We have shown that a parsimonious mathematical model can be used to determine the optimal screening interval for high-risk groups relative to the optimal screening interval for the average-risk population. Despite the simplifying assumptions of the mathematical model, the validation exercise shows that it is sufficient to capture the effect of changes in disease incidence on optimal screening intervals. Consequently, the mathematical model presented here offers a quick and easy way to explore optimal screening intervals for specific risk groups without the use of complex simulation models.

ACKNOWLEDGMENTS

We would like to kindly thank the 3 anonymous reviewers for their detailed and helpful comments.

REFERENCES

1. Alexander FE, Anderson TJ, Brown HK, et al. The Edinburgh randomised trial of breast cancer screening: results after 10 years of follow-up. Br J Cancer. 1994;70(3):542–8.

2. Andersson I, Aspegren K, Janzon L, et al. Mammographic screening and mortality from breast cancer: the Malmo mammographic screening trial. Br Med J. 1988;297(6654):943–8.

3. Nyström L, Wall S, Rutqvist LE, et al. Breast cancer screening with mammography: overview of Swedish randomised trials. Lancet. 1993;341(8851):973–8.

4. Shapiro S, Strax P, Venet L. Periodic breast cancer screening in reducing mortality from breast cancer. JAMA. 1971;215(11):1777–85.

5. Baker R. Use of a mathematical model to evaluate breast cancer screening policy. Health Care Manag Sci. 1998;1(2):103–13.

6. Brandt A, Bermejo JL, Sundquist J, Hemminki K. Age of onset in familial breast cancer as background data for medical surveillance. Br J Cancer. 2009;102(1):42–7.

7. Brandt A LBJ, Sundquist J, Hemminki K. Breast cancer risk in women who fulfill high-risk criteria: at what age should surveillance start? Breast Cancer Res Treat. 2010;121(1):133–41.

8. Knudsen AB, McMahon PM, Gazelle GS. Use of modeling to evaluate the cost-effectiveness of cancer screening programs. J Clin Oncol. 2007;25(2):203–8.

9. Ramsey SD, Wilschut J, Boer R, van Ballegooijen M. A decision-analytic evaluation of the cost-effectiveness of family history-based colorectal cancer screening programs. Am J Gastroenterol. 2010;105(8):1861–9.

10. Habbema JDF, Lubbe JTN, van Oortmarssen GJ, van der Maas PJ. A simulation approach to cost-effectiveness and cost-benefit calculations of screening for the early detection of disease. Eur J Oper Res. 1987;29(2):159–66.

11. Habbema JDF, van Oortmarssen GJ, Lubbe JTN, van der Maas PJ. The MISCAN simulation program for the evaluation of screening for disease. Comput Methods Programs Biomed. 1985;20(1):79–93.

12. van Oortmarssen GJ, Habbema JDF, van der Maas PJ, et al. A model for breast cancer screening. Cancer. 1990;66(7):1601–12.

13. Alagoz O, Ayer T, Erenay FS, et al. Operations Research Models for Cancer Screening. Wiley Encyclopedia of Operations Research and Management Science. New York: John Wiley & Sons; 2010.

14. Ivy JS. Can we do better? *Optimization models for breast cancer screening*. In: Pardalos P, Romeijn HE, eds. Handbook of Optimization in Medicine. New York: Springer; 2009. p 1–28.

15. Zelen M, Feinleib M. On the theory of screening for chronic diseases. Biometrika. 1969;56(3):601–14.

16. Prorok PC. The theory of periodic screening II: doubly bounded recurrence times and mean lead time and detection probability estimation. Adv Appl Probab. 1976;8(3):460–76.

17. Day NE, Walter SD. Simplified models of screening for chronic disease: estimation procedures from mass screening programmes. Biometrics. 1984;40(1):1–13.

18. Shwartz M. A mathematical model used to analyze breast cancer screening strategies. Oper Res. 1978;26(6):937–55.

19. Weiss GH, Lincoln TL. Analysis of repeated examinations for the detection of occult disease. Health Serv Res. 1966;1(3):272–86.

20. Kirch RLA, Klein M. Examination schedules for breast cancer. Cancer. 1974;33(5):1444–50.

21. Zelen M. Optimal scheduling of examinations for the early detection of disease. Biometrika. 1993;80(2):279–93.

22. Lee SJ, Zelen M. Scheduling periodic examinations for the early detection of disease: applications to breast cancer. J Am Stat Assoc. 1998;93(444):1271–81.

23. Parmigiani G. On optimal screening ages. J Am Stat Assoc. 1993;88(422):622–8.

24. Parmigiani G. Timing medical examinations via intensity functions. Biometrika. 1997;84(4):803–16.

25. Parmigiani G, Kamlet M. A cost-utility analysis of alternative strategies in screening for breast cancer. In: Gatsonis C, Hodges J, Kass RE, Singpurwalla N, eds. Bayesian Statistics in Science and Technology: Case Studies. New York: Springer; 1993. p 390–402.

26. Eddy DM. Screening for Cancer: Theory, Analysis, and Design. Englewood Cliffs (NJ): Prentice-Hall; 1980.

27. Eddy DM. ACS report on the cancer-related health checkup. CA Cancer J Clin. 1980;30(4):193–240.

28. Maillart LM, Ivy JS, Ransom S, Diehl K. Assessing dynamic breast cancer screening policies. Oper Res. 2008;56(6):1411–27.

29. Ayer T, Alagoz O, Stout NK. OR forum—a POMDP approach to personalize mammography screening decisions. Oper Res. 2012; 60(5):1019–34.

30. de Gelder R, Bulliard J-L, de Wolf C, et al. Cost-effectiveness of opportunistic versus organised mammography screening in Switzerland. Eur J Cancer. 2009;45(1):127–38.

31. Walter SD, Day NE. Estimation of the duration of a pre-clinical disease state using screening data. Am J Epidemiol. 1983;118(6): 865–86.

32. Gao F, Chia K-S, Ng F-C, Ng E-H, Machin D. Interval cancers following breast cancer screening in Singaporean women. Int J Cancer. 2002;101(5):475–9.

33. Jansen JTM, Zoetelief J. MBS: a model for risk benefit analysis of breast cancer screening. Br J Radiol. 1995;68(806):141–9.

34. Jansen JTM, Zoetelief J. Optimisation of mammographic breast cancer screening using a computer simulation model. Eur J Radiol. 1997;24(2):137–44.

35. Weinstein MC, Stason WB. Foundations of cost-effectiveness analysis for health and medical practices. N Engl J Med. 1977; 296(13):716–21.

36. Shen Y, Zelen M. Screening sensitivity and sojourn time from breast cancer early detection clinical trials: mammograms and physical examinations. J Clin Oncol. 2001;19(15):3490–9.

37. Bulliard JL, De Landtsheer JP, Levi F. Results from the Swiss mammography screening pilot programme. Eur J Cancer. 2003; 39(12):1761–9.

38. Boersma C, Broere A, Postma MJ. Quantification of the potential impact of cost-effectiveness thresholds on Dutch drug expenditures using retrospective analysis. Value Health. 2010;13(6):853–6.

39. Tyrer J, Duffy S, Cuzick J. IBIS Breast Cancer Risk Evaluation Tool. 7th ed. London: Wolfson Institute of Preventive Medicine; 2013. Available at www.ems-trials.org/riskevaluator/. Accessed August 1, 2013.

40. Tyrer J, Duffy SW, Cuzick J. A breast cancer prediction model incorporating familial and personal risk factors. Stat Med. 2004; 23(7):1111–30.

41. Brown ML. Economic considerations in breast cancer screening of older women. J Gerontol. 1992;47(S):51–8.

42. D'Souza AL, Tseng JR, Pauly KB, et al. A strategy for blood biomarker amplification and localization using ultrasound. Proc Natl Acad Sci U S A. 2009;106(40):17152–7.