



Johns Hopkins University, Dept. of Biostatistics Working Papers

12-1-2003

Optimization of Breast Cancer Screening Modalities

Yu Shen

M.D. Anderson Cancer Center, University of Texas, yshen@mdanderson.org

Giovanni Parmigiani

Johns Hopkins University, Departments of Oncology, Biostatistics & Pathology, gp@jimmy.harvard.edu

Suggested Citation

Shen, Yu and Parmigiani, Giovanni, "Optimization of Breast Cancer Screening Modalities" (December 2003). *Johns Hopkins University, Dept. of Biostatistics Working Papers*. Working Paper 18.
<http://biostats.bepress.com/jhubiostat/paper18>

This working paper is hosted by The Berkeley Electronic Press (bepress) and may not be commercially reproduced without the permission of the copyright holder.

Copyright © 2011 by the authors

Optimization of Breast Cancer Screening Modalities

BY YU SHEN

*Department of Biostatistics
M. D. Anderson Cancer Center
University of Texas, Houston, TX 77030
yshen@mdanderson.org*

GIOVANNI PARMIGIANI

*Departments of Oncology, Biostatistics and Pathology
Johns Hopkins University, Baltimore MD 21205
gp@jhu.edu*

Keywords: Breast cancer screening; Clinical breast examination; Cost-effectiveness; Decision analysis models; Mammography; Natural history model



December 1, 2003

ABSTRACT

Mathematical models and decision analyses based on microsimulations have been shown to be useful in evaluating relative merits of various screening strategies in terms of cost and mortality reduction. Most investigations regarding the balance between mortality reduction and costs have focused on a single modality, mammography. A systematic evaluation of the relative expenses and projected benefit of combining clinical breast examination and mammography is not at present available. The purpose of this report is to provide methodologic details including assumptions and data used in the process of modeling for complex decision analyses, when searching for optimal breast cancer screening strategies with the multiple screening modalities. To systematic evaluate the relative expenses and projected benefit of screening programmes that combine the two modalities, we build a simulation model incorporating age-specific incidence of the disease, age-specific pre-clinical duration of the disease, age-specific sensitivities of the two screening modalities, and competing causes of mortality. Using decision models, we can integrate information from different sources into the modeling processes, and assess the cost-effectiveness of a variety of screening strategies while incorporating uncertainties.



1 Introduction

Breast cancer is the most frequently diagnosed cancer among women. Its rate of incidence in the United States has continued to increase since 1986, (1) while breast cancer mortality has decreased overall in the United States, Canada and the United Kingdom (2; 3; 1). Plausible explanations for this decrease in mortality include progress in treatment, as well as widespread participation in early detection programs that contribute to increased cure rates and reduced disease-specific mortality. Many studies have indicated that early detection through screening can lead to more advantageous treatment options, and often leads to an increase in survival rates and improvement in the quality of life for women who develop breast cancer (4; 5). The development of new technologies and further improvement of the existing modalities for disease detection may increasingly make screening for cancer a routine part of secondary prevention.

The goals of early detection are to reduce breast cancer morbidity and mortality. Optimal screening strategies are expected to carefully balance these goals against the associated burden to women and cost to health care systems. Several issues regarding the optimal choice of breast cancer screening strategies remain open. For example, debate surrounds the question of whether regular mammographies are beneficial to women in their forties. Evidence of benefit varies across the relevant randomized clinical trials (6), and there is controversy on the relevance of the suggested benefits for individual women. Consensus panels (7; 8) who reviewed the evidence did not find it sufficiently strong to make general recommendations, emphasizing that “women should be informed of the potential benefits and risks of screening mammography and assisted in deciding at what age they wish to initiate the manoeuvre” (8). In addition to the issue of the appropriate age at which screening should begin, complex open issues include the appropriate frequency of screening examinations; whether women who are at increased risk of breast cancer would benefit from more frequent screening; and what would be the impact of combining multiple screening modalities.

Evaluating alternative screening strategies is difficult because the benefits of screening depend on complex interaction among several factors, including the ability of various screening tests to detect cancer sufficiently early; the time window during which such detection can take place, and its relation to the interval between screening exams; the relative advantage of an early detection compared to waiting for symptoms to arise; the age distribution of onset of pre-symptomatic cancer; competing causes of mortality; and others.

Simulation-based decision models have proved to be an effective way to evaluate health care interventions whose consequences are complex and depend on the interaction of many

factors. They can provide a formal structure for supporting optimal choice of screening strategies, cost-effectiveness analysis of specific interventions, and formal optimization of utility functions of interest. These models often generate simulated individual histories by drawing evidence from several sources, including the epidemiology and genetics of risk factors, relevant clinical trials of secondary prevention and treatment, and studies of tumor growth. A decision model can also support realistic assessments of uncertainty about the relative merits of alternative choices, an aspect that is often underappreciated in policy making (9). The literature on model-based evaluation of screening strategies is now extensive (10; 11).

In this article we consider the model by Parmigiani (12; 11), and generalize it by incorporating the possibility of using two breast cancer screening modalities in concert: mammography (MM) and clinical breast examination (CBE). We also update the model inputs to reflect recent contributions to the literature. Existing investigations regarding the balance between mortality reduction and costs have focused on mammography only, and have paid less attention to the combined use of periodic mammography with clinical breast examination. (13; 14; 15; 16; 17; 15; 18; 19; 20; 21; 22; 23; 24; 25; 26; 27; 28). Recent studies have shown that periodic clinical breast examinations combined with mammograms improve the overall sensitivity of the screening exam compared with mammography alone, (29; 30; 31; 32), and can be particularly valuable among younger women for whom the sensitivity of mammography alone is relatively low. Logistically, a regular clinical breast examination is easy to administer as part of a routine physical examination, and is less expensive compared to mammography.

To promote more efficient and cost-effective breast cancer early detection programs, we will explore optimal screening strategies in terms of the costs and the quality-adjusted years of life saved. The analyses focus on strategies that combine the use of both mammography and clinical breast examination. The other factors we investigate include age group and screening interval. The primary objective of this article are to discuss modeling issues arising in optimization of screening strategies with multiple modalities, and to provide methodologic justifications for models and sources of data used in the analyses reported Shen and Parmigiani (33).

The results from our investigation will help in the design of more efficient and near optimal early detection programs, thereby maximizing the survival benefit for breast cancer patients while also considering the associated societal costs. This study focuses on breast cancer, but the methods are also applicable to early detection programs for other types of cancer. The proposed research will provide a basis to guide health policy makers in designing optimal and cost-effective screening programs, and in extending such benefits to a large population.

2 Model

2.1 Natural History of Breast Cancer

The basis of our investigation of optimal combinations of screening modalities is a simulation model that can generate individual health histories. It is useful to distinguish the natural history model, which refers to the health histories of women without early detection screening, from the intervention model, which refers to the effects of screening. For a patient with preclinical disease, the natural history model provides a way of simulating age of onset and preclinical sojourn time, or, equivalently, growth rate. Conditional on these, it then simulates the age of the woman and the tumor size at the time of diagnosis. These variables can then be used in turn as covariates in predicting a woman's survival and quality adjusted survival. This multi-stage prediction can be repeated for various screening strategies, by superimposing a history of examinations to the natural history, appropriately simulating results of screening tests based on assumed sensitivity, and appropriately modifying age and size at detection when early detection takes place. Thus, given women's risk factors, a decision model using Monte Carlo simulations can be employed to jointly model the disease histories and screening interventions, and predict the outcomes of interest.

In the natural history model, breast cancer events are simulated according to the age-specific incidence of preclinical disease and mortality from other causes. For a woman with breast cancer, the natural history model also provides a way of generating her history of disease over time. The natural history of the disease over time requires a description of the transition between different states of the disease. Using the same notation as in Parmigiani (11), we assume that there are four relevant states: H , women who are either disease-free or asymptomatic; P , women who have detectable pre-clinical disease; C , women with clinical manifestation of the disease; and D , women who have died. For women in the cohort who have breast cancer, we generate their ages at the onset of pre-clinical breast cancer, P , ages at the onset of clinical breast cancer, C , via tumor growth, and ages at death D according to corresponding models.

Because the age-specific incidence of pre-clinical disease cannot be directly observed, we have to estimate such a quantity from the sojourn time distribution and age-specific incidence of the clinical disease. Specifically, we can derive the incidence of pre-clinical breast cancer backward from the following deconvolution formula:

$$I_c(y) = \int_0^y w_{hp}(t)w_{pc}(y-t|t)dt, \quad (1)$$

where $I_c(y)$ is the age-specific incidence of clinical breast cancer, w_{hp} is the instantaneous

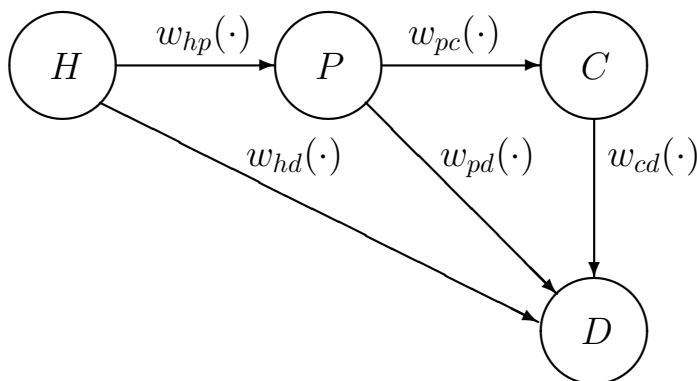


Figure 1: Summary of states, possible transitions, and transition densities for the natural history model. This scheme describes the progress of breast cancer in the absence of screening. All instantaneous probabilities of transition are indicated next to the corresponding transition. The two subscripts correspond to the origin and destination states, respectively.

probability of making a transition from H to P, and w_{pc} is age-specific sojourn time density.

Note that the age-specific incidence of clinical breast cancer can be observed and is often well documented in cancer registries or from the control arms of early detection trials. We use the age-cohort-specific breast cancer incidence estimates developed by Moolgavkar et al. (34). With a given distribution for the sojourn time of the pre-clinical disease state, the age-specific incidence of pre-clinical breast cancer (w_{hp}) can be estimated using the method of Parmigiani and Skates (35).

However, the estimation of the sojourn time distribution is not straightforward in general (36; 37; 38; 39; 40; 41). In this study, we focus on three commonly used parametric distributions for the sojourn time of the preclinical disease state, which are further modified to incorporate the effect of age at the onset of the preclinical disease.

We first consider a smoothed age-specific exponential sojourn time distribution:

$$w_{pc}(x|\lambda(t)) = \lambda^{-1}(t) \exp(-\lambda^{-1}(t)x),$$

where the mean sojourn time $\lambda(t)$ depends on the woman's age. To incorporate the uncertainty of the parameter into the model, we introduce an inverse gamma prior to the parameter λ , where the two parameters of the inverse gamma distribution are age-specific and are chosen to match the mean and standard deviation of sojourn times estimated from the Canadian National Breast Screening Studies (CNBSS) trials in Shen and Zelen (32).

An alternative assumption for the sojourn time distribution is the log-normal assumption. We consider a modified version taking into account the womans age at the onset of the preclinical disease, while generalizing the model by Spratt et al (42):

$$w_{pc}(x|t) = \frac{1}{\sqrt{2\pi}\sigma(t)x} \exp \left\{ -\frac{1}{2\sigma^2} (\log(x) - \mu(t))^2 \right\},$$

where the logarithm of the mean, $\mu(t)$ is specified to be a linear function of the womans age, t . An inverse gamma prior is used to incorporate the uncertainty for the variance, which does not show an age effect (11). The parameters of the inverse gamma are chosen to match the moments of the reported age-specific variances in Spratt et al (1986).

A modified tumor growth distribution of Peer et al (43; 44) is also used in our simulation for sensitivity analyses. Specifically, the sojourn time of the preclinical duration is modeled by the tumor growth rate or, equivalently, by the tumor doubling time. In particular, the relationship between sojourn time and tumor doubling time can be expressed as

$$X = \ln(V_p/V_c)DT / \ln(2),$$

where X is the sojourn time, DT is the tumor doubling time, and V_p and V_c are the volumes of a tumor at onset of detectable preclinical disease and at onset of clinical disease, respectively. We assume that the smallest tumor detectable by screening exam is 5mm, and that the average diameter at which breast cancer manifests is 20 mm (11). Tumor doubling times are assumed to follow an age-dependent log-normal distribution. The parameters in the predictive sojourn time distribution are estimated to match with the median and 95% quantile of the tumor doubling time based on findings from the Nijmegen trial (43). It is worth noting that there is a direct relationship between the tumor growth rate (or doubling time) and the sojourn time in the preclinical duration.

2.2 Survival Distributions and Mortality

One primary interest of the study is to evaluate the length of survival after diagnosis of breast cancer with various screening strategies. The survival distribution for women with breast cancer is determined by their age and tumor characteristics at diagnosis, and by the treatment they receive following diagnosis. Women in the cohort who receive periodic screenings are more likely to have breast tumors detected early and thus are more likely to have better prognoses than women who do not receive such screening. However, due to imperfect screening sensitivities and heterogeneity in pre-clinical durations, some breast cancer may still be clinically diagnosed between exams (interval cases). The survival distribution depends on screening only through the tumor characteristics and age at diagnosis.

Based on the natural history model, the tumor size and age at diagnosis are generated for a woman diagnosed to have breast cancer in the cohort. It is well known that lymph node involvement (nodal status) and the estrogen receptor (ER) status of the tumor (positive or negative) are also important risk factors, and are related to treatment options and survival. To estimate the number of positive nodes at diagnosis, a predictive model was developed using the data of a woman's age and tumor size at diagnosis from the SEER registries (11; 45). A constraint via the truncated Poisson distribution is given to ensure that the number of positive nodes for a screening- detected breast tumor is less than or equal to that for the same woman if her tumor is clinically detected. Without enough evidence to connect ER status with other risk factors, the ER status of a woman's breast tumor is simulated independently of the other risk factors, but according to the distribution for the general population. It is estimated that roughly 70% of breast tumors are ER positive (46).

As expected, the tumor characteristics at diagnosis will determine the treatment received thereafter. We assume that women in the cohort are treated according to the guidelines established by the NIH Consensus Conference on Early Breast Cancer (1991), given their risk factors including age, tumor ER status, tumor size, and nodal status at diagnosis. Whether a woman receives tamoxifen depends on her age and tumor ER status. The survival distribution for length with quality of life adjustment after diagnosis of breast cancer is estimated using a Cox regression model with covariates of treatment, age, tumor ER status, primary tumor size, and number of nodes involved. The predictive survival model was established based on a combined analysis of four CALGB trials (11; 47; 48; 49), as described in (50).

For a woman in the cohort, her age-specific mortality due to causes other than breast cancer is obtained from actuarial tables, using a 1960 birth cohort from the census database. If the breast-cancer-specific survival time for a woman is shorter than her simulated natural lifetime, then we assume that she died from breast cancer and contributed to the breast cancer mortality. Otherwise, we assume that she died from a competing cause.

2.3 Sensitivities of Mammography and Clinical Breast Examinations

The sensitivity of a screening program for the early detection of breast cancer plays a critical role in its potential for the reduction of disease-specific mortality. When a screening program involves more than one modality, it is important to obtain the sensitivity of each individual screening modality and the dependence structure among the multiple diagnostic tests (41; 51). This knowledge provides a basis to guide health policy makers in designing optimal and cost-effective screening programs.

Some recent studies reveal that the sensitivity of a screening exam is likely to depend on tumor size and age at the time of diagnosis (44; 32). Based on literature in the area of breast cancer screening and the estimates of screening sensitivities for both MM and CBE, we consider a model to relate the sensitivity of each modality with age and tumor size at diagnosis, respectively (11; 32). In particular, a logit function is employed to model the effects of age and tumor size at diagnosis on the sensitivities of mammography and clinical breast exam, respectively. We assume the sensitivity of each modality satisfying the following equation:

$$\beta_k(t, d) = \frac{\exp\{\alpha_{k0} + \alpha_{k1}(t - 45) + \alpha_{k2}(d - 2)\}}{1 + \exp\{\alpha_{k0} + \alpha_{k1}(t - 45) + \alpha_{k2}(d - 2)\}},$$

where t is the age at diagnosis, d is the diameter in centimeters of the primary tumor at diagnosis, $k = 1$ corresponds to mammography, and 2 is for CBE.

The coefficients in the logit models are determined based on the corresponding sensitivity estimates from the CNBSS trials (32) as follows. A sensitivity of mammography of 0.61 corresponds to a woman at age 45 with a tumor diameter of 2cm; a sensitivity of 0.1 corresponds to a woman at the same age but with a tumor size of 0.1 cm; and a sensitivity of 0.66 corresponds to a woman of age 55 with a tumor size of 2cm: $\beta_1(45, 2) = 0.61$, $\beta_1(45, 0.05) = 0.1$ and $\beta_1(55, 2) = 0.66$. Thus, the coefficients in the logit model are solved to be, $\alpha_{10} = 0.447$, $\alpha_{11} = 0.216$ and $\alpha_{12} = 1.36$ for mammography. In the same vein, we can solve the coefficients for the sensitivity of CBE: $\alpha_{20} = 0.364$, $\alpha_{21} = -0.077$ and $\alpha_{22} = 1.31$. Moreover, because the sensitivity can vary from subject to subject even when given the same age and tumor size (52), we use a beta distribution to reflect such a random variation for each sensitivity, while matching the corresponding mean and variance for the estimated sensitivity from the CNBSS trials, as reported in Shen and Zelen (51).

The Health Insurance Plan of Greater New York (HIP) trial and the CNBSS both offered independent annual clinical breast exams and mammograms to women in their study arms, which gave us an opportunity to assess the dependence between the two screening modalities. The analyses based on data from these trials indicate that mammography and clinical breast examinations contribute independently to the detection of breast cancer (51). Therefore, given the sensitivity of each individual screening modality, the overall sensitivity of a screening program using both MM and CBE is as follows:

$$\beta(t, d) = \beta_1(t, d) + \beta_2(t, d) - \beta_1(t, d)\beta_2(t, d),$$

when the two modalities are independent to each other.

2.4 Costs of Screening Programs

As expected in screening practices, the primary costs of a screening program is proportional to the total number of mammograms and clinical breast examinations given. Although there are additional costs related to follow-up confirmative tests such as a biopsy, and costs for the treatment of breast cancer at various stages after diagnosis, we will focus only on the cost of screening examinations in the current study. On its website, the National Cancer Institute lists the estimated cost of mammography in 2002 at \$100-200, and acknowledges that the cost can vary widely among different centers and hospitals. Since it is frequently part of a routine physical examination, the cost of a CBE is often less than that of mammography. In a public website promoting cancer prevention, the estimated cost for an annual CBE is \$45-55, whereas the cost of MM is \$75-150 (53). In the decision analysis, it is clear that the cost ratio of MM and CBE determines the results in the comparison of different screening strategies. Therefore, we investigate the effects of two cost ratios (1.5 and 2) between MM and CBE, and allow the cost for a CBE to be \$100. For simplification, we will not adjust for the type of currency, or for inflation over the years.

3 Optimization of Screening Strategies and Sensitivity Analyses

The focus of this investigation is to compare the effects of different breast cancer screening policies and the costs directly related to these policies, based on the models introduced in the last section. The health outcome of interest is the expected gain in quality-adjusted survival. We interpret this quality adjustment to be relative to a typical health history rather than that of a state of perfect health (50). Quality adjustments are important because they allow, with certain limitations, to account for the effects of medical intervention on morbidity as well as mortality. In screening this is especially important because of the so-called overdiagnosis problem. While beneficial to many women, screening leads to discovering cancer that would have not otherwise affected certain women's health. While length of life may be unaffected, this is a considerable loss of quality of life. Also, early detection can prolong the portion of one's life spent as a cancer survivor. The specific quality adjustments used in our model are the same as Parmigiani (11).

The marginal effectiveness for each screening strategy is calculated based on the difference between the expected quality-adjusted life in years for women in a cohort undergoing screening versus the same cohort of women without screening. The summaries of interest are the expected gain in quality life years (QALYS) and the expected total monetary cost for each screening strategy. Marginal cost is the difference in total cost between the screened and

unscreened cohorts. The marginal effectiveness for each screening strategy is the difference between the expected QALYS in the screened and unscreened cohorts. The ratio is marginal cost per year of quality-adjusted life saved (MCYQLS).

Three important issues to consider for screening policies are the age at which a woman should start a screening program, the screening frequency, and what screening modalities are to be used. In this study, we will evaluate a total of 48 screening strategies with the following combinations:

- The age to begin and end periodic screening: 40-79, 45-79, and 50-79 years;
- The interval between consecutive examinations: 0.5, 1, 1.5 and 2 year(s);
- The combined use of MM and CBE: whether mammogram or CBE is given for every one or every two exams.

Using the model described earlier, we generate a cohort of women and their natural histories of disease, and assess how the screening strategies interact with the disease process and the survival after diagnosis. The quantities of interest are estimated using the 100,000 Monte Carlo replicates, for each of the screening strategies.

In summary, we simulate a birth cohort of 100,000 women and follow them through the years. A fraction of them will develop breast cancer according to the age-specific incidence of pre-clinical breast cancer. For those women, we generate the natural histories of their disease, which include their ages at the onset of the preclinical disease, the pre-clinical durations (via tumor growth rates), and ages at the clinical onset of the disease. When a screening strategy is provided to a woman during a pre-clinical disease state, the probability that her cancer will be detected by this screening strategy is generated using the equations in Section 2.3, based on her age and tumor size at the time of the exam. If the diagnosis is missed during the exam, her breast cancer may be detected at her next scheduled exam or it may clinically manifest before the next exam depending on the sojourn time of the woman's preclinical disease state. Once a woman is diagnosed to have breast cancer, we obtain her tumor size and age at the time of detection. The information is then used to predict the woman's survival and quality-adjusted survival after the detection using models developed in Section 2.2. The expected cost is estimated based on the average cost of screening exams from the 100,000 women for each screening strategy in the simulation.

A balance sheet is a summary of the expected benefits and harms of an intervention. Its goal is to inform decision makers, and enable them to weigh benefits and harms according to their individual values (54; 55). Table 1 is a balance sheet for evaluating two alternative

	Screening Strategy	
	MM/1, CBE/1	MM/2, CBE/1
Additional number of MM per woman	33	17
Additional number of CBE per woman	33	33
Additional number of false positives per woman	5.2	4.3
Additional years of life per woman	0.144	0.124
Additional women detected in preclinical state	867	810
Women treated unnecessarily	55	51

Table 1: Balance sheet for two alternative screening strategies: annual MM and CBE screening and biannual MM and annual CBE. In both cases screening starts at 40 years of age and stops at age 79. Values are increments compared to no screening for a cohort of 10000 breast cancer women.

screening strategies, based on the model of this chapter. We consider annual MM and CBE screening and biannual MM and annual CBE. Differences between the two columns can inform decision makers about whether annual or biannual MM are to be preferred once annual CBE is planned. Elmore and colleagues (56) collected data on a retrospective cohort study of breast cancer screening and diagnostic evaluations among 2400 women who were 40 to 69 years old at study entry. False positive results occurred in 6.5% of the mammograms, an estimate that was used here to translate the estimated number of additional tests into estimated false positives. In addition we assume that positive CBE's would be followed by a mammography, that 10% of CBE are false positive, and the two tests are independent of each other. Then the overall false positive number per woman for the 1st strategy is: $(0.065 + 0.1 - 0.1 * 0.065) * 33 = 5.2$; and the overall false positive number per woman for the 2nd strategy is $0.065 * 17 + 0.1 * 33 - 0.1 * 0.065 * 17 = 4.3$.

In Section 2.1, three model specification are discussed for the distribution of sojourn times in the preclinical state of the disease. It is of interest to investigate how these different models may impact the QALYs and expected cost of each screening strategy reported by (33). We find that the analyses are fairly robust for the three model assumptions. The marginal QALYS is slightly higher (about 1-2%) for the lognormal model than for the exponential model for a given screening strategy. The relative marginal costs and QALYS among the screening strategies under evaluation are similar for the three model choices.

4 Discussion

Much attention has been focused on the early detection capabilities of new breast cancer screening technologies, including advances in mammography and MRI. The importance of clinical breast examination in breast cancer screening programs seems to be unclear. Even

though some recent studies have indicated that regular CBE in addition to MM can be important in the early detection of breast cancer, few studies have investigated the optimal use of both mammography and clinical breast exam to reduce the mortality of breast cancer while balancing the associated burdens and costs to women and to the health care system.

Developing early detection guidelines and making public health policy requires careful consideration of the long-term benefits, costs, and feasibility associated with the screening strategies. In Shen and Parmigiani (33), we explore the trade-off between the QALYS and costs related to each screening strategy among several combinations of starting ages of screening, frequencies of screening, and the use of two screening modalities. The study indicates that starting from 40 years of age, a biennial mammogram is often cost-effective for women who undergo annual clinical breast exams. Given the cost to women who are already receiving care for other health issues or regular check-ups in a clinic, an annual CBE as part of their routine examination should not add much burden. Our analyses also indicate that CBE alone cannot replace regular mammography in screening practice, but can be used complementarily or alternatively in a screening program.

The decision analysis methodology and simulation techniques we have developed for this study can be directly applied to investigate other screening strategies, and even to other chronic diseases with certain modifications to the models. We have modeled screening sensitivity for MM and CBE, respectively, through age and tumor size at diagnosis. We have also introduced random variations for the parameters to incorporate uncertainty of data input and population heterogeneity. We have considered various sojourn time distribution models, and have derived them based on data from the large randomized breast cancer screening trials of the HIP (57), CNBSS (58), and the Nijmegen Trial (43), while taking random variations into consideration. We have performed sensitivity analyses to assess the robustness of the patterns of benefit and cost with the alternative models.

Our study has several limitations. The cost of a biopsy following a CBE or MM that is positive for breast cancer has not been considered in the analysis. Moreover, we have not included the potential costs of false-positive exams, such as the anxiety, fear and discomfort that are associated with a biopsy. In fact, it is often difficult to convert these factors into dollar amounts (56). In addition, we have not included important cost components, which are the costs of follow-up procedures undertaken after the detection of breast cancer. This is in part due to the great variation in treatment protocols and in the cost of treating breast cancer that has existed over the years. Finally, we have used a hypothetical birth cohort of women with 100% compliance in the simulations for each screening strategy. In reality, it is rare to have 100% compliance for any screening program, and a real cohort would be

dynamic, which would include changes in the cohort due to migration.



References

- [1] HK W; MJ T; BF H; et al. Annual report to the nation on the status of cancer, 1975-2000, featuring the uses of surveillance data for cancer prevention and control. *Journal of the National Cancer Institute*, 95(17):1276–1299, 2003.
- [2] Cancer Surveillance Research Program NCI. The surveillance, epidemiology, and end results (seer) program. *www.seerims.nct.nih.gov*, 1998.
- [3] International Agency for Research on Cancer (IARC) IA. The cancer-mondial website. *www.depiarc.fr*, 1999.
- [4] Fletcher SW; Elmore JG. Mammographic screening for breast cancer. *New England Journal of Medicine*, 348(17):1672–1680, 2003.
- [5] Wang L. Mammography and beyond: Building better breast cancer screening tests. *Journal of the National Cancer Institute*, 94(18):1346–1347, 2002.
- [6] Berry DA. Benefits and risks of screening mammography for women in their forties: A statistical appraisal. *J Natl Cancer Inst*, 90:1431–1439, 1998.
- [7] Gordis L; Berry D; Chu S; et al. Breast cancer screening for women ages 40-49. *J Natl Cancer Inst*, 89:1015–1026, 1997.
- [8] Canadian Task Force on Preventive Health Care. Preventive health care, 2001 update: screening mammography among women aged 40–49 years at average risk of breast cancer. *CMAJ*, 164(4):469–476, 2001.
- [9] Parmigiani G. Measuring uncertainty in complex decision analyses models. *Statistical Methods in Medical Research*, 11(6):513–37, 2002.
- [10] van Oortmarseen G; Boer R; Habbema J. Modeling issues in cancer screening. *Statistical Methods in Medical Research*, 4:33–54, 1995.
- [11] Parmigiani G. *Modeling in Medical Decision Making*. Wiley, Chichester, 2002.
- [12] Parmigiani G. On optimal screening ages. *Journal of the American Statistical Association*, 88:622–628, 1993.
- [13] De Koning H. Breast cancer screening; cost-effectiveness in practice. *Eur J Radiol*, 33:32–7, 2000.
- [14] Lindfors KK; Rosenquist CJ. The cost-effectiveness of mammographic screening strategies. *Journal of American Medical Association*, 274:881–884, 1995.
- [15] Brown ML. Sensitivity analysis in the cost-effectiveness of breast cancer screening. *Cancer*, 69 (7 Suppl):1963–1967, 1992.

- [16] Mushlin AI; Fintor L. Is screening for breast cancer cost-effective? *Cancer*, 69 (7 Suppl):1957–1962, 1992.
- [17] van Ineveld BM; van Oortmarssen GJ; de Koning HJ; Boer R; van der Maas PJ. How cost-effective is breast cancer screening in different EC countries? *European Journal of Cancer*, 29:1663–1668, 1993.
- [18] Brown ML; Fintor L. Cost-effectiveness of breast cancer screening: preliminary results of a systematic review of the literature. *Breast Cancer Res Treat*, 25:113–118, 1993.
- [19] Elixhauser A. Costs of breast cancer and the cost-effectiveness of breast cancer screening. *Int J Technol Assess Health Care*, 7:604–615, 1991.
- [20] Clark RA. Economic issues in screening mammography. *American Journal of Roentgenology*, 158:527–534, 1992.
- [21] Parmigiani G; Kamlet M. Cost-utility analysis of alternative strategies in screening for breast cancer. In C Gatsonis; J Hodges; RE Kass; N Singpurwalla, eds., *Case Studies in Bayesian Statistics*, 390–402. Springer, New York, 1993.
- [22] Eddy DM; Hasselblad V; McGivney W; Hendee W. The value of mammography screening in women under age 50 years. *Journal of the American Medical Association*, 259:1512–1519, 1989.
- [23] Carter R; Glasziou P; van Oortmarssen G; de Koning H; Stevenson C; Salkeld G; Boer R. Cost-effectiveness of mammographic screening in Australia. *Australian Journal of Public Health*, 17:42–50, 1993.
- [24] Saltzmann P; Kerlikowske K; Phillips K. Cost-effectiveness of extending screening mammography guidelines to include women 40 to 49 years of age. *Ann Intern Med*, 127:955–965, 1997.
- [25] Burnside E; Belkora J; Esserman L. The impact of alternative practices on the cost and quality of mammographic screening in the United States. *Clinical Breast Cancer*, 2(2):145–152, 2001.
- [26] Fett M. Computer modelling of the Swedish two country trial of mammographic screening and trade offs between participation screening interval. *J Med Screen*, 8(1):39–45, 2001.
- [27] Yasmeeen S; Romano P; Pettinger M; Chlebowski R; Robbins J; Lane D; Hendrix S. Frequency and predictive value of a mammographic recommendation for short-interval follow-up. *J Natl Cancer Inst*, 95(6):429–436, 2003.

- [28] Kerlikowske K; Smith-Bindman R; Sickles E. Short-interval follow-up mammography: Are we doing the right thing? *J Natl Cancer Inst*, 95(6):418–419, 2003.
- [29] Barton MB; Harris R; Fletcher SW. Does this patient have breast cancer? The screening clinical breast examination: Should it be done? How? *JAMA*, 282:1270–80, 1999.
- [30] Baines CJ; Miller AB; Bassett AA. Physical examination. its role as a single screening modality in the canadian national breast screening study. *Cancer*, 63:1816–22, 1989.
- [31] Bobo J; Lee N; Thames SF. Findings from 752081 clinical breast examinations reported to a national screening program from 1995 through 1998. *JNCI*, 92:971–6, 2000.
- [32] Shen Y; Zelen M. Screening sensitivity and sojourn time from breast cancer early detection clinical trials: mammograms and physical examinations. *JCO*, 19:3490–9, 2001.
- [33] Shen Y; Parmigiani G. A model-based comparison of breast cancer screening strategies: Mammograms and clinical breast examinations. *Journal of the National Cancer Institute*, submitted, 2003.
- [34] Moolgavkar SH; Stevens RG; Lee JAH. Effect of age on incidence of breast cancer in females. *Journal of the National Cancer Institute*, 62:493–501, 1979.
- [35] Parmigiani G; Skates S. Estimating the age of onset of detectable asymptomatic cancer. *Mathematical and Computer Modeling*, 29:in press, 2001.
- [36] Albert A; Gertman P; Louis T. Screening for the early detection of cancer: I. the temporal natural history of a progressive disease state. *Mathematical Biosciences*, 40:1–59, 1978.
- [37] Day NE; Walter SD. Simplified models of screening for chronic disease: Estimation procedures from mass screening programmes. *Biometrics*, 40:1–13, 1984.
- [38] Brookmeyer R; Day NE; Moss S. Case-control studies for estimation of the natural history of preclinical disease from screening data. *Statistics in Medicine*, 5:127–138, 1986.
- [39] Etzioni RD; Shen Y. Estimating asymptomatic duration in cancer: the AIDS connection. *Statistics in Medicine*, 16:627–644, 1997.
- [40] Straatman H; Peer PG; Verbeek AL. Estimating lead time and sensitivity in a screening program without estimating the incidence in the screened group. *Biometrics*, 53:217–229, 1997.

- [41] Shen Y ZM. Parametric estimation procedures for screening programmes: Stable and nonstable disease models for multimodality case finding. *Biometrika*, 86:503–515, 1999.
- [42] Spratt JS; Greenberg RA; Heuser LS. Geometry, growth rates, and duration of cancer and carcinoma in situ of the breast before detection by screening. *Cancer Research*, 46:970–974, 1986.
- [43] Peer P; van Dijck JAAM; Hendriks J; Holland R; Verbeek ALM. Age-dependent growth rate of primary breast cancer. *Cancer*, 71:3547–3551, 1993.
- [44] Peer P; Verbeek A; Straatman H; Hendriks J; Holland R. Age-specific sensitivities of mammographic screening for breast cancer. *Breast Cancer Research and Treatment*, 38:153–160, 1996.
- [45] National Cancer Institute: Surveillance, Epidemiology, and End Results (SEER) Program. Seer homepage. <http://www-seer.ims.nci.nih.gov>, 1997.
- [46] National Cancer Institute: Chemoprevention of Estrogen Receptor (ER) Negative Breast Cancer Preclinical Studies. Nih homepage. <http://grants1.nih.gov/grants/guide/rfa-files/RFA-CA-03-005.html>, 2002.
- [47] Wood W; Weiss R; Tormey D; Holland J; Henry P; Leone L; et al. A randomized trial of CMF versus CMFVP as adjuvant chemotherapy in women with node-positive stage ii breast cancer: a CALGB study. *World J Surg*, 9:714–718, 1985.
- [48] Perloff M; Norton L; Korzun A; Wood W; Carey R; Gottlieb A; et al. Postsurgical adjuvant chemotherapy of stage ii breast carcinoma with or without crossover to a non-cross-resistant regimen: a cancer and leukemia group b study. *J Clin Oncol*, 14:1589–98, 1996.
- [49] Wood W; Budman D; Korzun A; Cooper M; Younger J; Hart R; et al. Dose and dose intensity of adjuvant chemotherapy for stage ii, node-positive breast carcinoma. *N Engl J Med*, 330:1253–1259, 1994.
- [50] Parmigiani G; Berry DA; Winer EP; Tebaldi C; Iglehart JD; Prosnitz L. Is axillary lymph node dissection indicated for early stage breast cancer—a decision analysis. *Journal of Clinical Oncology*, 17(5):1465–1473, 1999.
- [51] Shen Y; Wu D; Zelen M. Testing the independence of two diagnostic tests. *Biometrics*, 57:1009–1017, 2001.

- [52] Kerlikowske K; Grady D; Barclay J; Frankel SD; Ominsky SH; Sickles EA; Ernster V. Variability and accuracy in mammographic interpretation using the american college of radiology breast imaging reporting and data system. *J Natl Cancer Inst*, 90(23):1801–1809, 1998.
- [53] PREVENTION. Cancer tests worth paying for. <http://www.prevention.com/cda/feature/0,1204,876,00.html>, 2002.
- [54] Matchar DB; Samsa GP. Using outcomes data to identify best medical practice: the role of policy models. *Hepatology*, Jun;29(6 Suppl):36S–39S, 1999.
- [55] Barratt A; Irwig L; Glasziou P; et al. Users' guides to the medical literature: XVII. how to use guidelines and recommendations about screening. evidence-based medicine working group. *JAMA*, 281(21):2029–2034, 1999.
- [56] Elmore JG; Barton MB; Mocerri VM; Polk S; Arena PJ; Fletcher SW. Ten-year risk of false positive screening mammograms and clinical breast examinations. *New England Journal of Medicine*, 338:1089–1096, 1998.
- [57] Shapiro S. Periodic screening for breast cancer: The HIP randomized controlled trial. *Monogr Natl Cancer Inst*, 22:27–30, 1997.
- [58] Miller A; To T; Baines C; Wall C. The Canadian national breast screening study: Update on breast cancer mortality. *Monogr Natl Cancer Inst*, 22:37–41, 1997.

