Chapter 8:

The SPECTRUM Population Model of the Impact of Screening and Treatment on U.S. Breast Cancer Trends From 1975 to 2000: Principles and Practice of the Model Methods

Jeanne Mandelblatt, Clyde B. Schechter, William Lawrence, Bin Yi, Jennifer Cullen

Objective: **This stochastic simulation model was developed to estimate the impact of screening and treatment diffusion on U.S. breast cancer mortality between 1975 and 2000.** ***Modeling Approach:*** **We use an event-driven continuous-time state transition model. Women who are destined to develop breast cancer may be screen detected, present with symptoms, or die of other causes before cancer is diagnosed. At presentation, the cancer has a stage assigned on the basis of mode of detection. Cancers are assumed to be estrogen receptor (ER) positive or negative. Data on screening and treatment diffusion are based on national datasets; other parameters are based on a synthesis of the evidence available in the literature.** ***Model Methods:*** **The model is calibrated to predict incidence and stage distribution (in situ, local, regional, and distant). Other than screening or treatment, background events that affect mortality are not explicitly modeled but are captured in the deviation between model projections of mortality trends and actual trends. We assume that: 1) tumors progress more slowly in older age groups, 2) screen- and clinically detected disease have the same survival conditional on age and stage, 3) women do not die of breast cancer within the "lead time" period, 4) screening benefits are captured by shifts in stage at diagnosis, 4) tamoxifen benefits only ER-positive women, and 5) preclinical sojourn time and dwell times in each of the clinical stages are stochastically independent.** ***Model Results:*** **Dissemination of screening and therapeutic advances had a substantial impact on mortality trends. We estimate that, by the year 2000, diffusion of screening lowered mortality by 12.4% and treatment improvements and dissemination lowered mortality by 14.6%.** ***Conclusions:*** **Models such as this one can be useful to translate clinical trial findings to general populations. This model can also be used inform policy debates about how to best achieve targeted reductions in breast cancer morbidity and mortality. [J Natl Cancer Inst Monogr 2006;36:47–55]**

After lung cancer, breast cancer has been the leading cause of cancer morbidity and mortality among women and has affected up to 200 000 women in the United States each year *(1–3)*. Data from randomized controlled trials suggest that, at least among women participating in research, this critical cancer burden may be reduced though the application of screening *(4–8)* and new therapeutic modalities *(9–11)*. However, it is more difficult to evaluate the effectiveness of the introduction and dissemination of mammography and new treatment advances in broader, diverse populations and to understand how these cancer control interventions have actually affected the dynamic trends in breast cancer incidence and mortality over the past 30 years.

In such situations, modeling can be useful to evaluate the relative contributions of screening and improved treatment to the observed mortality trends over time *(12,13)*. Models can be used to synthesize information from diverse sources to make

Affiliations of authors: Departments of Oncology (JM, WL, BY, JC) and Medicine (JM), and Lombardi Comprehensive Cancer Center (JM, WL, BY, JC), Georgetown University Medical Center, Washington, DC; Department of Family and Social Medicine, Albert Einstein School of Medicine, Bronx, New York (CBS).

Correspondence to: Jeanne Mandelblatt, MD, MPH, 2233 Wisconsin Ave., Ste. 317, Washington, DC 20007 (e-mail: mandelbj@georgetown.edu).

See "Notes" following "References."

DOI: 10.1093/jncimonographs/lgj008

inferences about sequences of health events in the target population. Models can also be used to extend the time horizon of observation from that included in research studies *(14,15)* and test hypotheses that would be unethical or infeasible to test in a clinical trial (e.g., varying compliance rates) *(16)*. Models can also be used to inform public-policy decisions. Finally, model assumptions can be tested in sensitivity analysis to understand how changes in population behavior and other model parameters can affect population forecasts of breast cancer trends.

The SPECTRUM project has developed a continuous time stochastic population simulation model to estimate the impact of mammography and evolving treatment paradigms on breast cancer mortality in the United States between 1975 and 2000. This model was developed within the framework of a collaborative group modeling effort funded by the National Cancer Institute known as the Cancer Intervention and Surveillance Modeling Network (CISNET). In this paper, we describe the model structure and data sources used for reporting trends from 1975 to 2000, present a summary of our estimates of the relative impact of screening and treatment advances on the observed mortality, and discuss the context for interpretation of model results.

Methods

Modeling Philosophy

In developing this model, we made the decision to create a model based on the best-quality published data that were relevant to the disease process. As such, we do not estimate parameters. When calibration was necessary (e.g., to match to observed incidence of disease), we chose to make adjustments only to those parameters for which direct observation is not possible (e.g., screening test sensitivity and preclinical sojourn time) and to constrain our values to lie within the range of published estimates.

To maximize comparability with other CISNET models, we also chose to use the "base case" data set provided by the National Cancer Institute to initiate the model (see below). Beyond simple arithmetic transformations, we did not alter these data in our model simulation. Finally, we decided to calibrate the model to observed incidence and stage distributions, rather than mortality rates, since this allowed us to capture unmodeled background trends by comparing the observed and model predicted mortality.

Model Overview

We use an event-driven continuous-time state transition model. Women from different birth cohorts are simulated one at a time, and the times at which relevant events occur are determined by sampling from prespecified time-interval distributions. We simulate 55 million women to obtain reasonably smooth estimates of the mortality curves. Using U.S. Census data, we begin with women born in 1890 to simulate the population distribution of adult women alive in 1975.

The simulation of each woman begins by generating a date of birth and a date of death from causes other than breast cancer. If that date precedes 1975, no further action is taken and no output generated for that woman. The birth cohort–specific breast cancer incidence function, one of the model inputs, is then sampled and a date of clinical diagnosis of breast cancer is identified for those women who will develop the disease. A stage at clinical presentation, ER status, and a sojourn time for the lesion are then sampled from the distribution of ER statuses and stages for unscreened women with breast cancer and the sojourn time distribution (which are inputs to the program). From the sojourn time we calculate the earliest date after which a mammogram could result in detection. When the impact of screening is being modeled, a schedule of mammograms is also generated. The life history of the woman is then simulated by "carrying out" these events in chronological order. If a mammogram results in detection before clinical presentation, a stage at screen detection is calculated. Whenever a breast cancer is detected, whether clinically or through screening, treatment is assigned (based on input distributions of treatment choices) and then a date of death from breast cancer is determined based on treatment and age and stage at diagnosis.

The model contains several key components used to capture these events. The demographic component generates a population of simulated women having the age distribution of the female population of the United States in 1975. The breast cancer incidence component randomly selects which simulated women will develop breast cancer, with what estrogen receptor status, and at what time and in what stage, if the cancer presents clinically. The screening impact component governs the performance characteristics of screening, including screening test sensitivity and specificity. This portion of the model also calculates a stage shift for the tumor (based on sampling from age-specific stage distributions among screened women) conditional on the lead time realized by the screening test that detects it. The screening usage component determines when simulated women undergo breast cancer screening on the basis of a model of the observed diffusion through the population between 1975 and 2000. The treatment component is activated whenever a tumor is diagnosed (clinically or by screening) and selects a treatment and a corresponding breast cancer survival time based on Surveillance, Epidemiology, and End Results (SEER) data for age, stage, estrogen receptor status, and treatment-specific survival. Competing mortality is estimated using actuarial methods. The next sections describe our approach to modeling the natural history of breast cancer and detail each component of the model. Data input parameters and sources for each component are also summarized in Table 1.

Natural History of Disease

The model makes minimal assumptions about the natural history of breast cancer. All aspects of breast cancer are modeled in terms of stage (SEER historical stages of in situ, local, regional, and distant), estrogen receptor (ER) status (positive or negative), and the age of the woman at diagnosis, and treatment selected. These assumptions imply, in particular, that any effect of screening on survival is the result of stage shift (and, to a lesser extent, age shift in presentation). Screen-detected lesions are assigned the same ER status they would have had if they had presented clinically *(20)*.

Projected age-specific incidence rates in the absence of screening for each birth cohort from 1890 through 1970 were provided by the National Cancer Institute and include secular trends in incidence for each birth cohort [see chapter 5 of this monograph *(17)*]. These incidence rates were estimated using an age–period–cohort model that is described elsewhere *(18)*. These incidence rates are used as the hazard in a survival process, where failure consists of incident breast cancer. The corresponding

Table 1. Model parameters

Model component	Parameter description			Parameter value			Ref
Natural history of disease	Age-specific breast cancer incidence rates by birth cohort			Base case data parameter			*(17,18)*
	Age-specific stage distribution for clinically diagnosed tumors			Base case data parameter			*(17)*
	Age-specific stage distribution for screen-detected tumors	Age range	% In situ	% Local	% Regional	% Distant	*(19)*
		26–39	21.8	51.8	21.5	5.0	
		40–44	24.0	51.0	20.5	4.5	
		45–49	22.8	57.4	14.9	5.0	
		50–54	19.0	62.0	14.0	5.0	
		55–59	15.0	64.0	16.0	5.0	
		60–64	16.0	66.0	14.0	4.0	
		65–69	16.2	66.7	12.6	4.5	
		70–74	12.0	69.0	13.5	5.5	
		75–79	10.9	69.3	14.4	5.4	
		80–84	8.9	70.3	13.9	6.9	
		≥85	10.9	70.3	12.4	6.4	
	Age-specific fraction of incident breast cancers that are estrogen receptor positive	Age range					*(20)*
		26–44	60.0%				
		45–54	65.0%				
		55–64	74.0%				
		65–74	77.0%				
		≥75	83.0%				
	Average dwell times	Stage					(See "Methods" section)
		DCIS	3.0 y				
		Local	5.3 y				
		Regional	11.4 y				
	Age-specific mean preclinical sojourn time	Age range					*(21–24)*
		26–49	1.6667 y				
		50–54	2.0667 y				
		55–59	3.2667 y				
		60–69	3.8667 y				
		≥70	5.20 y				
Demographic Screening	Distribution of birth years for simulated women			Base case data parameter			*(25)*
	Diffusion of mammography			Base case data parameter			*(17)*
	Age-specific sensitivity and specificity of initial and subsequent mammography	Age range					*(26–29)*
		26–39*		Initial sensitivity	.773		
				Subsequent sensitivity	.850		
		40–49		Initial sensitivity	.867		
				Subsequent sensitivity	.85		
		50–69		Initial sensitivity	.941		
				Subsequent sensitivity	.85		
		≥70		Initial sensitivity	.912		
				Subsequent sensitivity	.85		
	Projected incidence without screening–age–period–cohort model			Base case data parameter			*(18)*
Treatment and Survival	Age–stage–ER specific distribution of treatments by year of diagnosis			Base case parameter			*(30)*
	Probability that tam, if used, will be given for 5 y			Base case parameter			*(30)*
	Dissemination of nonhormonal chemotherapy and tam			Base case data parameter			*(31)*
	Age–stage–ER specific survival without adjuvant treatment			Base case data parameter			*(17)*
	Annual reduction in odds of death associated with adjuvant treatment			Base case data parameter			*(7,8,30)*
	Mortality from all causes other than breast cancer						*(32)*

*Since incidence is near zero from ages 20 to 25, we do not include a sensitivity range for this age group. DCIS = ductal carcinoma in situ; Tam = tamoxifen; ER = estrogen receptor.

survival function is sampled for each woman, given her year of birth, to determine when she will develop clinical breast cancer. Because the survival function does not go to zero, or even near zero, most women will never develop breast cancer.

For those women for whom a date of incident clinical breast cancer is ascertained, a preclinical sojourn time is also simulated. Sojourn times are assumed to be exponentially distributed, with an age-dependent mean, based on published data *(21–24)*. The appropriate distribution is sampled to determine the preclinical sojourn time. The screening module then determines the actual date of preclinical incidence (if it occurs).

We assume that the dwell time in each stage is exponentially distributed, with mean stage dwell times as input to the program. Because few breast cancer lesions are left untreated for long, direct observation of dwell time in any clinical stage is not feasible. There is not widespread agreement on the best means of estimating these parameters. We estimated mean dwell times in the stages as the reciprocals of the annual rates of transition from each stage to the next (e.g., from local to regional disease or from regional to distant). These transition rates were themselves estimated using screened and unscreened stage distribution data from randomized controlled trials of breast cancer screening *(4,5,27,33–36)*, and fitting the transition probabilities of an underlying Markov stage transition process to that data, incorporating the studies' estimated lead times.

When a tumor is diagnosed by screening, a stage at screen detection is simulated as follows. First, the lead time is calculated by subtracting the screening date from the date at which the tumor was scheduled to present clinically. The marginal distribution of stages of screen-detected tumors is taken from the observed distribution of age-specific stages among tumors reported to SEER during the most recent period when screening was widespread (1995–1999) *(19)*. Among women who were reported as "stage unknown," we reassigned 50% as regional and 50% as distant stage on the basis of survival.

Because we have assumed that the dwell time in each stage follows an exponential distribution, standard formulas *(37)* can be used to calculate the probability that a tumor in any given stage at the time of screening would, during the calculated lead time, make the number of stage transitions that places it in the simulated stage at clinical diagnosis. Bayes' theorem is then applied to calculate the distribution of stage at screening conditional on the stage at clinical presentation and the lead time. This distribution is then sampled to identify the simulated stage at screen detection as a model input.

For example, suppose that a simulated woman who would have presented clinically with regional disease at age 68 is detected through screening 5 years earlier. For all women with lesions that are screen detected at age 63 (irrespective of lead time or subsequent stage), the stage probabilities for ductal carcinoma in situ (DCIS), local, regional, and distant disease are 0.16, 0.66, 0.14, and 0.04, respectively. We also know that her lesion would have been in the regional stage 5 years later but for the screening that detected it earlier. The probability that a distant lesion at age 63 would be regional 5 years later is 0, as regression is disallowed in our model. From our assumed stage dwell time distributions, we calculate that the probability that a regional lesion would remain regional 5 years later is approximately 0.645 *(37)*. (The remaining 35.5% of such lesions would progress to distant.) The probability that a local lesion would advance to regional (but no further) over 5 years is 0.478. And the probability that a DCIS lesion would be in the regional stage 5 years later is 0.295. Applying Bayes' theorem, given that the lesion will be regional 5 years later, the probability that the screened stage is DCIS is $0.16 \times 0.295/(0.16 \times 0.295 + 0.66 \times 0.478 + 0.14 \times 0.645 + 0.04 \times 0.00)$, or, to three decimal places, 0.104. Similarly, the probabilities that the screened stage is local, regional, or distant are 0.696, 0.199, and 0 respectively. (These do not add up precisely to 1.0 because of rounding.) This stage distribution is then randomly sampled to identify the stage at which this woman's screen-detected lesion actually presents.

One implication of the lead time is that a woman who is screen detected several years earlier than she would have presented clinically may end up getting less intensive treatment because the more intensive treatment, such as multidrug chemotherapy regimens, was not yet in sufficiently widespread use. Such a woman could actually end up with a worse prognosis as a result of screening, because she got her diagnosis in an earlier era when chemotherapy was not being widely used. Although such events do occur in our model, they occur with a low frequency and probably do not have a substantial impact on the results.

We model the incidence of only first breast cancers. Accordingly, the correct denominator for an incidence rate should be the number of women alive who have never had breast cancer. However, in compliance with the procedures adopted by the CISNET collaboration for calculating incidence rates, we actually use the count of all women alive for this denominator. This approach results in a slight underestimate of the incidence rates. The extent of this underestimate increases with a woman's age, reflecting both the rising rates of breast cancer and the falling surviving population denominator, and tends to increase over time among those over age 50. However, the underestimate of incidence never exceeds 1% for women under age 50 and reaches only 5.4% for women aged 75–79 after 1994.

The preclinical sojourn time is one of the "tunable" parameters of our model. That is, unlike, for example, incidence rates that are directly observable and for which excellent data exist, the sojourn time is a latent variable and can be estimated only by fitting models to population screening programs. Thus, we varied our estimate within the range of published estimates to generate simulated incidence and stage distribution in screened women that best corresponded to observed incidence.

Our original values for sensitivity and sojourn time were in the center of the range of published estimates. Because our output resulted in undergeneration of early-stage disease, we revised our values for both of these parameters upward. This was done by partitioning the intervals between our initial values and the highest published estimates in thirds and trying all nine combinations. This calibration was performed using a few simulations (5 million women per run). The resulting output was compared to the observed incidence and stage distribution. The combination of values that came closest (by chi-square statistic) to reproducing the observed distribution was retained and used for all later runs. This combination, however, still fell far short of producing the observed number of DCIS lesions.

Demographic Component of the Model

Each birth year is selected in our model with a frequency proportional to its prevalence among the U.S. female population in 1975 and inversely proportional to the probability of survival to 1975 *(25)*. This method ensures that the 1975 age distribution we simulate matches that of the given 1975 U.S.

female population. Each simulated person's life is modeled from birth, including the application of cancer incidence functions. Thus, a woman may develop breast cancer before 1975, and if she does not die before 1975, she will be a prevalent case at the start of the model. Women born between 1890 and 1975 but who die before 1975 are also simulated, but we do not include their data in the output.

Screening Utilization Component of the Model

We use existing data from a model of screening use over time to reflect the dissemination of mammography in the U.S. population [see chapter 5 *(17)*]. These data were based on fitting parametric frailty models to national screening data. However, these data covered only patterns occurring for women born between 1891 and 1970. Women born in 1890 were assumed to obtain screening at the same ages as women born in 1891.

Screening Impact Component of the Model

Each screening event (i.e., obtaining a mammography in a given year or not) is simulated by drawing a random number from a uniform distribution between zero and one. If screening occurs, the test sensitivity and specificity and the presence or absence of a tumor during its preclinical sojourn time are used to generate a test result (true positive, false positive, true negative, or false negative). True-positive test results trigger a diagnosis of breast cancer and calculation of the stage at presentation (and assignment of an ER status). We assume that screen-detected and clinically detected interval cancers (false negatives and clinically detected in the absence of screening) have similar tumor characteristics (i.e., distribution of ER) and that conditional on age and stage at diagnosis, ER status, and treatment, they have the same survival functions. To the extent that screen-detected tumors are less virulent than interval cases, then mortality reductions associated with screening may be slightly overestimated.

The sensitivity of mammography is the other tunable parameter in our model. In our program, sensitivity is a ratio, with the numerator consisting of positive test results among those with a tumor, and the denominator consisting of those with a tumor. In our model, "with a tumor" is implemented as "occurring during the preclinical sojourn time of a lesion." "With a tumor" therefore is an abstract, unobservable construct whose value cannot be directly measured but can be estimated by fitting statistical models to the data from large screening programs. As a starting point, we relied on published age-specific estimates of sensitivity from different points in time *(26–29)*. We assume that sensitivity is greatest for the first screen and then decreases over time with repeated screenings, but we do not vary the sensitivity according to tumor size or tumor growth over the preclinical sojourn time. Sensitivity is assumed to be age dependent for the first screening but age independent thereafter. We made this choice because there were good data on test performance as a function of age for first screening examinations, but fewer data on the results for later screens over time by age. We also model test sensitivity as a constant over the period of simulation. Although test performance is likely to have improved between 1975 and 2000, there was enough variability in published estimates from large screening trials by period *(26–29)* that no reasonable time- dependent curve could be fitted to the observed data.

Treatment and Survival Component

Adjuvant treatments gradually disseminated into practice after 1975. We used data from Mariotto and colleagues to estimate the dissemination of nonhormonal chemotherapy and tamoxifen (see chapter 2) *(30,38)*. Since the two main surgical options—mastectomy and breast conservation—have equivalent survival *(39,40)* we do not include any changes in local treatment approaches over time. For cancers diagnosed before 1975, the 1975 treatment distributions were used.

Data from 1975 were used to estimate survival in the absence of adjuvant treatment with multiagent chemotherapy or tamoxifen [see chapter 5 *(17)*]. Women receiving tamoxifen or chemotherapy were assigned a survival time on the basis of a modification of the 1975 survival curve using data from large meta-analyses (see chapter 2) *(7,8,30)*. For each therapy, the survival function for the base 1975 data is adjusted using the annual reduction in the odds of death associated with each modality. We then sample from the modified survival function to project survival given each therapy; the survival function is the same regardless of mode of detection. Only women with ER-positive tumors are assumed to have survival benefits associated with tamoxifen.

Because survival is calculated conditional on age at diagnosis, stage at diagnosis, ER status, and treatment, stage shifts can result in improved prognosis. We calculate survival from the date of clinical presentation, even if the lesion was screen detected. As a consequence, death from breast cancer cannot occur during the lead time. Death from other causes, however, can occur in the lead time. We do not present quality-adjusted survival, as the base model was designed to estimate the potential impact of screening and treatment on observed incidence and mortality in the period of interest.

Competing Mortality Component

Death from causes other than breast cancer was estimated using birth cohort–specific annual mortality data [see chapter 3 *(32)*].

Model Validation

We used several approaches to assess the internal reliability of the model and the validity of the results. For example, during programming, as each model component was coded, the code was tested to verify that key inputs were reproducible as outputs. For example, when the demographics module was programmed, the program was run, outputting the distribution of ages in 1975. This output was then compared to the known distribution of ages of U.S. women in 1975 to verify an adequate match. When the incidence of breast cancer was set to an arbitrary fixed constant, we verified that the output incidence followed suit. In runs where screening was not modeled, the simulated outputs of incidence and stage distribution of cancers in the absence of screening by birth cohort and age-matched these inputs. We also verified that the results with screening turned off were identical those obtained when the sensitivity of screening was set to zero. A range of similar scenarios was explored, and all produced the expected results.

Except for DCIS, our model with screening produced a reasonable approximation the observed distribution of stages of cancer (see "Results" below). The overall age-adjusted incidence of breast cancer over time also tracked the observed curve, except

for a systematic underestimate corresponding to DCIS. Also, the average lead times were reasonable, ranging from 0.2 to 4.7 years depending upon year and age group, with the expected increases with age and over calendar time.

Role of the Funding Source

The National Cancer Institute provided data, technical assistance, and guidance on the interpretation of the results of this simulation under a cooperative agreement. The model and model results are the sole responsibility of the investigators.

Results

Stage Distribution

The model projections of stage distributions of invasive tumors diagnosed at the end of the period of the simulation closely approximate those reported to SEER in the same period. However, the model underestimates rates of DCIS (Fig. 1).

Impact of Screening and Treatment on Mortality Trends

With the diffusion of breast cancer screening, our model estimates a 12.4% reduction in age-adjusted mortality by the year 2000, assuming no changes in therapy from that practiced in 1975. In the absence of screening, use of tamoxifen and nonhormonal chemotherapy separately, or in combination, also reduces mortality. For instance, if both modalities are used, together they account for a 14.6% mortality reduction. Overall, actual screening and treatment patterns together appear to have reduced underlying mortality rates by 24.9% by the year 2000.

Sensitivity Analyses

As noted above, we adjusted mammography sensitivity estimates by using values within the range of previously reported values for each period, along with our mean preclinical sojourn time, to attempt to reach results that corresponded to an appropriate overall incidence and stage distribution of screened cancers in each period. In this process we found that no combination of sensitivity and sojourn time within the published bounds could achieve the high incidence levels observed in the middle and late 1990s. Although observed incidence continued to rise during that period, particularly DCIS, our highest achievable incidence rates were nearly flat during that same period. Our best output also substantially underestimated the incidence of DCIS during the 1990s. In fact, the discrepancy between our best outputs and the observed data is due almost entirely to undergenerating cases of DCIS.

To further explore this matter, we tested values beyond the bounds of published estimates. Even with a mammography sensitivity of 100% at all ages, and with preclinical sojourn times extending to 12 years on average, the model-predicted incidence rates continue to fall short of the observed trend, remaining almost flat during the 1990s and still undergenerating DCIS. If, as is widely suspected, many DCIS lesions never progress or surface clinically, our approach, which only generates screen-detected lesions as the precursors of later clinical disease, is inherently incapable of generating all screen-detected DCIS. We take our inability to find any point in parameter space that leads to the observed incidence of DCIS as indirect evidence supporting the view that a substantial fraction of screen-detected DCIS represents overdiagnosis. Thus, our final model uses the sensitivity values and sojourn times depicted in Table 1.

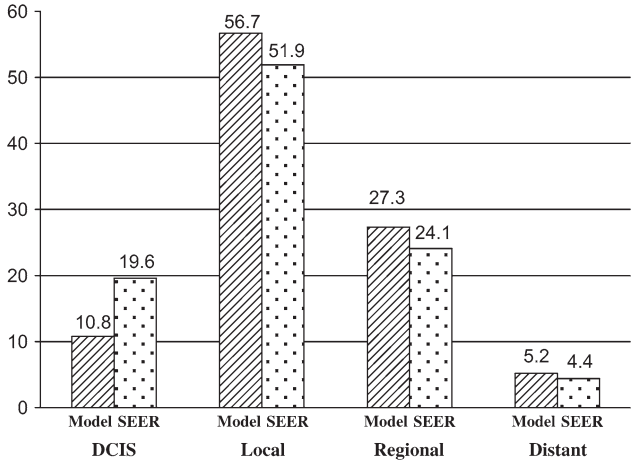

Fig. 1. This figure shows the comparison of the stage distribution of breast cancers generated by the model for the year 2000 and the actual distribution reported to Surveillance, Epidemiology, and End Results (SEER) in the same year *(19,41)*. DCIS = ductal carcinoma in situ.

Discussion

The dynamic trends in breast cancer incidence and mortality are likely to be the result of the interaction of several complex phenomena, including evolution of mammography screening, access to screening *(42–44)*, rates of diffusion of therapeutic advances, adequacy of treatment services available to different groups of women *(45–48)*, trends in life expectancy *(49)*, cohort trends in risk factors, and tumor biology *(50–53)*.

This model was developed within a framework of a collaborative group to estimate the relative contributions of screening, advances in therapy, and "other" factors on the observed declines in mortality over the past two and a half decades. Having several groups with different modeling philosophies and approaches estimate the same endpoints by using a common set of data provides an unprecedented opportunity to validate and cross-replicate data generated from simulation modeling. The resulting conclusions should be robust and have greater credibility than the inferences based on one model alone.

Our model was developed to be as transparent as possible, with a minimum of data manipulation. This model, and ones like it, can be used for policy decisions about investments in the "war on cancer." For instance, if it were estimated that most mortality declines seen over the past several decades were due to improvements in treatment, and not screening, then policy makers and insurers might consider enhancing access to comprehensive cancer care, expanding practice guidelines, and conducting large educational programs for consumers. Alternatively, if most of the benefit were attributed to screening, more investments in reaching women not yet participating would be warranted. If both screening and treatment contribute about equally to reducing morbidity and mortality, then decisions about priorities for interventions to meet targeted mortality reductions will depend more on preferences and relative costs of different cancer control approaches.

There are several issues that should be considered in evaluating our modeling approach and how our approach differs from those of others in the collaborative project. We make minimal explicit assumptions about the natural history of disease. The only biological characteristic that is included in our model is ER status. Unfortunately, many of the processes that occur in breast carcinogenesis are unobservable (i.e., preclinical detectable time) and must be inferred. Other factors, such as prognostic biomarkers, have yet to be sufficiently characterized and recorded for use in population models.

Given our approach to the disease process, screening can decrease mortality in our model only by accurately detecting disease within the preclinical detectable phase and producing a shift from one full summary stage to another (e.g., distant to regional, or regional to local and, to a lesser extent, age shift). We assume that within a given stage at diagnosis, screen and clinically detected cancers have the same survival. However, it is possible that within a given stage, screening detects smaller tumors than those diagnosed clinically, and on the basis of size alone, screen-detected tumors could have a better prognosis than the average stage-specific survival rate. To the extent that this assertion is true, our model would underestimate screening benefits relative to models that use tumor size and growth. Conversely, since slow-growing screen-detected tumors may have a better prognosis than faster-growing clinically surfacing lesions, even in the absence of screening (length bias) (see below), our assumption of comparable survival for screen and clinically diagnosed disease could overestimate the benefits of screening on mortality. Also, since our model underestimates cases of small preinvasive tumors (DCIS), we may not fully capture any decreases in mortality due to screen detection and treatment of DCIS cases that are destined to progress to invasive cancer.

Thus, it is difficult to know, on balance, if we are over- or underestimating the "true" unobservable benefits of screening, compared with models that use tumor–node–metastasis staging and tumor growth. However, models that incorporate biologic processes, such as tumor size as the main determinant of prognosis, may fail if the biological assumptions about tumor growth are not correct. By restricting ourselves to observable epiphenomena, our model gains robustness. As the science evolves, our model can be recalibrated to provide more refined estimates of outcome.

Screening benefits must also be disentangled from lead time and length bias. Whereas lead time is unobservable in real life, in the simulation both the date of diagnosis and screening and the date that the tumor would have arisen clinically are known. Thus, lead time is an output of our model, arising from the interaction of the screening schedule and the sojourn time. The model captures length bias in the preclinical sojourn time. However, we do not account for length bias in survival once a cancer is detected, since we assume that dwell time in any given stage is equal for screened and clinically detected tumors. Although it is common sense to assume that tumors with longer sojourn times also progress more slowly through their clinical stages (i.e., have longer dwell times), we are unaware of any evidence to support or refute this assumption. Without compelling evidence to the contrary, we chose to treat dwell times as independent of sojourn time because it greatly simplifies the calculation of stage shift. As noted above, to the extent that we have not fully captured length bias, we will overestimate the effects of screening on mortality reductions.

Any decreases in mortality that have been achieved through screening have been achieved at the expense of some increase in incidence of in situ disease *(54)*. In our model, we treat DCIS as a cancer state like all the others and assume that all disease passes through an in situ phase before becoming invasive. In theory, all DCIS cases could evolve into invasive cancer. That is, we do not explicitly assume that subpopulations of DCIS are nonprogressive. As noted in the results, our simulated incidence rates beyond 1990 are lower than observed by an amount that is accounted for almost entirely by a shortfall in cases of DCIS. We conclude that this discrepancy springs from the age–period–cohort model used to represent incidence in the absence of screening. In that model, age-adjusted breast cancer incidence rates (which represent extrapolation of trends that existed before 1980) show little increase during the 1990s. When screening is introduced into the population, the apparent rise in breast cancer incidence represents a forward shift in the time of diagnosis of cancers. Such a rise should be sustained only for a period approximating the mean sojourn time—thereafter the observed incidence should return to its prescreening level, unless there is an ongoing increase in the actual incidence of disease. Thus, it appears that during the 1990s, breast cancer incidence was subject to forces that are not mere extrapolations of what occurred before 1980. Rather, some additional increase in the incidence of breast cancer, especially DCIS, was taking place. Screening alone, even under the most aggressive assumptions, cannot explain the continuing rise in the observed incidence during the 1990s. Our model was not designed to explain what this increase might be due to, though many scenarios such as dietary changes, increased obesity, delayed childbearing, or overdiagnosis can be imagined. Our model does suggest that the increase seems to be almost entirely in cases of DCIS.

There is an alternative explanation for this phenomenon. It is possible that the incidence of DCIS has always been increasing just as in the 1990s but that prior to the advent of screening many of these lesions were not detected. This explanation would imply that this population of tumors often does not progress beyond the DCIS stage and remains, in general, too small to be found on physical examination or with older mammography techniques. This will be an important area for more research; a better understanding of the "true" nature of DCIS will be critical to simulations of future trends.

Next, the model was developed to estimate overall population trends and does not capture differences in outcomes among certain subgroups of the U.S. population. For instance, whereas mortality rates have declined overall, death rates have failed to decline to the same extent in older, compared with younger, women, and in black women compared with white women *(55,56)*. In fact, mortality had actually increased among black women 65 and older between 1985 and 1995 *(57)*. At present, our model cannot capture these types of differential trends. We also do not consider population immigration or emigration. Since the demographic profile of the United States is rapidly evolving, with minority groups such as Latinos expected to account for close to half of the population by 2030 *(58)*, breast cancer risk and access to care may produce different results from those expected based on the current population structure. Another demographic force that may impact future results is the "graying of America" *(59)*. By 2030, one in five women will be 65 years or older. This group carries more than half of the breast cancer burden and is expected to account for a large absolute number of cases into the future. We have assumed that screening is somewhat less useful in older women and that breast cancer is a more "benign" disease in this

age group than in younger women *(60)*. However, there is a paucity of data on the true natural history of disease in older hosts. Overall, incorporating demographic forces into future modeling efforts will be important for using this type of model to guide and measure progress in reducing age-, race-, and ethnicity-related disparities in breast cancer outcomes.

Also, our results do not account specifically for secular trends in risk factors, such as obesity *(61)*, use of hormone replacement therapy *(62)*, or prescription of tamoxifen for primary disease prevention *(63)*. Likewise, our model focuses on the period from 1975 to 2000, the period for which our input data are best calibrated. Future trends may differ as a result of treatment advances, improvements in early detection technology, and/or changes in risk factor prevalence.

Despite these caveats, our model methodology appears robust. The net outcomes of the multitude of forces affecting breast cancer risk and course will ultimately be determined by how populations behave in response to health information, how technology diffuses into practice, and the adherence of women and their physicians to cancer control guidelines.

References

(1) Silverberg E, Boring CC, Squires TS. Cancer statistics. CA Cancer J Clin 1990;40:9–26.

(2) Baquet C, Ringen K. Cancer control in blacks: epidemiology and NCI program plans. Prog Clin Biol Res 1986;216:215–27.

(3) American Cancer Society. Breast cancer facts and figures, 2003-2004. Available at: http://www.cancer.org/downloads/STT/CAFF2003BrFPWSecured.pdf. [Last accessed: January 2004.]

(4) Shapiro S, Venet W, Strax P, Venet L. Periodic screening for breast cancer: The Health Insurance Plan project and its sequelae, 1963–86. Baltimore (MD): Johns Hopkins University Press; 1988.

(5) Andersson I, Aspegren K, Janzon L, Landberg T, Lindholm K, Linell F, et al. Mammographic screening and mortality from breast cancer: The Malmo mammographic screening trial. BMJ 1988;297:943–8.

(6) Roberts MM, Alexander FE, Anderson TJ, Chetty U, Donnan PT, Forrest P, et al. Edinburgh trial of screening for breast cancer: Mortality at seven years. Lancet 1990;335:241–6.

(7) Early Breast Cancer Trialists' Collaborative Group. Polychemotherapy for early breast cancer: an overview of the randomised trials. Lancet 1998;352:930–42.

(8) Early Breast Cancer Trialists' Collaborative Group. Tamoxifen for early breast cancer: an overview of the randomised trials. Lancet 1998;351:1451–67.

(9) Fisher B, Dignam J, Wolmark N, DeCillis A, Emir B, Wickerham DL, et al. Tamoxifen and chemotherapy for lymph node-negative, estrogen receptor-positive breast cancer. J Natl Cancer Inst 1997;89:1673–82.

(10) Fisher B, Anderson S, Tan-Chiu E, Wolmark N, Wickerham DL, Fisher ER, et al. Tamoxifen and chemotherapy for axillary node-negative, estrogen receptor-negative breast cancer: findings from National Surgical Adjuvant Breast and Bowel Project B-23. J Clin Oncol 2001;19:931–42.

(11) Fisher B, Redmond C, Legault-Poisson S, Dimitrov NV, Brown AM, Wickerham DL, et al. Postoperative chemotherapy and tamoxifen compared with tamoxifen alone in the treatment of positive-node breast cancer patients aged 50 years and older with tumors responsive to tamoxifen: results from the National Surgical Adjuvant Breast and Bowel Project B-16. J Clin Oncol 1990;8:1005–18.

(12) Mandelblatt JS, Fryback DG, Weinstein MC, Russell LB, Gold MR, Hadorn D. Assessing the effectiveness of health interventions. Gold MR, Seigel J, Russell LB, Weinstein MC, editors. Cost-effectiveness in health and medicine. New York (NY): Oxford University Press; 1996. pp. 135–75.

(13) Mandelblatt JS, Fryback DG, Weinstein MC, Russell LB, Gold MR. Assessing the effectiveness of health interventions for cost-effectiveness analysis. Panel on Cost-Effectiveness in Health and Medicine. J Gen Intern Med 1997;12:551–8.

(14) Russell LB, Siegel JE, Daniels N, Gold MR, Luce BR, Mandelblatt JS. Cost-effectiveness analysis as a guide to resource allocation in health: roles and limitations. In: Gold MR, Siegel JE, Russell LB, Weinstein MC, editors. Cost-effectiveness in health and medicine. New York (NY): Oxford University Press; 1996. pp. 3–24.

(15) Gold MR, Siegel JE, Russel LB, Weinstein MC. Cost-effectiveness in health and medicine. New York (NY): Oxford University Press; 1996. pp. 54–81.

(16) Michaelson JS, Halpern E, Kopans DB. Breast cancer: computer simulation method for estimating optimal intervals for screening. Radiology 1999;212:551–60.

(17) Cronin KA, Mariotto AB, Clarke LD, Feuer EJ. Additional common inputs for analyzing impact of adjuvant therapy and mammography on U.S. mortality. J Natl Cancer Inst Monogr 2006;36:26–9.

(18) Holford TR, Cronin KA, Mariotto AB, Feuer EJ. Changing patterns in breast cancer incidence trends. J Natl Cancer Inst Monogr 2006;36:19–25.

(19) Surveillance, Epidemiology, and End Results (SEER) Program. Available at: https://www.seer.cancer.gov. SEER*Stat Database: Incidence—SEER 11 Regs Public Use, Nov. 2001 Sub (1992–1999). Age-Specific Stage Distribution. National Cancer Institute, DCCPS, Surveillance Research Program, Cancer Statistics Branch. [Last accessed: September 2005.]

(20) Ashba J, Traish AM. Estrogen and progesterone receptor concentrations and prevalence of tumor hormonal phenotypes in older breast cancer patients. Cancer Detect Prev 1999;23:238–44.

(21) Brekelmans CT, Westers P, Faber JA, Peeters PH, Collette HJ. Age specific sensitivity and sojourn time in a breast cancer screening programme (DOM) in the Netherlands: a comparison of different methods. J Epidemiol Community Health 1996;50:68–71.

(22) Paci E, Duffy SW. Modelling the analysis of breast cancer screening programmes: sensitivity, lead time and predictive values in the Florence district programme (1975-1986). Int J Epidemiol 1991;20:852–8.

(23) Boer R, de Koning H, Threlfall A, Warmerdam P, Street A, Friedman E, et al. Cost effectiveness of shortening screening interval or extending age range of NHS breast screening programme: computer simulation study. BMJ 1998;317:376–9.

(24) Boer R, de Koning HJ, van der Maas PJ. A longer breast carcinoma screening interval for women age older than 65 years? Cancer 1999;86:1506–10.

(25) Woods & Poole Economics Inc. 2001 Regional database estimated July 1 population by race, sex and single year and 5-year age groups based on 1990 census and post censal census bureau estimates. Available at: http://cisnet.flexkb.net/files/basecaseparams/yr1975_femal_pops_v2xls. Copyright 2001.

(26) Kerlikowske K, Grady D, Barclay J, Sickles EA, Ernster V. Likelihood ratios for modern screening mammography. Risk of breast cancer based on age and mammographic interpretation. JAMA 1996;276:39–43.

(27) Fletcher SW, Harris RP, Gonzalez JJ, Degnan D, Lannin DR, Strecher VJ, et al. Increasing mammography utilization: a controlled study. J Natl Cancer Inst 1993;85:112–20.

(28) Baines CJ, Miller AB, Wall C, McFarlane DV, Simor IS, Jong R, et al. Sensitivity and specificity of first screen mammography in the Canadian National Breast Screening Study: a preliminary report from five centers. Radiology 1986;160:295–8.

(29) Lawrence WF, Liang W, Mandelblatt J, Gold K, Freedman M, Ascher S, et al. Serendipity in diagnostic imaging: magnetic resonance imaging of the breast. J Natl Cancer Inst 1998;90:1792–800.

(30) Mariotto AB, Feuer EJ, Harlan LC, Abrams J. Dissemination of adjuvant multiagent chemotherapy and tamoxifen for breast cancer in the United States using estrogen receptor information: 1975–1999. J Natl Cancer Inst Monogr 2006;36:7–15.

(31) Cronin KA, Yu B, Krapcho M, Miglioretti DL, Fay MP, Izmirlian G, et al. Modeling the dissemination of mammography in the United States. Cancer Causes Control 2005;16:701–12.

(32) Rosenberg MA. Competing risks to breast cancer mortality. J Natl Inst Monogr 2006;36:15–9.

(33) Shapiro S, Venet W, Strax P, Venet L. Current results of the breast cancer screening randomized trial: the Health Insurance Plan (HIP) of Greater New York study. In: Day NE, Miller AB, editors. Screening for breast cancer. Toronto (Canada): Hans Huber; 1988. pp. 3–15.

(34) Chamberlain J, Coleman D, Moss, et al. Sensitivity and specificity of screening in the UK trial of early detection of breast cancer. In: Miller AB,

Chamberlain J, Daye NE, et al, editors. Cambridge: Cambridge University Press; 1991. pp. 3–17.
(35) Miller AB, Baines CJ, To T, Wall C. Canadian National Breast Screening Study: 1. Breast cancer detection and death rates among women aged 40 to 49 years. Can Med Assoc J 1992;147:1459–76.
(36) Miller AB, Baines CJ, To T, Wall C. Canadian National Breast Screening Study: 2. Breast cancer detection and death rates among women aged 50 to 59 years. Can Med Assoc J 1992;147:1477–88.
(37) Kleinrock L. Queueing Systems Vol. 1: Theory. John Wiley & Sons; 1975. p. 72.
(38) Mariotto A, Feuer EJ, Harlan LC, Wun LM, Johnson KA, Abrams J. Trends in use of adjuvant multi-agent chemotherapy and tamoxifen for breast cancer in the United States: 1975-1999. J Natl Cancer Inst 2002;94:1626–34.
(39) Fisher B, Anderson S, Redmond C, Wolmark N, Wickerham D, Cronin W. Reanalysis and results after 12 years of follow-up in a randomized clinical trial comparing total mastectomy with lumpectomy with or without irradiation in the treatment of breast cancer. N Engl J Med 1995;333:1456–61.
(40) NIH Consensus Conference. Treatment of early stage breast cancer. JAMA 1991;265:391–5.
(41) National Cancer Institute, SEER. Surveillance research program. Incidence-SEER 9 Regs Public Use. Available at: http://www.seer.cancer.gov/seerstat. [Last accessed: April 2004.]
(42) Costanza ME. The extent of breast cancer screening in older women. Cancer 1994;74(7 suppl):2046–50.
(43) Oakar MR. Legislative effect of the 102nd Congress. Cancer prevention, detection, treatment, and research. Cancer 1992;69(7 suppl):1954–6.
(44) Rimer BK, Keintz MK, Kessler HB, Engstrom PF, Rosan JR. Why women resist screening mammography: patient-related barriers. Radiology 1989; 172:243–6.
(45) McWhorter WP, Mayer WJ. Black/white differences in type of initial breast cancer treatment and implications for survival. Am J Public Health 1987;77:1515–7.
(46) Allen C, Cox EB, Manton KG, Cohen HJ. Breast cancer in the elderly—current patterns of care. J Am Geriatr Soc 1986;34:637–42.
(47) Chu J, Diehr P, Feigle P, Glaefke G, Begg C, Glicksman A, et al. The effect of age on the care of women with breast cancer in community hospitals. J Gerontol 1987;42:185–90.
(48) Newcomb PA, Carbone PP. Cancer treatment and age: patient perspectives. J Natl Cancer Inst 1993;85:1580–4.
(49) Statistical Abstracts of the United States. Available at: http://www.census.gov/prod/2005pubs/06statab/vitstat.pdf. [Last accessed: August 2006.]
(50) Mohla S, Sampson CC, Khan T, et al. Estrogen and progesterone receptors in breast cancer in black Americans: correlation of receptor data with tumor differentiation. Cancer 1982;50:552–9.
(51) Ownby HE, Frederick J, Russo J, Brooks SC, Swanson GM, Heppner GH, et al. Racial differences in breast cancer patients. J Natl Cancer Inst 1985;75:55–60.
(52) Crowe JP Jr, Gordon NH, Hubay CA, Pearson OH, Marshall JS, McGuire WL. The interaction of estrogen receptor status and race in predicting prognosis for stage II breast cancer patients. Surgery 1986;100:599–605.
(53) Elledge RM, Clark GM, Chamness GC, Osborne CK. Tumor biologic factors and breast cancer prognosis among white, Hispanic, and black women in the United States. J Natl Cancer Inst 1994;86:705–12.
(54) Feurer EJ, Wun LM. How much of the recent rise in breast cancer incidence can be explained by increase in mammography utilization—a dynamic population model approach. Am J Epidemiol 1992;136:1423–36.
(55) National Center for Health Statistics, National Cancer Institute. Groups at high risk for cancer deaths. Available at: http://progressreport.cancer.gov/doc/asp?pid=1&did=21&chid=14&coid=35&mid=vpco#groups. [Last accessed: January 2004.]
(56) Trends of US mortality age-adjusted rates: Total US, 1990–2000. Surveillance, Epidemiology, and End Results (SEER) Program (http://www.seer.gov); SEER*Stat Database: Mortality—All COD, Public-use with State, Total US, 1969–2003; National Cancer Institute, DCCPS, Surveillance Research Program, Cancer Statistics Branch, released April 2006. Underlying mortality data provided by NCHS (http://www.cdc.gov/nchs).
(57) Chu KC, Tarone RE, Brawley OW. Breast cancer trends of black women compared with white women. Arch Fam Med 1999;8:521–8.
(58) U.S. Census Bureau. Projections of the resident population by race, Hispanic origin, and nativity: Middle series, 2025 to 2045. Available at: http://www.census.gov/population/projections/nation/summary/np-t5-f.pdf. [Last accessed: July 2003.]
(59) Soldo BJ, Agree EM. America's elderly. Washington (DC): Population Reference Bureau; 1988.
(60) Mandelblatt J, Saha S, Teutsch S, Hoerger T, Siu A, Atkins D, et al. The cost-effectiveness of screening mammography beyond age 65: a systematic review for the U.S. Preventive Services Task Force. Ann Intern Med 2003;139:835–42.
(61) Centers for Disease Control and Prevention. U.S. obesity trends, 1985–2002. Available at: http://www.cdc.gov/nccdphp/dmpa/obesity/trend/maps/index.htm. [Last accessed: July 2003.]
(62) Colditz G, Hankinson SE, Hunter DJ, Willett WC, Manson JE, Stampfer MJ, et al. The use of estrogen and progestins and the risk of breast cancer in postmenopausal women. N Engl J Med 1995;352:1589–93.
(63) Fisher B, Costantino JP, Wickerham DL, Redmond CK, Kavanah M, Cronin WM, et al. Tamoxifen for prevention of breast cancer: report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. J Natl Cancer Inst 1998;90:1371–88.

Notes

Supported by NCI Cooperative Agreement # UO1-CA88293A (J. Mandelblatt, W. Lawrence, B. Yi, J. Cullen) and grant #KO5 CA96940 (J. Mandelblatt).

Present address: W. Lawrence, Center for Outcomes and Evidence at the Agency for HealthCare Research and Quality, Rockville, MD.

Present address: J. Cullen, Uniformed Services University of the Health Sciences (USUHS), Department of Surgery and Henry M. Jackson Foundation Center for Prostate Disease Research, Gaithersburg, MD.

We thank K. Robin Yabroff for assistance with literature review and Trina McClendon for manuscript preparation.