A Bayesian Simulation Model for Breast Cancer Screening, Incidence, Treatment, and Mortality

Xuelin Huang, PhD, Yisheng Li, PhD, Juhee Song, PhD, Donald A. Berry, PhD

Background. *The important but complicated research questions regarding the optimization of mammography screening for the detection of breast cancer are unable to be answered through any single trial or a simple meta-analysis of related trials. The Cancer Intervention and Surveillance Network (CISNET) breast groups provide answers using complex statistical models to simulate population dynamics. Among them, the MD Anderson Cancer Center (Model M) takes a unique approach by not making any assumptions on the natural history of breast cancer, such as the distribution of the indolent time before detection, but simulating only the observable part of a woman's disease and life.* ***Methods.*** *The simulations start with 4 million women in the age distribution found in the year 1975, and follow them over several years. Input parameters are used to describe their breast cancer incidence rates, treatment efficacy, and survival. With these parameters, each woman's history of breast cancer diagnosis, treatment, and survival are generated and recorded each year. Research questions can then be answered by comparing the outcomes of interest, such as mortality rates, quality-adjusted life years, number of false positives, differences between hypothetical scenarios, such as different combinations of screening and treatment strategies. We use our model to estimate the relative contributions of screening and treatments on the mortality reduction in the United States, for both overall and different molecular (ER, HER2) subtypes of breast cancer.* ***Results.*** *We estimate and compare the benefits (life-years gained) and harm (false-positives, over-diagnoses) of mammography screening strategies with different frequencies (annual, biennial, triennial, mixed) and different starting (40 and 50 years) and end ages (70 and 80 years).* ***Conclusions.*** *We will extend our model in future studies to account for local, regional, and distant disease recurrences.* ***Key words:*** *adjuvant treatments; approximate Bayesian computation; Bayesian simulation; beyond stage-shift; breast cancer; cancer screening; mammography.* ***(Med Decis Making XXXX;XX:xx–xx)***

Challenging questions arise in research on the effects of breast cancer screening and treatments, such as the identification of the best mammography screening strategy or the evaluation of the impact of new molecular pathways and genomic tumor profile-targeted treatment paradigms in the adjuvant setting and at disease recurrence. These tasks are critical to informing policy makers when allocating resources and designing public health guidelines and recommendations. However, it is impossible to use simple calculations or find some meta-analyses to answer these questions. Randomized trials may also be infeasible or unethical to conduct. The Cancer Intervention and Surveillance Network (CISNET) groups were formed to address these challenging questions by complex statistical models that simulate population dynamics.[1–10] The CISNET simulation model assembly results from relevant meta-analyses and large-scale databases as the input parameters, and yield outputs to answer various questions of interest through intensive computation.

The MD Anderson Model (Model M), 1 of the 6 CISNET breast cancer models, is a Bayesian simulation model.[11–16] It simulates populations of 4 million women with the age distribution that existed in

Received 9 August 2016 from Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, TX, USA (XH, YL, JS, DAB). This work was supported by the National Institutes of Health through funding from the National Cancer Institute: grants U01 CA152958 and U01 CA19921801. Revision accepted for publication 25 January 2017.

Address correspondence to Donald Berry, PhD, Department of Biostatistics, The University of Texas MD Anderson Cancer Center 1400 Pressler Street, Unit 1411, Houston, TX, 77230-1402, USA; Phone: 713-794-4141; Fax: 713-563-4243; e-mail: dberry@mdanderson.org.

DOI: 10.1177/0272989X17714473

the United States in 1975. The sample size of 4 million is found to be generally sufficient to give an accurate parameter estimation. For each virtual woman, the model simulates a natural course of her life separate from the possible occurrence of breast cancer. Moving forward in time, in each year, each woman is diagnosed with breast cancer or not, depending on the incidence of the disease for women of her age in that year, whether she had a screening mammogram that year, and her history of mammography. The model keeps track of which women are diagnosed with breast cancer each year and which women die of breast cancer and of other causes. It is implemented as follows. For each virtual woman who is diagnosed with breast cancer, we generate 2 potential death times for her: one due to breast cancer, the other due to all other causes combined. Her breast cancer specific survival time depends on the characteristics of her tumor, the mode of detection, and the treatment she received. Her death time due to all the other causes is simulated by using the life tables obtained from census data. We compare these 2 potential survival times, and use the earlier date as her death time, and record the corresponding cause of death.

The parameters of interest in our model include those that affect the diagnosis of breast cancer and its course once the disease is diagnosed. Treatment parameters include those for the effects of adjuvant treatments, including hormone therapy, trastuzumab, the combination of cyclophosphamide, methotrexate, and fluorouracil (CMF) chemotherapy, anthracycline, and taxanes. For mammography screening, we use a slope parameter to describe the linear changing pattern over time of breast cancer incidence rates in a hypothetical scenario of no screening. Screening-detected breast cancers tend to have earlier stage distributions than clinically detected cases. Besides that, it has been found that, even among cases of the same stage, screening-detected ones have better survival.[17] We use 2 parameters for these ''beyond stage-shift'' effects for tumor stages I/II and stages III/IV, respectively.

We use Bayesian methodology, which is a statistical inference alternative to the classical frequentist method. Suppose we are interested in a parameter θ of research interest, which could be, for example, the incidence rate of breast cancer in a specific population, or the hazard ratio of a test drug v. a reference drug. In Bayesian analysis, before we have data for estimating θ, we have a prior distribution $p(\theta)$ for θ. This prior distribution can be non-informative (uniform on a specific range) or informative (using historical data). Based on the observed data D, for each value of θ, the likelihood function is $l(D|\theta)$. Then, the principle of Bayesian data analysis is that the posterior distribution of θ is equal to $l(D|\theta)\ p(\theta)$, up to the difference of a proportionality constant. All the inference about θ, such as its mean and variation, can be obtained from this posterior distribution.

To illustrate our Model, we start by looking at the age-adjusted US breast cancer incidence and mortality rates from 1975 to 2010 (Figure 1, solid lines, obtained from the Surveillance, Epidemiology, and End Results (SEER) Program[18]). The incidence rates had been mostly increasing over these years except some decreasing from 2000 to 2010, and the mortality rates had been slightly increasing before 1990 and substantially decreasing after that. What caused these changing patterns? Potential factors include: environmental and lifestyle changes, mammography screening, the use and then halting of menopausal hormone therapy (MHT),[19] improvements in chemotherapy over the years, improvements in hormonal therapies for estrogen receptor-positive and/or progesterone receptor-positive (ER+/PR+, or simply ER+) breast cancer subtypes, and the discovery of trastuzumab for human epidermal growth factor receptor 2-positive (HER2+) breast cancer subtypes. We must consider the effects of these intertwining factors, which started in different years, and work on different subtypes of breast cancer. Thus, it is extremely difficult to determine the effects of all these factors on breast cancer incidence and mortality rates. Together with other CISNET models, we have previously evaluated the relative contributions of screening and adjuvant treatments on overall breast cancer mortality reduction in the US from 1990 to 2000.[11] Here, we update this analysis to the year 2010 and extend to include breast cancer subtypes based on each ER/HER2 subgroup.

This aim of this article is to describe the methodology of our statistical model, not to report research results. The next sections in this article are organized as follows: First, we will briefly describe our modeling procedure, with details deferred the subsequent section, which also illustrates the estimation of relative contributions of screening and adjuvant treatments to mortality reduction. After this, we consider model validation and uncertainty assessment. In the next section, we briefly introduce the comparison and optimization of mammography screening strategies, a task conducted for the US Preventive Services Task Force. And in the final section, we summarize our method, discuss its features, and describe some future research projects.

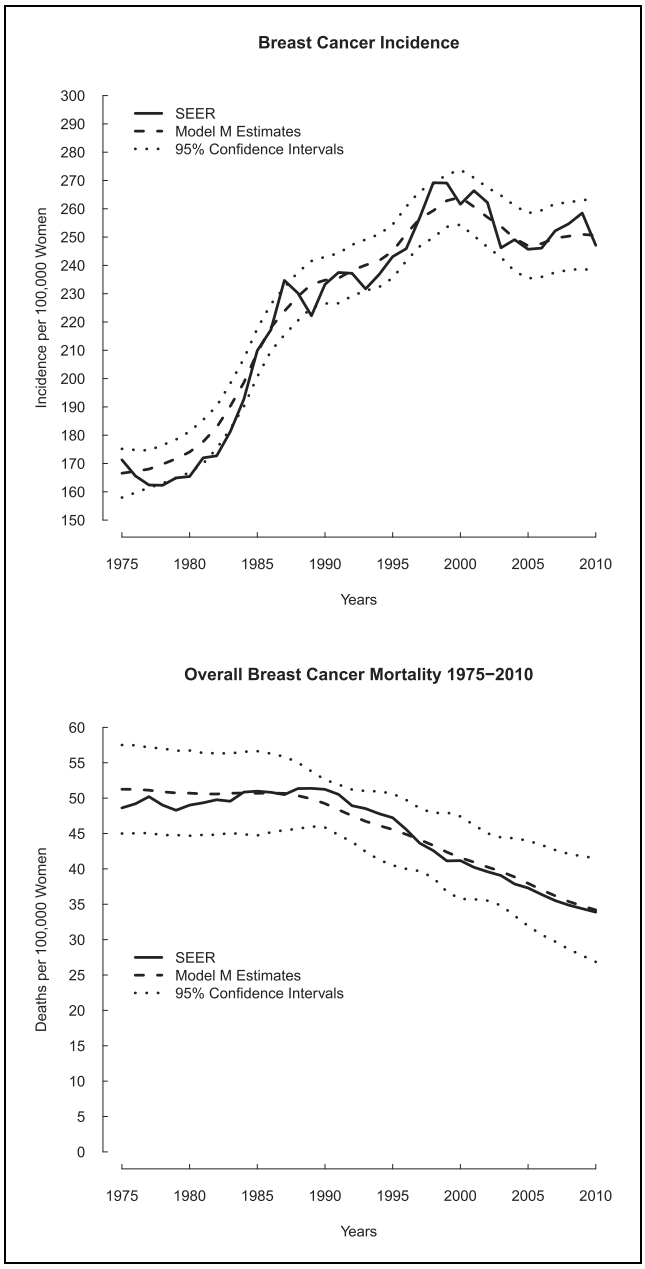

Figure 1 Incidence and mortality rates generated by Model M, matching with the Surveillance, Epidemiology, and End Results (SEER) data.[18]

MODELING PROCEDURE

We briefly describe our modeling procedure in this section and provide details in later sections.

The input of our model consists of 2 parts: known parameters and unknown parameters. The known parameters include the following list, with details described elsewhere[2], except for point 3, the parameters for which were obtained from the Center for Disease Control and Prevention and National Breast and Cervical Cancer Early Detection Program.[16]

1. Life tables and census data of the general US population from 1975 to 2011.
2. Survival distributions in the absence of screening and treatment, depending on age, tumor stage, and ER/HER2 status.
3. Breast cancer incidence table, according to screening type (film or digital), schedule (frequency), and age.[16]
4. Screening-detected breast cancer stage distributions, depending on age, screening type, and schedule.
5. Clinically detected breast cancer stage distributions, depending on age.
6. ER/HER2 status distributions, depending on age and tumor stage.
7. Mammography screening dissemination patterns over 1975 to 2010.
8. Treatment dissemination patterns over 1975 to 2010.

Besides the above known parameters, we also used some unknown parameters, shown in Table 1, to describe the effects of screening and adjuvant therapies. With these input parameters, known and unknown, our modeling proceeds as described below; for ease of understanding, we also illustrate the procedure using a flow chart in Figure 2.

Step 1: Simulate data under real-world screening and treatment disseminations, and estimate unknown parameters by matching simulated data with SEER summary data.

1.1 For the unknown parameters, specify a prior distribution for each. Then, using a random draw of values from their prior distributions, together with known parameters, generate a data set of 4 million women beginning in the year 1975, who are followed each year with respect to their breast cancer history and survival information until 2010. The sample size of 4 million can be adjusted, depending on applications, to allow for an accurate estimation.

1.2 Repeat Step 1.1 80,000 times, each time starting with a different random draw of values for the unknown parameters, taken from their prior distributions. This process generates 80,000 data sets. The number of data sets can be increased to provide more accurate and stable estimation.

Table 1 Screening and Treatment Efficacy Parameters Used by Model M

Symbol	Parameter	Description
Factors related to screening		
θ_1	NS slope	Slope of incidence increasing without screening
θ_2	BSS12	Beyond stage-shift stages I and II
θ_3	BSS34	Beyond stage-shift stages III and IV
Factors related to treatment efficacy		
θ_4	Tamoxifen	Hazard reduction (HR) by tamoxifen
θ_5	Chemotherapy CMF	HR by chemotherapy CMF
θ_6	Trastuzumab stages 0-III	HR by trastuzumab for cases with stages 0-III
θ_7	Trastuzumab stage IV	HR by trastuzumab for cases with stage IV
θ_8	Taxanes	HR by taxanes
θ_9	Anthracyclines	HR by anthracyclines
θ_{10}	Aromatase inhibitor	HR by aromatase inhibitors
Survival distribution without adjuvant therapies		
θ_{11}	Baseline survival factor	Shift the survival distribution

1.3 For each of the 80,000 data sets, compare the incidence and mortality rates from 1975 to 2010 with SEER data, and accept those that closely match SEER data, with the maximum differences over time less than specific thresholds. The values for the ''unknown'' parameters in these accepted data sets form the posterior distribution of these parameters. If the number of accepted data sets is less than 100, increase the value of 80,000 in Step 1.2 to an even larger number, and redo Steps 1.2 to 1.3.

Step 2: Simulation data again: Using estimated values of ''unknown'' parameters and assuming hypothetical scenarios of interest for screening and treatment disseminations.

Use each set of accepted values for ''unknown'' parameters to generate data for scenarios, assuming different screening schedules or different adjuvant treatment patterns, and use these simulated data to compute the resulting incidence and mortality rates for each year from 1975 to 2010. For each assumed scenario, for each year, compute the average or weighted average incidence and mortality rates over the multiple accepted parameter value sets.

Step 3: Answer questions of interest by comparing the results under different hypothetical scenarios of screening and treatment disseminations.

Compare the average mortality rates in Step 2 obtained under different scenarios, and use the magnitude of their differences to answer questions regarding apportioning the contributions of screening and treatments to breast cancer mortality

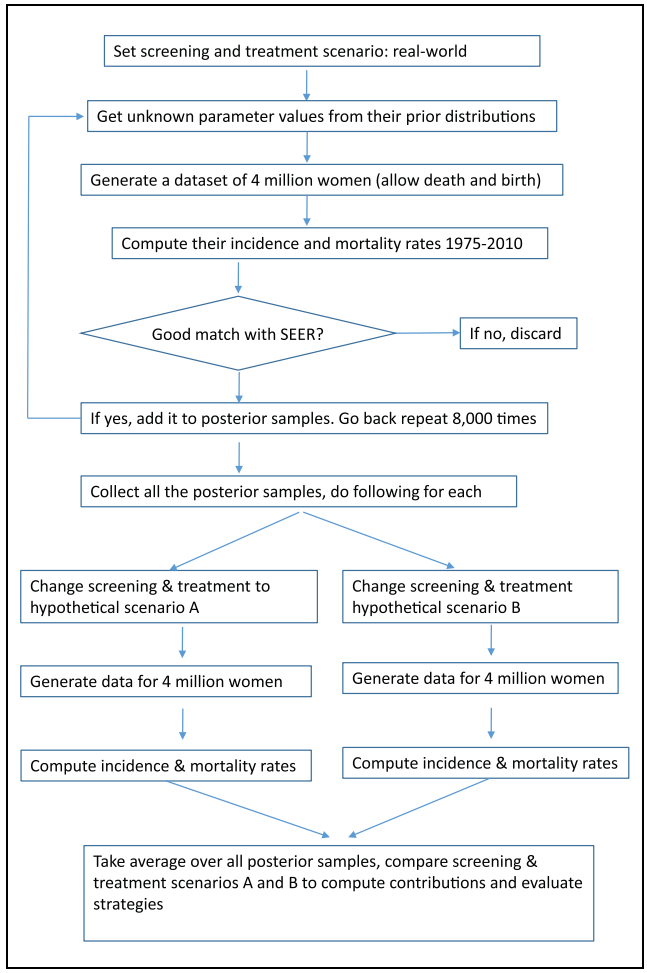

Figure 2 *Modeling procedure.*

reduction from 1990 to 2010, considering the benefits and harm of different screening schedules, and the benefit of different treatment patterns.

The details for the above steps are illustrated through an example in the next section.

ESTIMATING THE CONTRIBUTIONS OF SCREENING AND ADJUVANT TREATMENT TO MORTALITY REDUCTION

Age-adjusted breast cancer mortality per 100, 000 women in the United States declined from 51.2 in 1990 to 41.2 in 2000, and further declined to 33.9 in 2010, yielding a relative 33.8% decline from 1990 to 2010 in breast cancer mortality (Figure 1, data from SEER[18]). It is reasonable to believe that both mammography screening for early detection and improvement in adjuvant therapies have resulted in better survival for women with breast cancer, and thus contributed to these declines. One of the goals of the CISNET Models is to validate this belief by estimating the contributions of these 2 factors. The demographic changes in US women over the past 35 years has already been incorporated into all of these calculations. The US census data over the years were used to standardize the incidence and mortality rates in Figure 1 and throughout this manuscript. Therefore, it was meaningful to consider the changing trends over time using these standardized values. Without this standardization procedure, the changing trends seen in raw incidence and mortality rates could be simply due to the shifting of population age distributions.

We use input files common to all 6 CISNET breast cancer models.[2] They specify the distribution of ER/HER2 status by age and stage, the baseline survival function by age, stage, and ER/HER2 status ("baseline" means for the scenario in which patients did not receive any adjuvant therapy), and mammography screening and treatment dissemination patterns over the years, among many other factors related to breast cancer risks, screening, treatments, and survival. See the cited reference[2] for details.

Our Bayesian model includes the parameters in Table 1. For the parameters that describe the efficacy of adjuvant therapies, we use the results from the Oxford Review to assume their prior distributions.[25] These prior distributions are made less informative than the Oxford Review to allow for a more flexible fitting of the models. Some of the priors are even made flat (uniform) on a specific range, such as from 0 to 0.8.

Using all the above parameters, we conduct simulations. We randomly draw a number for each parameter $\theta_1, \ldots, \theta_J$, $J = 11$, from its prior distribution, and call these J numbers a parameter set. Since we need to perform this random drawing $K = 80,000$ times, we denote all these draws by

$$\theta_1^{(1)} \ldots, \theta_J^{(k)},\ k = 1 \ldots, K \qquad (1)$$

Thus, we generate 80,000 data sets, with each data set recording information during each year from 1975 to 2010, for 4 million women, starting with the age distribution that existed in the US population in the year 1975.[24]

We denote the breast cancer incidence rate in the year 1975 as R_{1975}, and use a parameter θ_1 for the annual increase in breast cancer incidence in the absence of mammography screening. This may have been caused by environmental and lifestyle changes or other unknown factors not explicitly accounted for in our model. Then, for any year x after 1975, the probability of detecting breast cancer for a breast cancer-free woman who has never received mammography screening is $R_{1975} + \theta_1(x - 1975)$ if she has also never received MHT. Otherwise, her chance of having screening-detected breast cancer is double the above probability.[19] This incidence rate is used to generate data for women who never receive mammography screening. We do not use the Age-Period-Cohort (APC) model, which is described elsewhere.2

We also generate data for scenarios in which women undergo mammography screening. In any year, depending on whether a woman has never been screened in the past 3 years or has been screened annually, biennially, or triennially, we use the corresponding tables obtained from the Center for Disease Control and Prevention and National Breast and Cervical Cancer Early Detection Program to determine that woman's chance of having breast cancer detected by screening mammography.16 Women receiving screening may also be clinically detected with breast cancer during the interval between 2 screenings. If detected, it is unknown whether this represents an over-diagnosis, i.e., the true disease status is unknown. The computation of over-diagnosis rates by model M is discussed later. If breast cancer is detected for a given woman, depending on whether it is detected by screening or clinical findings, we use our common input files[2] to assign a tumor stage for this case (ductal carcinoma in situ (DCIS), stage I, IIa, IIb, III, or IV, according to the system of the American Joint Committee on

Cancer [AJCC v.6][26]). Using the corresponding common input files, we also assign the hormone receptor and HER2 status of the tumor, which depends on patient age and tumor stage.

We use separate stage distributions for screening and clinically detected breast cancer.[2,16] For screening-detected breast cancer, the benefit of early detection is reflected by the shifted-to-early-stage distribution. Even within the same stage, women with breast cancer detected through screening may have a longer survival than those with clinically detected breast cancer.[16,17] Table 1 includes 2 additional parameters, θ_2 and θ_3, to describe these "beyond-stage-shift" benefits for breast cancer of stages I-II and III-IV, respectively. Applying the method by Shen and others,[17] we estimate the beyond-stage-shift benefits using the data sets from the Health Insurance Plan Project[27] and the Canadian National Breast Cancer Screening Study,[28–30] and then use these estimates as the basis of our prior distributions for these parameters. Through the above simulations to match SEER incidence and mortality data, these prior distributions are updated to become the posterior distributions, which provides all aspects of our final parameter estimation, such as the mean, mode, quantiles, etc.

Our screening and treatment dissemination patterns are obtained from the CISNET common input files.[2] Our model assumes that all women with detected breast cancer receive surgery and radiation. Their adjuvant treatments are determined by the treatment dissemination file. This file specifies, for each treatment, the fraction of women in the population to receive it, depending on their age, tumor stage and ER/HER2 status. After treatments are assigned, their efficacy levels are determined by the parameters $\theta_1^{(k)}, \ldots, \theta_J^{(k)}$ (which apply to only those women who receive the treatments). These parameters modify the baseline survival function (the survival function without adjuvant treatments). Suppose the baseline survival function is $S_0(t)$ and a patient receives taxane therapy, then her survival function is $S^{(1-\theta_8)\theta_{11}}(t)$, where θ_8 is the breast cancer death hazard reduction by the use of taxanes, and θ_{11} is Model M's adjustment made on top of $S_0(t)$, which is the survival distribution in the absence of screening and treatment, provided by common CISNET input.[2] Here, the notation $S_0(t)$ is a simplification; it actually depends on the woman's age, tumor stage, ER and HER2 status.

For each data set k with 4 million women, after generating the disease status and survival time for each woman, we plot the breast cancer incidence and mortality rates from 1975 to 2010, and the mortality rates for ER-positive and -negative subgroups from 1990 to 2010. We match these rates against the SEER data counterparts (SEER has ER data since 1990). If the matches are sufficiently close for all years, we accept the parameter set k. Otherwise, we reject it. This is the "acceptance/rejection" method for updating Bayesian posterior distributions.[14,15] Details about the matching criteria are provided below. The acceptable parameter sets form empirical posterior distributions for $\theta_1, \ldots, \theta_J$. They are presented in Figure 3 (dotted lines), together with their prior distributions (solid lines).

We set 2 criteria for a good match between simulated data and SEER data for incidence and mortality rates, respectively. The first is the maximum of the absolute differences between simulated and SEER data over the years 1975 to 2010. The second is the maximum of the absolute differences between simulated and SEER data regarding the slopes of the 5-year incidence and mortality rate changes, from 1980 to 1985, . . ., and from 2005 to 2010. The third is the mortality changing slope of the generated data from 1975 to 1990. For these criteria, we set and adjust boundary values (windows), and accept the simulated data that lie within the specified windows. The half-width for incidence matching window is 12 cases out of 100,000 women, and that for mortality matching is 5 breast cancer specific deaths out of 100,000 women. The half-width for the slope window is 2 cases for incidence, or 2 deaths for mortality, per year among 100,000 women. After applying these criteria, the accepted data sets match SEER data well. The results are reported in Figure 1, which shows the SEER incidence and mortality rates and their estimators obtained from the simulated data, together with their 95% confidence intervals.

After we determine acceptable values for the parameters specified in Table 1, we use them to generate data under different scenarios of screening and adjuvant therapy (chemotherapy, hormone therapy and trastuzumab). We use the data generated with our model to evaluate the relative contributions of screening and adjuvant treatments on breast cancer-related mortality reduction in the US in the respective years 2000 and 2010. This task involves simulating breast cancer incidence and mortality data under the 10 scenarios listed in Table 2. For example, in the first scenario, there is no screening, no chemotherapy, no hormone therapy, and no trastuzumab. Scenario 5 includes screening, chemotherapy, hormone therapy and trastuzumab.

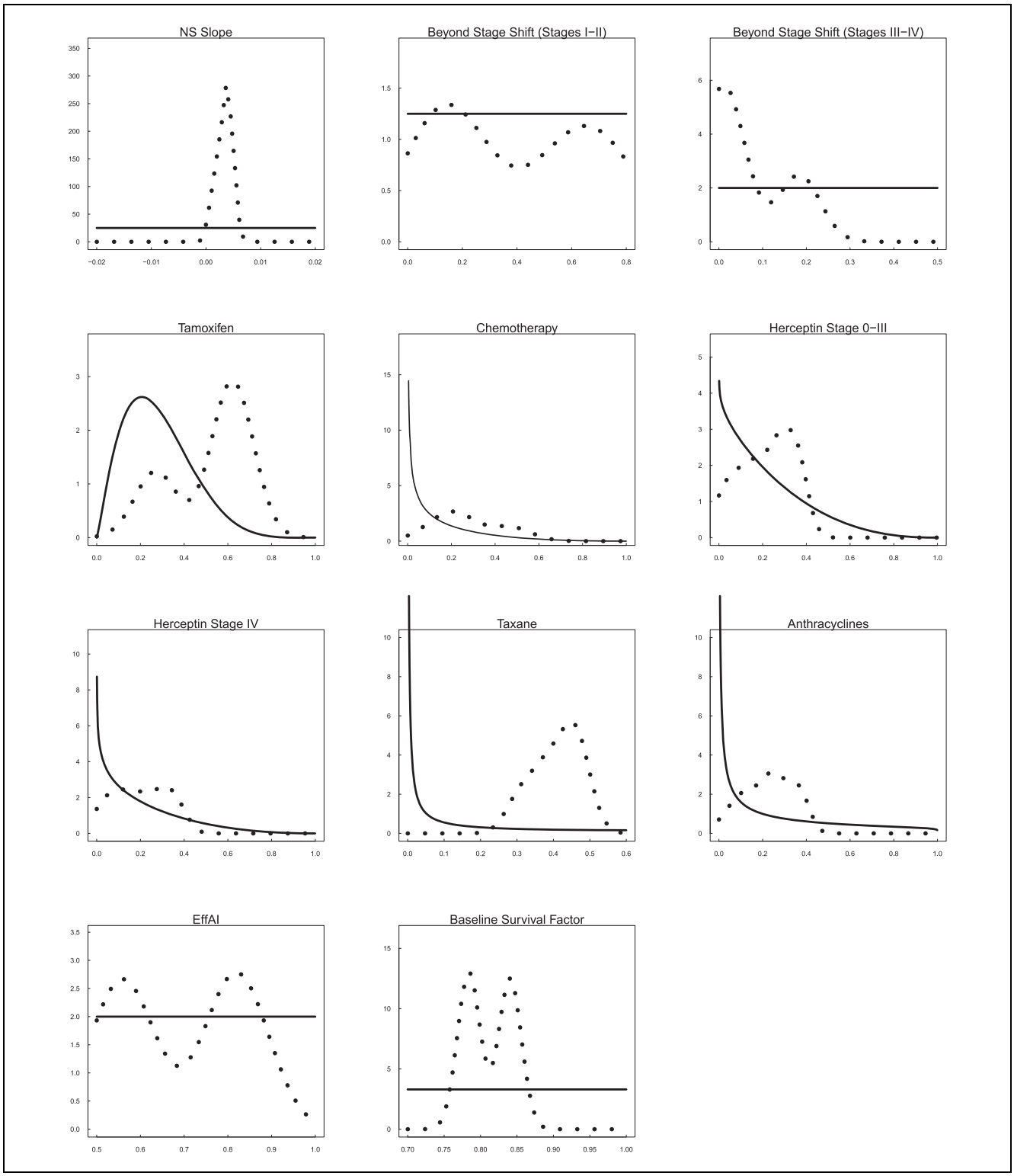

Figure 3 Prior (solid) and posterior (dotted) probability density distributions of parameters in Model M. NS, no screening; EffAI, aromatase inhibitors.

Table 2 Screening and Treatment Scenarios

	Screening	Chemotherapy	Hormone Therapy	Trastuzamab
Scenario 1	No	No	No	No
Scenario 2	Yes	No	No	No
Scenario 3	Yes	Yes	No	No
Scenario 4	Yes	Yes	Yes	No
Scenario 5	Yes	Yes	Yes	Yes
Scenario 6	No	Yes	No	No
Scenario 7	No	Yes	Yes	No
Scenario 8	No	Yes	Yes	Yes
Scenario 9	Yes	No	Yes	No
Scenario 10	No	No	Yes	No

Other scenarios are specified by turning on or off some of the screening and treatment options. Then, the difference in the mortality rates for a particular year (such as 2010) between scenarios 1 and 5 is the effects of screening and all adjuvant therapies combined. The benefit of screening alone or its combination with one or more adjuvant therapies can be estimated by looking at the difference in mortality rates under the corresponding scenarios. The above comparisons of mortality rates are completed for the overall breast cancer mortality, as well as the mortality rates attributed to the following 4 subtypes of breast cancer: ER+/HER2−, ER+/HER2+, ER−/HER2+, and ER−/HER2−. Details of these results will be reported separately, along with that by other CISNET breast models.

Using the above simulations and comparing the mortality rates under different scenarios, we estimate that the overall BC mortality reduction achieved by screening and all adjuvant therapies (hereafter simply called "treatment") is 39%, that by screening alone is 16%, and that by treatment alone is 27%, resulting in relative contributions of 38% (=16/(16+27)) and 62%, respectively. The contributions to mortality reduction by screening and treatment are 17% and 31% for the ER+/HER2− subgroup, with their relative contributions as 35% and 65%, respectively. Listing these contributions as 17 (35%) and 31 (65%), the contributions to mortality reduction by screening and treatments are respectively 11 (26%) and 32 (74%) for the ER+/HER2+ subgroup, 20 (45%) and 24 (55%) for the ER−/HER2+ subgroup, and 12 (50%) and 12 (50%) the ER−/HER2− subgroup. By our modeling, we answer these important questions regarding the contributions by treatment and screening on the breast cancer mortality reduction seen since 1990. These have profound implications on public health policy, screening strategy optimization, treatment dissemination, and health resource allocation. Our modeling approach provides a statistical framework for evaluating public health intervention programs.

MODEL VALIDATION AND UNCERTAINTY

We validate our model in 2 different ways. First, we apply it to the setting of UK AGE trial, a randomized controlled trial of mammographic screening for women aged between 40 and 50 years.[35,36] Our model yields results that match the observed data in the trial reasonably well. Details about this validation will be reported separately.[7] Second, we cross-validate our model with other 5 CISNET breast models.[3–6,8,9]

The above comparisons between the 6 models provide not only cross-validations, but also a measure of uncertainty of the models' results. Besides that, our model itself provides another degree of uncertainty. The Bayesian method, through its posterior distributions, naturally provides an evaluation of uncertainty in the estimation results. Model M gives distribution estimates (as opposed to point estimates only) of treatment effects, mortality reductions, and relative contributions by treatments and screening. We report results that are averages from the posterior samples. The uncertainty on these mean results can be assessed by the quantiles of these posterior samples. While the results for years from 1975 to 2010 will be reported elsewhere, we have provided a good illustration previously for the years from 1975 to 2000 (Figure 4 in Berry and others[11]).

SCREENING SCHEDULE OPTIMIZATION AND ESTIMATION OF OVER- DIAGNOSIS RATES

The US Preventive Services Task Force (USPSTF) asked the CISNET groups to evaluate the benefits and harm of different mammography screening schedules. These schedules include annual, biennial, and triennial screenings, and a mixture or hybrid screening schedule, with varying age ranges for the beginning and end of screening. In these studies, an important assumption is that women diagnosed with breast cancer will receive the best treatment currently available. In other words, we assume perfect treatment dissemination (as opposed to real-world dissemination) when we

conduct simulations to evaluate the benefits and harm of different screening strategies. Moreover, we also assume women have perfect adherence to their assigned screening schedules. These results have been reported in a separate publication.[34]

A recent topic of interest in the breast cancer community is over-diagnosis of cancers, which was also an important aspect of our report to USPSTF. A potential use of CISNET breast cancer models including Model M is their ability to estimate the rates of over-diagnosis, which is still an open research problem. An over-diagnosis is the detection of breast cancer (DCIS or invasive) that would not have been detected in a woman's lifetime in the absence of screening. Since our model is not a natural history model, we cannot identify over-diagnosis at the individual level, and the computation of over-diagnosis rate is not straightforward. In our screening simulations, when breast cancer is detected in a specific woman, her true disease status is unknown; thus, we cannot determine whether her diagnosis of breast cancer represents a case of over-diagnosis. Our computation of the over-diagnosis rate for a specific screening schedule, such as annual screening, is based on a comparison of the total number of cases between this screening scenario (including both screening-detected and interval cases) and the scenario of no screening. We generate data of 4 million women under each of these 2 scenarios, and denote the total number of breast cancer cases detected as N_{annual} and N_{ns}, respectively. Then, we estimate the over-diagnosis rate as $(N_{annual} - N_{ns})/N_{ns}$. Note that, without knowing the true disease status of each woman, we do not use sensitivity parameters. Our computation of over-diagnosis is based on the above comparison, which may be viewed as using data from a virtual trial randomizing women to receive screening or not.

SUMMARY, DISCUSSION AND FUTURE WORK

It can be seen from above applications that our Model M, together with other CISNET models, provides a useful tool for answering important research questions related to breast cancer screening and treatment strategy evaluation and optimization. These models are especially useful for situations where clinical trials are either unethical, infeasible, or too expensive.

Contrasting other models, Model M does not make assumptions on the indolent phase (unobservable part) of a tumor's history. These assumptions are not directly verifiable using observed data, so they may be biased. We only simulate the observable part of the history, including tumor detection, treatment, and patient survival. In this sense, our model is an empirical model, not a natural history model. Natural history models assume the tumor status is known during the indolent phase, and use screening sensitivity parameters to determine whether a screening will result in a true-positive or false-negative (when a tumor is present), or use specificity parameter to give a true-negative or false-positive (when a tumor is not present). We do not conduct simulations in this way. We generate incidences according to an incidence table, which specifies the numbers of screening and clinically detected breast cancer cases per 1,000 women under different screening schedules. After a tumor is detected, it will be assigned a stage and ER/HER2 status, and the woman's time to breast cancer death is generated according to the survival distributions that depend on age at diagnosis, tumor characteristics, detection mode, and treatments received.

Another feature of our M is that, while other models take treatment effects as known and fixed parameters, we let these effects be flexible. We use our knowledge on treatment efficacy from meta-analyses of clinical trials to specify a wide-range of prior distributions for treatment effects in the population. Then these wide prior distributions are narrowed down by comparing their resulting mortality rates with SEER data (accepting good matches and rejecting poor matches). It is well-known that the effects of a treatment in the general population may be quite different from its efficacy estimated from clinical trials, due to the eligibility criteria for these trials and many other factors. Simply applying clinical trial results to the general population may cause bias. The advantage of our Bayesian acceptance/rejection method is to let the observed data obtained from SEER determine the treatment effects in the general population.

We will continue to expand our model to include more components and cover broader applications. First, we will evaluate the impact of new molecular pathways and genomic tumor profile-targeted treatment paradigms in the adjuvant setting and at recurrence. Second, we will evaluate the effect of therapeutic advances at recurrence on overall and subtype-specific, disease-free survival and overall population breast cancer mortality, quality-adjusted life years, and costs of care. Third, we will predict the reduction in population-level breast cancer mortality if new therapeutic strategies proven to be

effective at recurrence are transferred into the primary adjuvant treatment setting. Meanwhile, we will also continuously update our methodology over time to accommodate these new applications.

REFERENCES

1. Alagoz O, Berry D, Feuer EJ, et al. Introduction to the Cancer Intervention and Surveillance Modeling Network (CISNET) Breast Working Group (BWG) Models. Med Decis Making. 2016; Submitted.

2. Mandelblatt JS, Near AM, Miglioretti DL, et al. Common model inputs in collaborative breast cancer modeling. Med Decis Making. 2016; Submitted.

3. Jackson H, Vanness D, Stout NK, et al. Contribution of breast cancer to overall mortality for U.S. women by age, cohort, race and body mass index. Med Decis Making. 2016; Submitted.

4. Lee SJ, Li X, Huang H. Models for breast cancer screening strategies updated for ductal carcinoma in situ and subgroups. Med Decis Making. 2016; Submitted.

5. van den Broek JJ, van Ravesteyn NT, Heijnsdijk EA, de Koning H. Estimating the effects of risk-based screening and adjuvant treatment using the MISCAN-Fadia continuous tumor growth model for breast cancer. Med Decis Making. 2016; Submitted.

6. Schechter CB, Near AM, Jayasekera J, Chang Y, Mandelblatt JS. Structure, function, and applications of the Georgetown-Einstein (GE) Breast Cancer Simulation Model. Med Decis Making. 2016; Submitted.

7. van den Broek JJ, Nicolien T. van Ravesteyn NT, et al. Comparing simulation model predictions to observed clinical trial results for the impact of mammography screening from ages 40 to 49 on breast cancer mortality. Med Decis Making. 2016; Submitted.

8. Alagoz, Oguzhan, Ergun MA, et al. The University of Wisconsin Breast Cancer Epidemiology Simulation Model: An update. Med Decis Making. 2016; Submitted.

9. van den Broek JJ, van Ravesteyn NT, Cevik M, et al. Comparison of model structure across CISNET breast cancer simulation models using MCLIR methodology. Med Decis Making. 2016; Submitted.

10. van Ravesteyn NT, van den Broek JJ, Li X, et al. Modeling ductal carcinoma in situ (DCIS) an overview of CISNET model approaches. Med Decis Making. 2016; Submitted.

11. Berry DA, Cronin KA, Plevritis SK, et al. Effect of screening and adjuvant therapy on mortality from breast cancer. N Engl J Med. 2005;353:1784–92.

12. Berry DA. Statistics: a Bayesian Perspective. Belmont (CA): Duxbury Press; 1996.

13. Berry DA, Stangl DK. Bayesian Biostatistics. New York (NY): Marcel Dekker; 1996.

14. Spiegelhalter DJ, Abrams KR, Myles JP. Bayesian Approaches to Clinical Trials and Health-Care Evaluation. London (UK): John Wiley & Sons; 2004.

15. Tanner MA. Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions. 3rd ed. New York (NY): Springer; 1996.

16. Berry DA, Inoue L, Shen Y, et al. Modeling the impact of treatment and screening on U.S. breast cancer mortality: A Bayesian approach. J Natl Cancer Inst Monogr. 2006;36:30–6.

17. Shen Y, Yang Y, Inoue LY, Munsell MF, Miller AB, Berry DA. Role of detection method in predicting breast cancer survival: analysis of randomized screening trials. J Natl Cancer Inst. 2005; 97:1195–203.

18. Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov) SEER*Stat Database: Incidence - SEER 9 Regs Research Data, Nov 2015 Sub (1973- 2013) (Katrina/Rita Population Adjustment) -Linked to County Attributes - Total U.S., 1969-2014 Counties, National Cancer Institute, DCCPS, Surveillance Research Program, Surveillance Systems Branch, released April 2016, based on the November 2015 submission.

19. Chlebowski RT, Hendrix SL, Langer RD, et al. Influence of estrogen plus progestin on breast cancer and mammography in healthy postmenopausal women the Women's Health Initiative Randomized Trial. J Am Med Assoc. 2003;289:3243–53.

20. 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.

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

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

23. 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.

24. Early Breast Cancer Trialists' Collaborative Group. Effects of chemotherapy and hor- monal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet. 2005;365:16871717.

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-censual Census Bureau estimates. Available at: https://www.woodsandpoole.com/.

26. American Joint Committee on Cancer (AJCC) Cancer Staging Manual. 6th edition. Philadelphia, PA: Lippincott Raven Publishers; 2002.

27. Shapiro S, Venet W, Strax P, Venet L. Periodic screening for breast cancer: The Health Insurance Plan Project and its sequelae, 1963-1986. Baltimore (MD): Johns Hopkins University Press; 1988.

28. 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.

29. 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.

30. Cronin KA, Yu B, Kraphcho M, et al. Modeling the dissemination of mammography in the United States. Cancer Causes Control. 2005;16:701–12.

31. Hakkama M, Holli K, Isola J, et al. Aggressiveness of screen-detected breast cancers. Lancet. 1995;345:221–4.

32. Frisell J, Eklund G, Hellstrom L, Somell A. Analysis of interval breast carcinomas in a randomized screening trial in Stockholm. Breast Cancer Res Treat. 1987;9:219–25.

33. 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.

34. Mandelblatt JS, Stout NK, Schechter CB, et al. Collaborative modeling of the benefits and harms associated with different U.S. breast cancer screening strategies. Ann Int Med. 2016;164:215–25.

35. Moss SM, Cuckle H, Evans A, Johns L, Waller M, Bobrow L. Trial Management Group. Effect of mammographic screening from age 40 years on breast cancer mortality at 10 years' follow-up: a randomised controlled trial. Lancet. 2006;368:2053–60.

36. Moss SM, Wale C, Smith R, Evans A, Cuckle H, Duffy SW. Effect of mammographic screening from age 40 years on breast cancer mortality in the UK Age trial at 17 years' follow-up: a randomised controlled trial. Lancet Oncol. 2015;16:1123–32.