FLEXKB DOCUMENT
Version: HI.001.07242013.3523
Document generated: 07/24/2013

CISNET Breast Cancer Collaborators

Important note: This document will remain archived as a technical appendix for publications. New versions will be added periodically as model refinements and updates are completed. The most current version is available at http://cisnet.cancer.gov/profiles. Note that unlike most PDF documents, the CISNET model profiles are not suitable for printing as they are not typically written or read in sequential fashion.

We recommend you let your interests guide you through this document, using the navigation tree as a general guide to the content available.

The intent of this document is to provide the interested reader with insight into ongoing research. Model parameters, structure, and results contained herein should be considered representative but preliminary in nature.

We encourage interested readers to contact the contributors for further information.

	University of Wisconsin
	Dana-Farber University
	Standford University
Erasmus MC	Erasmus MC (Breast)
THE UNIVERSITY OF TEXAS MD ANDERSON CANCER CENTER Making Cancer History	University of Texas MDACC
	Georgetown University

FLEXKB DOCUMENT
Version: HI.001.07242013.3518
Document generated: 07/24/2013

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

University of Wisconsin

Important note: This document will remain archived as a technical appendix for publications. New versions will be added periodically as model refinements and updates are completed. The most current version is available at http://cisnet.cancer.gov/profiles. Note that unlike most PDF documents, the CISNET model profiles are not suitable for printing as they are not typically written or read in sequential fashion.

We recommend you let your interests guide you through this document, using the navigation tree as a general guide to the content available.

The intent of this document is to provide the interested reader with insight into ongoing research. Model parameters, structure, and results contained herein should be considered representative but preliminary in nature.

We encourage interested readers to contact the contributors for further information.

Go directly to the: Reader's Guide.

READERS GUIDE

Core Profile Documentation

These topics will provide an overview of the model without the burden of detail. Each can be read in about 5–10 minutes. Each contains links to more detailed information if required.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

Model Purpose

This document describes the primary purpose of the model.

Model Overview

This document describes the primary aims and general purposes of this modeling effort.

Assumption Overview

An overview of the basic assumptions inherent in this model.

Parameter Overview

Describes the basic parameter set used to inform the model, more detailed information is available for each specific parameter.

Component Overview

A description of the basic computational building blocks (components) of the model.

Output Overview

Definitions and methodologies for the basic model outputs.

Results Overview

A guide to the results obtained from the model.

Key References

A list of references used in the development of the model.

Further Reading

These topics will provide a intermediate level view of the model. Consider these documents if you are interested gaining in a working knowledge of the model, its inputs and outputs.

JNCIMonograph Outline

This topic provides links to profile content organized according to the JNCI Monograph Outline for Model Description Chapters. Use this outline for comparisons focused on the CISNET Base Case simulations.

Advanced Reading

These topics denote more detailed documentation about specific and important aspects of the model structure

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

MODEL PURPOSE

SUMMARY

The Wisconsin model simulates breast cancer in a population over time generating cancer registry–like data sets. By manipulating parametric input assumptions about natural history, screening, and treatment the model can be used to address a number of important policy questions.

PURPOSE

The purpose of the NCI Cancer Intervention and Surveillance Modeling Network (CISNET) is to promote simulation modeling as a tool that, in conjunction with the nation's cancer surveillance systems, can help to explain observed changes in cancer incidence and mortality. Wisconsin offers a unique population laboratory to develop and test breast cancer simulation models to meet this goal. We propose a collaboration among simulation and statistics experts, and surveillance and epidemiology experts at the University of Wisconsin, and the state of Wisconsin's Cancer Reporting System to study the use of simulation modeling to better understand trends in breast cancer epidemiology and to enhance the use of simulation modeling for this purpose.

The Wisconsin model evolved from a simulation model constructed by Polun Chang a decade ago for his Ph.D. dissertation[1].Chang asked whether the observed breast cancer incidence and mortality in the state of Wisconsin over the years 1982–92 could be represented by a mechanistic simulation model comprised of reasonable sub–models of population demography, biologic onset and progression of breast cancer, screening for and detection of breast cancer, and breast cancer treatment effectiveness. He programmed a deterministic model which attempted to replicate the Wisconsin Cancer Reporting System (WCRS) data on the annual age– and stage–specific incidence of breast cancer from 1980 to 1992[2]. Chang concluded that a substantial fraction (9%–23%) of all breast cancers are pre–destined from their occult biologic onset to grow only to a limited size (~1 cm diameter), would not present a lethal threat to the woman, and would be indistinguishable from potentially lethal tumors of similar size. He termed these indolent tumors "limited malignant potential (LMP)" tumors.

We had two objectives in redesigning the Chang model as part of the Cancer Intervention and Surveillance Modeling Network (CISNET). The first objective is to answer a question similar to Chang's:

Is it possible to generate a realistic virtual Wisconsin cancer registry of incident breast cancers for women residing in Wisconsin from 1975 to 2000, and to simultaneously replicate age–specific breast cancer mortality in this population during the same time period, with a micro–simulation model comprised of realistically modeled processes representing breast cancer

biologic onset and progression, detection by mammogram screening and case finding outside of screening, and evolving treatment effectiveness over the same time period?

Thus we ask whether observed cancer registry data are compatible at a relatively fine scale over time with the joint product of dynamic processes which most epidemiologists and physicians would agree underlie observed breast cancer data, when those processes are constrained to behave in manner and scale as we think they should. Chang found that he had to add a class of tumors, LMPs, which are indistinguishable from small invasive breast cancers but which in fact do not represent a threat to the host. We began the modeling process prepared to add such assumptions reluctantly, instead exploring many plausible combinations of parameters to improve fit of the virtual cancer registry before resorting to unobservable assumptions about the underlying systems being modeled.

The second objective is to produce a model which can be used to explore ramifications of alternative programs of screening and treatment for breast cancer. Once a simulation model is constructed, it is quite flexible. The model allows output of both simulated cancer registry data and also similar data about breast cancer latent in the population at any given time. This provides the means to answer "What if?" questions about changes in tumor detection and improvements in therapy.

REFERENCES:

[1] Chang, P., "A Simulation Study of Breast Cancer Epidemiology and Detection since 1982: The Case for Limited Malignant Potential Lesions [Ph.D.]" 1993;

[2] Bureau of Health Information, Division of Health Care Financing, Wisconsin Department of Health and Family Services "Wisconsin Cancer Incidence and Mortality, 1999" 2002;

Model Overview

Summary

This document overviews the modeling effort, the problems it addresses and previous work relevant to this model.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

Purpose

Model Purpose

Background

The National Cancer Institute's Cancer Intervention and Surveillance Modeling Network (CISNET) was formed to use simulation modeling of surveillance data to better understand cancer incidence and mortality. The state of Wisconsin offers a unique population laboratory to develop and test breast cancer simulation models to meet national goals of improving cancer surveillance methods. As part of this consortium we developed and calibrated the Wisconsin Breast Cancer Epidemiology Simulation Model, a discrete–event, stochastic simulation model designed to replicate breast cancer incidence and mortality rates in the Wisconsin female population and applicable to the US population from 1975–2000. The simulation was developed using a systems–science, process modeling approach.

Model Description

We have taken a systems engineering approach to construction of our simulation model for breast cancer incidence and mortality in a population. The complex, dynamic biologic and sociodemographic system which results in observed breast cancer statistics is comprised of models of subsystems and specifying the interactions among them in a process analogous to what has been described as "reverse engineering" of complex biologic systems[1]. Our model is a discrete–event simulation with a fixed cycle time of 6 months beginning in calendar year 1950. The model is populated by 2.95 million women, divided into birth cohorts, and making up the female population aged 20–100 years of age living in Wisconsin between 1950 and 2000. Women in each birth cohort are individually simulated from calendar year 1950 (or the year in which they were age 20) until they die a simulated death, achieve age 100, or the simulated year 2000 is reached. The processes simulated are:

A. the natural history of breast cancer from inception to breast cancer death;
B. detection of breast cancer by screening mammography or clinical surfacing;
C. improvements in treatment of breast cancer and diffusion of treatments over time; and
D. death from non–breast cancer causes.

Each of these four major processes is stochastic, unfolding over time in the population, and they jointly result in the observed cancer registry data. These processes form a delicately balanced, interacting system within the population over time. They result in observable consequences unfolding over time as embodied in the statistics collected by a comprehensive cancer surveillance system. When referred to a specific population and time period, these processes result in observed counts of incident breast cancers in each of four distinct stages of disease, in women with known ages, year by year in the reporting system. The model processes also result in counts of deaths in women with known ages across the same years.

CONTRIBUTORS

Dennis G. Fryback, Ph.D.
Natasha K. Stout, Ph.D.
Marjorie A. Rosenberg, Ph.D., F.S.A.
Amy Trentham–Dietz, Ph.D.
Patrick L. Remington, M.D., M.P.H.
Vipat Kuruchittham, M.S.

Acknowledgments

We wish to thank Polun Chang, Ph.D., Associate Professor, National Yang Min University, Taipei, Taiwan, for his assistance in initiating reprogramming of his original model, and Prof. Michael Shwartz has assisted with the tumor progression model. Prof. Miron Livny and his graduate students have provided valuable assistance implementing our model on CONDOR. Discussions with Prof. Polly Newcomb, Dr. Richard Love, Dr. Elizabeth Burnside, Dr. Tara Breslin, and with CISNET collaborators at other institutions have been invaluable in many aspects of the process. Laura Stephenson of the Wisconsin Cancer Reporting System was very helpful in providing WCRS data and responding to questions. We have had programming assistance from Thotsaporn Thanatipanonda, Lorne Tappa, Sriram Ganesan, and Vivek Puttabuddhi.

REFERENCES:

[1] Csete, ME, Doyle, JC "Reverse Engineering of Biological Complexity" in Science 2002; 295: 5560: 1664-1669

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

ASSUMPTION OVERVIEW

SUMMARY

Our modeling approach assumes that observed incidence and mortality from breast cancer over time can be replicated mechanistically by assembling simulation modules representing basic processes whose general nature and operation is known, but which may be governed by poorly known parameters. This document lists these processes and assumed structure.

BACKGROUND

We assume four major processes, each ultimately stochastic, and each unfolding over time in a poulation, result in the statistics observed over time in breast cancer surveillance systems:

A. Natural history of breast cancer—

1. Onset: A primary breast cancer may be initiated at some point in a woman's life.
2. Progression/regression: The cancer will grow over time generally progressing in size and spreading to other tissues; depending on their character, tumors may or may not be a lethal threat to the host, and some tumors may regress.
3. Death: Breast cancer deaths occur as an endpoint of a process of uncontrolled growth and spread of the tumor.

D. Detection of breast cancer—

1. Mammogram screening: Breast cancer may be detected by a screening mammogram. A woman's participation in screening may be stochastic or systematic over time. Mammography screening has diffused, and possibly improved over time.
2. Clinical surfacing: Breast cancer may be detected other than by screening mammogram. We do not differentiate these pathways which include diagnosis after symptomatic presentation, self–breast examination, clinical breast examination, or incidental pre–symptomatic detection, terming them collectively as "clinical surfacing."

C. Treatment of breast cancer—

- Potentially lethal breast cancer may in some cases be arrested or retarded by medical intervention. Interventions include surgery of varying extent with or without subsequent radiation, which we take as baseline treatment. Adjuvant therapy with tamoxifen and/or polychemotherapy has been introduced over the time period covered by the model. In the simulation at the individual woman level we use an all or nothing "cure" model for treatment; this sub–model approximates population–level treatment effectiveness statistics.
- Non–breast cancer mortality—
- Women die of non–breast cancer causes, with or without breast cancer present.

Each of these processes is modeled with mathematical functions and stochastic processes. The joint model is referred to as the simulation model or "the model." The parameters governing these functions and processes are constrained by general knowledge, by published data, and by systematic experiments in which the simulation computes estimates of observed surveillance data over time and deviations between the simulation results and observed data are reduced by changing the parameters (we term this latter process calibration, or fitting of the model). Because these processes form a complex and interacting system, some individual parameters may not be identifiable in which case the goal of simulation modeling is to identify feasible sets of values in the parameter space.

Assumption Listing

Breast Cancer Natural History Assumptions

1. The probability of breast cancer onset in any given time interval is a function of the woman's individual risk factors and a residual secular trend. (Cancer Incidence Component)
2. Tumor growth is a function of a random intial growth parameter. (Natural History Component)
3. Some tumors have limited malignant potential and will never be a lethal threat to the woman host. (Natural History Component)
4. Breast cancer death can occur only after a tumor has reached the distant stage. (Survival And Mortality Component)

Breast Cancer Detection Assumptions

5. Breast cancer is detected by either screening (characterized in the model as screening mammography) or by other means (characterized in the model as "surfacing"). (Screening Component)
6. The probability an undetected tumor surfaces in a given cycle of the model is an increasing function of tumor diameter. (Screening Component)
7. The sensitivity of a screening mammogram is an increasing function of the woman's age (=50), an increasing function of the diameter of the tumor, and an increasing function of the calendar year of the screening mammogram. (Screening Component)
8. The probability a woman receives a screening mammogram is characterized as increasing over calendar time and her age. (Screening Component)

Treatment Effectiveness Assumptions

9. Treatment outcome is modeled as "cured" or "not cured".
    a. A cured cancer's growth is arrested instantaneously at the time of treatment.
    b. An uncured cancer continues to progress at the rate governed by its initially assigned growth parameter to the distant stage and eventual breast cancer death if death from other causes does not intervene. Spherical diameter increase in the model due to "progression" in this case is obviously metaphorical reference to tumor load and not physical diameter. (Natural History Component)

3. The probability of cure is a function of the treatment given and stage of the tumor at detection and possibly the woman's age. (Treatment Component)
4. Treatment modality is a probabilistic function of the calendar year of detection, as treatments have changed and diffused over time. (Treatment Component)

Non Breast Cancer Mortality

12. Non breast cancer mortality occurs independently of whether or not the woman has breast cancer and is a function of the woman's current age and the calendar year of her birth. (Survival And Mortality Component)

PARAMETER OVERVIEW

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

SUMMARY

The Wisconsin model has many input parameters. This document lists these inputs and their general nature—constants, tables, functions, etc.

BACKGROUND

The Wisconsin model is not estimated from any one data source. We have identified processes that are inputs to the epidemiology of breast cancer or which affect statistics that are collected by cancer surveillance in Wisconsin (and nationally) and have attempted to characterize each of these processes mathematically, either with parameters from published or available data sources or parameters estimated by expert judgment.

For demographic data we generally used census data for the US and Wisconsin. Mortality is derived from the Berkeley tables. Cancer surveillance data from the Wisconsin Cancer Reporting System and from SEER were used. Mammography dissemination was from data published in Wisconsin and from the NCI CISNET basecase analysis. Treatment dissemination was supplied by NCI and treatment effectiveness was from the EBCTCG meta–analyses. Breast cancer natural history was patterned after the Shwartz model[1]. Mammography characteristics were based on published literature supplemented by expert judgment.

SEER data for Iowa were excluded to allow cross–validation of the Wisconsin model using Iowa data.

Also see Model Calibration Procedures

PARAMETER LISTING OVERVIEW

The following lists the logical clusters of inputs to the Wisconsin breast cancer simulation model.

Simulation Control

- Starting Simulated Year
- Ending Simulated Year
- Number of Burn–in years
- Termination age
- Cycle Length
- Number of replications

Population Component

- Birth Cohorts
- Number of women in Birth Cohort
- Non–breast cancer mortality tables
- Breast cancer mortality tables

- Population for age–adjusted output

Screening Component

- Mammography screening dissemination model
- Clinical surfacing probabilities
- Mammography operating characteristics

Cancer Incidence Component

- Tumor onset function

Natural History Component

- Tumor type
- Tumor growth rate (exponential growth)
- Positive node probability parameters
- Time to BC death distribution given stage 4
- Tumor size–to–historical–stage translation table

Treatment Component

- Treatment effectiveness by stage
- Adjuvant therapy dissemination model

REFERENCES:

[1] Shwartz, M. "A Mathematical Model Used to Analyze Breast Cancer Screening Strategies." in Operations Research 1978; 6: 26: 937-955

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

Component Overview

Summary

The Wisconsin model simulates life histories of individual women from age 20 to age 100, death, or the year 2000. The model utilizes components comprising mortality from non–breast cancer causes, breast cancer onset and progression and mortality, screening and detection, and treatment. These are represented in the flowchart below.

Overview

Component Listing

Population Component – Starting the model in 1950, we simulate a number of women equal to each single–year age cohort of women in the Wisconsin population aged 20–99. In each year 1951–1999 we add to the simulation model the number of women aged 20y in that year. The total number of women simulated is approximately 2.95 million.

Cancer Incidence Component – We use "incidence" to mean detection of a tumor; we use "onset" to refer to biological initiation of the tumor. Since tumors grow over time in our model we assumed that tumors which are incident (detected) at one time in fact were biologically onset earlier, and that some women will die of non–breast cancer causes with occult, undiagnosed breast cancer.

Natural History Component – We model breast cancer as a progressive disease, starting with biologic onset at a small focus within the breast and growing spherically in size over time, with probabilistic spread to lymph nodes. The growth model is a simplification of a complicated biologic process, assuming that most breast cancers start as small non–invasive entities and if detected at this time would likely be classified as having historical stage "in situ."

Screening Component – We specify tumor detection probabilities, whether by

mammography or clinical surfacing, as a function of the diameter of the tumor in centimeters, the age of the woman, and the calendar year being simulated. These probabilities were originally specified by a priori expert judgment, then refined by calibrating model outputs to observed surveillance incidence data by tumor stage, not size. The dependence within the model on size is due to an interaction in the model between the growth sub–model and the detection sub–model and the output of these as counts of detected simulated tumors graded into the four stages.

Treatment Component – For simplicity we model treatment as a cure/no–cure process. When a breast cancer is detected, regardless of mode of detection, we assume it is treated. The result of simulated treatment is either "cure" with total arrest of progression at that time and hence no possibility of progressing to a breast cancer death, or the result is "no cure" in which case the tumor continues to progress as if it were undetected and the woman may die of breast cancer, competing causes, or achieve age 100 depending on her individual circumstances.

Survival And Mortality Component – Breast cancer death occurs if the time to a woman's death from non–breast cancer causes is longer than her time to death from breast cancer. The time until death from non–breast cancer causes is chosen at the start of the simulation.

OUTPUT OVERVIEW

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

SUMMARY

Because the Wisconsin model simulates life histories for individual women, in principle these entire histories are available at the end of the run. In practice the main outputs are age– and historical stage–specific incidence rates of breast cancer and age–specific breast cancer mortality for each year from simulated 1975–2000.

OVERVIEW

The usual outputs saved from a run of the model are as follows:

Means, and standard deviations across N simulations (of the population of interest) for —

1. 5–yr age group age–, and historical–stage–specific incidence rates of breast cancer in each of the simulated years 1975–2000.
2. 5–yr age–specific breast cancer mortality rates in each of the years simulated.
3. Prevalence of breast cancer in simulated year 1975 (an output constructed for the breast base case)

Alternatively, the same output information can be displayed as age–adjusted rates over calendar years.

With modest reprogramming, the simulation also is capable of outputting the following sorts of quantities —

4. Rates or counts of true positive, false positive, true negative, and false negative mammograms conditioned on screening history and age.
5. Age–specific and stage–specific prevalence of undetected breast cancer in any given calendar year from 1975–2000.

OUTPUT LISTING

See Output Description for more detail.

RESULTS OVERVIEW

SUMMARY

This document describes results generated by the model.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

OVERVIEW

THE FIT OF THE FINAL MODEL AGAINST SEER AND WCRS.

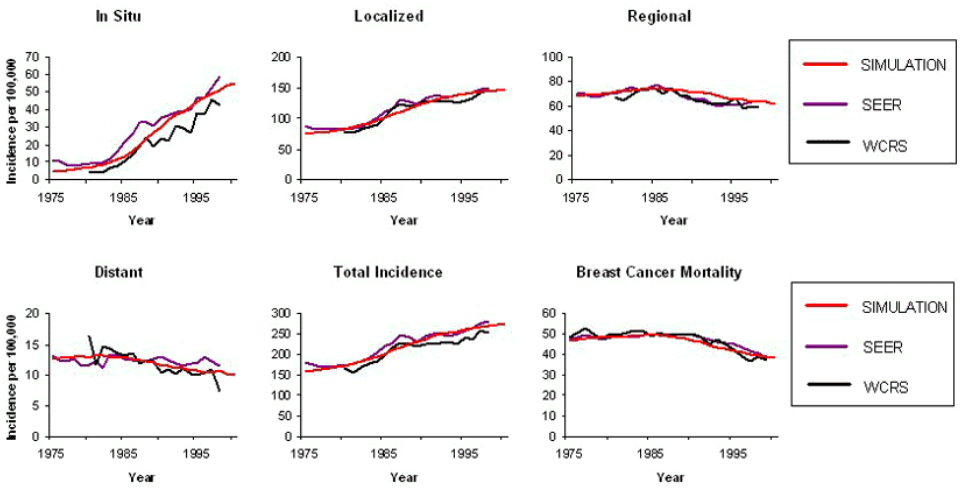


FIGURE 1. The fit of the final model against SEER and WCRS. Incidence rates are age–adjusted to the U.S. population aged 30–79 in year 2000.

RESULTS LIST

To fit observed data, the model required the following assumptions about breast cancer.

1. Breast cancer is a heterogeneous disease, varying in growth rates and in aggressiveness. It is not necessary to assume the heterogeneity is related to age to reproduce many aspects of observed epidemiology of breast cancer with respect to age.
2. There is a class of breast cancer with limited malignant potential, constituting a relatively prevalent latent pool of cancer in the population. If modeled as solid spherical tumors LMP tumors will grow to approximately 1 cm diameter, persist at that size for up to 2 years, and then recede. LMP tumors constituted the followings in the year 2000 (Figure 2).

- 42% of all cancers at biological onset (= 126÷ (126+174))
- 28% of all incident tumors (= (24+45)÷(24+45+30+101+62+10))
- 44% of all incidence in situ tumors (= 24÷(24+30))
- 31% of all incident localized tumors (= 45÷(45+101))
- There is a small population of breast cancers which are metastatic almost from their beginning. These constitute approximately 4% of non–LMP tumors (or 2% of all tumors).

- The growth rates of tumors which are neither LMP nor the early metastatic type are described by a Gompertz distribution implicit in the growth curves in Figure 3, which shows size of tumors as function of time and percentile of the distribution of means for the gamma distribution of Gompertz growth rates.

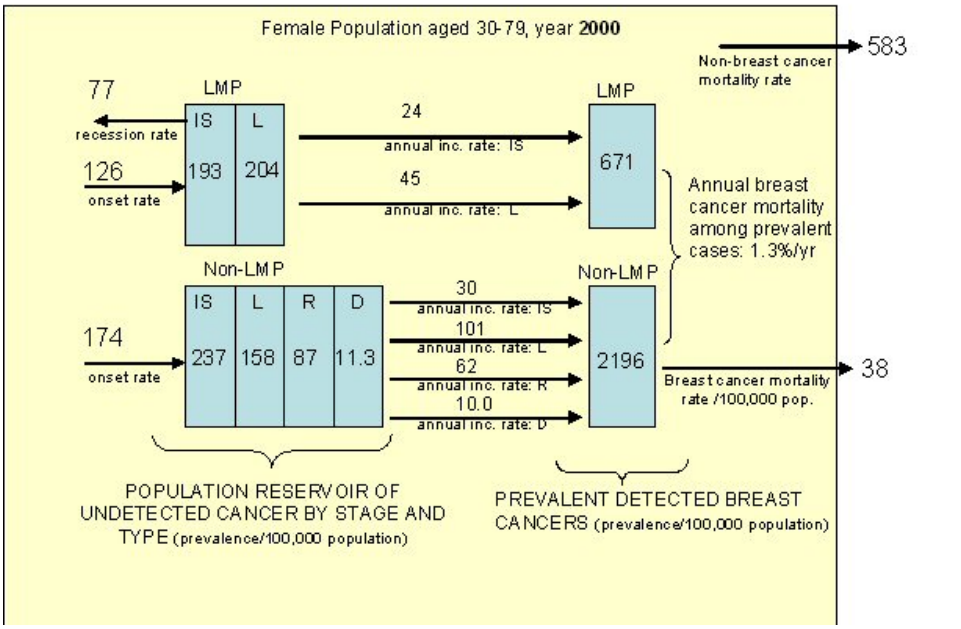


FIGURE 2. A snapshot of occult, incident, and prevalent breast cancer in the U.S. female population aged 30–79 in 2000 as predicted by the simulation model is depicted. Incidence and prevalence are shown as rates per 100,000 in the female population aged 30–79. The annual prevalent case mortality rate is computed as percent per year of all prevalent breast cancers (including LMPs) who die in that year. Of the 671+2196=2867 prevalent cases of breast cancer, 23% are LMP; these women were treated for breast cancer that was not a threat to them. **ABBREVIATIONS:** IS = in situ, L=local, R=regional, D=distant, LMP=Limited Malignant Potential.

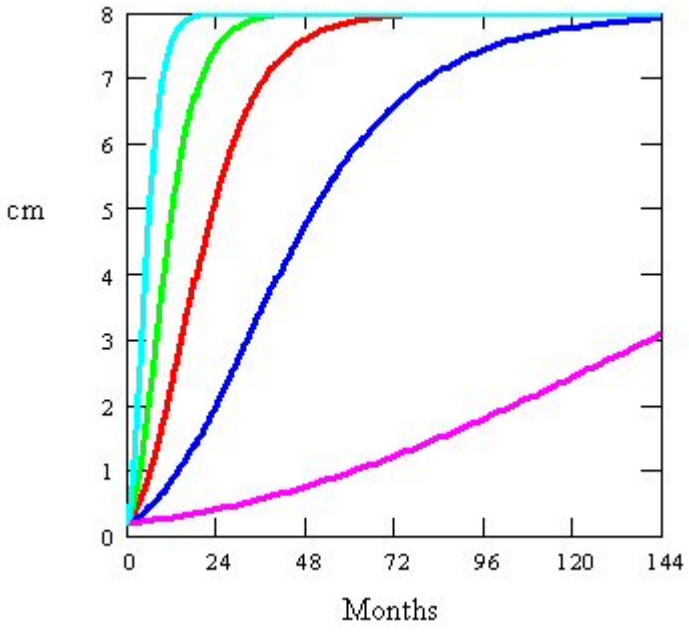


FIGURE 3. Sample growth curves (diameter of ideal spherical tumor as function of time since onset at 0.2 cm diameter) for tumors are shown. The five curves, left to right, represent values of the Gompertz growth rate at the 95th, 75th, 50th, 25th, and 5th percentiles of the gamma distribution for growth rate in the best fitting parameter vector.

Also see: Model Calibration Procedures

CANCER INCIDENCE COMPONENT

SUMMARY

This document describes the method by which tumor is initiated in the model.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

OVERVIEW

We use "incidence" to mean detection of a tumor; we use "onset" to refer to biological initiation of the tumor. Since tumors grow over time in our model we assumed that tumors which are incident (detected) at one time in fact were biologically onset earlier, and that some women will die of non–breast cancer causes with occult, undiagnosed breast cancer. With this consideration in mind, and allowing assumptions about the LMP and hyper–aggressive classes of tumors discussed above, the onset rate for breast cancer is computed in our model as follows.

DETAIL

To derive an onset rate for a woman aged N years in calendar year Y we suppose the observed incidence rate in the absence of screening for a woman in the same birth cohort, but age $N + l$ is I_{N+l} where l is an average lag in years between onset and incidence of tumors in the population. Let p be the proportion of incident (detected) tumors which are not LMP tumors, and let f be the fraction of all onset tumors which are LMP. Then the onset rate for the woman aged N years is given by $O_N = \frac{pI_{N+l}}{(1-f)}$. The numerator of the fraction is the rate of non–LMP tumors as a fraction of all incident tumors l years later, and these are $1 - f$ of the total number of tumors which are onset. The model uses age–period–cohort breast cancer incident rates inferred in the absence of screening, which incorporate an increasing secular trend, provided to the CISNET collaboration by the NCI (see Figure 4)[1]. We fit the lag parameter, l, empirically during calibration. The fitted values derived for the three parameters during model calibration are l=3, f=0.42, and p=0.95.

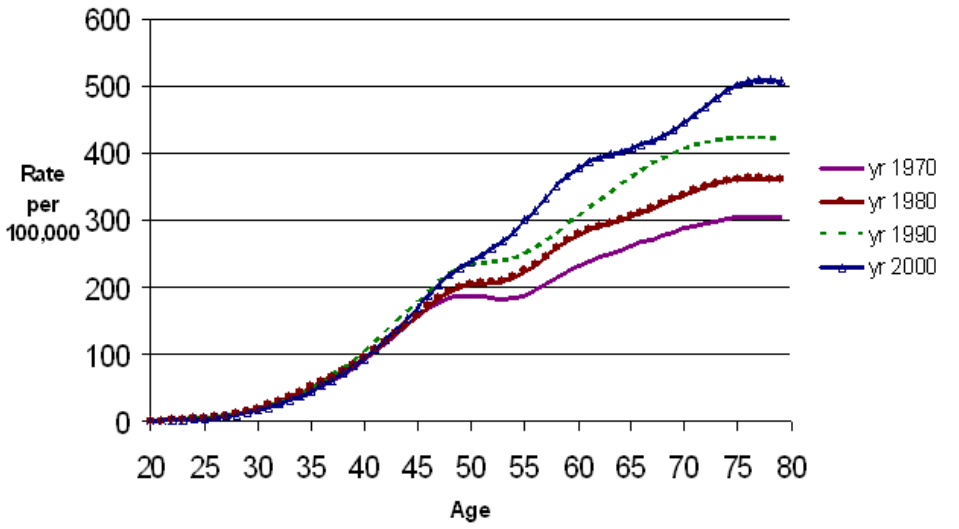


FIGURE 4. An age–period–cohort model for breast cancer incidence in absence of screening. Derived for CISNET consortium.[1]. The four lines depict incidence in the indicated years.

Actual incidence observed as output from the simulation model is a function of the underlying pool of tumors that are biologically onset, the nature of the tumors (in particular, tumor diameters), the probability of clinical surfacing as a function of size of tumor, and the operating characteristics for screening mammography as a function of size of the tumors (i.e., screening sensitivity).

With one exception, we do not model recurrence or second primary breast cancers. The exception is that a woman with an undetected LMP tumor which then disappears is again at risk for onset of any type of breast cancer. A woman may not have both an LMP and a "real" tumor simultaneously. Once a woman is diagnosed with breast cancer—LMP or otherwise—in the current simulation she cannot develop a second primary breast cancer. We do not model recurrence per se, as discussed later under "treatment."

REFERENCES:

[1] Holford, TR., Cronin, K., Feuer, EJ., Mariotto, A. "Changing patterns in breast cancer incidence trends, manuscript" 2003;

NATURAL HISTORY COMPONENT

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

SUMMARY

This document describes the model of tumor progression.

OVERVIEW

We model breast cancer as a progressive disease, starting with biologic onset at a small focus within the breast and growing spherically in size over time, with probabilistic spread to lymph nodes. The growth model is a simplification of a complicated biologic process, assuming that most breast cancers start as small non–invasive entities and if detected at this time would likely be classified as having historical stage "in situ."

DETAIL

Natural history of the disease.

We discuss tumor progression before biological tumor onset here since we found parameters controlling onset depend on the nature of the progression model.

The tumor progression model.

We initially began with a model proposed by Shwartz[4] which modeled tumor growth as an exponential doubling process. In Shwartz's model, every breast cancer is assumed to have a fixed growth rate once drawn at the tumor's biologic inception from a lognormal distribution. Shwartz modeled the number of involved lymph nodes at a given time as a cumulative Poisson process with rate parameter determined by current tumor diameter and the rate at which the diameter is changing so that the larger the tumor, the more likely it is to have involved nodes and the faster growing the tumor is the more likely it is to have metastatic spread. In Shwartz's model all tumors started at a diameter of 0.5cm.

We altered two aspects of Shwartz's model to better calibrate to surveillance data. Because modern screening is potentially capable of detecting breast cancers less than 0.5cm in size, tumors now enter the simulation with a size of 0.2cm. Second, we implemented a decelerating, Gompertz growth function[5] to replace Shwartz's exponential growth function. Shwartz had suggested a Gompertz function may fit equally well as the exponential model and did some exploration of this alternative.[6]; (M. Shwartz, personal communication, 2002) The exponential model, characterized by constant tumor volume doubling times, is plausible for early tumor growth, but implausible as tumors become larger in size. The Gompertz growth model is exponential growth with decelerating doubling time shown in equation (1).

$$V(t) = V_0 e^{\frac{\beta}{\alpha}(1-e^{-\alpha t})} \tag{1}$$

Here $V(t)$ is the tumor volume at time t, V_0 is the initial tumor volume, the parameter β is the initial tumor growth rate, and a governs deceleration in growth rate over time. We fix one of these parameters by fixing the maximum asymptotic tumor volume as $t \to \infty$, $V_{max} = V_0 e^{\beta/\alpha}$. Solving for β and substituting into equation (1) we get the volume at time t expressed in terms of one free parameter, a.

$$V(t) = V_0 e^{ln\left(\frac{V_{max}}{V_0}\right)(1-e^{-\alpha t})} \tag{2}$$

Finally, assuming spherical tumors, tumor volume and diameter are related by $V = \frac{\pi}{6} d^3$. Substituting this expression where appropriate in (2), taking the cube root of both sides, and solving for $d(t)$ as a function of corresponding constants d_{max}, d_0, and a, we get an equation for tumor diameter at time t.

$$d(t) = d_0 e^{ln\left(\frac{d_{max}}{d_0}\right)(1-e^{-\alpha t})} \tag{3}$$

The minimum and maximum diameters were specified arbitrarily at 0.2 and 8.0 cm, to be reasonably within bounds observed clinically. The growth parameter α is modeled as a gamma–distributed random variable with a mean of 0.12 and a variance of 0.012, values derived in the model calibration process. The mean and variance of the gamma distribution were determined by model calibration. Figure 1 shows tumor diameter as a function of time (in months) for tumors with growth parameters from various percentiles of the growth rate distribution.

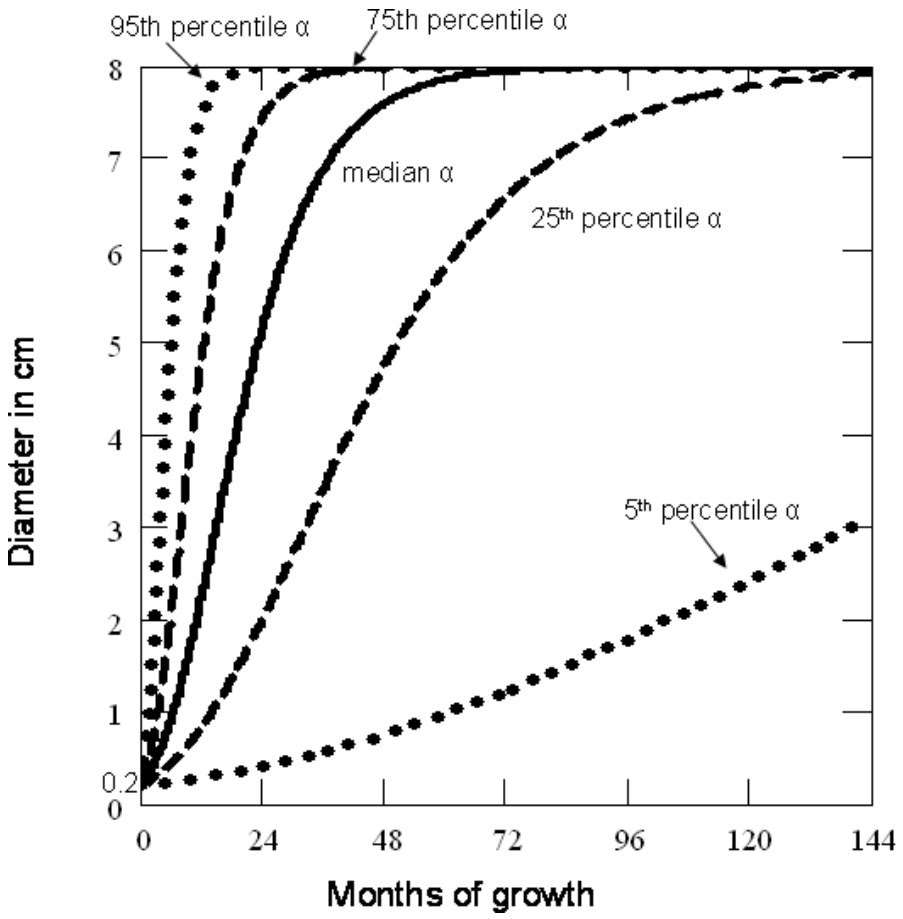


FIGURE 1. Sample growth curves for tumors. The ordinate is diameter in cm, and the abscissa is the number of months since tumor inception. The five curves, left to right, represent values of α at the 95th, 75th, 50th, 25th, and 5th percentiles of the gamma distribution for growth rate (smaller values of represent slower growth)

The rate of additional involved nodes was modeled based on Shwartz's model[6]. In Shwartz's model the instantaneous rate at time t was given by $n(t) = b_1 + b_2V(t) + b_3V(t)$ where the bi, are constants with values 0.0058, 0.0053, and 0.0002 respectively, i=1,2,3,. Integrating $n(t)$ from time $t-1$ to time t, we derive the additional nodes in that interval,

$$N(t)-N(t-1) = \int_{t-1}^{t} (b_1+b_2\times V(u)+b_3\times V'(u))du$$

In calibrating the model to observed incidence surveillance data we found that this formula gave slightly too fast a rate of tumor spread. For a tumor of a given diameter, $d(t)$, and corresponding volume $V(t)$, we used Shwartz's equation for a tumor with a volume corresponding to a diameter 25% smaller, $0.75d(t)$, based on fitting this parameter during calibration. Figure 2 shows the resulting empirical rate of additionally involved lymph nodes in a 6 month period as a function of simulated tumor diameter and growth rate.

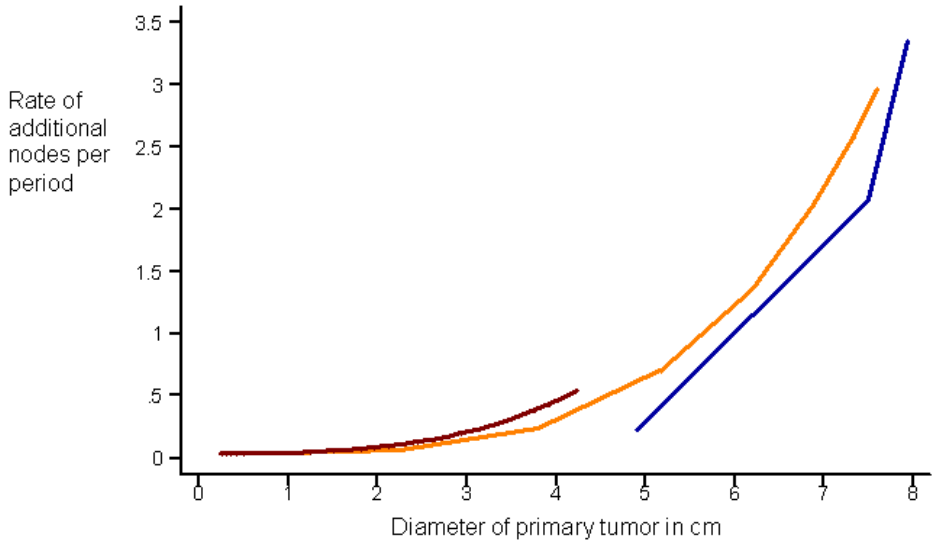


FIGURE 2. Poisson rate of additional involved nodes in the next 6 month period as a function of current tumor size. The three curve segments represent (from left to right) tumors in the 5th , 50th, and 95th percentiles of the growth rate distribution.

Tumors with zero nodes that are less than 0.95 cm in diameter are defined as having in situ stage in our model. Tumors with zero nodes and greater than 0.95 cm are considered to be localized. Tumors with 1 to 4 nodes are considered to be regional stage tumors. Any tumor with 5 or more involved nodes is considered to be in the distant stage and proxy for more widespread involvement.

Additional assumptions about tumor natural history.
Preliminary model results revealed that far too many occult tumors were required to achieve the rise in incidence (detection) that occurred after the widespread use of screening mammography during the late 1980s. In order for breast cancer incidence to be sufficiently high during this time period the model required unrealistically elevated breast cancer mortality rates prior to the dissemination of mammography. For this reason, we incorporated the same assumption that Chang included in his original model: A substantial fraction of all incident breast cancers prior to screening must have be of limited malignant potential (LMP), i.e., of no lethal threat to the host woman. Furthermore, we inferred there must be a reservoir of these occult LMP tumors that would be discovered with the advent of screening programs. The following characterization of LMP tumors evolved during model fitting and calibration:

LMP tumors have the same growth rate distribution as other tumors (the gamma distribution described above), and grow according to the same Gompertz growth function as other tumors. However they cease growth at 1 cm in size, and they disappear after 2 years dwell time at this size. LMP tumors never exhibit metastatic spread, and thus do not lead to breast cancer death.

Each of these characteristics was needed to obtain satisfactory calibration of the model dynamically across time. Since LMP tumors were needed as biologically prevalent but

occult tumors to feed the dramatic rise in incidence of localized breast cancer with the advent of screening, they could not progress beyond the localized stage and do not cause breast cancer death. To remain largely occult in the absence of screening they could not grow too large, hence the 1 cm upper limit. Too large a pool of undiagnosed, prevalent tumors developed in the simulation if the LMP tumors remained indefinitely after reaching the 1 cm limit, so we imposed the 2–year dwell time (this parameter being fit during calibration); future versions of the model will include regression over time, but currently disappearance is modeled as instantaneous at 2 years. The fraction of all tumors that were defined as LMP tumors was found through calibration to incidence surveillance data and the corresponding parameters are discussed below under "onset." The existence of LMP tumors is at present hypothetical. We must assume they are histologically indistinguishable from "real" breast cancer or their existence would be already known. It is our claim, resulting from our modeling efforts, that assuming their existence is necessary to explain the dynamics of the observed facts about breast cancer in the population over the past 25 years.

It is reasonable to ask whether the role that LMP tumors play in the model might be simulated by skewing the growth distribution toward slow–growing tumors. In fact the upper growth limit of about 1 cm is also needed in the characterization of LMP tumors, as is the regression of these tumors. While they may represent an "indolent" end of the growth distribution of breast cancer they appear to be a distinct subpopulation of cancers statistically, with the entire population being a mixture of LMP type and lethal type tumors. An extensive discussion of the need for the LMP tumor type is presented elsewhere[7].

In the same way that LMP tumors represent an indolent end of the tumor growth spectrum, we found we needed to assume there were hyper aggressive tumors at the other end of the spectrum as well. Since the mid–1990s, the incidence rates for breast cancer diagnosed at the regional or distant stage have leveled off[8]. Our simulation model assumes that a small fraction of tumors have rapid metastatic spread. At initiation in the growth model, when the primary tumor diameter is assumed to be 0.2 cm, 1% of non–LMP tumors are assumed to have 4 positive nodes, and 2% has 5 or more nodes, i.e., these tumors are either regional or distant stage tumors right at initiation and cannot be detected in an earlier stage in the model.

REFERENCES:

1 Shwartz, M. “A Mathematical Model Used to Analyze Breast Cancer Screening Strategies.” in Operations Research 1978; 6: 26: 937-955

2 Shwartz, M. “An Analysis of the Benefits of Serial Screening for Breast Cancer Based Upon a Mathematical Model of the Disease.” in Cancer 1978; 4: 41: 1550-1564

3 Shwartz, M. “Estimates of lead time and length bias in a breast cancer screening program” in Cancer 1980; 46: 4: 844-51

4 Shwartz, M. “Validation and Use of a Mathematical Model to Estimate the Benefits of Screening Younger Women for Breast Cancer.” in Cancer Detection and Prevention 1981; 4: 595-601

5 Spratt, JS., Spratt, JA. “Chapter 21 Growth Rates. In: Donegan WL, editor” in Cancer of the Breast 2002; : 443-476

6 Shwartz, M. “Analysis of Preventive Screening to Detect Early Breast Cancer [Ph.D.]” 1975;

[7] Fryback, DG, Stout, NK, Rosenberg, MA, Trentham-Dietz, A, Remington, PL, Kuruchittham, V "The case for limited malignant potential breast tumors: Implications of the Wisconsin breast cancer epidemiology simulation model" 2004;

[8] Ries, LAG., Eisner, MP., Kosary, CL., Hankey, BF., Miller, BA., Clegg, L "SEER Cancer Statistics Review, 1975-2000" in National Cancer Institute 2003;

Survival And Mortality Component

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

Summary

This document describes the methods by which survival and mortality are determined in the model.

Overview

Breast cancer death occurs if the time to a woman's death from non–breast cancer causes is longer than her time to death from breast cancer. The time until death from non–breast cancer causes is chosen at the start of the simulation using mortality probabilities derived from birth cohort–specific, U.S. mortality tables published at UC Berkeley. From these probabilities we have removed breast cancer as a cause of death using breast cancer mortality rates in SEER data, and standard actuarial procedures.

Detail

Breast cancer death.

We assume that breast cancer death results only from disease which has progressed to the distant stage. (In our model, tumors continue to grow larger and involve an increasing number of lymph nodes even after clinical detection if they are not "cured." This is obviously a simplistic shadow model for more complicated biologic processes of recurrence and spread of treated breast cancers.) At the time a woman's cancer reaches the distant stage in the simulation, whether detected or not, she is assigned a time of death from breast cancer. Time until death is drawn from an empirical distribution based on survival times for women in the SEER registry diagnosed with distant stage breast cancer between 1975 and 1982 and who died of breast cancer[1]. We selected this time period to be prior to the advent of widespread mammography screening. Presumably tumors found in the distant stage after the advent of screening are from the increasingly aggressive end of the growth spectrum and we wished to use survival times to breast cancer death from a more representative sample of the spectrum. The 1975–82 time period also pre–dates modern treatment protocols for distant stage breast cancer. The empirical distribution (figure 3) was estimated by smoothing a life table of years of life from age at diagnosis created for these women from SEER data[1]. The median time to breast cancer death after arriving at the distant stage in our model is 1.95 years, and mean time 5.22 years.

so death from distant cancer could exceed 10 years. consequently, death from early stage cancer could also exceed 10 years because of the time for tumor growth from small to distant cancer.

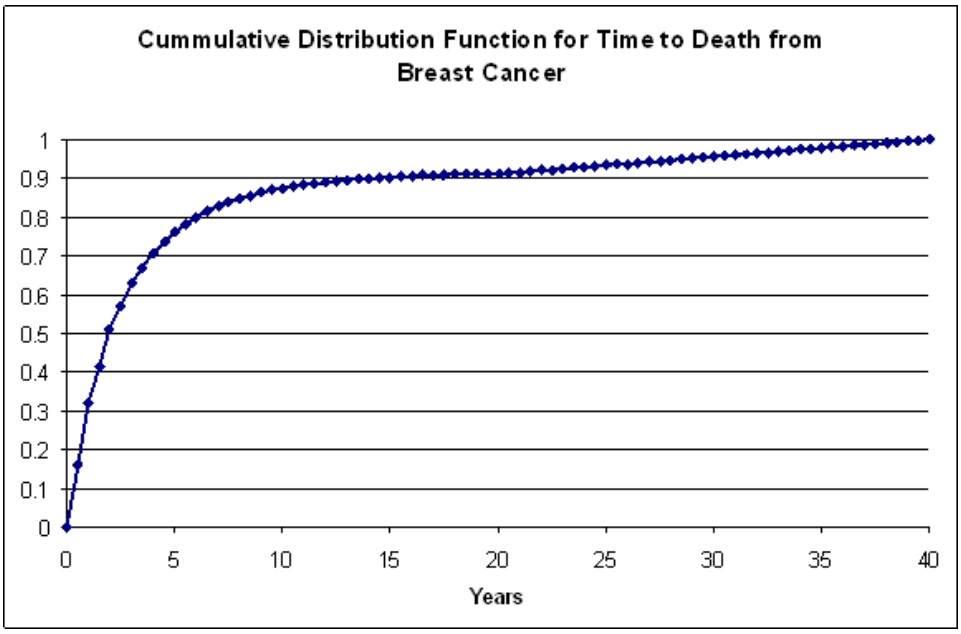


FIGURE 3. Empirical distribution of time from diagnosis to time of breast cancer death for women diagnosed with distant stage breast cancer in SEER between 1975 and 1982.

Mortality from non–breast cancer causes

We actuarially adjusted age–specific all–cause life tables of female mortality by birth cohort to develop mortality rates from non–breast cancer causes. The all–cause mortality by birth cohort from 1891 to 2000 is published by the Berkeley Human Mortality Data Base[2], while the age–specific breast cancer mortality is from National Center for Health Statistics, Centers for Disease Control and Prevention, vital statistics(see[3] for details). The resulting non–breast cancer mortality rates are common input to all the CISNET breast cancer collaboration models.

REFERENCES:

[1] Ries, LAG., Eisner, MP., Kosary, CL., Hankey, BF., Miller, BA., Clegg, L "SEER Cancer Statistics Review, 1975-2000" in National Cancer Institute 2003;

[2] Wilmoth, JR., Shkolnikov, V., [Directors] "Human Mortality Database. US Female Death Rates by Year of Birth" 2003;

[3] Rosenberg, MA "Competing risks" in JNCI Monographs 2004;

SCREENING COMPONENT

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

SUMMARY

This document describes the methods by which screening is simulated in the model.

OVERVIEW

We specify tumor detection probabilities, whether by mammography or clinical surfacing, as a function of the diameter of the tumor in centimeters, the age of the woman, and the calendar year being simulated. These probabilities were originally specified by a priori expert judgment, then refined by calibrating model outputs to observed surveillance incidence data by tumor stage, not size. The dependence within the model on size is due to an interaction in the model between the growth sub–model and the detection sub–model and the output of these as counts of detected simulated tumors graded into the four stages.

DETAIL

Clinical Surfacing

Clinical surfacing is expressed as an annual probability, and converted to a 6–month interval probability in each cycle of the simulation. We presume the probability of detecting an existing breast cancer in the absence of screening is low for small tumors and the probability should increase with diameter of the tumor. We also believe that women's self–detection of breast lumps has improved, particularly over the past decade, as a result of increasing public awareness of breast cancer. Under these assumptions, the annual probabilities for clinical surfacing of a tumor are shown in Table 1. These are the result of fitting and smoothing during calibration of the model subject to constraints that they are increasing in tumor diameter and over time, and that the probability at ≤ 0.3 cm is zero and at 8 cm is 1.0. Notice that the improvement in detection in the decade of the 1990s appears to be mostly for small to mid–size tumors.

TABLE 1. Annual Probabilities of Clinical Surfacing

	Diameter of the tumor (cm)*							
Year*	≤ 0.3	1.0	1.5	2.0	3.0	4.0	5.0	8.0
≤ 1990	0.0	0.06	0.07	0.15	0.40	0.70	0.80	1.0
2000	0.0	0.06	0.07	0.30	0.55	0.75	0.80	1.0

*Probabilities are linearly interpolated for years between 1990 and 2000, and between tumor diameters.

The one exception to this table is that if at the beginning of a simulated 6 month interval a woman with an as–yet undetected tumor is determined to die of breast cancer, we force detection at that time regardless of size of the primary tumor on the presumption that the cause of death would almost surely be diagnosed ante mortem from other signs and symptoms of metastatic disease even though the primary may be small. This proviso is necessitated because of the discrete time steps of the simulation and the fact that we do not model symptoms.

Sensitivity of Mammography Screening

Table 2 shows detection probabilities of a mammogram for an existing breast cancer. These are probabilities of detecting a tumor of a given diameter, in a woman of a given age, in a given calendar year on a given screening mammogram. (These are not rates or annual probabilities.) The values in Table 2 are the result of calibration to incidence data given constraints of monotonicity in tumor diameter, woman's age, and calendar year, and that the sensitivity to a tumor 5 cm is .99, and 8 cm in diameter is 1.0. Note the probability of detecting a tumor between 0.2 and 0.5 cm in diameter is constant in tumor diameter as detection of small tumors is mostly dependent on factors other than size such as calcification.

TABLE 2. Sensitivity of Mammography

	Diameter of the tumor (cm)**						
Year*	≤0.20	0.2–0.5	0.75	1.5	2.0	5.0	8.0
	Women aged						
≤1984	0.00	0.06	0.35	0.65	0.85	0.99	1.0
≥2000	0.00	0.20	0.60	0.65	0.85	0.99	1.0
	Women aged ≥50y**						
≤1984	0.00	0.1	0.45	0.65	0.85	0.99	1.0
≥2000	0.00	0.30	0.65	0.80	0.90	0.99	1.0

*Probabilities are linearly interpolated for years between 1984 and 2000, and between tumor diameters.

**Sensitivity of mammography is presumed to increase after menopause when breast tissues become more radiolucent with fatty replacement. Our current model uses an abrupt transition at age 50 to represent what is likely a more gradual process.

Diffusion of mammography utilization

Screening mammography before 1982 was sporadic and rare. Originally, we used mammography dissemination data from Wisconsin to specify the probability that a woman of a given age would receive a screening mammogram in a given calendar year from 1982–1991[2]. More recently we have incorporated an age–period model of annual screening probability[3]. The output of this model is random ages representing the age at first screen for a woman and the ages of subsequent screens up to age 100. The marginal frequencies match national data from the US for screening rates from 1975 to 2000. If a woman with an occult tumor is screened in the simulation, the probabilities in Table 2 are used to determine whether or not the tumor is detected at that screen. When a tumor is detected, the underlying state of the tumor according to the natural history model determines the stage of detection. Note that detection probabilities are not a function of stage, but of simulated tumor size and mammography sensitivity given the woman's age and the calendar year. Mammography detection rates by tumor stage are emergent properties of the simulation.

REFERENCES:

[1] Lantz, PM., Weier, A., Remington, PL. "Mammography use in Wisconsin: 1987-1992" 1993; 92: 10: 580-2

[2] Lantz, P., Bunge, M., Remington, PL. "Trends in mammography in Wisconsin" in Wis Med J 1990; 89: 6: 281-2

[3] Cronin, K "The dissemination of mammography in the US" in JNCI in press

TREATMENT COMPONENT

SUMMARY

This document describes the treatment of cancer and how its effectiveness is modeled.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

OVERVIEW

Treatment and treatment effectiveness

For simplicity we model treatment as a cure/no–cure process. When a breast cancer is detected, regardless of mode of detection, we assume it is treated. The result of simulated treatment is either "cure" with total arrest of progression at that time and hence no possibility of progressing to a breast cancer death, or the result is "no cure" in which case the tumor continues to progress as if it were undetected and the woman may die of breast cancer, competing causes, or achieve age 100 depending on her individual circumstances. Continued simulated growth in this case is used to mark a time–line for progression and size per se is not meant to be biologically representative.

This method for modeling treatment outcomes models survival given treatment as a mixture of two survival curves conditioned on the woman's age and the tumor characteristics at the time of detection. One curve is the survival curve of women that age without breast cancer; the other curve is women that age and with tumors of the same stage but untreated, which is a function of the underlying natural history model. The mixture probability to be fit is termed the "cure" fraction. Although this method gives up fitting the shape of the survival curve given stage, and treatment, it avoids having to make direct assumptions about survival given mode of detection, treatment, etc., which we wish to be emergent properties of the model rather than inputs.

The treatment submodel has three logical parts. First, we specified treatment effectiveness—cure fractions—in the pre–tamoxifen pre–adjuvant polychemotherapy era for tumors treated at different stages with a standard, baseline therapy. These "baseline" cure probabilities metaphorically represent overall mastectomy with or without radiation as was common in the pre–1975 era. Second, we specified the relative improvement in survival with the various combinations of adjuvant therapies added to the baseline therapy. Third, we specified the diffusion of these adjuvant treatments over time as a function of characteristics of the woman and the stage of tumor at diagnosis.

DETAIL

The model assumes that all women receive baseline treatment consisting of standard therapy such as surgery and/or radiation and that the effectiveness of this baseline treatment has not changed over time. In addition women may receive one of five modes of adjuvant therapy. The different modes of adjuvant therapy, which are determinants of the effectiveness of treatment, are chemotherapy only, tamoxifen only for two years, tamoxifen only for five years, chemotherapy in combination with Tamoxifen for two years or chemotherapy in combination with tamoxifen for five years.

Women with breast cancer detected in the localized or regional stages are assigned a

mode of adjuvant treatment based on the calendar year, her current age, tumor size/stage and revealed estrogen receptor (ER) status. Tumors diagnosed in the in situ or distant stages are not assigned adjuvant therapy. The likelihood of each mode of treatment is based on observed use of the treatment. Data describing the likelihood of treatment were provided by NCI based on analysis of data from the Patterns of Care study as well as combined data from numerous cancer registries[1]. These data show increasing use of chemotherapy for large localized tumors and regional tumors over time and increasing use of a 5–year course of tamoxifen for ER positive tumors. All CISNET collaboration models use these data as input.

Revealed ER status is modeled as a function of the true ER status of the tumor as well as the calendar year. True ER status is based on the age of the woman at the time of tumor onset (Table 3)[2]. In the simulation, the treatment probabilities are determined in part by whether the ER status is known. We used SEER data from 1990 forward (the first year this was recorded in the SEER data) to estimate the proportion of tumors with ER status determined; probabilities before this time were based on assessment of a local expert oncologist–breast cancer researcher (Table 4). The treatment administered is in part determined by whether the ER status is known and if so whether it is positive or negative. The treatment effectiveness is determined as a function of the 'true' underlying ER status of the tumor and the treatment given.

TABLE 3. Probability that a tumor is ER Positive by Age

Age	Pr(ER+)
	0.6
45–54	0.65
55–64	0.74
65–74	0.77
75+	0.83

TABLE 4. Likelihood that the True ER status of a tumor will be known

Year	Pr(ER Status Known)
	0.1
1975–1979	0.2
1980–1984	0.5
1985–1989	0.63
1990	0.68
> 1991	0.69

No data directly describe the effectiveness of treatment as implemented in the simulation model. Data from randomized clinical trials are usually reported as relative survival gains or decreases in annual mortality odds. Thus the derivation of the likelihood that a particular mode of treatment would be effective was a several step process.

The estimation of treatment effectiveness started with baseline effectiveness for standard treatment by stage at detection in the absence of adjuvant therapy. These likelihoods were estimated by expert opinion and were refined during the calibration process. Observed survival was approximated by a mixture of the survival given treatment is completely effective and the survival given treatment is completely ineffective where the mixture proportion is the probability that treatment is effective.

Using this relationship, 10–year survival probabilities (considering all causes of death) were estimated from the simulation model under conditions that all treatment is completely effective and again assuming all treatment is completely ineffective, and modified 10–year survival probabilities were computed by age group and stage. Annualized mortality rates were calculated assuming that the annual mortality rate was constant over the 10–year interval.

The annual odds of mortality for adjuvant treatment were calculated by applying trial results about the performance of treatment reported as the annual reduction in the odds of mortality to the computed annual odds of mortality for baseline treatment. These annual odds were further adjusted based on the length of the course of treatment for tamoxifen. These adjustments to the baseline odds of mortality assume that the effects of chemotherapy and tamoxifen are independent and that tamoxifen is effective for women with tumors that are ER positive[3](pp1–15,71–85).

These adjusted annual odds of mortality by the current age of the woman, tumor stage (localized and regional only), true ER status, and mode of adjuvant treatment were converted back to the annualized mortality rates. Again assuming constant annual all–cause mortality, implied 10–year survival probabilities were computed. Adjusted mixture proportions that treatment is curative were recomputed based on the adjusted 10–year survival probabilities and the two simulated survival probabilities under conditions that all treatment is either completely effective or completely ineffective. These adjusted cure fractions are reported in table 5.

TABLE 5. Cure probabilities used in the model.

Stage	Age	ER status	No Adj Tx	Tam 2yr	Tam 5yr	Chemo only	Tam 2yr +chemo	Tam 5yr +chemo
In Situ		–	0.990	0.990	0.990	0.990	0.990	0.990
In Situ	50–59	–	0.990	0.990	0.990	0.990	0.990	0.990
In Situ	60–69	–	0.990	0.990	0.990	0.990	0.990	0.990
In Situ	70+	–	0.990	0.990	0.990	0.990	0.990	0.990
In Situ		+	0.990	0.990	0.990	0.990	0.990	0.990
In Situ	50–59	+	0.990	0.990	0.990	0.990	0.990	0.990
In Situ	60–69	+	0.990	0.990	0.990	0.990	0.990	0.990
In Situ	70+	+	0.990	0.990	0.990	0.990	0.990	0.990
Localized		–	0.820	0.820	0.820	0.882	0.882	0.882
Localized	50–59	–	0.820	0.820	0.820	0.867	0.867	0.867
Localized	60–69	–	0.820	0.820	0.820	0.864	0.864	0.864
Localized	70+	–	0.820	0.820	0.820	0.957	0.957	0.957
Localized		+	0.820	0.861	0.884	0.882	0.913	0.931
Localized	50–59	+	0.820	0.881	0.916	0.867	0.921	0.952
Localized	60–69	+	0.820	0.921	0.980	0.864	0.959	1.000
Localized	70+	+	0.820	1.000	1.000	0.957	1.000	1.000
Regional		–	0.400	0.400	0.400	0.527	0.527	0.527
Regional	50–59	–	0.400	0.400	0.400	0.470	0.470	0.470
Regional	60–69	–	0.400	0.400	0.400	0.446	0.446	0.446
Regional	70+	–	0.200	0.200	0.200	0.268	0.268	0.268
Regional		+	0.400	0.482	0.532	0.527	0.599	0.642
Regional	50–59	+	0.400	0.492	0.549	0.470	0.558	0.611
Regional	60–69	+	0.400	0.509	0.578	0.446	0.554	0.621
Regional	70+	+	0.200	0.370	0.493	0.268	0.448	0.575
Distant		–	0.050	0.050	0.050	0.050	0.050	0.050
Distant	50–59	–	0.050	0.050	0.050	0.050	0.050	0.050
Distant	60–69	–	0.050	0.050	0.050	0.050	0.050	0.050
Distant	70+	–	0.025	0.025	0.025	0.025	0.025	0.025
Distant		+	0.050	0.050	0.050	0.050	0.050	0.050
Distant	50–59	+	0.050	0.050	0.050	0.050	0.050	0.050
Distant	60–69	+	0.050	0.050	0.050	0.050	0.050	0.050
Distant	70+	+	0.025	0.025	0.025	0.025	0.025	0.025

The baseline cure fractions are shown in the "No Adj Tx" column. Initially we held these constant within stage regardless of the age of the woman. However breast cancer mortality among older women was consistently too low. Accordingly, we reduced the cure fractions of the two more advanced stages for women aged 70 years and older. This is consistent with observations that older women appear to be less aggressively treated than younger women [6].

Treatment of LMP tumors is assumed to be 100% curative since these tumors are, by our definition, not lethal.

REFERENCES:

[1] Mariotto, A., Feuer, EJ., Harlan, LC., Wun, L-M., Johnson, K., Abrams, J. "The dissemination of adjuvant therapy for breast cancer in the US: 1975-1997" in The National Cancer Institute 2002;

[2] Ashba, J., Traish, AM. "Estrogen and progesterone receptor concentrations and prevalence of tumor hormonal phenotypes in older breast cancer patients" in Cancer Detect Prev 1999; 23: 3: 238-44

[3] EBCTCG (Early Breast Cancer Trialists' Collaborative Group) "Systemic treatment of early breast cancer by hormonal, cytotoxic, or immune therapy. 133 randomised trials involving 31,000 recurrences and 24,000 deaths among 75,000 women" in Lancet 1992; 339: 8785: 71-85

[4] Newcomb, PA., Carbone, PP. "Cancer treatment and age: patient perspectives" in J Natl Cancer Inst 1993; 85: 19: 1580-4

[5] Kemeny, MM., Peterson, BL., Kornblith, AB., Muss, HB., Wheeler, J., Levine, E. "Barriers to clinical trial participation by older women with breast cancer" in J Clin Oncol 2003; 21: 12: 2268-75

[6] Wyld, L., Reed, MW. "The need for targeted research into breast cancer in the elderly" in Br J Surg 2003; 90: 4: 388-99

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

MODEL CALIBRATION PROCEDURES

The inputs to the simulation model are numerous. We took the following inputs as given:

1. Estimated breast cancer incidence in the absence of screening;
2. The patterns of mammography dissemination over time;
3. The patterns treatment dissemination over time;
4. Dissemination of ER status determination over time, and the relationship between the woman's age and ER status of her tumor;
5. The relative effectiveness of various modes of adjuvant therapy.

Other inputs to the model were calibrated to make the model outputs conform to surveillance data from 1975–2000. These calibrated inputs include the tumor detection probabilities in Tables 1 and 2 in Screening Component , the baseline treatment effectiveness probabilities in the no adjuvant therapy ("No Adj Tx") column of Table 5 in Treatment Component , and the variables listed in Table 6. Table 6 is entitled "core parameters" because these 10 turned out to be the parameters about which the least is known and which apparently control model output with the most sensitivity.

TABLE 6. Calibrated core input parameters.

Parameter	Use in the model	[Sampled Range] (increment size for discrete sampling)		Final value
		Wide ranges*	Focused ranges*	
1. LMP Fraction	Proportion of all biologically incident tumors assumed to be Limited Malignant Potential (LMP)	[0% – 55%] (1%)	[30% – 50%] (1%)	42%
2. Max LMP Size	LMP tumors assumed to grow no larger than this diameter (cm) (this variable was fixed as it is entangled with In Situ Boundary, the Gompertz growth parameters, and LMP Dwell Time)	1 cm	1 cm	1 cm
3. LMP Dwell Time	Maximum sojourn time (years) for LMP tumor after reaching Max LMP Size; after this time without discovery, the LMP tumor disappears next simulation cycle.	[1–3] (0.5)	[1.5–2.5] (0.5)	2 y
4. In Situ Boundary	The diameter (cm) below which the tumor is classified as in situ stage in the simulation if there are no associated positive lymph nodes	[0.75 – 1.0] (0.01)	[0.85 – 0.99] (0.01)	0.95 cm
5. Onset Proportion	Ratio of assumed age–specific biologic onset rate divided by age–specific incidence rate (the latter specified by age–period–cohort model estimated in absence of screening – see text).	0.85 – 1.2 (0.01)	0.8 – 1.0 (0.01)	0.9
6. Onset Lag	Time interval (years) between year of index onset rate and incidence rate used in Onset Proportion. This is to "fill the pipeline" with biologically onset tumors which will be discovered at a given incidence rate some years later. Because the cycle time of the model in 0.5 years, this was taken to be step size.	[1–8] (0.5)	[1.5–4] (0.5)	3 y
7. Percent 4 nodes	Percent of biologically onset, non–LMP tumors which are assigned 4 positive lymph nodes at onset. (This places these tumors at the upper limits of simulated regional tumors, which are presumed to have 1–4 positive nodes.)	[0 – 5%] (1%)	[0 – 1%] (1%)	1%
8. Percent 5 nodes	Percent of biologically onset, non–LMP tumors which are assigned 5 positive lymph nodes at onset. (This simulates these tumors in the distant stage from their initiation in the model.)	[0 – 5%] (1%)	[2 – 4%] (1%)	2%
9. Mean Gamma	The Gompertz growth rate is assumed to have a gamma distribution across all onset tumors. This parameter is the mean of this gamma distribution (see text).	[0.01 – 0.2] (0.01)	[0.08 – 0.18] (0.01)	0.12
10. Var Gamma	The variance of the gamma distribution of Gompertz growth rates	[0.006 – 0.1] (0.001)	[0.01 – 0.05] (0.001)	0.012

*The wide range was used for initial sampling of parameter space; the focused range was used to focus sampling in a sub–region of parameter space closer to where the best solution was thought to be. See text for details.

Our objective in constructing this simulation model of breast cancer epidemiology was to have a computer model whose output mimics the contents of a cancer surveillance registry. Thus, we calibrated to breast cancer incidence data for 10 5–year age groups from 30–34 years to 75–79 years across the 26 calendar years 1975–2000, and 4 historical stages at detection (in situ, local, regional, distant) appearing in the Wisconsin Cancer

Reporting System and in the national SEER data base, exclusive of Iowa[1]. Jointly, these comprise a surface of (10x26x4) = 1040 points at which model output should approximate data from the two (non–equal) surveillance data sets. Manipulating the inputs in order to maximize fit to this surface is a complex optimization problem with no unique objective function and no closed form solution. Accordingly, we approached calibration heuristically in several stages.

First, to reduce the dimensionality of the problem of comparing model output to known registry data we adjusted incidence rates to the year 2000 standard age population (for ages 30–79 years) and subjectively judged the fit of the model by comparing the 4 stage–specific incidence rate curves and the age–adjusted breast cancer mortality curve over the single years from 1975 to 2000. We concentrated not only on minimizing error in predicting rates (demanding approximately the same relative fit for each curve despite the large variation in absolute rates by stage), but also on the qualitative shapes of these curves compared to WCRS and SEER data.

We began calibration with a set of initial values for input parameters in Tables 1, 2 in Screening Component , and 6. These tables jointly summarize all calibrated inputs to the simulation model and are the parameters controlling incidence rates. Initial values were based on literature searches and expert judgments. Factorial experiments were devised to examine the effects of varying combinations of parameters in reasonable ranges. The output of each combination of parameters was represented as a "calibration plot" such as shown in the first five panels of Figure 5. We assessed these plots "by eye", judging how closely model output fit WCRS and SEER data. A typical factorial experiment varying 3 parameters in a 2 x 2 x 4 design took on the order of 3 days to run and analyze. This process led to modifications of the model structure and inputs. Once model output appeared to fit incidence curves well, the 4 baseline cure fractions (one for each stage) in the "No Adj Tx" column of Table 5 in Treatment Component were adjusted to bring age–adjusted breast cancer mortality rates predicted by the model to the levels of the observed data.

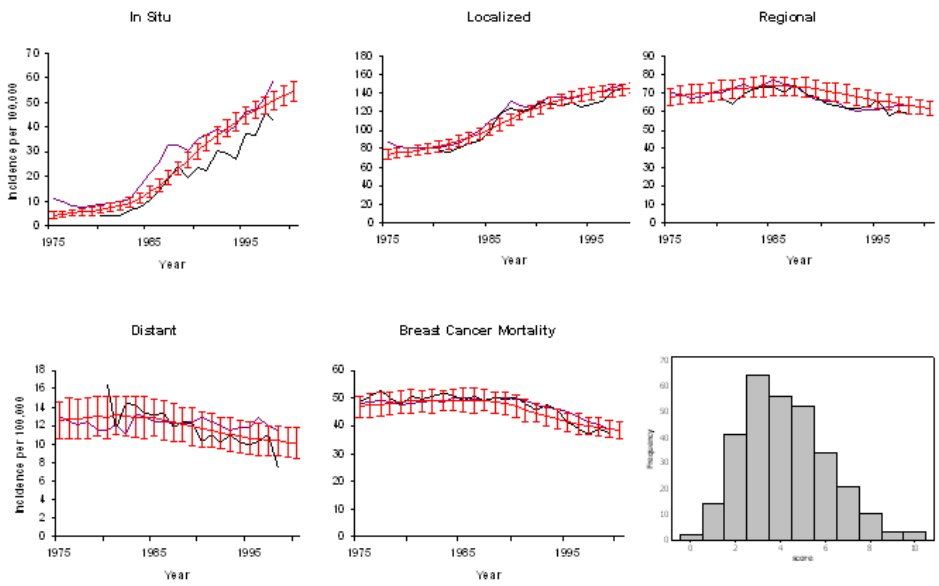


FIGURE 5. This is a calibration plot for the final model inputs. Four panels of the figure show incidence rates per 100,000 women (adjusted to population aged 30–79, US standard population for year 2000) for each stage, in situ, local, regional, or distant, and one panel shows breast cancer mortality rates. The scales have been adjusted to provide comparable visual assessment of fit across the panels; the absolute rates vary considerably across the panels. The two dark lines are data from SEER (excluding data from Iowa) and from the Wisconsin Cancer Registry. Iowa has been held out for later cross–validation as a state similar geographically and demographically to Wisconsin. Vertical bars indicate plus or minus 2 standard deviations for each annual point value across 300 replications using the fixed input parameters. The final panel is a histogram of the evaluation scores (as described in the text) for the 300 replications with the model inputs fixed.

For the next phase of calibration, we performed parameter sampling experiments. The purposes of parameter sampling were to explore whether a better fitting parameter combination(s) may be found than the one at which we arrived heuristically, to assess the likelihood of parameter combinations yielding good fitting solutions, and to be able to make uncertainty statements about our fitted parameter values. In this phase we specified wide ranges around the parameter values resulting from the first phase of calibration and sampled trial parameter values from uniform distributions in these ranges.

It was necessary to develop auxiliary computational tools to carry out the sampling experiments. The major tool was a function to automatically screen the model's "fit" resulting from a particular set of input parameters. Each unique combination of input parameters is a vector in the model parameter space (a 10–dimensional space if we consider only parameters in Table 6, or a higher dimension space if parameter values from Tables 1 and 2 in Screening Component are also considered). Let $\vec{v}$ be a vector of input parameter values. We wished to evaluate a set of parameter vectors, $\vec{V} = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_N\}$, where N is a large number (on the order of tens or hundreds of thousands), sampled uniformly in a pre–specified hypercube of parameter space (ranges for sampling are shown in Table 6). To do this we evaluated the fit of the

simulation output to observed surveillance data using each $\vec{v}_i$ as input. Because there were many parameter vectors to evaluate, a method was required to automatically evaluate "fit", the evaluation we had been doing by "eye" in phase 1. Denote the fit of the simulation at $\vec{v}_i$ by the function, $f(\vec{v}_i)$. The function f must evaluate the closeness of the simulation output to approximately 1040 non–independent points defined by 26 years of surveillance incidence data. Further, given $\vec{v}_i$, the output of the simulation is complex stochastically. Each execution of the simulation, a "replication", mimics detailed surveillance data from 1975–2000. We did not wish to fit observed data exactly because observed data are only one "sample" from a stochastic real world, but rather to find underlying parameters leading to behavior stochastically similar to observed surveillance data.

Accordingly, we were reluctant to define f solely in terms of an error function such as a sum of squared deviations across the 1040 points. Such an error function de–emphasizes incidence rates with small numerical values thus obscuring the importance of the trends from low to high rates in early stage cancers. Using a squared error penalty function resulted in representing age–adjusted incidence curves which do not "flatten" enough in later years compared to observed surveillance data. We instead elected to define f using acceptance envelopes around the four age–adjusted stage–specific incidence curves shown in Figure 6 and to count number of points at which the simulation output fell outside the envelopes. This equalizes importance of matching general shapes of the observed data curves regardless of the absolute numerical size of the rates in a given year. Envelope widths were defined using replication–to–replication variation in the simulation output (the variation is largely independent of the specific parameter values) such as shown in Figure 5. We set envelope widths to screen for potentially "good" input parameter combinations. That is, given an input parameter vector $\vec{v}_i$ whose *mean* output across many replications would fall near the centers of the envelopes, we expected the 4 stage–specific output incidence curves from a _single _replication run would be unlikely to exceed the envelope often and a vector whose mean output was near the edges of the envelope was likely to generate data exceeding the envelope in many places on any given replication. We defined a scoring function, $f(\vec{v}_i)$, as the sum of the number of points at which the 4 curves from one replication executed using the input parameters $\vec{v}_i$ fell outside the envelopes. With 26 years from 1975–2000 and 4 curves, there is a total of 26x4=104 points at which the output from one replication could exceed the envelopes, hence the range of the evaluation function was $0 \leq \vec{v}_i \leq 104$.

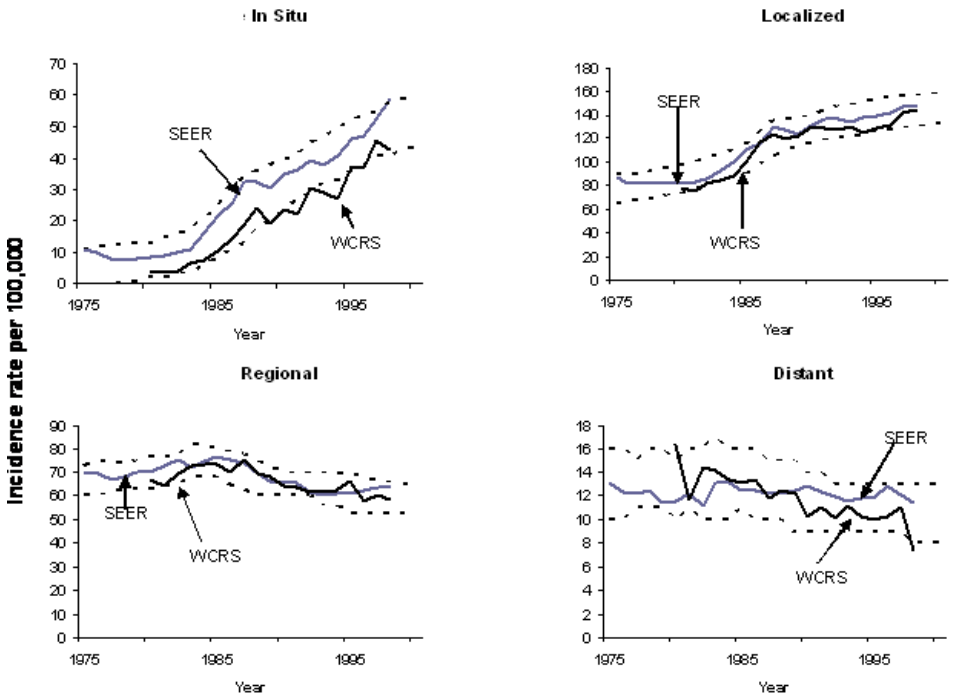


FIGURE 6. The dotted boundaries show the acceptance envelopes used for parameter sampling experiments. Incidence rates are age–adjusted to the U.S. population aged 30–79 in year 2000. The fit of a proposed input parameter vector was scored by counting excursions of one replication's output of the simulation beyond the envelope bounds at any of the 26x4=104 possible annual points from 1975 to 2000. The envelope for in situ cancer is biased toward the SEER data since those represent a much larger number of cases. Generally envelope widths vary with wider envelopes being associated with lower, and hence more variable, rates.

We used the CONDOR environment for simultaneous execution of the simulation by a large pool of networked computers. With approximately 120 computers in the CONDOR pool to which we had access, we could evaluate approximately 1000–1500 sampled $\vec{v}_i$s per day.

Three "experiments" were conducted to calibrate and evaluate the model. Experiment 1 was designed to sample the parameter space broadly across all variables from Tables 1, 2 in Screening Component , and 6. The curve cut–points in Tables 1 and 2 in Screening Component were constrained so that probabilities of detection increased with size of tumor. Table 6 shows ranges sampled. All initial sampling was from uniform distributions (or equal probability distributions for variables sampled discretely). After approximately 57,000 $\vec{v}_i$s were evaluated with the result being no score under 10 (i.e., $f(\vec{v}_i) \geq 0$ for all $\vec{v}_i$s evaluated), we constrained the LMP fraction to be under 10% for the next 15,000 samples to ensure dense sampling near zero for this potentially contentious variable. A total of 72,335 $\vec{v}_i$s were evaluated in this fashion with none scoring less than 10 and only 10 scoring under 16. For the next 30,000 samples we constrained LMP fraction to range from 30–55%, closer to our initial solution of 42%. Of the 30,188 uniformly sampled $\vec{v}_i$s a total of 363 yielded scores of 10 or less, 91 with a score of 5 or less, 15 with a score of 3, 4 scoring a 2, 1 scoring 1, and none scoring 0.

From Experiment 1, it appeared $\vec{v}_i$s leading to acceptably calibrated simulated incidence curves are rare in the parameter space defined by plausible marginal ranges of inputs. By sampling, we improved only slightly on our initial "by eye" solution, making small changes in parameters to match a $\vec{v}_i$ with score of 2. We did not select the vector with a score of 1 because of the 4 vectors with score of 2, one of them had slightly better age–specific results than the one scoring 1 point better on age–adjusted evaluation. We also conclude that models with LMP fraction less than 30%, and especially those with LMP fraction near 0%, can be excluded for lack of fit to observed incidence, and this was in spite of testing a very wide range of other inputs to compensate.

Experiment 2 asked whether $\vec{v}_i$ which were identified with low (good) scores in Experiment 1 generally had good solutions also existing in neighborhoods near them, and if those identified with higher (poorer) scores generally had poor solutions in their neighborhoods – i.e., does our scoring function f appear to separate neighborhoods well. We picked four of the $\vec{v}_i$s with scores in the 0–3 range ("good"), and four $\vec{v}_i$s with scores in the 11–15 range ("poor"), and sampled and scored an additional 500 $\vec{v}_i$ in the neighborhood of each of these eight vectors. A "neighborhood" was defined by freezing the detection probabilities for each vector (parameters from Tables 1 and 2 in Screening Component), and sampling remaining parameters (from Table 6) within a range of ±5% of the original vector's values. Examining the distributions of scores from samples around these vectors we generally found about 30% of samples in the neighborhood of a "good" vector to have scores equal to or less than 10, and far fewer than 1% of samples around "poor" vectors to score at 10 or less. We concluded that using our scoring function, *f*, was a reasonable method to automate exploring the parameter space.

Although broadly speaking our acceptable solutions—i.e., the set of $\vec{v}_i$ with scores less than 10—appear somewhat concentrated in the parameter space, they do not form a connected set. We believe there may be clusters of mutually non–identifiable parameters, hence we see a solution set with some marginal distributions being bi–modal or multimodal. Figure 7 presents marginal plots of selected parameters in the set of $\vec{v}_i$ with scores less than 10. These may be interpreted roughly as posterior marginal distributions for the input parameter values.

Experiment 3 ran 300 replications using our "best fit" solution, denoted $\vec{v}_i^*$. These replications, when scored using **f** yielded a score range of 0–10; the distribution of scores is shown in the lower right hand panel of Figure 5. The calibration plot in Figure 5 was drawn using these replications of $\vec{v}_i^*$. Tables 1, 2 in Screening Component , and 6 present the parameter values for $\vec{v}_i^*$. Note that $\vec{v}_i^*$ is not simply a joint mean of the marginal distributions in Figure 7 because we have selected the this vector based on having a score at the lower end of the range deemed "acceptable" for purposes of producing figure 7.

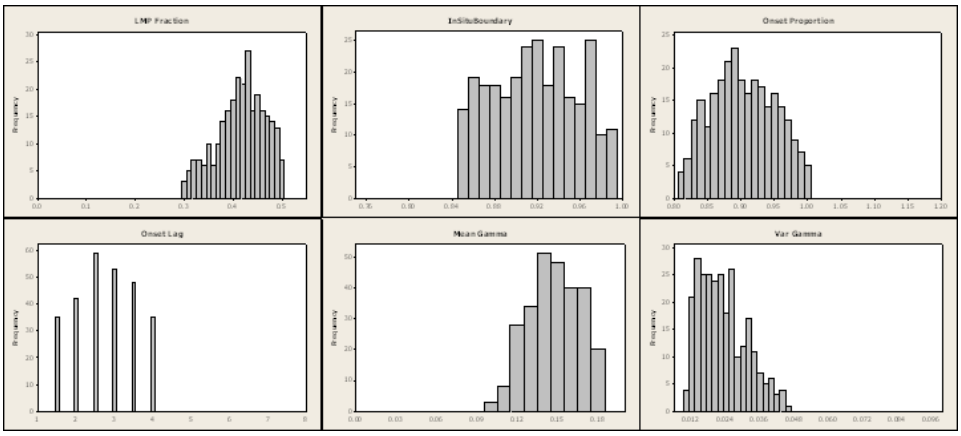


FIGURE 7. Marginal posterior distributions for six of the input parameters in Table 6. These histograms show values of the named parameters for input parameter vectors with scores of 10 or less in the sampling experiment. The ranges shown correspond to the ranges sampled for these parameters. Note that the parameters are not independent.

Additional outputs of the model have been compared to data where they exist (results not shown). For example the survival curves implied by the baseline cure probabilities are a good match to long–term survival published by Fisher, et al, in a follow–up study of modified mastectomy and lumpectomy treatments[3]. The stage distribution at detection as a function of tumor size is similar to SEER data[1]. The age–specific prevalence of breast cancer in the year 1975 appears to match prevalence rates derived by NCI statisticians from SEER data (data not shown).

REFERENCES:

[1] Ries, LAG., Eisner, MP., Kosary, CL., Hankey, BF., Miller, BA., Clegg, L "SEER Cancer Statistics Review, 1975-2000" in National Cancer Institute 2003;

[2] Fisher, B., Jeong, JH., Anderson, S., Bryant, J., Fisher, ER., Wolmark, N. "Twenty-five-year follow-up of a randomized trial comparing radical mastectomy, total mastectomy, and total mastectomy followed by irradiation" in N Engl J Med 2002; 347: 8: 567-75

[3] Fisher, B., Anderson, S., Bryant, J., Margolese, RG., Deutsch, M., Fisher, ER. "Twenty-year follow-up of a randomized trial comparing total mastectomy, lumpectomy, and lumpectomy plus irradiation for the treatment of invasive breast cancer" in N Engl J Med 2002; 347: 16: 1233-41

POPULATION COMPONENT

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

SUMMARY

This document describes the method by which population is simulated in the model.

OVERVIEW

The model was designed to produce counts of incident tumors and breast cancer deaths over time to reproduce corresponding counts in the WCRS from 1980–1998, this being the interval for which WCRS is available and stable. Hence we scaled the simulation to replicate the life histories of a population of women from whom these counts would arise. Starting the model in 1950, we simulate a number of women equal to each single–year age cohort of women in the Wisconsin population aged 20–99. In each year 1951–1999 we add to the simulation model the number of women aged 20y in that year. (We chose age 20 under the assumption that breast cancer is so rare in younger women that we can ignore its occurrence before that age.)

DETAIL

The total number of women simulated is approximately 2.95 million. Each complete simulation of these 2.95 million women is one replication of the simulation and results in data equivalent to the breast cancer cases in the WCRS from 1978–1998 plus the years 1950–1977, and 1999–2000. One replication on a desktop computer with a Pentium 4 processor and 384 Mb of RAM requires approximately 10 minutes. We submit runs consisting of multiple replications, typically 10–50, using a large set of networked computers utilizing the CONDOR sharing software[3].

We started the model in 1950 assuming no breast cancer is present in any woman living at that time. The breast cancer onset and progression submodels described above are invoked in 6 month cycles from simulated year 1950 to 1975 as "burn–in" for our model under the assumption that the prevalence has stabilized after the 25 year run–up.

For comparison of output to other CISNET collaborators, when we computed age–adjusted rates, we adjusted results to the US standard population (male and female) aged 30–79 in the year 2000.

REFERENCES:

[1] Livny, M, Basney, J, Raman, R, Tannenbaum, T “Mechanisms for high throughput computing” in SPEEDUP Journal 1997; 11:

[2] Basney, J., Livny, M. “Chapter 5: Deploying a high throughput computing cluster” in In: Buyya R, editor

[3] Thain, D, Tannenbaum, T., Livny, M. “Condor and the grid. In: Berman F, Hey AJG, Fox G, editors” in Grid Computing: Making the Global Infrastructure a Reality

Output Description

When a woman meets the exit criteria for the simulation (i.e., she dies, she reaches age 100, or the ending calendar year of the simulation is reached) her data are tallied in a series of event counters.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

The primary counters accumulate counts relevant to numerators and denominators of output quantities of interest. Example counters are:

1. indexed by age and calendar year, counts instances of being alive and a given age in a given calendar year. This count is used as a denominator.
2. similarly indexed, counts instances of being alive and breast–cancer free at a given age in a given calendar year. This too is a potential denominator.
3. indexed by historical stage, age, and calendar year, and mode of detection (mammogram or clinical surfacing), counts instances of incidence breast cancer.
4. indexed by historical stage, age, and calendar year, and nature of the tumor (agressive breast cancer or breast cancer of limited malignant potential) tallies undetected breast cancers

Similar ad hoc counters can be implemented for special runs of the simulation.

The counts in these arrays are converted to rates per 100,000 at the end of 1 simulation replication of the entire population. These rates are then stored for N complete replications and the median, mean, and standard deviations are computed across the N replications.

These single–year outputs can be combined to show age–adjusted outputs across time (see Results Overview)

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

KEY REFERENCES

Abrams, K., Ashby, D. , et al. (1996) A Bayesian Approach to Weibull Survival Models – Application to a Cancer Clinical Trial. in *Lifetime Data Analysis* 2, p 159–174

Ackerman, E. (1994) Simulation of Micropopulations in Epidemiology: Tutorial 3. Simulation Model Evaluation Methods. in *International Journal of Bio–Medical Computing*:37, p 195–204

Ackerman, E. (1994) Simulation of Micropopulations in Epidemiology: Tutorial 1. Simulation: An Introduction. in *International Journal of Bio–Medical Computing*:36, p 229–238

Ackerman, E. (1994) Simulation of Micropopulations in Epidemiology: Tutorial 2. Analytic Forms of Event Probabilities. in *International Journal of Bio–Medical Computing* 2:37, p 139–149

Ackerman, E. (1994) Simulation of Micropopulations in Epidemiology: Tutorial 4. Evaluations of Simulation Models. in *International Journal of Bio–Medical Computing* 2:39, p 219–229

Albertsen, PC., Hanley, JA., Gleason, DF., Barry, MJ. (1998) Competing risk analysis of men aged 55 to 74 years at diagnosis managed conservatively for clinically localized prostate cancer in *Jama* 280:11, p 975–80

Armstrong, K., Eisen, A., et al. (2000) Assessing the Risk of Breast Cancer. in *New England Journal of Medicine* 8:342, p 564–571

Ashba, J., Traish, AM. (1999) Estrogen and progesterone receptor concentrations and prevalence of tumor hormonal phenotypes in older breast cancer patients in *Cancer Detect Prev* 23:3, p 238–44

Basney, J., Livny, M. Chapter 5: Deploying a high throughput computing cluster in *In: Buyya R, editor*,

Bernstein, L. , Ross, R. K. (1993) Endogenous Hormones and Breast Cancer Risk. in *Epidemiologic Reviews* 1:15, p 48–65

Black, W. C., Nease, R. F. J. , et al. (1996) Determining Transition Probabilities from Mortality Rates and Autopsy Findings. in *Medical Decision Making* 1:17, p 87–93

Black, W. C., Welch, H. G. (1997) Screening for Disease. in *American Journal of Radiology*:168, p 3–11

Blumenson, L. E. Bross, I. D. J. (1969) Mathematical Analysis of the Growth and Spread of Breast Cancer. in *Biometrics*, p 95–109

Brooks, J. M., Chrischilles, E., et al. (2000) Information Gained From Linking SEER Cancer Registry Data to State–Level Hospital Discharge Abstracts. in *Medical Care* 11:38, p 1131–1140

Brown, M. L., Fintor, L. (1993) Cost–Effectiveness of Breast Cancer Screening: Preliminary Results of a Systematic Review of the Literature. in *Breast Cancer Research and Treatment*:25, p 113–118

Buchanan, J. B., Spratt, J. S. , et al. (1983) Tumor Growth, Doubling Times, and the Inability of the Radiologist to Diagnose Certain Cancers. in *Radiologic Clinics of North America* 1:21, p 115–126

Bureau of Health Information , D. O. H. C. F. (2000) Wisconsin Cancer Incidence and Mortality in *Wisconsin Department of Health and Family Services*, p 1–79

Bureau of Health Information, Division of Health Care Financing, Wisconsin Department of Health and Family Services (2002) Wisconsin Cancer Incidence and Mortality, 1999,

Burke, W., Olsen, A. H. , et al (2001) Misleading Presentation of Breast Cancer in Popular Magazines. in *Effective Clinical Practice* 2:4, p 58–64

Cameron, DA., Ritchie, AA., Miller, WR. (2001) The relative importance of proliferation and cell death in breast cancer growth and response to tamoxifen in *Eur J Cancer* 37:12, p 1545–53

Chang, P. (1993) A Simulation Study of Breast Cancer Epidemiology and Detection since 1982: The Case for Limited Malignant Potential Lesions [Ph.D.],

Clemons, M. , Goss, P. (2001) Estrogen and the Risk of Breast Cancer in *New England Journal of Medicine* 4:344, p 279–285

Colditz, G. A., Rosner, B. (2000) Cumulative Risk of Breast Cancer to Age 70 Years According to Risk Factor Status: Data From the Nurses' Health Study. in *American Journal of Epidemiology* 10:152, p 950–964

Cronin, K The dissemination of mammography in the US in *JNCI in press*,

Cronin, K. A., Legler, J. M. , et al. (1998) Assessing Uncertainty in Microsimulation Modelling With Applications to Cancer Screening Interventions. in *Statistics in Medicine*:17, p 2509–2523

Crowley, TJ (2000) Causes of Climate Change Over the Past 1000 Years in *Science* 289, p 270–7

Csete, ME, Doyle, JC (2002) Reverse Engineering of Biological Complexity in *Science* 295:5560, p 1664–1669

Daley, J., Delbanco, T. L. , et al. (1996) A 47–Year–Old Woman With Ductal Carcinoma In Situ, 1 Year Later [comment]. in *JAMA* 6:276, p 491

DeVita (1997) Staging of Breast Cancer. in *Cancer Principles & Practice of Oncology*, p 1569

Draper, D. (1995) Assessment and Propagation of Model Uncertainty. in *Journal of the Royal Statistical Society Series B* 1:57, p 45–97

Drinkwater, M. (2000) Superclusters—the Largest Structures in the Universe in *Science* 287:5456, p 1271–1218

EBCTCG (Early Breast Cancer Trialists' Collaborative Group) (1992) Systemic treatment of early breast cancer by hormonal, cytotoxic, or immune therapy. 133 randomised trials involving 31,000 recurrences and 24,000 deaths among 75,000 women in *Lancet* 339:8785, p 71–85

Falkowski, P., Scholes, RJ., Boyle, E., Canadell, J., Canfield, D., Elser, J. (2000) The Global Carbon Cycle: A Test of Our Knowledge of Earth as a System in *Science* 290:5490, p 291–296

Feuer, E. J. , Wun, L.–M. (1992) How Much of the Recent Risk in Breast Cancer Incidence Can Be Explained by Increases in Mammography Utilization? in *American Journal of Epidemiology* 12:136, p 1423–1436

Fisher, B., Anderson, S., Bryant, J., Margolese, RG., Deutsch, M., Fisher, ER. (2002) Twenty–year follow–up of a randomized trial comparing total mastectomy, lumpectomy, and lumpectomy plus irradiation for the treatment of invasive breast cancer in *N Engl J Med* 347:16, p 1233–41

Fisher, B., Jeong, JH., Anderson, S., Bryant, J., Fisher, ER., Wolmark, N. (2002) Twenty–five–year follow–up of a randomized trial comparing radical mastectomy, total mastectomy, and total mastectomy followed by irradiation in *N Engl J Med* 347:8, p 567–75

Fisher, B., Slack, N. H. , et al. (1969) Cancer of the Breast: Size of Neoplasm and Prognosis. in *Cancer* 5:24, p 1071–1080

Fournier, D. V., Weber, E. , et al. (1980) Growth Rate of 147 Mammary Carcinomas. in *Cancer*:45, p 2198–2207

Fryback, DG, Stout, NK, Rosenberg, MA, Trentham–Dietz, A, Remington, PL, Kuruchittham, V (2004) The case for limited malignant potential breast tumors: Implications of the Wisconsin breast cancer epidemiology simulation model,

Gail, M. H., Brinton, L. A. , et al. (1989) Projecting Individualized Probabilities of Developing Breast Cancer for White Females Who Are Being Examined Annually. in *Journal of the National Cancer Institute* 24:81, p 1879–1886

Garber, J. (1999) A 40–Year–Old Woman with a Strong Family History of Breast Cancer. in *JAMA* 20:282, p 1953–1960

Gustafson, P. (1998) Flexible Bayesian Modeling for Survival Data. in *Lifetime Data Analysis*:4, p 281–299

Habbema, J. D. F., van Oortmarssen, G. J. , et al. (1984) The MISCAN Simulation Program for the Evaluation of Screening for Disease. in *Computer Methods and Programs in Biomedicine*:20, p 79–93

Haber, D. (2000) Roads Leading to Breast Cancer. in *New England Journal of Medicine* 21:343, p 1566–1568

Holford, TR., Cronin, K., Feuer, EJ., Mariotto, A. (2003) Changing patterns in breast cancer incidence trends, manuscript,

Kemeny, MM., Peterson, BL., Kornblith, AB., Muss, HB., Wheeler, J., Levine, E. (2003) Barriers to clinical trial participation by older women with breast cancer in *J Clin Oncol* 21:12, p 2268–75

Kerlikowske, K., Carney, P. A. , et al. (2000) Performance of Screening Mammography Among Women With and Without a First–Degree Relative With Breast Cancer. in *Annals of Internal Medicine* 11:133, p 855–863

Klawansky, S., Fox, M. C. (1984) A Growth Rate Distribution Model for the Age Dependence of Human Cancer Incidence: A Proposed Role for Promotion in Cancer of the Lung and Breast. in *Journal of Theoretical Biology* 3:111, p 531–587

Kopans, D. B. (1996) Detection and Treatment of Ductal Carcinoma In Situ of the Breast [Letter]. in *JAMA* 11:276, p 896

Koscielny, S., Tubiana, M. , et al. (1985) A Simulation Model of the Natural History of Breast Cancer. in *British Journal of Cancer*:52, p 515–524

Kusama, S., Spratt, J. S. , et al. (1972) The Gross Rates of Growth of Human Mammary Carcinoma. in *Cancer* 2:30, p 594–599

Lantz, P., Bunge, M., Remington, PL. (1990) Trends in mammography in Wisconsin in *Wis Med J* 89:6, p 281–2

Lantz, P. M., Remington, P. L. , et al. (1991) Mammography Screening and Increased Incidence of Breast Cancer in Wisconsin. in *Journal of the National Cancer Institute* 21:83, p 1540–1546

Lantz, PM., Weier, A., Remington, PL. (1993) Mammography use in Wisconsin: 1987–1992 92:10, p 580–2

Levin, D. L., Gail, M. H. , et al. (1986) A Model for Projecting Cancer Incidence and Mortality in the Presence of Prevention, Screening, and Treatment Programs. in *NCI Monographs 2(Cancer Control Objectives for the Nation: 1985–2000)*, p 83–93

Li, C. I., Anderson, B. O. , et al. (2000) Changing Incidence Rate of Invasive Lobular Breast Carcinoma among Older Women. in *Cancer* 11:88, p 2561–2569

Livny, M, Basney, J, Raman, R, Tannenbaum, T (1997) Mechanisms for high throughput computing in *SPEEDUP Journal* 11,

Loeve, F., Boer, R. , et al. (1999) The MISCAN–COLON Simulation Model for the Evaluation of Colorectal Cancer Screening. in *Computers and Biomedical Research*:32, p 13–33

Mariotto, A., Feuer, EJ., Harlan, LC., Wun, L–M., Johnson, K., Abrams, J. (2002) The dissemination of adjuvant therapy for breast cancer in the US: 1975–1997 in *The National Cancer Institute*,

Markowska, J, Markowska, A. (1998) Spontaneous tumor regression in *Ginekol Pol* 69:1, p 39–44

McKay, M. D. (1995) Evaluating Prediction Uncertainty in *Nuclear Regulatory Commission*,

McPherson, K., Steel, C. M. , et al. (2000) Breast Cancer — Epidemiology, Risk Factors, and Genetics. in *British Journal of Medicine* 7261:321, p 624–628

Mezzetti, M., Robertson, C. (1999) A Hierarchical Bayesian Approach to Age–Specific Back–Calculation of Cancer Incidence Rates. in *Statistics in Medicine*:18, p 919–933

Michaelson, J. S., Halpern, E., et al. (1999) Breast Cancer: Computer Simulation Method for Estimating Optimal Intervals for Screening. in *Radiology*:212, p 551–560

Newcomb, P. A., Lantz, P. M. (1993) Recent Trends in Breast Cancer Incidence, Mortality and Mammography. in *Breast Cancer Research and Treatment* 2:28, p 97–106

Newcomb, PA., Carbone, PP. (1993) Cancer treatment and age: patient perspectives in *J Natl Cancer Inst* 85:19, p 1580–4

Osnes, K., Aalen, O. O. (1999) Spatial Smoothing of Cancer Survival: A Bayesian Approach. in *Statistics in Medicine*:18, p 2077–2086

Patz, E. F., Goodman, P. C. , et al. (2000) Screening for Lung Cancer. in *New England Journal of Medicine* 22:343, p 1627–1633

Piccart, MJ. (1998) The changing landscape of breast cancer clinical research. ESMO–Award Lecture, ECCO–9 Hamburg, 18 September 1997 in *Ann Oncol* 9:2, p 133–8

Pike, M. C., Krailo, M. D. , et al. (1983) 'Hormonal' Risk Factors, 'Breast Tissue Age' and the Age–Incidence of Breast Cancer. in *Nature*:303, p 767–770

Ries, LAG., Eisner, MP., Kosary, CL., Hankey, BF., Miller, BA., Clegg, L (2003) SEER Cancer Statistics Review, 1975–2000 in *National Cancer Institute*,

Rosenberg, MA (2004) Competing risks in *JNCI Monographs*,

Ross, CM. (1991) DYNASIM2 and PRISM: Examples of dynamic modeling in *In: Citro, CF. Hanushek EA, editors. Improving Information for Social Policy Decisions: The Uses of Microsimulation Modeling* II,

Rutter, C. M., Mandelson, M. T. , et al. (2001) Changes in Breast Density Associated with Initiation, Discontinuation, and Continuing Use of Hormone Replacement Therapy. in *JAMA* 2:295, p 171–176

Shwartz, M. (1978) A Mathematical Model Used to Analyze Breast Cancer Screening Strategies. in *Operations Research* 6:26, p 937–955

Shwartz, M. (1978) An Analysis of the Benefits of Serial Screening for Breast Cancer Based Upon a Mathematical Model of the Disease. in *Cancer* 4:41, p 1550–1564

Shwartz, M. (1981) Validation and Use of a Mathematical Model to Estimate the Benefits of Screening Younger Women for Breast Cancer. in *Cancer Detection and Prevention*:4, p 595–601

Shwartz, M. (1980) Estimates of lead time and length bias in a breast cancer screening program in *Cancer* 46:4, p 844–51

Shwartz, M. (1975) Analysis of Preventive Screening to Detect Early Breast Cancer [Ph.D.],

Smith, B. L. (2000) Approaches to Breast Cancer Staging. in *New England Journal of Medicine* 8:342, p 580–581

Soden ,BJ., Wetherald, RT., Stenchikov, GL., Robock, A. (2002) Global Cooling After the Eruption of Mount Pinatubo: A Test of Climate Feedback by Water Vapor in *Science* 296:5568, p 727–730

Speer, J. F., Petrosky, V. E. , et al. (1984) A Stochastic Numerical Model of Breast Cancer Growth That Simulates Clinical Data. in *Cancer Research*:44, p 4124–4130

Spiegelman, D., Colditz, G. A., et al. (1994) Validation of the Gail et al. Model for Predicting Individual Breast Cancer Risk. in *Journal of the National Cancer Institute* 8:86, p 600–607

Spratt, JS., Spratt, JA. (2002) Chapter 21 Growth Rates. In: Donegan WL, editor in *Cancer of the Breast*, p 443–476

Stout, NK (2004) Quantifying the benefits and risks of mammography for women, researchers, and policy makers: An analysis using a simulation model.,

Tabar, L., Duffy, S. W. , et al. (1999) The Natural History of Breast Carcinoma: What Have We Learned for Screening? in *Cancer* 3:86, p 449–462

Thain, D, Tannenbaum, T., Livny, M. Condor and the grid. In: Berman F, Hey AJG, Fox G, editors in *Grid Computing: Making the Global Infrastructure a Reality*,

Tsodikov, A. D., Asselain, B. , et al. (1997) A Distribution of Tumor Size at Detection: An Application to Breast Cancer Data. in *Biometrics*:53, p 1495–1502

Urban, N., Drescher, C. , et al. (1997) Use of a Stochastic Simulation Model to Identify an Efficient Protocol for Ovarian Cancer Screening. in *Controlled Clinical Trials*:18, p 251–270

Walter, L. C., Covinsky, K. E. (2001) Cancer Screening in Elderly Patients: A Framework for Individualized Decision Making. in *JAMA* 21:285, p 2750–2756

Welch, H. G., Black, W. C. (1997) Using Autopsy Series to Estimate the Disease "Reservoir" for Ductal Carcinoma in Situ of the Breast: How Much More Breast Cancer Can We Find? in *Annals of Internal Medicine* 11:127, p 1023–1028

Welch, H. G., Black, W. C. (1997) Evaluating Randomized Trials of Screening. in *Journal of General Internal Medicine*:12, p 118–124

Welch, H. G., Shwartz, L. M. , et al. (2000) Are Increasing 5–Year Survival Rates Evidence of Success Against Cancer? in *JAMA* 22:283, p 2975–2978

Welch, H. G. (2001) Informed Choice in Cancer Screening. in *JAMA* 21:285, p 2776–2778

Welch, HG., Black, WC. (1997) Using autopsy series to estimate the disease "reservoir" for ductal carcinoma in situ of the breast: how much more breast cancer can we find? in *Ann Intern Med* 127:11, p 1023–8

White, E. (1987) Projected Changes in Breast Cancer Incidence Due to the Trend Toward Delayed Childbearing. in *American Journal of Public Health* 4:77, p 495–497

Wilmoth, JR., Shkolnikov, V., [Directors] (2003) Human Mortality Database. US Female Death Rates by Year of Birth,

Woloshin, S., Shwartz, L. M. , et al. (1999) Women's Perceptions of Breast Cancer Risk: How You Ask Matters. in *Medical Decision Making* 3:19, p 221–229

Wyld, L., Reed, MW. (2003) The need for targeted research into breast cancer in the elderly in *Br J Surg* 90:4, p 388–99

van den Akker–van Marle, E., de Koning, H. , et al. (1999) Reduction in Breast Cancer Mortality Due to the Introduction of Mass Screening in the Netherlands: Comparison with the United Kingdom. in *Journal of Medical Screening*:6, p 30–34

FLEXKB DOCUMENT
Version: HI.001.07242013.3519
Document generated: 07/24/2013

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

DANA-FARBER CANCER INSTITUTE

Important note: This document will remain archived as a technical appendix for publications. New versions will be added periodically as model refinements and updates are completed. The most current version is available at http://cisnet.cancer.gov/profiles. Note that unlike most PDF documents, the CISNET model profiles are not suitable for printing as they are not typically written or read in sequential fashion.

We recommend you let your interests guide you through this document, using the navigation tree as a general guide to the content available.

The intent of this document is to provide the interested reader with insight into ongoing research. Model parameters, structure, and results contained herein should be considered representative but preliminary in nature.

We encourage interested readers to contact the contributors for further information.

Go directly to the: Reader's Guide.

READERS GUIDE

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

Core Profile Documentation

These topics will provide an overview of the model without the burden of detail. Each can be read in about 5–10 minutes. Each contains links to more detailed information if required.

Model Purpose

This document describes the primary purpose of the model.

Model Overview

This document describes the primary aims and general purposes of this modeling effort.

Assumption Overview

An overview of the basic assumptions inherent in this model.

Parameter Overview

Describes the basic parameter set used to inform the model, more detailed information is available for each specific parameter.

Component Overview

A description of the basic computational building blocks (components) of the model.

Output Overview

Definitions and methodologies for the basic model outputs.

Results Overview

A guide to the results obtained from the model.

Key References

A list of references used in the development of the model.

Further Reading

These topics will provide a intermediate level view of the model. Consider these documents if you are interested gaining in a working knowledge of the model, its inputs and outputs.

JNCIMonograph Outline

This topic provides links to profile content organized according to the JNCI Monograph Outline for Model Description Chapters. Use this outline for comparisons focused on the CISNET Base Case simulations.

Advanced Reading

These topics denote more detailed documentation about specific and important aspects of the model structure

MODEL PURPOSE

SUMMARY

This document gives the general purpose of the model and other typical applications it might be used in.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

PURPOSE

The purpose of the model is to predict the mortality associated with female breast cancer. The predictions may be by chronological year and/or age. Mortality may change by advances in treatment and/or changing dissemination of screening. The model incorporates the possibility that these latter two factors will change by chronological time and age. The model is general and enables the prediction of changes in mortality if technical advances are made by radiology or the discovery of other disease markers.

The probability model was developed to describe the early detection process for any chronic disease. The application to breast cancer requires knowledge of the relevant parameters associated with the natural history, diagnosis and treatment of breast cancer. Application to other chronic disease requires similar specialized inputs.

In addition to using the model to predict national mortality trends there are a number of other important applications of the model; i.e. (i) prediction of the outcome of early detection clinical trials without the necessity of long–term follow–up; (ii) evaluation of service programs on early detection; (iii) investigation of different screening schedules to compare mortality benefit. The screening schedules are a function of: age at first exam, number of exams, spacings between examinations and modality of diagnosis (physical exam, mammogram or both).

See Model Overview for deeper details and some limitations inherent in the model.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

MODEL OVERVIEW

SUMMARY

This document provides an overview of the modeling effort including the reasons it was undertaken and the work it builds upon. It also contains a summary of the methodologies employed.

PURPOSE

The purpose of the model is to predict the mortality associated with female breast cancer. The predictions may be by chronological year and/or age. Mortality may change by advances in treatment and/or changing dissemination of screening. The model incorporates the possibility that these latter two factors will change by chronological time and age. The model is general and enables the prediction of changes in mortality if technical advances are made by radiology or the discovery of other disease markers. See Model Purpose for more details.

BACKGROUND

The model is a stochastic model of the natural history of the disease. A series of equations are derived that predicts the age specific probability of death which is the mortality rate. The introduction of screening in the model makes the mortality rate equations more complex as it is necessary to distinguish among screen detected and interval detected cases. Screen detected cases are those in which the woman is asymptomatic and the disease is diagnosed by an early detection screening examination; interval detected cases are those cases not detected at a screening examination, but there is a history of at least one negative screening examination. The model takes into account both lead time and length biased sampling biases. The assumption for the effect of screening assumes that diagnosis of screen detected cases changes the distribution of staging beyond what would be expected due to length biased sampling. We refer to this as a stage shift and is in the direction of having a higher proportion of more favorable prognostic cases.

The basic probability model requires the choice of a reference time point in chronological time. The model predicts the cumulative mortality relative to this reference time conditional on being a specific age at the reference time point. If the time point is chosen as a birth cohort year, then the model can predict the age specific mortality rate for a specific birth cohort year. The age specific mortality for any point in chronological time may be calculated by choosing a collection of birth cohort years. These may be averaged with respect to a weight function to give the overall mortality rate for a specific chronological year. Due to the generality of the model, predictions may be made for populations and sub populations.

MODEL DESCRIPTION

The basic assumption of the model is that breast cancer is a progressive disease. Four or possibly five states of health are envisioned. These states are:

- S_0: A woman is disease free or has disease but it is asymptomatic and cannot be diagnosed by any modality;

- S_p: A woman has breast cancer, but it is asymptomatic and may be diagnosed by a special examination or examination program;
- S_c: A woman, having usual care, is diagnosed with invasive breast cancer;
- S_d: Death attributed to breast cancer;
- S_d^*: Death, not attributed to breast cancer

The progressive disease model may be described by the path:

$$\begin{array}{ccccccc} \mathrm{S}_0 & \rightarrow & S_p & \rightarrow & S_c & \rightarrow & S_d \\ & \searrow & & & & \swarrow & \\ & & & S_d^* & & & \end{array}$$

The main output of the model is breast cancer specific mortality. Hence the transition into S_d^* may be ignored. See Component Overview for more details on the model's building blocks.

Inputs to the Model

The philosophy of the model is to have a probability model in which the parameters can be observed or can be directly estimated from existing data. The inputs required of the model are:

a. Age dependent incidence rates for a time period in which breast cancer screening was not widely used;
b. Age dependent transition rates into Sp (pre–clinical state);
c. Stage distribution for usual care cases and cases diagnosed by screening or having a history of screening exams. The stage distribution may be age related.
d. Survival distribution conditional by stage, chronological time and age. The reason for specifying chronological time is to account for advances in therapy. The dependence on chronological time will be a function of the dissemination of treatment advances in the general population.
e. Dissemination or pattern of screening;
f. Sensitivity of mammograms and physical exams by age,
g. Birth cohort year(s) to which mortality predictions will be made.

Some of these parameters may be estimated from the eight randomized trials investigating mammography; e.g. stage distribution by modality of diagnosis, sensitivity. Others can be obtained from databases such as SEER (survival conditional on stage, stage distributions with usual care). The age dependent transition rates into S_p may be obtained from the age incidence rates using the methods earlier derived by Lee and Zelen. See Parameter Overview for more details.

Outputs

The outputs of the model are: overall breast cancer mortality for chronological time and reduction in mortality relative to some base. We believe that the reduction in mortality may be the most accurate prediction. Our reasoning is that if there are other factors influencing breast cancer mortality, which do not interact with treatment and/or early detection modalities, their effect on mortality reduction will be negligible as their contribution to the hazard function will be additive. The reductions in mortality may

be relative to a base year, a screening strategy (before and after) or a reduction in lag time between a clinical trial showing the benefit of a new therapy and the time when the new therapy is widely adopted.

The mortality outputs may also be age specific. Furthermore, the model outputs may be cause specific or total mortality. See Output Overview for more details.

Limitations of the Model

The basic assumptions of the model are: (i) breast cancer is a progressive disease and (ii) the benefit of early detection is through a stage shift in diagnosis. See Assumption Overview for a more detailed list of assumptions. Current views of the natural history of the disease agree that breast cancer is a progressive disease. However there is less agreement on the reasons earlier diagnosis may result in reducing mortality. There is a group of investigators who believe that early detection by mammography confers no benefit. However it is now generally accepted that the criticisms of the scientific evidence have been satisfactorily answered, discredited or are peripheral. Our model, suitably modified, can be used to predict the outcome of the early detection clinical trials. It was applied to the eight randomized early detection breast cancer clinical trials and was able to predict the outcome of seven of the trials. The agreement with the breast cancer early detection randomized trials indicates that the stage shift assumption may be a valid assumption to explain benefit. Our findings were presented at the Global Summit on Mammography held in Milan, Italy in early June 2002.

Another possible criticism of our model is that survival depends on modality of detection as well as disease stage. However there is no clinical data to support this conjecture.

CONTRIBUTORS

We gratefully acknowledge the collaboration with Dr. Diana Miglioretti of the Breast Cancer Surveillance Consortium for making the stage shift data (Table 2) available. We very much appreciate the many discussions with Drs. Kathy Cronin, Angela Mariotto, Rocky Feuer and Rebecca Gelman in clarifying the nature of the input data used in our model. Finally we are very much indebted to Ms. Hui Huang for carrying out the difficult calculations required by the model to obtain numerical results. This investigation was supported by the NCI CISNET project funded under grant CA88270.

ASSUMPTION OVERVIEW

SUMMARY

In this section we summarize the main assumptions fo the model

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

BACKGROUND

The basic assumptions of the model are: (i) breast cancer is a progressive disease and (ii) the benefit of early detection is through a stage shift in diagnosis. A more detailed list follows.

It is widely believed that breast cancer is a progressive disease.The stage shift assumption is the only reasonable hypothesis as to why early detection may be beneficial.

The prediction of the mortality reduction in the eight randomized early detection trials indicated no benefit for the two canadian trials. These were the only two trials that did not show a stage shift.

There is no intreaction between the progressive disease model and stage shift. However there are interactions with age.

ASSUMPTION LISTING

a. Progressive disease model is basic to the model. The natural history of disease progresses from no disease (or disease which cannot be detected) to pre–clinical disease to clinical disease.The natural history will depend on age.
b. The process by which early detection changes prognosis is by a stage shift in that a higher proportion of screened detected cases will have disease stages with better prognosis. The stage shift may be age dependent.
c. Women who are interval–detected cancers have the same stage distribution as those not participating in a screening exam.This is an observation from the eight early detection trials.However this assumption is not necessary.
d. The survival distribution consists of a mixture of survival distribution conditional on stage. The weights correspond to the probability of being diagnosed in a particular stage. They will change according to a stage shift for screen–detected cases. The conditional survival distributions will change with chronological time corresponding to the introduction of advances in treatment.
e. The sensitivity of the exams (mammograms, physical and the combination) will be age–related with lower mammogram sensitivities for younger women.
f. The sojourn time distribution is assumed to follow the exponential distribution with a mean which is age dependentNOte that this distribution is not observed. Older women are assumed to have longer mean sojourn times than younger women.The expinential assumption is based on the results of the HIP trial in which full data is available.We have shown that a necessary and suficient condition for the sojourn time to be exponential is that the mean ages of those diagnosed at the first exam is equal to the mean age of a diagnosed control group. This condition was true for the HIP study.

g. We have assumed that the sensitivity of a mammogram is age dependent with higher sensitivities for the women over 50. This age dependence has been illustrated in many of the early detection trials.

PARAMETER OVERVIEW

SUMMARY

This document describes the basic parameters used by the model as well as provides current estimates for each.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

BACKGROUND

The key assumptions are outlined in Assumption Overview . These are that (i) breast cancer is a progressive disease and (ii) the benefit from early detection is due to a stage shift in screen–detected cases. The stage shift is by definition beyond that expected from length biased sampling.

Our model requires input data which may come from various sources. They include: survival conditional on stage, sensitivity of mammograms, sojourn time distribution in the pre–clinical stage, stage distribution with and without screening, dissemination of screening and therapy in the 1975–1999 period and the estrogen receptor (ER) status. Many of these inputs may be age–related. In this section we discuss the values of these inputs.

The philosophy of the model is that the input data may be observed or can be estimated from existing data. Examples of the latter are the sensitivity of the screening modality and the transition probabilities into S_p. The model does not contain parameters which are estimated to fit existing mortality. In this section the sources of the basic input data and applications to our model are described. The notation used in this section is previously defined in Natural History Component . We have used the software, Mathcad 2001i from Mathsoft Inc., to carry out the calculations of the breast cancer mortality in the U.S. women in the chronological time period 1975–1999.

PARAMETER LISTING OVERVIEW

The model requires numerical values of various parameters. In this section we summarize our estimates and discuss various data sources.

Survival, Sensitivity, Sojourn Time in the Pre–clinical State and Stage Distribution

The SEER database provides breast cancer incidence, staging and survival for the period 1975–1979. We have chosen this time period for the input data as breast cancer screening was not common at that time. Choosing a later period would result in these data sources being influenced by screening. The estimate of the age–specific breast cancer mortality for birth cohort year v without screening history ($d_v(T)$), defined in equation (2.5) has utilized the input data $S_v(t)$ and $I_v(\tau)$ for birth cohort v which was provided by the CISNET NCI group. In our model, Ductal Carcinoma In Situ (DCIS) cases were not included.

Table 1 summarizes the stage distribution without screening (θ_i) based on the SEER data. The CISNET NCI group has estimated the AJCC stage distributions using the SEER extent of disease data for the years 1975–1979. Age–specific breast cancer survival, conditional on the AJCC stage, has also been provided by the CISNET NCI group. We estimated the annual hazard rate and cumulative survival conditional on stage and age. By multiplying these two quantities, the p.d.f of age–specific breast cancer survival conditional on stage was estimated. Then the p.d.f of breast cancer specific survival as defined in equation (2.1) was generated using the stage distributions (θ_i) and the p.d.f. of breast cancer survival conditional on stage.

TABLE 1. Summary on Stage Distribution without Screening

AJCC Stage Distribution from SEER 1975–1979					
Age	**Stage I**	**Stage II–**	**Stage II+**	**Stage III**	**Stage IV**
30–39	0.31	0.23	0.31	0.09	0.06
40–49	0.30	0.23	0.31	0.10	0.06
50–59	0.29	0.22	0.31	0.10	0.08
60–69	0.30	0.22	0.27	0.10	0.11
70–84	0.32	0.27	0.22	0.10	0.10

Our model requires further input data to incorporate screening history and advances in treatment over chronological time. Using the age–specific incidence data $I_v(\tau)$ for birth cohort v, we have estimated transition probabilities between S_0 to S_p and between S_p to S_c. We have further assumed that the pre–clinical sojourn time follows an exponential distribution with an age dependent mean. The mean sojourn times ($m(t)$) serving as input to the model are:

$$m(t) = \begin{cases} 2 & \text{for } t \leq 40 \\ -6 + 0.2t & \text{for } 40 < t \leq 50 \\ 4 & \text{for } t > 50 \end{cases}$$

These values are based on data from the early detection in randomized clinical trials. The Breast Cancer Surveillance Consortium (BCSC) published the age–dependent sensitives of screening mammograms in the U.S. administered in 1996–1998[1]. We used their estimates in the model. (The BCSC project was founded by NCI in 1994 to evaluate mammogram screening practices in the U.S. population.) The BCSC database currently contains mammogram screening data and follow–up for approximately one million U.S. women beginning in 1994. Specifically the age–dependent estimated sensitivities ($\beta(t)$) for screening exams from the BCSC data are:

$$\beta(t) = \begin{cases} 0.55 & \text{for } t < 40 \\ 0.65 & \text{for } 40 \leq t < 45 \\ 0.70 & \text{for } 45 \leq t < 50 \\ 0.75 & \text{for } 50 \leq t < 70 \\ 0.8 & \text{for } t \leq 70. \end{cases}$$

Shen and Zelen[2] estimated sensitivities of screening examinations and mean pre–clinical sojourn times from the randomized clinical trials evaluating the benefit of mammography. In their calculations, the mammogram sensitivities had an improving trend over time for screening clinical trials conducted in 1963 – 1990's. Therefore the sensitivities presented above were applied to screening exams conducted in 1995–1999 and the sensitivities for the previous years were lowered. For 1985–1995, the sensitivities were lowered by 0.10 for ages Finally the assumption of exponential sojourn times in the pre–clinical state can be justified from two sources. Zelen and Feinleib[3] have proved that the necessary and sufficient condition for the sojourn time to follow an exponential distribution is that the mean age of diagnosis for the initial early detection exam be the same age as those diagnosed in a control group. This condition was verified in the HIP randomized trial. The second source is the empirical study carried out by Day and Walter[4] in which they investigated various distributions for the pre–clinical sojourn time in the HIP trial and found that the exponential distribution gave the best fit.

Stage Shift

The BCSC has provided data on AJCC stages at diagnosis for screen–detected and interval cases. In the BCSC data, a screen detected cancer was defined as cancer diagnosed within 4 months of a positive screening mammogram (bilateral mammograms indicated by the radiologist to be done for routine screening). An interval cancer was defined as cancer diagnosed within 4 months of a diagnostic mammogram (mammogram indicated by the radiologist to be done for evaluation of a breast problem). A mammogram was considered positive if it was given a final BI–RADS assessment code of 0 (need additional imaging evaluation), 4 (suspicious abnormality), 5 (highly suggestive of malignancy), or 3 (probably benign finding) with a recommendation for immediate follow–up. Time since prior mammography was determined using dates of prior examinations in the mammography registry or self–reported information. We have categorized the time since prior mammogram as one–year and longer than one–year.

We estimated the distribution of AJCC stages for screen detected cases in age groups of

TABLE 2A. Summary of Stage Distribution for Screen Detected Cases

AJCC Stage Distribution with Annual Screening (from BCSC)					
Age	**Stage I**	**Stage II–**	**Stage II+**	**Stage II**	**Stage IV**
	0.62	0.11	0.21	0.04	0.01
50–59	0.67	0.11	0.19	0.03	
60–69	0.76	0.07	0.14	0.02	
70+	0.78	0.09	0.11	0.02	0.01

AJCC Stage Distribution with Screening Interval					
Age	**Stage I**	**Stage II–**	**Stage II+**	**Stage III**	**Stage IV**
	0.58	0.12	0.24	0.04	0.015
50–59	0.62	0.15	0.17	0.04	0.016
60–69	0.66	0.13	0.18	0.03	0.007
70+	0.73	0.13	0.11	0.01	0.014

TABLE 2B. Summary of Stage Distribution for Interval Cases

AJCC Stage Distribution with Annual Screening (from BCSC)					
Age	**Stage I**	**Stage II–**	**Stage II+**	**Stage III**	**Stage IV**
	0.46	0.19	0.26	0.07	0.02
50–59	0.45	0.17	0.30	0.07	0.01
60–69	0.54	0.15	0.23	0.06	0.01
70+	0.54	0.23	0.16	0.05	0.01

AJCC Stage Distribution with Screening Interval					
Age	**Stage I**	**Stage II–**	**Stage II+**	**Stage III**	**Stage IV**
	0.37	0.22	0.31	0.08	0.02
50–59	0.29	0.26	0.26	0.12	0.06
60–69	0.41	0.22	0.27	0.07	0.03
70+	0.43	0.29	0.18	0.07	0.03

The stage distribution in the absence of screening can be compared to the stage distribution estimated from the BCSC data. For example, for women under the age of 50 years, 54% were diagnosed with Stage I/II– disease when no screening was conducted. (Stage I/II–disease is essentially node negative or local disease stage). However 73% of the same age group were detected at screening with stage I/II– with annual mammograms and 70% with exams having longer than a one–year interval between exams. This shift of 54% to 73% in finding more cases at an earlier stage (stage I/II–), when women were screened annually, results in a mortality reduction.

This staging information compared to the SEER stage distribution presented in Table 1 allows comparison of stage shift data for screening versus usual care. There is a larger

proportion of women detected at earlier stages when diagnosed by screening. Furthermore the stage shift is slightly more pronounced with a shorter screening interval. Similar stage shift data are available from the eight randomized early detection clinical trials and are in close agreement (see Discussion). Finally, the p.d.f. of disease–specific survival for screening exam diagnosed groups were estimated using the stage distributions presented in Tables 2a and 2b combined with the 1975–79 SEER survival data.

It is interesting to note that the interval cases had a slightly better prognostic stage distribution than the control group. The stage shift distribution for interval cases depends slightly on the screening intervals.

Screening Dissemination

Screening patterns for each birth cohort year have been modeled by the CISNET NCI group using the data from the National Health Interview Survey (NHIS) and BCSC. This effort provides information on the probability of the first screening examination for birth cohorts 1891–1970 at chronological years 1975–1999. This information is directly incorporated into our model. In addition, the information on screening patterns, conditional on the age at the first screening examination, was available. The screening pattern was incorporated into our model using the age intervals 18–39, 40–49, 50–59, 60–69 and 70–79. In addition, the screening patterns are summarized using three idealized screening intervals; i.e., short (1 year), medium (2 years) and long (5 years). For women starting the first screening examination at ages 50–59, the possible screening patterns and probabilities of observing specific patterns are summarized in Table 3. If women die of breast cancer before age 70, screening patterns up to age 69 and the corresponding probabilities are utilized.

TABLE 3. Screening Patterns for Women with First Screening Exam at Ages 50–59 Years

Screening during Ages			
50–59	**60–69**	**70–79**	**Probability**
s	s	s	0.369
s	s	m	0.033
m	s	m	0.012
m	m	m	0.259
m	m	l	0.034
m	l	l	0.001
l	l	l	0.292

s=1 year, m=2 years, l=5 years

For the purpose of illustration, we have displayed only a summary of screening patterns for women who had their first screening examinations between ages 50–59. However we have created similar tables for all of the age categories described above. The combinations of various screening patterns (H_i for $i = 1, ..., k$) together with disease–specific survival data and stage shift information have been incorporated into equation (2.8) to assess the disease–specific mortality for the screened population. The

stage distributions used in the model correspond to the screening patterns summarized in Tables 2a and 2b. For women following screening exams with mixed intervals (1,2 and 5 years), the stage distribution associated with screening interval greater than one–year was used. When all the screening exams are 5 years apart, the stage shift associated with screening interval greater than one–year interval was lowered by combining it with the stage distribution of the control group in Table 1. These adjustments were made to take into account the empirical observation that the magnitude of the stage shift is associated with actual screening intervals.

Treatment Dissemination

The dissemination of adjuvant therapies for breast cancer has also been modeled by the NCI CISNET group. The patterns of care data has been utilized to model the dissemination of breast cancer treatments in the U.S. between the years 1975–1999[5]. The CISNET NCI group has provided the data on the proportion of women receiving Tamoxifen, multi–chemotherapy or both by age groups (69) and the AJCC stages for the years 1975–1999. For each treatment option, a median smoothing technique was applied to model the proportion of women receiving therapy as a function of chronological years 1975–1999. The smoothed function of the dissemination pattern for each treatment option has been directly incorporated into our model.

We have utilized the survival benefit of multi–chemotherapies reported by the Early Breast Cancer Trialists' Collaborative Group[6]. The EBCTCG reported the proportional reduction in the annual odds of death for multi–chemotherapies by age groups
A similar adjustment was made for the survival benefit attributed to Tamoxifen. The EBCTCG[7] reported the proportional reduction in the annual odds of death ratio for tamoxifen use of 2 years and 5 years of continuous use. Again the disease specific survival from the SEER 1975–1979 database has been appropriately adjusted using the reported annual odds of death for 2 year or 5 year Tamoxifen course of therapy. We have estimated the age specific ER positivity using the 1988–1993 SEER data (ER status data became available in the SEER database beginning in 1988). Table 4 summarizes the age specific ER status data used in our model. The benefit of Tamoxifen was applied only to ER+ women. In addition, the dissemination and benefit of Tamoxifen have been modeled separately for the 2 year vs. 5 year use of Tamoxifen.

TABLE 4. Distribution of ER Status by Age Group in SEER 1988–1993

Age	ER+	ER–
	63%	37%
50–69	77%	23%
≥ 70	85%	15%

REFERENCES:

[1] Carne, PA, Miglioretti, DL, et al "Individual and combined effects of age, breast density, and hormone replacement therapy on the accuracy of screening mammography" in Annals of Internal Medicine 2003; 138: 3: 168-175

[2] Shen, Y, Zelen, M "Screening sensitivity and sojourn time for breast cancer early detection trials: mammograms and physical exams." in J Clin Oncol 2002 2002; 19: 3490-3499

[3] Zelen, M, Feinleib, M "On the theory of screening for chronic diseases" in Biometrika 1969; 56: : 601-614

[4] Day, NE, Walter, SD. "Simplified models of screening of chronic disease: estimation procedures from mass screening programmes" in Biometrics 1984; 40: : 1-13

[5] 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" in JNCI 2002; 94: : 1626-1634

[6] EBCTCG "Polychemotherapy for early breast cancer: an overview of randomized trials" in Lancet 1998a; 352: : 930-942

[7] EBCTCG "Tamoxifen for early breast cancer: an overview of the randomized trials" in Lancet 1998b; 351: : 1451-1467

Component Overview

Summary

Describes major model components.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

Overview

In this section we describe the major components of the model. Since our models are probability models we will describe the elements in minimal technical language. The development of the model requires that individuals without a screening history be treated differently than those with a screening history. The equations for individuals with screening history are more complex.

Component Listing

Natural History Component
Survival And Mortality Component

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

Output Overview

Summary

A Stochastic model has been developed for predicting U.S. breast cancer mortality as a function of chronological time and/or age.The model takes into account the changinging disemination of new therapies and screening patterns using mammography.

Overview

The output from this model may consist of: annual breast cancer mortality for specified chronological times and age specific breast cancer mortality for specified chronological times. Reductions in breast cancer mortality can also be similarly generated relative to a base year. In addition the model will be able to partition the reduction in breast cancer mortality according to changes in treatment and changes in the dissemination of mammography use. In general the output for the model will will be mortality as a function of the inpput parameters; i.e.age distribution of population, screening schedules, survival conditional on stage and treatment, stage distribution. In many cases the output will be the proportional reduction in mortality relative to a control group.

Output Listing

Important outputs: mortality by chronological time and/or age, reduction in proportional mortality by age or chronological time. The mortality can be standardized to any base year. Finally, our overall model has been used to predict the outcomes of the eight randomized early detection breast trials. We have been able to verify the reduction in mortality reported by seven of the eight trials.

RESULTS OVERVIEW

SUMMARY

This contains the outputs of the model

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

OVERVIEW

Eventually the application of model will generate:

1. Mortality predictions
2. Predictions of mortality reduction for proposed screening schedules
3. Reductions in mortality associated with screening dissemination, advances in therapy and a combination of both.
4. Prediction of probability of over diagnosis by age.
5. Test of model by predicting outcomes of eight early detection breast cancer trials

RESULTS LIST

Several results have been generated by this modeling effort. We list a few of them below.

Model Validation Procedures

Describes model validation and sensitivity analysis used in this effort.

Predicted Mortality Reductions

A table of predictions of the outputs for the eight randomized breast cancer early detection trials.

NATURAL HISTORY COMPONENT

SUMMARY

This document overviews the models treatment of the natural history of the disease.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

OVERVIEW

The theoretical model builds on the natural history of the disease. The basic assumption of the natural history is that breast cancer is a progressive disease. Four or possibly five states of health are envisioned. The states are:

S_0 : A woman is disease free or has disease but it is asymptomatic and cannot be diagnosed by any modality;

S_p : A woman has breast cancer, but it is asymptomatic and may be diagnosed by a special examination;

S_c : A woman having usual care is diagnosed with invasive breast cancer;

S_d : Death attributed to breast cancer;

S_d^* Death, not attributed to breast cancer.

The progressive disease model may be described by the path:

$$\begin{array}{ccccccc} S_0 & \Rightarrow & S_p & & \Rightarrow & S_c & \Rightarrow & S_d \\ & \searrow & & \searrow & & \swarrow & & \\ & & & S_d^* & & & & \end{array}$$

The main interest is the reduction in breast cancer specific mortality. Women diagnosed with breast cancer who eventually die of other causes are regarded as right–censored observations. Hence the transition into S_d^* may be ignored.

The goal of a breast cancer screening program is to diagnose women who are asymptomatic for breast cancer. Hence by definition women diagnosed by a screening exam are in the pre–clinical state. It is necessary to distinguish among cases which are diagnosed: (i) by a screening exam, (ii) after a negative exam when the disease becomes symptomatic and (iii) by usual care. *Screen detected* cases are those in which the women are asymptomatic and the disease is diagnosed by an early detection examination. *Interval cases* are those not detected at a screening examination, but there is a history of at least one negative screening examination. An *incident case* refers to women who have no history of screening exams but are diagnosed by usual care; i.e., the disease has generated signs/symptoms which makes the women seek medical attention. Interval and incident cases are assumed to be diagnosed in the clinical state. Note that mammography and/or a physical exam may be used to aid in the diagnosis of breast cancer when there are signs/symptoms as well as being used to detect cases in which there are no signs/symptoms. The latter is referred to as a screening exam whereas the former is a diagnostic exam even though the same examination modality is used. In addition to the assumption that breast cancer is a progressive disease, the other basic assumption is that the potential reduction in breast cancer specific mortality from screening is due to a favorable stage shift in diagnosis relative to the distribution of stages when diagnosis is by usual care. We have used the AJCC classification for breast cancer staging. However any system of disease staging may be used in the model.

Detail

See Survival And Mortality Component for details on the modeling of mortality reduction.

Survival And Mortality Component

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

Summary

This document describes development of mortality reduction models for screening and non–screening individuals.

Overview

In this section we describe the major components of the model. The development of the model requires that individuals without a screening history be modeled differently than those with a screening history. Our formulation allows us to follow a specific birth cohort and predict the age–specific breast cancer mortality in any chronological year for the birth cohort.

Detail

No Screening History Model

Define:

v = year of birth cohort
τ = age of incidence
T = age at death
$S_v(t)$ = probability of normal population surviving to year t for birth cohort v

$I_v(\tau)$ = age specific disease incidence for birth cohort v

$g(t|\tau + v)$ = probability density function (p.d.f.) of disease specific survival for subject incident at age τ in chronological year $\tau + v$;

$d_v(T)$ = probability of disease–specific death at age T for birth cohort v.

$M_v(T)$ = age–specific mortality rate for birth cohort v.

The p.d.f. $g(t|\tau + v)$ is a mixture of distributions weighted by the probability of being diagnosed in a particular stage. Specifically

$$g(t|\tau + v) = \sum_{i=1}^{k} \theta_i g_i(t|\tau + v), \tag{2.1}$$

where θ_i is the probability of being diagnosed in stage i

$(i = 1, 2, \ldots, k)$

and $g_i(t|\tau + v)$ is the survival distribution p.d.f. for stage i for a subject diagnosed in chronological time $(\tau + v)$ for a subject having incidence at age τ.

In the chronological year of diagnosis $(\tau + v)$, there may be several treatment options which may have different survival outcomes.

In this case $g_i(t|\tau + v)$ may be written as the mixture

$$g_i^*(t|\tau + v) = \sum_{r=1}^{R} \phi_i(r|\tau + v) g_{ir}(t|\tau + v),$$

(2.2)

where $\phi_i(r|\tau + v)$ = probability of treatment r for a subject diagnosed in stage i at chronological time $\tau + v$ and $g_{ir}(t|\tau + v)$ is the corresponding survival p.d.f. Then the p.d.f. of the disease–specific survival for a subject diagnosed at age τ in chronological year $\tau + v$ and receiving available treatments at that time is

$$g^*(t|\tau + v) = \sum_{i=1}^{k} \theta_i g_i^*(t|\tau + v).$$

(2.3)

The age–specific mortality rate for a subject from birth cohort year v is defined as

$$M_v(T) = \frac{d_v(T)}{S_v(T)} \times 100,000,$$

(2.4)

where

$$d_v(T) = \int_T^{T+1} \int_0^y S_v(\tau) I_v(\tau) g(T - \tau|\tau + v) d\tau dy.$$

(2.5)

That is, the age–specific mortality rate represents the number of disease–specific deaths between ages (T, T+1) in a population of 100,000 from birth cohort year v.
One aim of the model is to estimate the age specific mortality by chronological year. If t refers to chronological year and T denotes the age of death, then $t = T + v$. Hence by choosing a birth cohort year, estimates can be made about age specific mortality corresponding to chronological time t.
Overall disease–specific mortality rate for a chronologic year t may be calculated with reference to some base year. Suppose $p_0(T)$ represents the distribution of ages for a chosen base year. Then the age–adjusted disease–specific mortality rate for chronological year t is

$$M(t) = \int M_{t-T}(T)p_0(T)dT. \tag{2.6}$$

The range of integration will be over the values of T in which $p_0(T)$ is non–negligible.

Screening History Model

Subjects undergoing screening require a more complex model than those without a screening history. Furthermore it is necessary to distinguish between cases diagnosed at a screening examination (screen detected cases) and those diagnosed at other than a screening exam (interval cases). Suppose a subject from cohort year v has a history of screening exams H_i at ages $t_1 < t_2 < \cdots < t_n$.

Screen detected cases get diagnosed at any exam given at ages $t_1, t_2, \cdots, t_n$. It is assumed that no further exams are given after a diagnosis. Interval cases get diagnosed in between exams (t_i, t_{i+1}) for $i = 1, \cdots, n-1$ or after the last exam at t_n.

The probability of disease–specific death at age T for birth cohort v who follows a screening pattern H_i has a more complicated expression than the probability expression (??) of the non–screened population. It can be written as

$$d_v(T|H_i) = \int_T^{T+1} \int_0^y \{D_v(\tau|H_i) + I_v(\tau|H_i)\} d\tau dy,$$

where

$D_v(\tau|H_i)$ = Probability of disease–specific death at age τ for screen detected cases with screening history of H_i,

$I_v(\tau|H_i)$ = Probability of disease–specific death at age τ for interval cases with a screening history of H_i

These probabilities are a function of many parameters involved in the case finding process and have complicated expressions. Details of the derivations and expressions will be published in another paper[1]. In calculating these probabilities, it is necessary to introduce a sensitivity parameter which may be age dependent and two new probability distributions. One of the distributions corresponds to the age–specific probability of entering the pre–clinical state and the other denotes the sojourn time in the pre–clinical state. Both may be age–related.

The survival distribution for screen–detected cases is assumed to be a mixture of distributions as described in equation (1), except that the probabilities (θ_i) of being diagnosed in the various disease stages have changed due a possible stage shift. In our model the stage shift is represented by the new values of θ_i. Generally larger values (θ_i) are expected for better prognostic stages when screening is involved. The lead time, which is defined to be the difference between the age transitioning into the clinical state and the age of earlier diagnosis, is a random variable, which is not observed. It is equivalent to having a random guaranteed survival time; i.e., the subject will live at least to the age at which clinical diagnosis is made. The model accounts for the guaranteed survival time. Otherwise there will be a lead time bias when compared to non–screened cases.

The age–specific mortality rate for a birth cohort v having a screening history H_i is

$$M_v(T|H_i) = \frac{d_v(T|H_i)}{S_v(T)} \times 100,000.$$

The quantity $M_v(T|H_i)$ is a basic element in estimating various screening scenarios. Weighted linear combinations of this quantity can be used to predict the age–specific mortality rate for birth cohort v having a variety of screening histories. The screening histories consist of various combinations of the age at the first screening, frequencies of the screening examinations and the total number of exams. Then the age–specific mortality rate of birth cohort v with screening histories H_i, for $i = 1, ..., k$ is defined as

$$M_v(T|H) = \sum_{i=1}^{k} h_i M_v(T|H_i),$$

(2.7)

where h_i is the probability of screening history H_i.
The age–specific mortality rate for chronological year t can be calculated using the relation $t = v + T$ to identify the appropriate birth cohort year. The overall disease–specific mortality rate for chronological year t for a population with screening histories H, standardized to a population having ages $p_0(T)$, is then

$$M(t|H) = \int M_{t-T}(T|H) p_0(T) dT,$$

(2.8)

where the limits of integration are over the range of T which has non–negligible probabilities $p_0(T)$.

Mortality Reduction

We have formulated the expressions for the overall disease–specific mortality rate at chronological year t. This formulation can be used to estimate the mortality reduction due to treatment, screening or to both treatment and screening disseminated in the population over years. Using the expressions in equations (2.6) and (2.8), one can estimate the overall disease–specific mortality reduction at chronological year t due to screening as

$$MR_{sc}(t|H) = \int \{M_{t-T}(T) - M_{t-T}(T|H)\} p_0(T) dT / \ M(t).$$

(2.9)

The mortality reduction due to treatment disseminated in the population is given by

$$MR_{rx}(t) = \int \{M_{t-T}(T) - M^*_{t-T}(T)\} p_0(T) dT / \ M(t),$$

(2.10)

where $M^*_{t-T}(T)$ can be estimated from equation (2.6) using the treatment incorporated survival p.d.f. described in (2.3).
Lastly, the mortality reduction due to both screening and treatment disseminated in the population is formulated as

$$MR_{sc*rx}(t|H) = \int \{M_{t-T}(T) - M^*_{t-T}(T|H)\} p_0(T) dT / \ M(t). \tag{2.11}$$

REFERENCES:

[1] Lee, SJ, Zelen, M "Mortality modeling of early detection programs" in (in manuscript) 2004;

MODEL VALIDATION PROCEDURES

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

SUMMARY

This document describes some preliminary validation and sensitivity analysis work done with the model.

RESULT TYPE

Validation

OVERVIEW

Model Validation and Sensitivity Analysis

The stochastic model we proposed has two basic assumptions: (i) natural history is progressive and (ii) gains from screening are attributed to a stage shift. In order to validate our model, we applied it to the eight randomized trials investigating the benefits of mammography. The application used input parameters from the trials; e.g., stage shift distribution, exam sensitivities, frequency and spacing of examinations, age distributions and mean sojourn time in the pre–clinical state. These parameters would generally be available during the first few years of the trial. The survival, conditional on stage, was obtained from the 1975–79 SEER data base. The follow–up period for the trials ranged from seven to nineteen years. The follow–up times coincided with the last published follow–up time. Our model predictions for mortality reduction were within the reported confidence intervals for seven of the trials (c.f. Lee and Zelen[1]). The other trial did not report a confidence interval for the reported mortality reduction.

A sensitivity analysis for the model has also been carried out. We have varied two of the input parameters specific to our model (mammogram sensitivities and stage distributions) to evaluate the impact on disease–specific mortality. In particular we have: (i) increased the sensitivity of mammograms to $\beta(t) = 1$ for all ages in the period 1975–1999, (ii) lowered the age–dependent sensitivities to 0.35–0.70 in the period 1975–1999, (iii) changed the stage shift for women following a 5 year screening pattern to the stage distribution from the BCSC for screening with more than one–year and (iv) lowered the stage shift of women following a 5 year screening pattern to the stage distribution of the control group. The results are displayed in Figures 1 and 2.

RESULT

FIGURE 1. Sensitivity Analysis for Mammogram Sensitivity

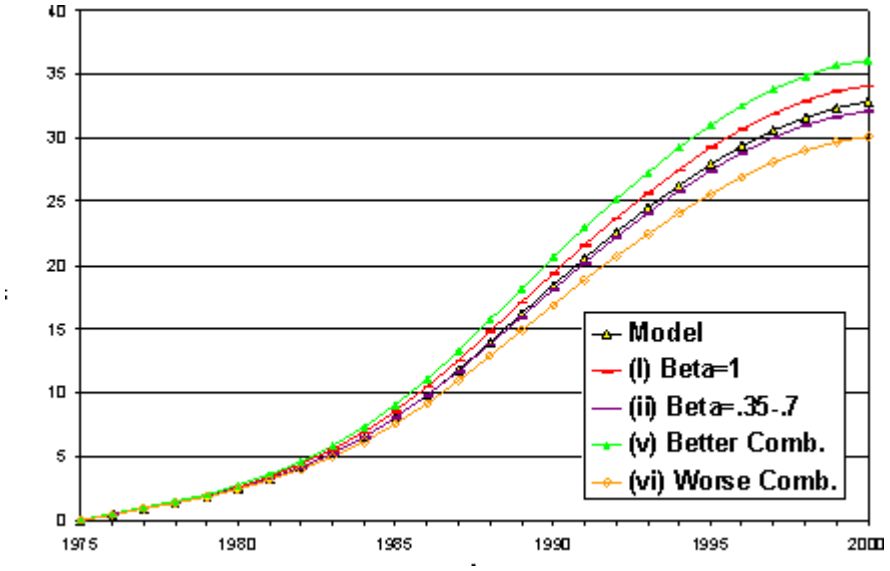
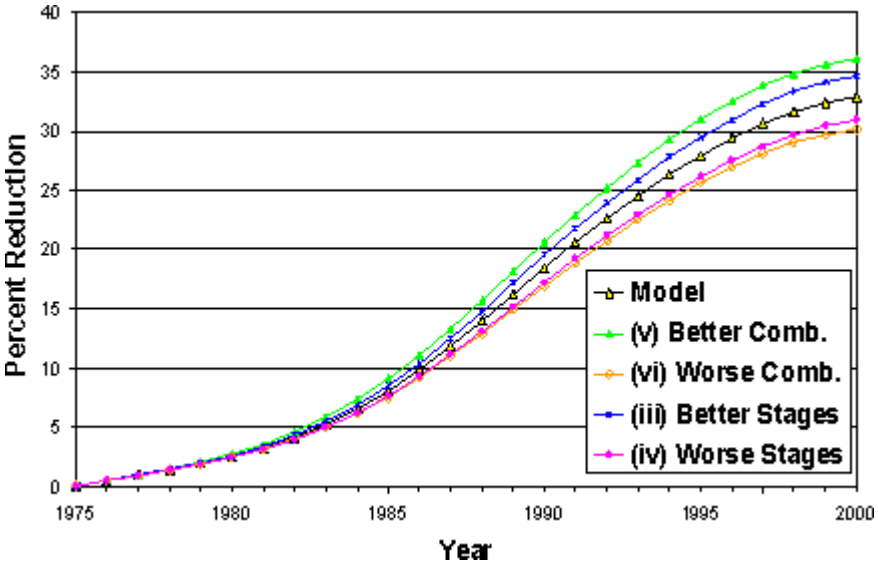


FIGURE 2. Sensitivity Analysis for Stage Shift



The curve labled "Model" in Figures 1 and 2 represents our final model prediction (equation 2.11). The other curves represent the changes by varying the input parameters. The curve labeled "Worse Comb" represents a combination of (ii) and (iv); and the curve labled "Better Comb" represents a combination of (i) and (iii). The magnitude of the reduction depends on the dissemination patterns. Generally the mortality reduction (MR) increases over time as screening and modern treatment become more widely disseminated in the population. Note that the MR ranged from 0% to 34% in the 25 year period 1975 to 1999 in the CISNET model.

As displayed in Figure 1, if the sensitivity of the screening examination was increased to 1, the MR increased. In the year 1999, it increased to 34% compared to 33% from the lower sensitivities in the base case model. When the mammogram sensitivities were

lowered, the MR in 1999 was lowered to 32%. Figure 1 also displays the better and worse combinations of mammogram sensitivities and stage shifts. In the year 1999, the maximum MR with a better combination of mammogram sensitivities and stage shift was 36% and the minimum MR was 30% with a poorer prognosis combination. Our sensitivity analysis indicates that the deviation from the model predictions was always less than 3%.

The mammogram dissemination patterns modeled by the NCI that indicated approximately 30% of U.S. women, who have started screening, followed a screening schedule of exams 5 years apart. In calculations, a combination of stage distributions from the BCSC estimate for screening with more than a one–year scheduling interval and SEER (1975–1979, no screening exams) was used for this group. The stage shift for this group has been changed to assess the impact of the stage distribution on MR. Figure 2 displays the results. A more favorable stage distribution was utilized by using the BCSC estimate of screening with a more than one–year screening interval; a less favorable stage distribution was utilized by using the stage distribution of SEER. The better stage shift improved the MR in the year 1999 to 35% and the control stage distribution lowered the MR in the year 1999 to 31%. Thus these sensitivity calculations show that there are small deviations between the model predictions and the predictions based on alternative stage shift distributions.

REFERENCES:

[1] Lee, SJ, Zelen, M. “Modeling the early detection of breast cancer” in Annals of Oncology 2003; 14: : 1199-1202

Predicted Mortality Reductions

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

Result Type

Validation

Result

Below is the predictions of the outputs for the eight randomized breast cancer early detection trials.

TABLE 1. Summary of Reported and Predicted Mortality Reductions (MR)

Trial	FU(yrs)	Reported	MR(CI)	Prediction MR
Malmo–1	19	19%	(0,34%)	17%
Stockholm	15	10%	(–28%,37%)	21%
Gothenburg	13.5	22%	(–7,43%)	20%
Ostergotland	17	11%	(–9,28%)	22%
Edinburgh	14	17%	(–18,42%)	11%
HIP	10	30%	(?)	3%
Canada–1 (40–49)	7	–36%	(–121,16%)	1%
Canada–2 (50–59)	7	3%	(–52,38%)	1%

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

KEY REFERENCES

Carne, PA, Miglioretti, DL, et al (2003) Individual and combined effects of age, breast density, and hormone replacement therapy on the accuracy of screening mammography in *Annals of Internal Medicine* 138:3, p 168–175

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

EBCTCG (1998a) Polychemotherapy for early breast cancer: an overview of randomized trials in *Lancet* 352, p 930–942

EBCTCG (1998b) Tamoxifen for early breast cancer: an overview of the randomized trials in *Lancet* 351, p 1451–1467

Lee, SJ, Zelen, M (2004) Mortality modeling of early detection programs in *(in manuscript)*,

Lee, SJ, Zelen, M. (2003) Modeling the early detection of breast cancer in *Annals of Oncology* 14, p 1199–1202

Mariotto, A, Feuer, EJ, Harlan, LC, Wun, LM, Johnson, KA, Abrams, J (2002) Trends in use of adjuvant multi–agent chemotherapy and tamoxifen for breast cancer in the United States: 1975–1999 in *JNCI* 94, p 1626–1634

Shen, Y, Zelen, M (2002) Screening sensitivity and sojourn time for breast cancer early detection trials: mammograms and physical exams. in *J Clin Oncol 2002*:19, p 3490–3499

Zelen, M, Feinleib, M (1969) On the theory of screening for chronic diseases in *Biometrika* 56, p 601–614

FLEXKB DOCUMENT
Version: HI.001.07242013.3520
Document generated: 07/24/2013

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

STANFORD UNIVERSITY

Important note: This document will remain archived as a technical appendix for publications. New versions will be added periodically as model refinements and updates are completed. The most current version is available at http://cisnet.cancer.gov/profiles. Note that unlike most PDF documents, the CISNET model profiles are not suitable for printing as they are not typically written or read in sequential fashion.

We recommend you let your interests guide you through this document, using the navigation tree as a general guide to the content available.

The intent of this document is to provide the interested reader with insight into ongoing research. Model parameters, structure, and results contained herein should be considered representative but preliminary in nature.

We encourage interested readers to contact the contributors for further information.

Go directly to the: Reader's Guide.

Stanford University
Readers Guide

READERS GUIDE

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

Core Profile Documentation

These topics will provide an overview of the model without the burden of detail. Each can be read in about 5–10 minutes. Each contains links to more detailed information if required.

Model Purpose

This document describes the primary purpose of the model.

Model Overview

This document describes the primary aims and general purposes of this modeling effort.

Assumption Overview

An overview of the basic assumptions inherent in this model.

Parameter Overview

Describes the basic parameter set used to inform the model, more detailed information is available for each specific parameter.

Component Overview

A description of the basic computational building blocks (components) of the model.

Output Overview

Definitons and methodologies for the basic model outputs.

Results Overview

A guide to the results obtained from the model.

Key References

A list of references used in the development of the model.

Further Reading

These topics will provide a intermediate level view of the model. Consider these documents if you are interested gaining in a working knowledge of the model, its inputs and outputs.

JNCIMonograph Outline

This topic provides links to profile content organized according to the JNCI Monograph Outline for Model Description Chapters. Use this outline for comparisons focused on the CISNET Base Case simulations.

Advanced Reading

These topics denote more detailed documentation about specific and important aspects of the model structure

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

MODEL PURPOSE

SUMMARY

The primary purpose of our simulation model is to explain the impact of breast cancer screening and treatment on US breast cancer incidence and mortality SEER trends from 1975–2000. We simulate breast–cancer specific events in the US female population and track related health outcomes using a natural history model of the disease.

PURPOSE

Our CISNET model (Model Overview) aims to reproduce population level breast cancer mortality rates from 1975 to 2000 by capturing breast cancer events that involve heterogeneity in disease progression, patient characteristics, compliance to screening and response to adjuvant treatment. The main purpose of our model is to quantify the impact of screening mammography and adjuvant therapy on breast cancer mortality trends from 1975 to 2000. In addition, our model can be extended to predict what the incidence and mortality trends would have been had alternative age–group been targeted for screening, had there been changes to the interval between screening examinations and/or changes to the groups targeted for adjuvant therapy.

Our specific aims are:

1. To develop and validate a stochastic natural history model of breast cancer.

We assume the tumor grows exponentially in the period that it can be screen detected. We assume that the growth rate is constant for a given individual but it can vary between individuals. We assume that the tumor progresses from local to regional to distant disease as it increases in volume. We aim to produce robust estimates of tumor growth rate distribution; the probability of the onset of the regional stage as a function of tumor volume; and the probability of the onset of the distant stage as a function of tumor volume.

2. To develop and validate a simulation model of the US female population undergoing screening and treatment.

We embed the natural history model described in Specific Aim 1 into a simulation

model of the US female population undergoing screening and treatment. For each individual, we simulate her age of birth, age and stage at breast cancer clinical detection, her age at breast cancer death and age at other–cause death. We assume that breast cancer death and other–cause death are independent, competing events. If the individual is screen detected, we will also simulate her age and stage at screen detection and age at breast cancer death from her screen detected tumor. We simulate individuals from every birth cohort since 1887 and aggregate the outcomes by calendar year. We estimate leadtime, lengthtime and overdiagnosis biases from screening.

3. To estimate the impact of breast cancer screening and treatment on SEER incidence and mortality rates.

We use the simulation model described in Specific Aim #2 together with a model of the dissemination of screening and treatment to estimate SEER trends in age–adjusted breast cancer incidence and mortality rates. We estimate: (i) the impact screening had on reducing breast cancer incidence and mortality, independent of treatment; (ii) the impact treatment had on breast cancer mortality, independent of screening; (iii) the impact treatment had as a result of early detection from screening programs.

* In this discussion, treatment refers to multi–agent chemotherapy and tamoxifen. It is assumed that breast cancer patients had mastectomy or breast conserving surgery followed by radiation.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

Model Overview

Summary

This document provides a background and basic description of the simulation model that we developed to estimate the impact of screening and treatment on US breast cancer incidence and mortality SEER trends since 1975.

Purpose

The purpose of this simulation model is to explain the impact of breast cancer screening and treatment on SEER trends in age–adjusted breast cancer incidence and mortality since 1975.

Background

In the United States, approximately one out of every eight women will develop breast cancer in their lifetime and one out of every 29 women will die from breast cancer. Cancer control programs on prevention, screening and treatment aim to reduce breast cancer mortality. Yet the impact of past and current cancer control programs are hard to interpret. SEER–based analyses demonstrate that age–adjusted breast cancer mortality rates have been approximately stable from 1975 to 1989. From 1990 to 2000, breast cancer mortality rates fell 1.7% per year. Breast cancer incidence rates climbed between 1977 and 1987, and have approximately been level between 1990 and 1996. DCIS incidence rates are increasing. There is no widely accepted explanation for these trends.

Views on the battle against breast cancer vary. It can be argued that our current breast cancer control programs in screening and treatment (with multi–agent chemotherapy and tamoxifen) are working. In particular, the increasing fraction of local disease, as well as DCIS, is a result of screening. The increased incidence of DCIS would be positive outcome if DCIS were known to be a precursor to invasive breast cancer. Other positive news is that 5–year and 10–year survival probabilities have increased and breast cancer mortality has started to decrease. On the other hand, it can be argued that the benefits of breast cancer control programs are not obvious. Survival benefits (measured from the time of diagnosis) may be due to leadtime, length time and overdiagnosis biases. The measured mortality decline since 1990 may be due to changes in treatment alone. This perspective would not say that the screening programs are not essential, but might suggest that the most substantial benefit of screening is local control of the primary disease as opposed to its life–threatening metastases. Given these two divergent perspectives, an analytically rigorous explanation of the trends is needed for effective, and cost–effective, cancer control programs in the future. We are developing a computer model to simulate breast cancer specific event in the US female population in order to explain the impact of screening and treatment on SEER–observed breast cancer incidence and mortality trends.

Model Description

Our basic simulation model can be described by the following algorithm:

For birth cohorts from 1887 to 1970

For each woman in the birth cohort

Generate her natural history of breast cancer
Compute her life history without screening and adjuvant treatment
Compute her life history with screening but without adjuvant treatment
Compute her life history without screening but with adjuvant treatment
Compute her life history with screening and adjuvant treatment

End
End

Our model provides estimates for population–level breast cancer mortality trends by simulating the life history of individual patients then aggregating the breast cancer related outcomes at the population level. Via the Monte Carlo method, the following characteristics are generated for an individual breast cancer patient: (1) the date of her birth, (2) the age of her death of causes other than breast cancer, (3) the ages she undergoes screening examinations, (4) the age she would be detected with invasive breast cancer in the absence of screening, (5) the age she would be detected with invasive breast cancer in the presence of screening, (6) her primary tumor size, extent of nodal and distant involvement and ER status at the time of detection in the presence and absence of screening, (7) the adjuvant treatment she received in the presence and absence of screening (it is assumed that she received primary therapy which would include surgery and possibly radiation) (8) her breast cancer survival time given her disease stage, size, age at detection and mode of detection, (9) her cause of death (i.e. breast cancer, other causes).

Our model generates information that could never be observed in a single patient. For example, if a woman was screen detected with invasive breast cancer in a given stage, we could not know when she would have been clinically detected and her disease stage at clinical detection. Similarly, if the patient treated died of other causes, we could not know when she would have died and her cause of death in the absence of screening and/or adjuvant therapy. Outputs such as these enable us to estimate the

survival and mortality benefit of screening and adjuvant therapy alone, as well as estimate leadtime and overdiagnosis effects of screening on breast cancer survival.
In order to generate the breast cancer outcomes for an individual breast cancer patient in the presence and the absence of screening and/or treatment, we model the natural history of the disease. In particular, we model the tumor size and SEER historic stage (defined as local, regional or distant) of patient's tumor at and before the moment that the tumor clinically surfaces. By "clinically surfaces," we mean the tumor is detected upon clinical examination because the patient experiences symptoms such as breast pain or nipple discharge. The simulation model traces the tumor from the moment it clinically surfaces "backwards" in time and provides estimates of the size and stage of the tumor at any time during the preclinical phase of the disease. A screening schedule that specifies the patient's age at the time of screening mammography is superimposed on the patient's disease history. A patient is screen detected only if the size of her tumor is at or above the tumor size detection threshold of mammography at the time of screening. Once the patient is detected, she is assigned a breast cancer specific survival time dependent on her age, tumor size, SEER historic stage mode of detection and her use of adjuvant treatment. Her age of death is the minimum age of breast cancer death and the age of other cause death. Individual level outcomes are aggregated and summarized as population level outcomes in terms of age–adjusted breast cancer incidence and mortality.

CONTRIBUTORS

PlevritisSylvia
RaiSanatan
GlynnPeter
SalzmanPeter
RosenbergJarret
SigalBronislava

ASSUMPTION OVERVIEW

SUMMARY

We describe the assumptions in our model.

BACKGROUND

In order for the model to be mathematically tracktable we need to make certain assumptions. We subdivide the assumptions we make into sevaral categories, i.e. the ones related to patients characteristics, tumor characteristics, screening, adjuvant treatment (Component Overview).
Some assumptions regarding background cancer incidence are required, since, for example, incidence for later years is contaminated by screening and hence background rates are unobservable. The natural history of breast cancer is also unobservable. Existing measurements of the growth rate are biased toward slower growing tumors. The onset of regional and metastatic disease is not known. Finally, the rate that the tumor advances in stage is not known. Size of the tumor at which it becomes detactable by mammography is unobservable as well and requires modeling.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

ASSUMPTION LISTING

Patient Characteristics

- The age– and cohort–specific incidence of breast cancer in the absence of screening is provided by NCI. It is assumed to be zero for women under 25 years of age and over 84 years. We treat this incidence as the hazard rate of the first cancer and apply it to our estimated cancer–free population.
- A woman is screen detected with breast cancer only if the size of her tumour at a screening exam is greater than the *screening threshold* (SD threshold). If this tumor would have been clinically detected after her natural death, then she is "overdiagnosed."
- Regardless of the mode of detection, breast cancer survival is based on SEER–derived breast cancer age, size and stage specific survival curves for cases detected in 1975–1979 (ie before screening).
- Death from causes other than breast cancer and death from breast cancer assumed to be two independent events.

Tumor Characteristics

- The tumour grows exponentially.
- The tumour volume doubling time has gamma distribution.

- The tumor begins in the local stage and progresses through the regional stage before enters the distant stage. We define the onset of the regional stage as the point at which nodal involvement first becomes detectable by techniques commonly used in clinical practice. Similarly, we define the onset of the distant stage as the point at which distant disease first becomes detectable by techniques commonly used in clinical practice. If the tumor is clinically detected before the onset of the regional or distant disease, it is staged as local disease. If the tumor is clinically detected after regional transition but before distant transition, it is staged as regional disease. If the tumor is clinically detected after the distant transition, it is staged as distant disease.
- The hazard of clinical detection at time t is proportional to the volume of the tumor at time $t, V(t)$. The onset of regional disease and the onset of distant disease are each modeled as the time to the first out of the two independent competing events. The hazard of the first event is constant over time and the hazard of the second event is proportional to the volume of the tumor at time $t, V(t)$. Clinical detection, and onset of the regional stage are independent of each other given the tumor volume doubling time. Onset of the distant stage is independent of the clinical detection given doubling time and tumor size at transition to regoinal stage.
- Tumor doubling time, hazard of clinical detection and stage transition are independent of the birth cohort.

Operating Characteristics of Mammography

- The detection threshold is defined as first tumor size that can be detected by the screening examination; tumors below this size will be missed and tumors above this size will be detected.
- We assume that detection threshold dose not depend on the year in which mammography is performed.

Mammography Dissemination

- Model for generating woman's ages at screening mammography is provided by NCI.

Treatment Efficacy

- We assume proportional benefits due to adjuvant treatment using hazard ratios published by Early Breast Cancer Trialists' Collaborative Group[2].

Treatment Dissemination

- Model for generating treatment (adjuvant chemotherapy, adjuvant tamoxifen, both or no treatment) recieved for given age, stage, size and ER status at detection is provided by NCI.

Breast Cancer Survival

- In the absence of screening and adjuvant treatment breast cancer cause specific survival stratified by tumor size, stage and patient's age at detection is assumed to be independent of the year of diagnosis.

REFERENCES:

1 Early Breast Cancer Trialists' Collaborative Group. “Tamoxifen for early breast cancer: an overview of the randomised trials.” in Lancet 1998; 351: 9114: 1451-67

2 Early Breast Cancer Trialists' Collaborative Group. “Polychemotherapy for early breast cancer: an overview of the randomised trials.” in Lancet 1998; 352: 9132: 930-42

PARAMETER OVERVIEW

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

SUMMARY

This document describes our modeling philosophy and the input parameters of our model.

BACKGROUND

Our model building process involves the following steps: we (1) decomposed the factors that drive population level breast cancer outcomes into distinct, less complex, physically meaningful components at the individual patient level; (2) analytically formulated the model components with a small number of parameters using basic biological, clinical and epidemiological principles and widely–held assumptions; (3) estimated the parameters from the best available national data; the parameters are classified as either: (i) observable parameters, meaning that they are directly observed, typically from clinical trials or (ii) unobservable parameters, meaning that they cannot be directly observed but may be estimated based on modeling assumptions; (4) merged all the components into a simulation algorithm that generates the population level breast cancer outcomes across a large range of calendar years; (5) validated the population level simulation model by measuring how well it reproduces data on breast cancer trends that were not used in parameter estimation or model calibration; (6) evaluated uncertainty in the estimated breast cancer trends due to uncertainty in the estimates of the parameters of the model components; (7) performed sensitivity analysis to modeling assumptions and (8) identified worthy refinements to the model should the appropriate data become available.

The step that is perhaps the most critical in our model building process is that of model validation. From the start of this project, we intended to validate the model by analyzing how well it reproduces the observed SEER incidence and mortality trends, particularly in the post–screening period. For this reason, we required that no parameter estimation should rely on calibrating to the observed population trends, particularly in the post–screening period. We needed to moderately relax this requirement in the current version of the model.

PARAMETER LISTING OVERVIEW

Patient Characteristics

- Age–specific incidence rate of breast cancer with secular trend across birth cohorts.
- Other–cause death by birth cohort
- Kaplan–Meier breast cancer specific SEER survival curves stratified by age, tumor size and stage.

Tumor Characteristics

- Mean growth rate is estimated together with mammography detection threshold by calibrating to SEER incidence and BCSC data on size distribution of screen detected cancers.

- Parameters of the natural history model, i.e. hazard of clinical detection, hazard of transition from local to regional and from regional to distant stages, are estimated from SEER 1975–1981 data (prescreening period) on tumor size and stage at detection using only cases which are the first malignant tumor in a patient.
- Probability of having ER+ tumor is estimated from SEER 1990–1994 data.

Mammography Operating Characteristics

- The distribution for the mammography threshold was modeled by assuming that the hazard function for "screen–detectability", i.e. the transition from a nonscreen detectable tumor to a screen detectable tumor, is proportional to the cross–sectional area of the tumor, which is in turn proportional to the tumor volume raised to the two–thirds power.

Mammography Dissemination

- Model for generating woman's ages at screening mammography is provided by NCI.

Treatment Dissemination

- Model for generating treatment (adjuvant chemotherapy, adjuvant tamoxifen, both or no treatment) recieved for given age, stage, size and ER status at detection is provided by NCI.

Treatment Efficacy

- We assumed proportional benefits due to adjuvant treatment using hazard ratios published by Early Breast Cancer Trialists' Collaborative Group (1998).

See also Component Overview .

COMPONENT OVERVIEW

SUMMARY

We describe the components of the model.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

OVERVIEW

The following briefly describes the model algorithm and the model components involved at each step.

We start by generating woman's date of birth and age at death from other causes given her year of birth. Then her age at clinical detection of breast cancer is generated. Given parameter estimates for the stage–shift natural history model and age at clinical detection, we sample from estimated distributions of the tumor volume doubling time, tumor size at detection and tumor size at onset of the regional and the distant stage conditioned on the doubling time. In the next step we compute tumor stage at clinical detection in the absence of screening. If the tumor is clinically detected before the onset of the regional disease, after the onset of the regional disease but before the onset of the distant disease, after the onset of the distant disease, it is staged as local, regional or distant respectively. Given age, size and stage at clinical detection we compute the age at breast cancer death following clinical detection using survival curves estimated from 1975–1979 SEER data. Screening schedule is generated using either model provided by NCI or custom schedule, i.e. screening interval can be kept constant or set arbitrarily. Tumor–size detection threshold of the screening examination is generated using approach in [ref]. Based on woman's screening schedule and detection threshold we compute age, tumor size and stage at screen detection. A patient is screen detected only if the size of her tumor is at or above the tumor size detection threshold of mammography at the time of screening. The tumor size at the time of screening is determined knowing the size of the tumor at clinical detection, the tumor volume doubling time and the difference in the woman's age at the time of screening and her age at the time of clinical detection in the absence of screening. Next, using the model provided by NCI we generate type of the adjuvant treatment recieved by the woman given her year of diagnosis, age at diagnosis, tumor size and stage. Once the patient is detected, she is assigned a breast cancer specific survival time dependent on her age, tumor size, SEER historic stage mode of detection and her use of adjuvant treatment. Her age of death is computed as the minimum of the age at breast cancer death and the age at death from other causes.

COMPONENT LISTING

Our model consists of six underlying components: population, breast cancer incidence, breast cancer survival, natural history model, screening intervention and adjuvant treatment intervention. The components themselves contain subcomponents, some of which are referred to as "CISNET base case inputs" because they were defined by CISNET Breast Cancer Working Group as common inputs that would facilitate comparison among the seven different models.

Below is the pseudocode for our simulation model.

For birth cohorts 1887:1970,
For individuals 1:2,000,000

Step 1: Generate date of birth (Population Component)
Step 2: Generate age at other–cause death given birth cohort (Population Component)
Step 3: Generate age at clinical detection (Cancer Incidence Component)
Step 4: Generate tumor volume doubling time (Natural History Component)
Step 5: Generate tumor size at clinical detection given the tumor volume doubling time (Natural History Component)
Step 6: Generate tumor size at the onset of regional and distant stage (Natural History Component)
Step 7: Compute stage of the tumor at clinical detection
Step 8: Generate age at breast cancer death following clinical detection given age and stage at clinical detection (Survival And Mortality Component)
Step 9: Generate the ages undergoing screening given birth cohort (Screening Component)
Step 10: Generate the tumor size detection threshold of mammography (Screening Component)
Step 11: Compute age, tumor size and stage at screen detection
Step 12: Generate type of adjuvant therapy (Treatment Component)
Step 13: Generate age at breast cancer death following screen detection given age and stage at detection (Survival And Mortality Component)
Step 14: Compute age of death as min{age of breast cancer death, age of other cause death} (Survival And Mortality Component)

Repeat for next individual
Repeat for next birth cohort

OUTPUT OVERVIEW

SUMMARY

We describe the output of breast cancer population simulation model.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

OVERVIEW

The following outputs are produced by our model to help answer CISNET questions: age specific annual breast cancer incidence, age specific annual breast cancer deaths, mid–year population. The different type of output is used to estimate operational mammography characteristics, and such unobservable quantities as lead time and overdiagnosis.

OUTPUT LISTING

Breast cancer incidence and mortality:

- Age–specific (5–year) annual counts of detected breast cancers.
- Age–specific (5–year)annual counts of breast cancer deaths.
- Age–specific (5–year) annual counts of breast cancer prevalence cases.
- Age–specific (5–year) annual counts of detected breast cancers by size and stage.
- Age–specific (5–year) annual mid–year population.

The above outputs could be used to compute age–specific and age–adjusted annual breast cancer incidence rates by year of diagnosis, breast cancer mortality rates by year of death and prevalence rates.

Screening program characteristics based on cancers generated in years 1975–2000:

- Mean leadtime by 5–year age groups.
- Mean overdiagnosis by 5–year age groups.
- Mammography detection rates by 5–year age groups for the first screen and for the all subsequent screens.
- Program sensitivity by 5–year age groups.

All of the above outputs could be produced under the four scenarios (where appropriate):

- Background risk only
- Treatment only
- Screening only
- Screening and treatment

RESULTS OVERVIEW

SUMMARY

In this section we present the results that could be derived from our program outputs as well as the results of sensitivity analysis.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

OVERVIEW

In order to answer CISNET base case question : "What are the contributions of screening and treatment to mortality reduction?", we predict age–adjusted breast cancer mortality trends from 1975 to 2000. We also perform one–way sensitivity analysis of the age–adjusted mortality trends with respect to changes in the various input parameters of our model.

RESULTS LIST

Fit to the Age–adjusted Mortality Trend for Years 1975–2000

Comparison of the predicted and actual age–adjusted breast cancer mortality rates from 1975 to 2000 as reported by the National Center for Health Statistics (NCHS) are shown in Figure 1 (Age Adjusted Mortality).

Uncertainty analysis

We assessed uncertainty in the annual breast cancer mortality due to the estimated uncertainty in the scaled parameters of the natural history model (see Uncertainty Sensitivity Analysis)

Sensitivity Analysis

Because our model predicts a higher mortality rate than observed and does not predict the continued decline in mortality after 1995, a sensitivity analysis was especially critical. We performed the following one–way sensitivity analyses:

1. Varying the secular trend in breast cancer incidence.
2. Adding a temporal improvement in mammography detection.
3. Adding a temporal trend to treatment efficacy.
4. Adding a temporal improvement in baseline survival.
5. Allowing a fraction of screen detected invasive tumors to be screen detected as DCIS.

For details see Uncertainty Sensitivity Analysis .

CISNET Base Case Result

In answering the CISNET base case question: "What are the contributions of screening and treatment to mortality reduction?", we predict age–adjusted breast cancer mortality trends from 1975 to 2000 under the following four scenarios:

1. in the absence of screening and adjuvant therapy

2. in the presence of screening only
3. in the presence of adjuvant therapy only
4. in the presence of both screening and adjuvant therapy.

See Base Case Results for details.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

POPULATION COMPONENT

DETAIL

Our population component specifies the birth cohorts for our population level simulation analysis. To reproduce the outcomes of women ages 30 to 84 years in the US from the years 1975 to 2000, a representative sample of women born in the US between the years 1887 and 1970 is generated. Each birth cohort consists of two million women, which we found was a sufficiently large number to reduce the variability associated with the Monte Carlo method. Even though factors such as population immigration and emigration are likely to vary the relative sizes of the birth cohorts, the size of each birth cohort is kept constant in our simulation because the incidence and mortality trends are reported as age–adjusted rates. Each woman is assigned a birth date and an age at death from other causes. Death from breast cancer and other–causes are assumed to be independent. The other–cause death rate is a NCI base–case input and based on the Berkeley Mortality Database which start with the 1900 birth cohort[2]. For birth cohorts before 1900, we assume 1900 other–cause mortality rates.

REFERENCES:

[1] The Berkeley Mortality Database. Department of Demography at the University of California, Berkely. "BerkeleyMortality"

[2] Rosenberg, M. "Annual probabilities of death from causes other than breast cancer." in Cancer Intervention and Surveillance Modeling Network; Base Case. Unpublished data.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

CANCER INCIDENCE COMPONENT

DETAIL

The breast cancer incidence component determines whether or not an individual from a particular birth cohort would become clinically detected with invasive breast cancer in the absence of the screening and other–cause mortality. The breast cancer patient is assigned the age at which her first primary invasive tumor clinically surfaces.
This component relies on a CISNET base case input commonly referred to as the "secular trend in breast cancer incidence" and is estimated from the historic Connecticut Tumor Registry (CTR) and SEER[1]. The base case input estimates breast cancer incidence (invasive and in situ) in the absence of screening for annual birth cohorts starting 1891 by single year of age, for ages 25 to 84 years. The base case incidence is assumed to be zero for women under 25 years old and over 84 years old.

We treat the base case incidence as the hazard rate of the first cancer and apply it to our estimated cancer–free population. For each birth cohort we interpret the given incidence per 100,000 women (h_a) as the hazard rate for age a . To reduce the execution time for multiple runs, we generate and sample from a distribution function for the clinical detection age (A_{BC}) at symptomatic detection of the first invasive breast cancer for each birth cohort in the following way:

$$P(A_{BC} \leq a) = 1 - \prod_{i=25}^{a} (1 - h_i/100,000)$$

where the age a is an integer. Because this is a discrete distribution function of the woman's age at the first symptomatic detection, we generate the exact age by assuming a uniform distribution within a year. The same calculation is made for all the birth cohorts.
Two limitations with the breast cancer secular trend exist. First, we are likely overestimating the true hazard of the first primary, particularly in the older age groups. The base–case incidence was approximated as the observed count of "new cancers" divided by the size of the mid–year population based on data from the CTR and SEER, whereas the true incidence is defined as the count of "first cancers" divided by the size of the cancer free population[2]. This approximation is made because SEER does not include the size of the cancer free population. It would produce the true incidence if women are equally at risk for breast cancer regardless of their history of breast cancer; however, the risk of breast cancer likely increases with a prior history of primaries as evidenced by the Gail model[3].
The second limitation with our use of base case incidence is that we had to modify it in order to estimate the incidence for clinically detected invasive cancers only. The base case input provides the estimated trend for the sum of clinically detected in situ and invasive disease. We adjusted it by removing an estimated proportion of in situ cases as a function of age, which was estimated from SEER 1975–1979 data. The same correction factor was applied to all birth cohorts.

REFERENCES:

[1] Holford, T.R., et al. "Changing Patterns in Breast Cancer Incidence Trends." in J Clin Epi. submitted.

[2] Merrill, R.M. and E.J. Feuer "Risk-adjusted cancer-incidence rates (United States)." in Cancer Causes Control 1996; 7: 5: 544-52.

[3] Gail MH,et al. "Projecting individualized probabilities of developing breast cancer for white females who are being examined annually." in J Natl Cancer Inst 1989; 81: 24: 1879-86.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

NATURAL HISTORY COMPONENT

OVERVIEW

Our natural history component specifies the size of a breast cancer patient's first primary invasive breast tumor and its SEER historic stage during the preclinical phase of the disease mechanism under which the first primary invasive tumor is clinically detected[1]. Herein, the preclinical phase is defined from the moment the tumor is invasive and 2mm in diameter to the moment it clinically surfaces. We trace the tumor size backwards in time from the time it would have been clinically detected. (We are not modeling the tumor forward in time by estimating the onset of the first malignant cell or the onset of a tumor mass of a fixed size.)

DETAIL

Doubling time distribution of the primary tumor

The tumor is assumed to be spherical and grow exponentially, at a constant rate, in the preclinical phase. The volume of the tumor at time t is expressed as $V(t) = c_0 \exp^{t/R}$, where the inverse growth rate R (which is doubling time divided by ln(2)) has gamma distribution with rate α and shape β.

Stage transition of the primary tumor

The disease is assumed to start in the local stage, and progress to regional and distant stages as it increases in size. We define the onset of the regional stage as the point at which nodal involvement first becomes detectable by methods commonly used in clinical practice. Similarly, we define the onset of the distant stage as the point at which distant disease first becomes detectable by techniques commonly used in clinical practice. If the tumor is clinically detected before the onset of the regional or distant disease, it is staged as local disease. If the tumor is clinically detected after regional transition but before distant transition, it is staged as regional disease. If the tumor is clinically detected after the distant transition, it is staged as distant disease.

The onset of regional disease and the onset of distant disease are each analytically modeled as the time to the first out of the two independent competing events. The hazard of the first event is constant over time and the hazard of the second event is proportional to the volume of the tumor at time $t, V(t)$. In mathematical terms the hazard of the time to onset of observable nodal involvement (T_N) has the form,

$$P(T_N \in [t, t+dt) | T_N \geq t) = (\eta_0 + \eta_1 V(t))dt + o(dt).$$

The hazard of the time to onset of observable distant metastasis (T_M) is

$$P(T_M \in [t, t+dt) \ | T_M \geq t, T_N = t_N)$$

$$= \begin{cases} (\omega_0 + \omega_1 V(t))dt + o(dt), & t \geq t_N \\ 0, & t < t_N. \end{cases}$$

We do not include a temporal trend in the stage–transitions due to the lack of data to support parameter estimation. Yet it is reasonable to assume that technology advancements have caused a stage migration[2].

Clinical detection function

The hazard of the time to clinical detection (T_D) is assumed to be proportional to the current tumor volume,

$$P(T_D \in [t, t+dt) | T_D \geq t) = \gamma V(t) dt + o(dt).$$

This clinical detection function was introduced previously for breast cancer[3]. We do not vary the clinical detection function by calendar year due to lack of data to support estimation. Yet the clinical detection function has probably changed over time, yielding smaller median tumor sizes as increasing numbers of women are becoming aware of early breast cancer symptoms through greater education and outreach programs.

Parameter estimation

Estimation of scaled rate parameters

In total, our natural history model has seven parameters: $\alpha, \beta, \gamma, \eta_0, \eta_1, \omega_0, \omega_1$ which are specified above. Maximum likelihood estimates for these parameters are based on SEER data of the tumor size and stage of invasive cancers that were clinically detected in the absence of screening[4]. Only tumors detected between 1975 and 1981, which represents a period of no to little screening, were considered, and of these only the breast tumor which is the first primary tumor in a women with multiple primaries were selected. Because we do not use any data that contains temporal information (such as age), the rate parameter estimates are dimensionless and scaled by the mean doubling time. The scaled natural history parameters estimates are stratified by age–groups (20 – 39 years old, 40 – 49 years old, 50 – 69 years old, 70 – 84 years old).

Estimation of the mean volume doubling time and screening detection threshold

We estimated two unobservable parameters, namely the median tumor size detection threshold of mammography and mean tumor volume doubling time simultaneously, by calibrating to the SEER incidence trends and data from the Breast Cancer Surveillance Consortium (BCSC)[5] using a two–step procedure. In the first step, each of five–year age–specific SEER incidence curves were smoothed with respect to the year of diagnosis using natural splines (SPLUS 6.1) in terms of the number of new cancers divided by the mid–year population. Using our simulation program we estimated incidence as a number of first cancers divided by the mid–year population minus the prevalence. Sum of squared difference between age–specific smoothed SEER incidence and simulated incidence was used as a goodness of fit measure, thus assuming the same weight for each age group and each calendar year. This measure was computed over the two–dimensional parameter grid with increments of 0.05 year for the mean doubling time and 0.05 cm for the median threshold. The mean doubling time was varied between 0.2 year and 1.1 year and median detection threshold was varied between 6mm to 12mm. Various combinations of the parameters produced similar

goodness of fit measures. In the second step, for each fixed threshold we selected the best mean doubling time. Using thus created "pairs" we selected the one that produced better fit to the median size at detection for screen detected cases in BCSC 1994–2000 data of cancers screen detected within three years of the previous screening mammogram for women 50–69 years old. The resulting estimates currently used in the program are 0.75 year for mean doubling time and 1.0 cm for the median threshold of screening mammography.

This estimation procedure counters our intention to avoid calibration to post–screening SEER trends. However, it is limited to two parameters and calibration to incidence trends only. Through sensitivity analysis, we found these two parameters largely impact the change in breast cancer incidence rate once screening was introduced. In future work, this calibration procedure can be avoided because, in theory, these parameters along with the other parameters underlying the natural history can be simultaneously estimated from screening trial data.

Natural history model validation

After estimating the scaled rate parameters, we generate the goodness–of–fit measure by comparing the observed versus estimated tumor size distribution and the stage distribution conditioned on tumor size, for each age–group.

REFERENCES:

[1] Plevritis, S.K., et al. "A Natural History Model of Stage Progression Applied to Breast Cancer." in (Submitted).

[2] Feinstein, A.R., D.M. Sosin, and C.K. "Wells, The Will Rogers phenomenon. Stage migration and new diagnostic techniques as a source of misleading statistics for survival in cancer." in N Engl J Med 1985; 312: 25: 1604-8

[3] Bartoszynski, R. "A modeling approach to metastatic progression of cancer Cancer Modeling." in Cancer Modeling, J.R.a.B. Thompson, B. W., Editor. 1987; : 237

[4] Surveillance, E., and End Results (SEER) Program (www.seer.cancer.gov) SEER*Stat Database: Incidence – SEER 9 Regs Public–Use, Aug 2000 Sub (1973–1998) "SEERStat" in National Cancer Institute, DCCPS, Surveillance Research Program, Cancer Statistics Branch.

[5] Ballard–Barbash, R., et al. "Breast Cancer Surveillance Consortium: a national mammography screening and outcomes database." in AJR Am J Roentgenol 1997; 169: 4: 1001–8

Survival And Mortality Component

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

Overview

Survival is determined by age, size, stage, mode of detection and use of adjuvant therapy. For clinically detected cancer cases (in the absence of screening and treatment), we use age, stage and cause–specific survival curves obtained from SEER for women detected between the years 1975–1979 for which breast cancer was the first primary tumor. This is an NCI base case input for tumor size categories

Detail

Baseline breast cancer survival

Breast cancer survival curves for patients detected in the absence of screening and adjuvant therapy are a CISNET base case input, referred to as baseline breast cancer survival curves. These are Kaplan–Meier estimates obtained from SEER for female breast cancer patients who were detected between the years 1975–1979 and for whom breast cancer was their first primary tumor. Because the period from 1975 to 1979 is associated with minimal levels of screening and adjuvant therapy, the baseline survival curves are assumed to only capture the effects of primary breast cancer treatment, namely surgery with the possibility of radiation. The curves are stratified by age at detection (30 – 39 years old, 40 – 49 years old, 50 – 59 years old, 60 – 69 years old, 70 – 84 years old), SEER historic stage (i.e. local, regional, distant), for local and regional stages curves are further stratified by tumor size (i.e. We do not include temporal variation in these curves, although improvements in primary treatment may have improved survival[1]. In addition, baseline survival may be impacted by temporal variation in the proportion of disease histology[2].

Breast cancer survival post screen detection without adjuvant therapy

The breast cancer survival curve post screen detection is taken to be the maximum of two curves: (1) the baseline survival curve that corresponds to the age, size and stage at screen detection and (2) the baseline survival curve that corresponds to the age, size and stage at clinical detection. Both survival curves are initiated at the corresponding age of detection and survival probability for curve that corresponds to clinical detection is set to be 100% during the leadtime. This approach rules out the possibility of death during the leadtime.

The assignment of breast cancer survival post–screen detection is arbitrary. Better breast cancer mortality outcomes would be obtained by using the baseline breast cancer survival curve that corresponds to the screen detected tumor characteristics initiated at the age of clinical detection. Worse outcomes would be obtained by using the baseline breast cancer survival curve that corresponds to the screen detected tumor characteristics initiated at the age of screen detection, because it would allow for death in the leadtime. In a sensitivity analysis, we found that these two extremes do not deviate significantly from the decision rule that we applied.

Breast cancer survival following adjuvant therapy with and without screening

We assume a proportional hazard reduction in breast cancer mortality due to adjuvant

treatment (Treatment Component). In the absence of screening, the hazard ratio is applied to the base case baseline breast cancer curves. In the presence of screening, the hazard ratio is applied to the resulting survival curve obtained earlier.

Other cause mortality

Other cause mortality and cause–specific mortality are assumed to be independent. Age and cause of death is then determined as the minimum of the cause–specific and other cause mortality.

What is the effect of not modeling in–situ disease?

DCIS is not included in our natural history model because there is little known about its progression. Some forms of in situ (in particular, high–grade DCIS) have been suggested to progress to invasive disease[3], but exactly what percent progresses and how fast it progresses is not known. For this reason, our model is limited to disease that would have been clinically detected as invasive. By not including clinically detected in situ disease we are implicitly making the assumption that DCIS does not contribute to breast cancer mortality. We are also not considering disease that would have been clinically detected as invasive but is screen detected as in situ. In our model, DCIS is likely screen detected as localized, small invasive tumor and as such would have good prognostic outcome. Should this assumption lack validity, we expect a poor prediction of breast cancer mortality since DCIS is a substantial fraction of incident breast cancer in the screening period[4].

REFERENCES:

[1] Bartelink, H., et al. “Recurrence rates after treatment of breast cancer with standard radiotherapy with or without additional radiation.” in N Engl J Med 2001; 345: 19: 1378-87

[2] Li, C.I., et al. “Trends in incidence rates of invasive lobular and ductal breast carcinoma.” in Jama 2003; 289: 11: 1421-4

[3] Kerlikowske, K., et al. “Characteristics associated with recurrence among women with ductal carcinoma in situ treated by lumpectomy.” in J Natl Cancer Inst 2003; 95: 22: 1692-702

[4] Ernster, V.L., et al. “Detection of ductal carcinoma in situ in women undergoing screening mammography.” in J Natl Cancer Inst, 2002. 94(20): p. 1546-54. 2002; 94: 20: 1546-54

SCREENING COMPONENT

OVERVIEW

Our screening component specifies the screening schedule of a given individual and the criterion upon which a patient is screen detected.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

DETAIL

Screening Dissemination Module

We use the CISNET base–case input for mammography dissemination to specify the ages at which a given individual undergoes screening, given her birth cohort.[1] The screening schedule is truncated at the age of clinical detection or death of other causes, whichever occurs first.

Screen Detection Mechanism

Each woman who receives at least one screening examination is randomly assigned a mammographic detection threshold. The mammography detection threshold is defined as smallest tumor diameter detectable on screening mammography. Tumors below this diameter will be missed and tumors above will be classified as screen detected if they have not clinically surfaced before the time of the screening examination. Because the tumor size increases between screening examinations, the probability of screen detection increases. Once a patient is screen detected, her age, tumor size, SEER historic stage at detection are recorded. A patient is classified as an "interval case" if her tumor is clinically detected between two scheduled screening examinations.
The distribution for the mammography threshold was modeled by assuming that the hazard function for "screen–detectability", i.e. the transition from a nonscreen detectable tumor to a screen detectable tumor, is proportional to the cross–sectional area of the tumor, which is in turn proportional to the tumor volume raised to the two–thirds power, i.e.

$$P(V_{TH} \in [v, v+dv) | V_{TH} \geq v) = \lambda v^{2/3} dv + o(dv)$$

In terms of the tumor diameter, the resulting cumulative distribution function is $\mathrm{F}_{TH}(d) = 1 - \exp^{-0.6\lambda d^5}$. In our simulations, the distribution was truncated at diameter d = 2mm i.e. we set $\mathrm{F}_{TH}(d) = 0$ for d
Our mammography detection function has the advantage that it is fully specified by one unobservable parameter, but the disadvantage that it produces a narrow distribution. A wider distribution is more plausible however, it would require an additional unobservable parameter that could not be identifiable from the available data. Temporal variation in the screening detection function was not modeled because its estimation would rely on data that are not available (other than data from the SEER post–screening period which is being reserved for model validation).

REFERENCES:

[1] Cronin, K.A., et al. "Modeling the dissemination of mammography in the United States. Cancer Causes Control, in press." in Cancer Causes Control, in press.

TREATMENT COMPONENT

OVERVIEW

Our adjuvant treatment component assigns adjuvant treatment and its corresponding survival benefit to a breast cancer patient.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

DETAIL

Treatment Dissemination

Each breast cancer patient is assigned adjuvant treatment (tamoxifen, multiagent chemotherapy, both, neither) depending on the patient's age, tumor size, disease stage, ER status, and year at detection, as specified by the NCI treatment dissemination base case input[1].

Tamoxifen dissemination targets ER–positive women in more recent years. Because ER status was not part of the natural history model, we assume that ER status does not vary over the preclinical course of the disease and does not impact the probability of screen detection. The probability that a given breast cancer patient is ER–positive was based on the proportion of women with ER–positive disease in the SEER data for years 1990–1994: the proportion of ER+ is 62%, 75%, and 83% for women

Survival Benefit from Adjuvant Treatment

We assumed proportional benefits due to adjuvant treatment using published hazard ratios[3]. For chemotherapy, the hazard ratio for the breast cancer specific survival depends on the age at detection: 0.72 for women 2 to convert it to a breast cancer specific mortality benefit. Ideally, this correction should depend an age and nodal status. If a woman receives both treatments, the product of hazard ratios is applied.

REFERENCES:

[1] Mariotto, A., et al. "Trends in use of adjuvant multi-agent chemotherapy and tamoxifen for breast cancer in the United States: 1975-1999." in J Natl Cancer Inst 2002; 94: 21: 1626-34

[2] Early Breast Cancer Trialists' Collaborative Group. "Tamoxifen for early breast cancer: an overview of the randomised trials." in Lancet 1998; 351: 9114: 1451-67

[3] Early Breast Cancer Trialists' Collaborative Group. "Polychemotherapy for early breast cancer: an overview of the randomised trials." in Lancet 1998; 352: 9132: 930-42

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

AGE ADJUSTED MORTALITY

A comparison of the predicted and actual age–adjusted breast cancer mortality rates from 1975 to 2000 as reported by the National Center for Health Statistics (NCHS) are shown in Figure 1. Because a measure for goodness of fit is not clear for our purposes, we proceed with a qualitative assessment. The general shape of the predicted mortality curve is similar to the actual curve. The mortality trend has a dominant downward trend in mortality approximately starting in the year 1990. However, two significant discrepancies exist between the modeled and actual curves. First, the predicted mortality rates are higher than the actual rates on an absolute scale. Second, the predicted mortality curve levels off starting in the year 1995 but the actual mortality curve shows a continued decrease. Both of these discrepancies were anticipated. The mortality is systematically higher than expected because the incidence for the first primary may be too high (see Cancer Incidence Component). The predicted trends are fairly flat after 1995 because there is little temporal variation in the model inputs just before and during this period. Also, as expected, we find that both of these discrepancies are most dominant among women over age 60 at death. The differences, if any, are minor among the younger women, for whom the incidence of breast cancer is relatively low.

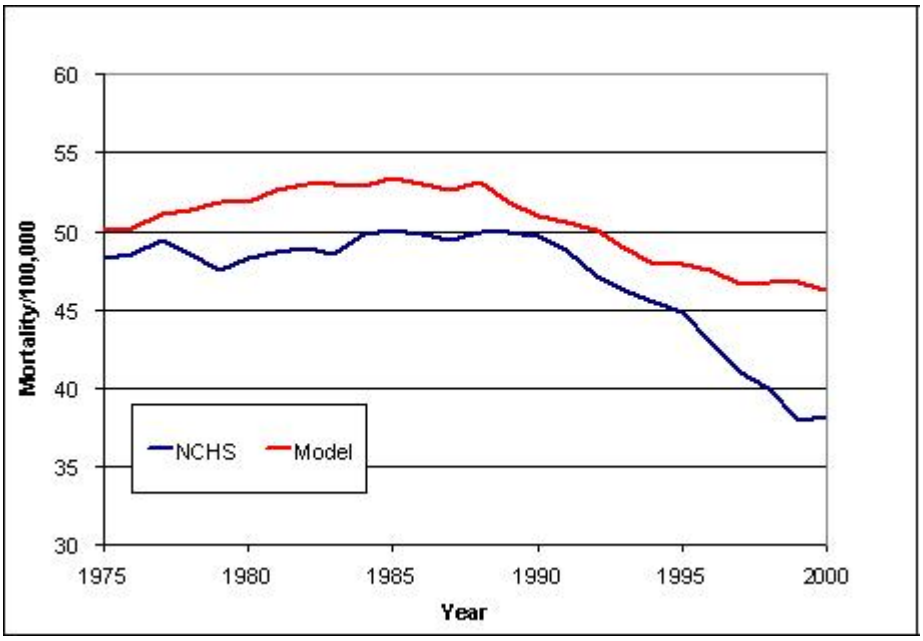


FIGURE 1. Comparison of age–adjusted US breast cancer mortality: NCHS data versus simulation model.

Uncertainty Sensitivity Analysis

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

Uncertainty Analysis

The uncertainty in the annual breast cancer mortality due to the estimated uncertainty in the scaled parameters of the natural history model (Natural History Component) is less than 1 death per 100,000 women. We estimated this uncertainty by bootstrapping the model with a new set of scaled natural history parameters. Parametric bootstrap was used to approximate joint distribution of the maximum likelihood estimates. The new parameter sets were generated by sampling from this joint distribution. This uncertainty does not fully account for differences between the modeled and actual breast cancer mortality.

Sensitivity Analysis

Because our model predicts a higher mortality rate than observed and does not predict the continued decline in mortality after 1995, a sensitivity analysis was especially critical. We proceed as follows:

1. *Varying the secular trend in breast cancer incidence:* Our breast cancer mortality rate may be too high because we may be overestimating the true hazard of first cancer (see Cancer Incidence Component). We adjust the number of new first cancer patients as follows: (1) we determine if a patient has been diagnosed previously given her age and stage, based on data from the CTR from the years 1975–1979; (b) if yes, then we return her to the healthy population. While this adjustment may underestimate incidence of first cancer, we find that it does not change the shape of the predicted mortality trend, but reduces its absolute level closer to the observed level, as shown in Figure 1(a). These desirable features warrant further investigation with alterations to the CISNET base case input for the secular trend in incidence.

1. *Adding a temporal improvement in mammography detection:* We introduced a stepwise change in the median tumor size detection threshold of mammography by reducing it from 1.0 cm to 0.5 cm at a specified calendar year and thereafter. The results from a step–wise change in the years 1985, 1990 and 1995 are shown in Figure 1(b). A noticeable reduction in breast cancer mortality begins approximately three years after the step–wise change. Despite the significant reduction in the detection threshold, the reduction in mortality is not large enough to account for continued mortality decline after the year 1995. For this reason, we do not suspect that we are significantly underestimating the benefit of screening mammography.

1. *Adding a temporal trend to treatment efficacy:* The efficacy in adjuvant treatment was assigned a stepwise change by improving the efficacy by 2 standard deviations based on published meta–analysis[2] in a specified calendar year and thereafter. The results following a stepwise change in 1985, 1990 and 1995 are shown in Figure 1(c). The response in mortality was immediate. If one were to consider a more gradual change in efficacy, such change in the 1990's could predict a continued mortality decline after 1995 without compromising the agreement between the predicted and actual mortality trends before 1995. Because a gradual improvement in multi–agent chemotherapy is likely due to changes in prescribed agents, it is possible that we are currently underestimating the benefit of adjuvant therapy in the later years. Future refinements to the treatment efficacy are warranted.

1. *Adding a temporal improvement in baseline survival:* We introduced a stepwise change in the baseline survival by forcing a 20% increase in a specific calendar year and thereafter. The results from a step–wise change in years 1985, 1990 and 1995 are shown in Figure 1(d). The response in mortality was immediate. Even a more modest improvement in baseline survival could significantly alter the shape of the predicted mortality curve before the year 1995 since it affects all cancer patients, so we do not expect it to be the major factor between 1975 and 1995 and hence after 1995.

1. *Allowing a fraction of screen detected invasive tumors to be screen detected as DCIS:* We assumed that a fraction p of tumors screen detected in local stage and below 1cm would be in situ disease with no risk of death from breast cancer. This was done by re–calibrating the remaining invasive cases to incidence while keeping mammography threshold fixed at 1 cm (see Natural History Component). As the percent of in situ disease varied from 5%, 10%, 20%, and 50%, the mortality reduction in the year 2000 varied from 0.2%, 1.1%, 1.4%, and 3.1% respectively. Even if 100% of all less than 1cm, local tumors were screen detected as in situ disease, breast cancer mortality in the year 2000 would decrease by only 6%, which is not large enough to explain the unaccounted for decline in mortality. We suspect that mortality is not greatly impacted because small, local tumors already have good prognosis.

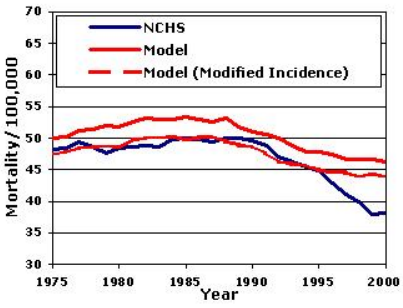


Figure 1(a). Comparison of the age-adjusted mortality: the observed data, the simulation model and the simulation model with the incidence modified.

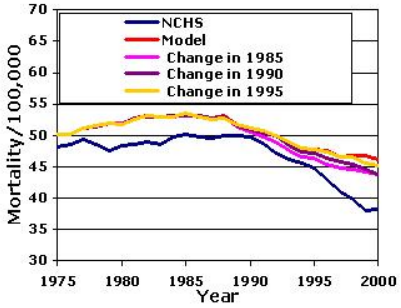


Figure 1(b). Age-adjusted mortality due to a stepwise change in mammography tumor size detection threshold in 1985, 1990 and 1995.

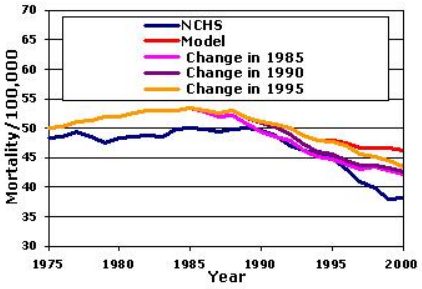


Figure 1(c). Age-adjusted mortality following a step-wise change in treatment efficacy in 1985, 1990 and 1995.

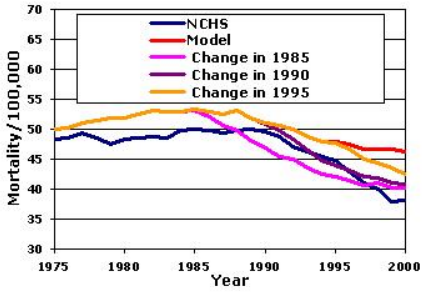


Figure 1(d) Age-adjusted mortality due to stepwise change in baseline breast cancer survival in 1985, 1990 and 1995.

REFERENCES:

[1] Early Breast Cancer Trialists' Collaborative Group. "Tamoxifen for early breast cancer: an overview of the randomised trials." in Lancet 1998; 351: 9114: 1451-67

[2] Early Breast Cancer Trialists' Collaborative Group. "Polychemotherapy for early breast cancer: an overview of the randomised trials." in Lancet 1998; 352: 9132: 930-42

BASE CASE RESULTS

CISNET Base Case Result

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

In answering the CISNET base case question: "What are the contributions of screening and treatment to mortality reduction?", we predict age–adjusted breast cancer mortality trends from 1975 to 2000 under the following four scenarios:

(1) in the absence of screening and adjuvant therapy,
(2) in the presence of screening only,
(3) in the presence of adjuvant therapy only, and
(4) in the presence of both screening and adjuvant therapy.

The results are shown in Figure 1. In the absence of screening and treatment, we predict a steady increase in age–adjusted breast cancer mortality due to the secular trend in incidence. Compared to the predicted mortality rate in the absence of screening and adjuvant therapy in the year 2000, the mortality rate in the presence of both screening and adjuvant therapy is reduced by a total of 29.9%, which is decomposed as follows: 16.9% due to screening, 6.9% due to chemotherapy and 8.9% due to adjuvant therapy. The estimated relative contributions of screening and adjuvant therapy to the mortality reduction were similar in magnitude: 53% due to screening versus 47% due to adjuvant therapy. Based on a sensitivity analysis, we found little difference in the relative contributions of screening and adjuvant therapy with the variation in the breast cancer secular trend (see Uncertainty Sensitivity Analysis). However, we may be underestimating the contribution due to adjuvant therapy given a likely temporal improvement in treatment efficacy (see Uncertainty Sensitivity Analysis). If we allow death in the leadtime (analysis not shown), we would be allowing the possibility that screening impacts only survival but not mortality, yet we still find a decline in breast cancer mortality.

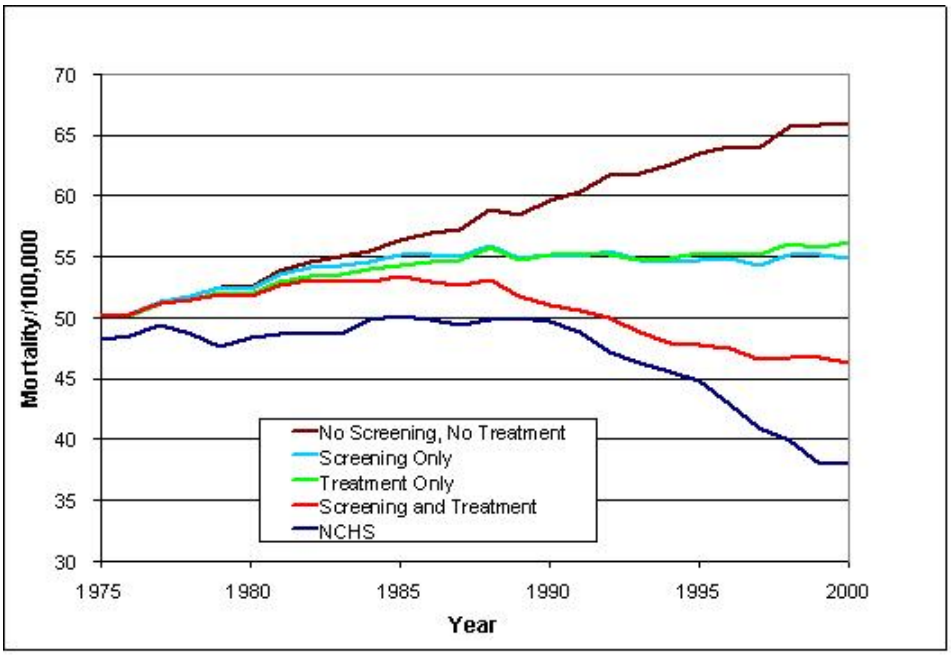


FIGURE 1. Simulated age–adjusted breast cancer mortality under four scenarios: (1) absence of screening and adjuvant therapy; (2) presence of screening only; (3) presence of adjuvant therapy only; and (4) presence of screening and adjuvant therapy. Observed NCHS age–adjusted breast cancer mortality is plotted for comparison.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

KEY REFERENCES

Ballard-Barbash, R., et al. (1997) Breast Cancer Surveillance Consortium: a national mammography screening and outcomes database. in *AJR Am J Roentgenol* 169:4, p 1001-8

Bartelink, H., et al. (2001) Recurrence rates after treatment of breast cancer with standard radiotherapy with or without additional radiation. in *N Engl J Med* 345:19, p 1378-87

Bartoszynski, R. (1987) A modeling approach to metastatic progression of cancer Cancer Modeling. in *Cancer Modeling, J.R.a.B. Thompson, B. W., Editor.*, p 237

Bartoszynski, R., et al. (2001) Modeling cancer detection: tumor size as a source of information on unobservable stages of carcinogenesis. in *Math Biosci* 171:2, p 113-42

Cronin, K.A., et al. Modeling the dissemination of mammography in the United States. Cancer Causes Control, in press. in *Cancer Causes Control, in press.*,

Early Breast Cancer Trialists' Collaborative Group. (1998) Polychemotherapy for early breast cancer: an overview of the randomised trials. in *Lancet* 352:9132, p 930-42

Early Breast Cancer Trialists' Collaborative Group. (1998) Tamoxifen for early breast cancer: an overview of the randomised trials. in *Lancet* 351:9114, p 1451-67

Ernster, V.L., et al. (2002) Detection of ductal carcinoma in situ in women undergoing screening mammography. in *J Natl Cancer Inst, 2002. 94(20): p. 1546-54.* 94:20, p 1546-54

Feinstein, A.R., D.M. Sosin, and C.K. (1985) Wells, The Will Rogers phenomenon. Stage migration and new diagnostic techniques as a source of misleading statistics for survival in cancer. in *N Engl J Med* 312:25, p 1604-8

Gail MH,et al. (1989) Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. in *J Natl Cancer Inst* 81:24, p 1879-86.

Holford, T.R., et al. Changing Patterns in Breast Cancer Incidence Trends. in *J Clin Epi. submitted.*,

Kerlikowske, K., et al. (2003) Characteristics associated with recurrence among women with ductal carcinoma in situ treated by lumpectomy. in *J Natl Cancer Inst* 95:22, p 1692-702

Li, C.I., et al. (2003) Trends in incidence rates of invasive lobular and ductal breast carcinoma. in *Jama* 289:11, p 1421-4

Mariotto, A., et al. (2002) Trends in use of adjuvant multi-agent chemotherapy and tamoxifen for breast cancer in the United States: 1975-1999. in *J Natl Cancer Inst* 94:21, p 1626-34

Merrill, R.M. and E.J. Feuer (1996) Risk-adjusted cancer-incidence rates (United States). in *Cancer Causes Control* 7:5, p 544-52.

Plevritis, S.K., et al. A Natural History Model of Stage Progression Applied to Breast Cancer. in *(Submitted).*,

Rosenberg, M. Annual probabilities of death from causes other than breast cancer. in *Cancer Intervention and Surveillance Modeling Network; Base Case. Unpublished data.*,

Surveillance, E., and End Results (SEER) Program (www.seer.cancer.gov) SEER*Stat Database: Incidence – SEER 9 Regs Public–Use, Aug 2000 Sub (1973–1998) in *National Cancer Institute, DCCPS, Surveillance Research Program, Cancer Statistics Branch.*,

The Berkeley Mortality Database. Department of Demography at the University of California, Berkely. ,

FLEXKB DOCUMENT
Version: HI.001.07242013.3520
Document generated: 07/24/2013

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

ERASMUS MC (BREAST)

Important note: This document will remain archived as a technical appendix for publications. New versions will be added periodically as model refinements and updates are completed. The most current version is available at http://cisnet.cancer.gov/profiles. Note that unlike most PDF documents, the CISNET model profiles are not suitable for printing as they are not typically written or read in sequential fashion.

We recommend you let your interests guide you through this document, using the navigation tree as a general guide to the content available.

The intent of this document is to provide the interested reader with insight into ongoing research. Model parameters, structure, and results contained herein should be considered representative but preliminary in nature.

We encourage interested readers to contact the contributors for further information.

Go directly to the: Reader's Guide.

READERS GUIDE

Core Profile Documentation

These topics will provide an overview of the model without the burden of detail. Each can be read in about 5–10 minutes. Each contains links to more detailed information if required.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

Model Purpose

This document describes the primary purpose of the model.

Model Overview

This document describes the primary aims and general purposes of this modeling effort.

Assumption Overview

An overview of the basic assumptions inherent in this model.

Parameter Overview

Describes the basic parameter set used to inform the model, more detailed information is available for each specific parameter.

Component Overview

A description of the basic computational building blocks (components) of the model.

Output Overview

Definitions and methodologies for the basic model outputs.

Results Overview

A guide to the results obtained from the model.

Key References

A list of references used in the development of the model.

Further Reading

These topics will provide a intermediate level view of the model. Consider these documents if you are interested gaining in a working knowledge of the model, its inputs and outputs.

JNCIMonograph Outline

This topic provides links to profile content organized according to the JNCI Monograph Outline for Model Description Chapters. Use this outline for comparisons focused on the CISNET Base Case simulations.

Advanced Reading

These topics denote more detailed documentation about specific and important aspects of the model structure

MODEL PURPOSE

SUMMARY

This page summarizes the model's purpose.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

PURPOSE

The MISCAN computer simulation program[2] has been developed for building models for cancer screening in a dynamic population, and for subsequently applying these models to analyze and explain results of cancer screening trials, to predict and compare the (cost–) effectiveness of different screening policies, and to monitor the results of population screening programs. MISCAN models have been made and applied for cancer of the cervix, breast, colon, and prostate[6]. In these standard MISCAN models, the natural history is described by defining discrete tumor stages, transition probabilities between these stages, and dwelling times in each stage. A problem of such a discrete disease stage model is that no clear distinction is made between local parameters that are specific to a situation in an area, and "biological" parameters that can be assumed to be equal in different areas. It also appeared to be difficult to explore assumptions about the natural history or other explanations for the differences between model results and observations, for example regarding the stage distribution of screen–detected cancer, interval cancers, and cancers diagnosed in case of no screening.

Therefore we decided to develop a more biologically oriented continuous tumor growth component as an alternative for the standard discrete stage natural history and screening component in MISCAN. In this alternative MISCAN breast cancer model, which is described here, a new component is used for the Natural History Component of invasive breast cancer and the effect of treatment and screening on survival. This 'Fadia' component simulates histories of tumors based on continuous tumor growth and the concept of a fatal diameter: each tumor has a size (the fatal diameter, which differs between tumors) at which diagnosis and treatment will no longer result in cure (reflecting the available treatment options), at this point the tumor enters the stage of fatal disease., i.e. one or more micro metastases exist for which treatment will not be

effective and will cause death from breast cancer for this woman. If the tumor is diagnosed (either on the basis of symptoms or by screening) and treated before the tumor reaches the fatal diameter the woman will be cured. In Fadia a distinction is made between tumor biology (tumor growth rate distribution) and model variables that may vary between areas and over time and / or age (diameter at clinical diagnosis, screening threshold diameter and fatal disease diameter, and survival). In the remainder "MISCAN–Fadia" will refer to the MISCAN version that includes this Fadia component, as described here; "standard MISCAN" will refer to the standard MISCAN model that can be used for simulating cancer screening as described by Loeve et al[2], and the "standard MISCAN breast cancer model" refers to the existing model for breast cancer with discrete tumor stages[4].

We also developed a cohort version of the MISCAN–Fadia population model and used it to estimate the parameters of the Fadia component using data from the Two County trial for breast cancer screening in Sweden[9]. See also Model Calibration Procedures

For the CISNET–Breast Base Case, the MISCAN–Fadia model was used to perform a series of model simulations for the years 1975–2000, including a background run assuming no screening and adjuvant treatment and runs with the assumed use of screening and / or adjuvant treatment during this period.

REFERENCES:

[1] Habbema JDF, van Oortmarssen GJ, Lubbe JTN, van der Maas PJ "The MISCAN simulation program for the evaluation of screening for disease" in Comput Methods Programs Biomed 1985; 20: : 79-93

[2] Loeve, F, Boer, R, van Oortmarssen, GJ, van Ballegooijen, M, Habbema, JDF "The MISCAN-COLON simulation model for the evaluation of colorectal cancer screening" in Comput Biomed Res 1999; 32: : 13-33

[3] van den Akker-van Marle, ME, van Ballegooijen, M, van Oortmarssen, GJ, Boer, R, Habbema, JDF "Cost-effectiveness of cervical cancer screening: comparison of screening policies" in J Natl Cancer Inst 2002; 94: : 193-204

[4] Boer, R, de Koning, HJ, van der Maas, PJ "A longer breast carcinoma screening interval for women age older than 65 years?" in Cancer 1999a; 86: : 1506-10

[5] Loeve, F, Brown, ML, Boer, R, van Ballegooijen, M, van Oortmarssen, GJ, Habbema, JDF "Endoscopic colorectal cancer screening: a cost-saving analysis" in J Natl Cancer Inst 2000; 92: : 557-63

[6] Draisma, G, Boer, R, Otto, SJ, van der Cruijsen, IW, Damhuis, RA, Schroder, FH, de Koning, HJ "Lead times and overdetection due to prostate-specific antigen screening: estimates from the European Randomized Study of Screening for Prostate Cancer" in J Natl Cancer Inst 95: : 868-78

[7] de Koning, HJ, Boer, R, Warmerdam, PG, Beemsterboer, PM, van der Maas, PJ "Quantitative interpretation of age-specific mortality reductions from the Swedish breast cancer-screening trials" in J Natl Cancer Inst 1995; 87: : 1217-23

[8] Tabar, L, Fagerberg, G, Duffy, SW, Day, NE, Gad, A, Grontoft, O "Update of the Swedish two-county program of mammographic screening for breast cancer" in Radiol Clin North Am 1992; 30: : 187-210

[9] Tabar, L, Vitak, B, Chen, HH, Duffy, SW, Yen, MF, Chiang, CF, Krusemo, UB, Tot, T, Smith, RA "The Swedish Two-County Trial twenty years later. Updated mortality results and new insights from long-term follow-up" in Radiol Clin North Am 2000; 38: : 625-51

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

Model Overview

Summary

This document provides an overview of the modeling effort, and describes the model itself in general terms.

Purpose

In the MISCAN–Fadia model knowledge on natural history, screening and adjuvant treatment practice and breast cancer risk derived from randomized controlled trials and observational studies will be integrated. In this way MISCAN–Fadia can be helpful in analyzing and explaining results of cancer screening trials, predicting the (cost–) effectiveness of different screening policies, and predicting the potential of present and new interventions on future national trends. See also Model Purpose .

Background

The MISCAN computer program has been used for building screening models for cancers of breast, cervix, colon en prostate [1,2,3,4,5]. In the standard MISCAN breast cancer model, the natural history is described as a semi Markov model. A problem of such a discrete disease stage model is that no clear distinction is made between local parameters that are specific to a situation in an area, and "biological" parameters that can be assumed to be equal in different areas. The Fadia natural history model is an alternative for the standard MISCAN breast cancer natural history model. It is based on continuous tumor growth instead of discrete tumor stages and on the concept of fatal diameter (a woman will die from breast cancer unless the tumor is detected before the tumor has reached the fatal diameter). In Fadia a distinction is made between tumor biology (tumor growth rate distribution) and model variables that may vary between areas and over time and / or age (diameter at clinical diagnosis, screening threshold diameter and fatal disease diameter, and survival).

Model Description

The MISCAN models use microsimulation: using the model inputs, independent life histories are generated including a possible cancer history and the effects of treatment and early detection by screening. Major differences between MISCAN–Fadia the standard MISCAN breast cancer model are: The MISCAN–Fadia model uses a continuous tumor growth model for the natural history of a tumor instead of a discrete stage natural history model;

- In the standard MISCAN models the screening test result depends on a stage–specific test sensitivity. In the MISCAN–Fadia model each tumor has a threshold diameter, which differs between tumors. If a tumor's diameter at the moment of screening is larger than this threshold the test result will be positive;
- In the standard MISCAN breast cancer model, the favorable effect of screening is relative to a woman's disease history without screening: a stage–specific proportion of screen–detected cancers will be cured. The MISCAN–Fadia model uses the fatal disease concept for modeling the survival of both clinically diagnosed and screen–detected cancers, and the diameter at which disease becomes fatal depends on the treatment given;

- The MISCAN–Fadia model allows for alternative adjuvant treatments that differ in the associated survival, and in their usage over time;
- In the standard MISCAN model it is possible to allow for multiple disease histories in a person, in the MISCAN–Fadia model each woman can only have one disease history;
- The MISCAN–Fadia model uses an external program for dissemination of screening.

Also see: Model Verification Procedures , Model Calibration Procedures , Model Validation Procedures

CONTRIBUTORS

- J Dik F Habbema a
- Rob Boer a,b
- Harry J de Koning a
- Gerrit J Oortmarssen a
- Sita Tan a

a Department of Public Health, Erasmus MC, University Medical Center Rotterdam
b RAND Corporation

ASSUMPTION OVERVIEW

SUMMARY

Summarizes the assumptions used in the MISCAN–Fadia model.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

BACKGROUND

The MISCAN–Fadia model can be used to simulate breast cancer screening and treatment policies in a dynamic population (see Model Purpose), based on assumptions on demography, natural history of breast cancer, screening and treatment. Compared to the other major model components (see Component Overview), the natural history component uses the most assumptions, as the natural history is modeled very detailedly.

ASSUMPTION LISTING

The MISCAN–Fadia model uses the following assumptions:

1. Demography Assumptions
2. Natural History Assumptions
3. Screening Assumptions
4. Treatment Assumptions

Limitations

The present version of the MISCAN–Fadia model has the following limitations:

1. only one tumor per woman
2. only one screening test
3. test result is completely determined by tumor size and the threshold for a screening test (no random variation in performance of the test or in reading the test result)
4. No ER status modeled

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

PARAMETER OVERVIEW

SUMMARY

Provides a complete overview of the parameters used to quantify the MISCAN–Fadia model.

BACKGROUND

The MISCAN–Fadia model consists of four basic components (see Component Overview): the Population Component , the Natural History Component , the Screening Component and the Treatment Component .

PARAMETER LISTING OVERVIEW

1. Demography Parameters (See also Population Component)
2. Natural History Parameters (See also Natural History Component)
3. Screening Parameters (See also Screening Component)
4. Treatment Parameters (See also Treatment Component)

COMPONENT OVERVIEW

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

SUMMARY

This document describes the major components of the model, their function and relative arrangement.

OVERVIEW

The MISCAN–Fadia model consists of four major components (see figure 1). The population component simulates the demography of the population, the natural history component simulates the natural history of a breast cancer tumor, the screening component simulates dissemination of mammography screening and its effects, and the adjuvant treatment component simulates dissemination of adjuvant treatment and its effects on survival and on breast cancer mortality

Figure 1 also shows the data used by the Cohort Model for estimation of the parameters of the Fadia natural history component, and data used by MISCAN–Fadia for producing the Base Case results.

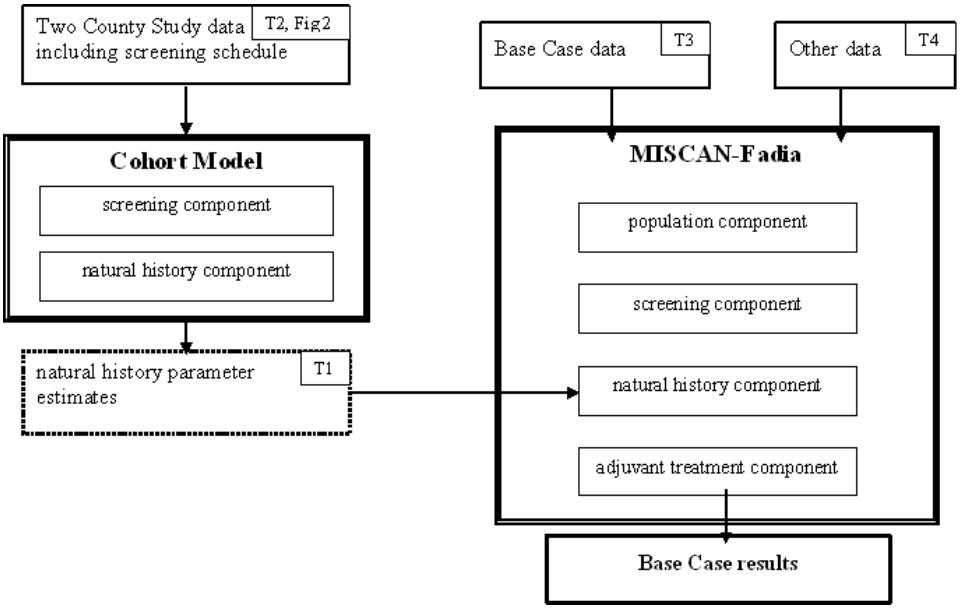


FIGURE 1: The two simulation models used for producing the Base Case results. The Cohort Model is used to estimate the parameters of the Fadia natural history of breast cancer, using the data from the Two County trial for breast cancer screening, by simulating the screening schedule of this trial. These natural history estimates are used in the MISCAN–Fadia population model, in combination with the Base Case data and other data, to run the simulations that produce the Base Case results for the US breast cancer incidence and mortality in the period 1975–2000. The labels T1...T4 refer to the tables that give an overview of the data used by the two models, Fig 2 refers to the survival data in Figure 2 (see Two County Study Result).

Component Listing

These are the primary components in the MISCAN–Fadia model:

- Natural History Component which simulates the natural history of a breast cancer tumor. In MISCAN–Fadia cancer incidence (see Cancer Incidence Component) and survival/mortality (Survival And Mortality Component) are a part of the Natural History Component .
- Population Component which simulates the demography of the simulated population
- Screening Component which simulates dissemination of mammography screening and its effect on the simulated population
- Treatment Component which simulates dissemination of adjuvant treatment and its effect on the simulated population

OUTPUT OVERVIEW

SUMMARY

Describes the outputs generated by the MISCAN–Fadia model.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

OVERVIEW

The MISCAN–Fadia model simulates the Base Case outputs.

OUTPUT LISTING

The output component produces the final output of the model:
(1) Incidence counts by calendar year (1975–2000), stage and age in five year age groups (30–84)
(2) Mortality counts by calendar year (1975–1999) and age in five year age groups (30–84)
(3) Population on July 1 of each calendar year (1975–1999) by age in five year age groups (30–84)
(4) Mean lead time by age (30–84, 30–39, 40–49, 50–59, 60–69, 70–84). Lead time is defined as the time from screen detection to the time a person would have been clinically detected in the absence of screening. Persons are excluded if they die of other causes during their lead time.
(5) Overdiagnosis percent by age (30–84, 30–39, 40–49, 50–59, 60–69, 70–84). Overdiagnosis percent is defined as the # of women who are screen detected who never would have been clinically detected / # of women who are screen detected .
(6) Overdiagnosis count by five year age group and calendar year
(7) Detection rate at first screen by age (30–84, 30–39, 40–49, 50–59. 60–69, 70–84). Detection rate is defined as cancers detected / # of women screened
(8) Detection sate at second and later screen by age (30–84, 30–39, 40–49, 50–59. 60–69, 70–84). Detection rate is defined as cancers detected / # of women screened
(9) Program sensitivity* by age at screening using one year interval (ages 30–84, 30–39, 40–49, 50–59, 60–69, 70–84).
(10) Program sensitivity* by age at screening using two year interval (ages 30–84, 30–39, 40–49, 50–59, 60–69, 70–84).

- Each case diagnosed within an age range can be classified as screen detected, clinically detected with a negative screening exam within the defined interval before detection (interval case), or clinically detected with no screening exam within the defined interval before detection (not included in the calculation). Program sensitivity = (# screen detected)/(#screen detected + # interval cases)

RESULTS OVERVIEW

SUMMARY

This document lists various results generated by the model.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

OVERVIEW

The main results from the MISCAN–Fadia model are the results for the breast Base Case analysis. Another important analysis using the Fadia natural history component was the Two County Study analysis (see Model Calibration Procedures). This analysis led to two important results. First, it gave us estimates for the parameters of the Fadia natural history component (see Natural History Component) that were used in order to produce the Base Case results. Second, the Two County Study analysis gave us more insight in the (dis)advantages of using a biogically grounded natural history component.

RESULTS LIST

- Two County Study Results
- Base Case Results

NATURAL HISTORY COMPONENT

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

SUMMARY

This document describes the Natural History generation portion of the microsimulation.

OVERVIEW

We will first describe the Fadia natural history component. Then we will describe how we estimated its parameters based on data from the Two County study using the Cohort Model and describe the results. Next, we will describe how the Fadia natural history component was adapted for the Base Case analysis.

The continuous tumor growth natural history model

The Fadia natural history component simulates invasive tumors, as well as ductal carcinoma in situ (dcis). In the sub–component for dcis, as in the standard MISCAN breast cancer model, three different types of dcis are assumed: regressive dcis, dcis that will be diagnosed clinically and dcis that will progress to invasive disease.

In Fadia, invasive breast tumors are initiated and are assumed to have a constant growth rate, which differs between tumors. Tumors also differ in the size (the fatal diameter) at which diagnosis and treatment will no longer result in cure (reflecting the available treatment options). At this point the tumor enters the stage of fatal disease. Clinical diagnosis of the tumor is triggered by two competing risks: by signs or symptoms resulting from the primary tumor, or by symptoms related to distant metastases. The probability of primary tumor related signs or symptoms is assumed to depend on the diameter of the primary tumor. The probability of distant metastases related signs or symptoms is assumed to depend on time since the disease became fatal. If the disease is already fatal at the moment of diagnosis of the tumor, the time of death from breast cancer is described by a probability distribution for the survival time since the start of fatal disease. This time between start of fatal disease and death from breast cancer applies both to the case in which the breast cancer is diagnosed clinically and to the case where this cancer is detected by screening.

DETAIL

The life course of a tumor is described by the following five variables, which are governed by probability distribution functions with two parameters each (scale and shape), and a sixth variable with one parameter:

1. Growth rate of the tumor (lognormal distribution with parameters μ_1 and σ_1);
2. Fatal diameter of the tumor (weibull distribution with a scale and a shape parameter);
3. Survival time after reaching the fatal diameter (lognormal distribution with parameters μ_3 and σ_3);
4. Threshold diameter of a tumor for a screening test, i.e. the tumor diameter at which a tumor becomes screen detectable (Weibull distribution with a scale and a shape parameter);

5. Tumor diameter at clinical diagnosis because of the primary tumor (lognormal distribution with parameters μ_2 and σ_2);
6. Moment at which distant metastases lead to clinical diagnosis of the tumor, modeled as a constant fraction of the survival time after reaching the fatal diameter (with this fraction as parameter).

In order to obtain a reasonable fit of the Two County Study screening trial data, we had to assume that three of the model variables—the tumor growth rate, the tumor diameter at clinical diagnosis, and the survival after inception of fatal disease—are correlated. This adds three more parameters to the model: the correlation ρ_1 between tumor growth rate and the survival time after reaching the fatal diameter, the correlation ρ_2 between tumor growth rate and the tumor diameter at clinical diagnosis because of the primary tumor, and the correlation ρ_3 between the tumor diameter at clinical diagnosis because of the primary tumor and the survival time after reaching the fatal diameter.

The tumor history model is thus characterized by the values of these 14 parameters (five pairs, one fraction, and three correlations). In MISCAN–Fadia, changes in the survival over time or as a result of improved treatment are modeled as a shift in the fatal diameter (parameter μ_3), and changes in the screening test sensitivity are modeled as a time dependent scale parameter of the threshold diameter distribution.

When a breast tumor is initiated in a simulated woman, values of the six tumor variables are generated. For each simulated tumor, the clinical diagnosis diameter is determined by the minimum of the diameter at clinical diagnosis because of the primary tumor and the diameter at clinical diagnosis because of metastases. The growth rate of the tumor then determines the times since its initiation at which it reaches the fatal diameter, the clinical diagnosis diameter, and the threshold diameter for a screening test. If the clinical diagnosis diameter is larger than the fatal diameter, then the sum of the time at which the fatal diameter is reached and the survival time after reaching the fatal diameter will give the time at which a woman will die of breast cancer. The woman will be cured if the cancer is detected, either clinically or by a screening test, before the fatal diameter is reached. For a woman with a tumor, the result of a screening test is completely determined by the tumor diameter and the threshold diameter for this test for this tumor, i.e. no allowance is made for random variation in performance of the test or in reading the test result.

Relevant Assumptions

See Natural History Assumptions

Relevant Parameters

See Natural History Parameters

Relevant Components

- Cancer Incidence is described in the Cancer Incidence Component
- survival mortality is decribed in the Survival And Mortality Component

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

Model Calibration Procedures

Overview

The main calibration procedures of the model are the one in which the parameters of the Cohort Model have been estimated from the Two County screening trial and the one in which th MISCAN–Fadia model is calibrated for the Base Case analyses. Other calibration procedures have been described in the Cancer Incidence Component .

Estimation of natural history model parameters using results from the Swedish Two County breast cancer screening trial

For estimation of the parameters of Fadia component, we used a simplified cohort version that allows for efficient estimation of model parameters using data from screening trials. This Cohort Model focuses on the natural history of invasive breast cancer tumors and the effect of screening in a cohort of women that participate in a screening trial; it does not include age (and thus neglects age–dependencies), it only includes tumor size as tumor attribute (e.g. neglecting distant metastases and lymph node status) and also neglects death from other causes, dcis, time trends in breast cancer incidence or survival, and the impact of adjuvant treatment. The Cohort Model also uses microsimulation, but it simulates screening in a cohort instead of in a full dynamic population. Moreover, it simulates tumor histories instead of life histories of women. The Cohort Model was used to estimate model parameters from the results reported by the Swedish Two County breast cancer screening trial (TCS)[3]. The tumor histories simulated by the Cohort Model are used to generate output on detection rates at successive screening rounds, interval cancer rates, tumor diameter distribution of screen detected cancers, of interval cancers and of cancers diagnosed in the control group, on survival by time since diagnosis for screen detected cancers, interval cancers and cancers diagnosed in the control group, and on survival by tumor diameter and by time since diagnosis.

The TCS started in October 1977 in Kopparberg and in May 1978 in Östergötland[1]. Women aged 40–74 were randomized to either the study group, consisting of 77,080 women, or control group, consisting of 55,985 women. Women in the study group were invited to mammography screening; women aged 40–49 were invited every 24 months and women aged 50–74 every 33 months. In our analysis, we used data from women aged 50–69 at entry ([1],[2], personal communication). The follow–up period after the last screening round ended on average 8 years after start of the study. At that moment women in the control group were invited for a screening examination too. Data on cancers detected at this screening are not included in the estimation of the model parameters.

Tumors are assumed to initiate with a diameter of 0.1 mm with constant onset rate of 2.2 per 1000 women years, which is the observed incidence rate in the control group[1]. Predicted detection and interval rates are corrected for the aging of the women during the trial, to adjust for the fact that age is not incorporated in the Cohort Model whereas breast cancer incidence increases by age. For given values of the model parameters, a single micro–simulation run will produce expected values (rates or proportions) for each of the results of the TCS study. Maximum likelihood estimates of the model parameters are derived by repeated evaluation of the simulated histories using the Score Function (SF) method in combination with a quasi–Newton optimization

procedure[5]. With respect to the likelihood of the model, the screening data considered are either very small proportions of breast cancer cases, e.g. detection rates at screening and interval cancer rates, or distributions of breast cancer cases over sub categories, e.g. tumor stage distribution of screen detected cancers and interval cancers. The observed numbers of cases were assumed to be governed by a Poisson and a multinomial distribution, respectively.

The goodness of fit of the model is calculated using the deviance, which is defined as minus two times the difference in log likelihood between the expected model and the saturated model (i.e. the best possible model which takes the observed numbers as expected values for the Poisson and multinomial distributions).

Initially, we fitted the Cohort Model to the TCS data assuming Weibull probability distribution functions for all variables. However, when it became apparent that correlation had to be assumed between some of the variables, a switch was made to lognormal distributions that are more convenient in this respect. Thus, lognormal distributions were used for correlated variables: the tumor growth rate, the tumor diameter at clinical detection and survival time after reaching the fatal diameter.
See also Two County Study Result .

Quantification of the MISCAN–Fadia model for the Base Case analyses

The Cohort Model quantification as based on the Two County Study data was used as starting point for the quantification of the natural history model parameters of MISCAN–Fadia. In order to quantify MISCAN–Fadia for the Base Case, we calibrated some of these natural history model parameters to Base Case data (see Table 3), described in *Calibration of natural history model parameters* below, and we made some extensions to the model, described in *Extensions of the MISCAN–Fadia natural history component* below. Table 4 gives an overview of how Base Case data were used for the quantification of MISCAN–Fadia.

TABLE 3. Overview of MISCAN–Fadia parameters that are based on other data than the Base Case data.

MISCAN–Fadia parameter	Method	Data used	value
tumor growth rate	Estimated with Cohort Model	TCS	Table 1
survival duration	Estimated with Cohort Model	TCS	Table 1
screening threshold	Estimated with Cohort Model + estimation of trend 1975–2000	TCS, HIP	Tables 1 and 6
Correlations between growth rate, diameter at clinical diagnosis because of primary tumor, and survival	Estimated with Cohort Model	TCS	Table 1
dcis duration and progression		Dutch	Table 8
dcis survival			100%

TABLE 4. Base Case Data Usage. U = Used as provided by Cisnet; P = Uses a processed version of the Base Case data; C = Model is calibrated to the Base Case data by varying a parameter of the continuous tumor growth model; O = is determined by other Base Case data used in the model

Base Case Data	Usage	Notes
Treatment Dissemination	U	Used as direct input
Mammography Dissemination	U	Base Case Dissemination model was directly used as external program
Other Cause Mortality	U	Used to calculate size of birth cohorts in the US population, see Table 11 in Population Component
SEER incidence	C	1975 size specific incidence was used to calibrate parameters for tumor diameter at diagnosis because of primary tumor, see Table 5.
1975 Breast Cancer Prevalence	O	Results from cohort risks, age specific distribution of incidence and calibration of fatal diameter to 1975 survival and time trend in fatal diameter prior to 1975
1975 Cause Specific Survival	C	Used to estimate 1975 fatal diameter parameters, see Table 5.
Historical Survival	C	Used to estimate time trend in fatal diameter prior to 1975, see Table 5
1975 Stage Distribution	C	Used to estimate AJCC stage distribution parameters, see Table 7
1975 Breast Cancer Mortality	O	Results from cohort risks, age specific distribution of incidence and calibration of fatal diameter to 1975 survival and time trend in fatal diameter prior to 1975
Breast Cancer APC Incidence	P	Converted to Age–Cohort model with cumulative incidences as cohort risks and one fixed age specific distribution of incidence of pre–clinical screen–detectable disease for all cohorts, see Tables 9 and 10.
Treatment Effect	C	Calibrated by a shift in fatal diameter, see Table 13 in Treatment Component
SEER 9 Mortality	O	Results from cohort risks, age specific distribution of incidence, calibration of fatal diameter to 1975 survival, and time trend in fatal diameter prior to 1975, dissemination of mammography and adjuvant treatment

TABLE 5. MISCAN–Fadia. Parameters of the distributions for the fatal diameter and for the diameter at clinical diagnosis because of the primary tumor.

Variable	distribution	year	par1	par2	mean	st.dev.
fatal diameter (cm)	Weibull (scale, shape)	1915	0.82	0.95	0.84	0.88
		1975	4.02		4.11	4.33
clinical diagnosis because of the primary tumor(diameter, cm)	Lognormal (ρ, σ)	(all)	0.97	0.63	3.22	2.25

Calibration of natural history model parameters

The diameter at clinical diagnosis because of the primary tumor was calibrated to the 1975 stage distribution as provided by the Base Case SEER incidence data, and the fatal diameter was calibrated to 1975 Cause Specific Survival Base Case data. These model parameters were calibrated simultaneously since they both influence the stage distribution as well as the survival, see Table 5.

Extensions of the MISCAN–Fadia natural history component

The quantification obtained by the Cohort Model was extended by including a calendar time dependency of the fatal diameter, and an age and calendar time dependency of the screening threshold diameter. The tumor diameter distribution was extended to an AJCC stage distribution, and the effect of adjuvant treatment was included as a change in the scale parameter of the fatal diameter distribution. The quantification of the dcis part is equal to that used in the standard MISCAN breast cancer model. The Base Case APC incidence data was simplified to an age–cohort model, and then used to calculate age–specific onset of pre–clinical screen–detectable disease.

The quantification of the fatal diameter has been extended in the MISCAN–Fadia model: the scale parameter of the Weibull distribution for the fatal diameter has been made dependent on the year of diagnosis, accounting for the improvement of treatment. Assuming no improvements in treatment other than adjuvant treatment after 1975, we only modeled a time dependency prior to 1975. The quantification of this time dependency was based on the hazard ratio between 1940–1949 and 1970–1974 20 years survival, using the *Historical Survival Base Case* data, which in turn is based on Connecticut data.

In the MISCAN–Fadia model, a hazard ratio cannot be applied directly because survival is described by lognormal survival distribution for women (proportion: 1–c_{1975}) in whom the diameter at diagnosis exceeds the fatal diameter, and cure for the proportion 1–c_{1975} of women in whom the diameter at diagnosis is smaller than the fatal diameter. So we translated the hazard ratio r between 1940–1949 and 1970–1974 into a shift in fatal diameter between 1975 and 1945. This shift in fatal diameter from 1975 to 1945 leads, in combination with the distribution of the clinical diagnosis diameter—which is modeled constant over time—to a new cure proportion c_{1945}. We approximated the 1945 cure proportion c_{1945}, using the hazard ratio r between 1940–1949 and 1970–1974, the 1975 cure proportion c_{1975} and the probability distribution function $F(t)$ for the survival time since the moment at which the tumor reached its fatal diameter:

$$c_{1945} = \frac{\{1 - (1 - c_{1975}F(t)\}^{1/r} + F(t) - 1}{F(t)}$$

Based on the 1945 cure proportion c1945 and the distribution of the clinical diagnosis diameter, we calculated the value of the scale parameter of the fatal disease diameter for 1945 that corresponds to the 1945 cure proportion. Linear interpolation is applied from 1975 to 1945 and this time trend is extrapolated backwards to 1915, which is the first possible year of onset for the oldest women. The quantification of the threshold diameter for screen detection as estimated from the Two County trial data using the Cohort Model has been extended in the MISCAN–Fadia model: the scale parameter of the Weibull distribution for the threshold for screen detection has been made dependent on the year and the age of the woman at the moment of screening, using four age groups The stage distribution of invasive tumors was extended to AJCC stages by adding three more variables to the model and then calibrated to the Base Case 1975 stage distribution data (see also Table 7):

- the tumor diameter at which a N1 lymph node disease becomes detectable by modern techniques
- the difference in tumor diameters at which a N1 and a N2 lymph node disease become detectable by modern techniques
- The time at which distant metastases become detectable by modern techniques, modeled as a fraction of the time between the moment at which the tumor reaches the fatal diameter and the death from the breast cancer

Note that in MISCAN–Fadia, detectable distant metastases are assumed to be fatal. MISCAN–Fadia does not model the moment of initiation of distant metastases; it only models the moment at which distant metastases become detectable by modern techniques and the moment at which distant metastases lead to diagnosis of the primary tumor—the first event always preceding the latter.

The MISCAN–Fadia model includes a discrete disease state for dcis. The submodel for dcis was taken from the standard MISCAN breast cancer model[7], which was based on data from the screening trials in Utrecht and Nijmegen (The Netherlands). In this submodel, there are three different types of dcis: regressive dcis, dcis that will be diagnosed clinically and dcis that will progress to invasive disease. All types of dcis have a mean duration of 5.22 years. The distribution among the different types of dcis depends on age; the durations do not depend on age. (see Table 8).

In the MISCAN–Fadia model, the onset of invasive disease is defined as the minimal value of the threshold diameter over ages and time, i.e. the value at age 80 and in the year 2000. The minimal value of the threshold diameter thus has a Weibull distribution with scale parameter 0.65 (Table 6; age 80, year 2000) and shape parameter 2.95 (see Table 1 in Component Overview). We assumed the duration of dcis not to change over time for regressive dcis and for dcis that will be diagnosed clinically. For the duration of dcis that progresses to invasive disease—which is equal to the duration from onset of dcis to the moment the tumor reaches the minimal threshold size—we assumed the MISCAN quantification to hold for 1975. To this end, we adapted the mean duration of dcis that progresses to invasive disease in order for the sum of the mean duration of dcis that progresses to invasive disease and the mean duration between the minimal value of the threshold diameter and 1975 threshold diameter to match the quantification of the standard MISCAN breast cancer model (5.22 years, see Table 8).

For screening of dcis, the MISCAN–Fadia model also uses the same mechanisms as the standard MISCAN breast cancer model. The probability of screen detection of dcis is modeled through a test sensitivity parameter. The standard MISCAN quantification for the sensitivity of dcis (0.4) is used for 1975, which is 0.4, and has been made time dependent, increasing linearly to 0.8 in 2000. Screen detected dcis is, just like clinically detected dcis, assumed to have a 100% survival.

REFERENCES:

1 Tabar, L, Fagerberg, G, Duffy, SW, Day, NE, Gad, A, Grontoft, O "Update of the Swedish two-county program of mammographic screening for breast cancer" in Radiol Clin North Am 1992; 30: : 187-210

2 Chen, HH, Duffy, SW, Tabar, L, Day, NE "Markov chain models for progression of breast cancer. Part 1: tumour attributes and the preclinical screen-detectable phase" in J Epidemioloy and Biostatistics 1997; 2: : 9-23

3 Tabar, L, Vitak, B, Chen, HH, Duffy, SW, Yen, MF, Chiang, CF, Krusemo, UB, Tot, T, Smith, RA "The Swedish Two-County Trial twenty years later. Updated mortality results and new insights from long-term follow-up" in Radiol Clin North Am 2000; 38: : 625-51

4 Rubinstein, R.Y., Shapiro, A. "Discrete event systems. Sensitivity analysis and stochastic optimization by the score function method." 1993;

5 Tan, SYGL, van Oortmarssen, GJ, Piersma, N "Estimating parameters of a microsimulation model for breast cancer screening using the Score Function method" in Ann Oper Res 2003; 119: : 43-61

6 de Koning, HJ, Boer, R, Warmerdam, PG, Beemsterboer, PM, van der Maas, PJ "Quantitative interpretation of age-specific mortality reductions from the Swedish breast cancer-screening trials" in J Natl Cancer Inst 1995; 87: : 1217-23

7 Boer, R, de Koning, HJ, van der Maas, PJ "A longer breast carcinoma screening interval for women age older than 65 years?" in Cancer 1999a; 86: : 1506-10

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

Model Verification Procedures

The standard MISCAN model has been rigorously tested during its development. A large number of test were designed , carried out, documented and evaluated to check all components of the MISCAN program. The results of these test have been evaluated in a group of primary users. Similar tests have been applied to later versions of the model, in particular when the MISCAN–Fadia version was created.

The Cohort Model was initially programmed in Pascal. The code was checked when it was reprogrammed in C++ (by another person) by comparing results of both versions.

The MISCAN extensions for the CISNET project involve implementation of the continuous tumor growth model in MISCAN, and creating additional model output. The continuous tumor growth model component in MISCAN was checked by inspecting individual histories (using Matlab for comparison), including checks of output. New output was also checked against existing MISCAN output with partially overlapping sub–classifications. Diagnostic runs with extreme assumptions were performed (0% and 100% adjuvant treatment effect, screening threshold diameter > clinical diagnosis diameter, etc.) and gave expected outcomes.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

MODEL VALIDATION PROCEDURES

Independent validation

No independent validation has been performed yet. Earlier versions of the continuous tumor growth model that did not include correlation between growth rate, tumor diameter at clinical diagnosis, and survival since the moment at which fatal diameter was reached, were fitted both to the TCS and HIP screening trials, but did not give acceptable results. We will again include the HIP trial in further analyses, by checking the TCS estimates on the HIP data, and fit the HIP data with the current structure of the model including the correlation.

DEMOGRAPHY ASSUMPTIONS

Assumptions in the MISCAN–Fadia model regarding demography: (See also Population Component)

a. The age distribution of the tumor initiation rate is the same for all birth cohorts
b. the life time breast cancer risk is the same for all women in a certain birth cohort
c. the life table is the same for all women in a certain birth cohort
d. death from breast cancer and death from other causes are independent

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

Natural History Assumptions

The MISCAN–Fadia model uses the following assumptions on natural history: (See also Natural History Component)

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

a.tumor initiation

In the MISCAN–Fadia model DCIS and invasive tumors are assumed to initiate with the same age specific initiation rate.

b.dCIS

dCIS is a preclinical discrete stage with a certain duration that precedes a proportion of the invasive tumors. There are three possible transitions from pre clinical dCIS:

1. regression
2. progression to an invasive tumor
3. clinical detection of dCIS

If a dCIS progresses to an invasive tumor, the invasive tumor starts growing from the smallest detectable stage. As soon as the dCIS has progressed to an invasive tumor, the invasive tumor will determine the stage of the tumor.

c.invasive breast cancer

Invasive tumors are assumed to grow exponentially, i.e. with constant growth rate. The natural history of an invasive breast cancer is characterized by the following variables.

1. tumor growth rate (governed by a lognormal distribution)
2. fatal diameter of the tumor (governed by a weibull distribution)
3. survival duration after reaching the fatal diameter (governed by a lognormal distribution)
4. tumor diameter at clinical diagnosis because of the primary tumor (governed by a lognormal distribution)
5. moment at which distant metastases lead to clinical diagnosis of the tumor, modeled as a constant fraction of the survival duration after reaching the fatal diameter (deterministic)
6. tumor diameter at inception of lymph node metastases N1 occur (governed by a weibull distribution)
7. difference between the tumor diameters at which N1 and N2 lymph node involvement occur (deterministic)
8. moment at which distant metastases occur, modeled as a constant fraction of the survival duration after reaching the fatal diameter (deterministic)

We assume that the tumor growth rate, the survival duration after reaching the fatal diameter and the tumor diameter at clinical diagnosis because of the primary tumor are correlated.
If the tumor is not detected before the tumor has reached the fatal diameter, the woman will die from breast cancer if the woman does not die from other causes before. If the tumor is detected before inception of detectable lymph node metastases, the stage of the detected tumor is node negative; otherwise it is node positive. If the tumor is detected before inception of detectable distant metastases, the stage of the detected tumor is distant metastases negative; otherwise it is distant metastases positive.

SCREENING ASSUMPTIONS

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

The MISCAN–Fadia model uses the following assumptions regarding mammography screening: (See also Screening Component)

a. effect of screening

1. dCIS: A preclinical dCIS may be detected by screening, depending on the sensitivity of the screening test for dCIS.
2. invasive tumor: The screen detectability of an invasive tumor is completely determined by its threshold size for screen detection. If a woman is screened before the threshold size is reached, the screening test will not detect the tumor; after the threshold size is reached a test will always detect the tumor. Each tumor has its own threshold size, which is governed by a weibull distribution with two parameters, mean and shape. The threshold size for screen detection is assumed to depend on age and year of diagnosis.

b. dissemination of screening

In the MISCAN–Fadia model there are two screening dissemination routines:

1. MISCAN screening dissemination routine, that simulates a regular invitation based screening schedule based on specified screening period, screening ages and attendance rates.
2. CISNET screening dissemination routine, that simulates the actual dissemination of mammography in the US during the period 1975–2000, given a woman's date of birth

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

TREATMENT ASSUMPTIONS

The MISCAN–Fadia model uses the following assumptions regarding treatment: (See also Treatment Component)

a. **Effect of adjuvant treatment**: A woman diagnosed with cancer may be given adjuvant treatment. There may be different kinds of adjuvant treatment. Each kind of adjuvant treatment has a certain probability to cure the woman, i.e. to eliminate the fatal metastasis if the tumor is diagnosed after inception of fatal metastasis.
b. **dissemination of adjuvant treatment**: In the MISCAN–Fadia model the dissemination of adjuvant treatment is simulated using the CISNET adjuvant treatment dissemination routine that simulates the actual dissemination of adjuvant treatment in the US during the period 1975–2000, given a woman's age at diagnosis, the tumor stage at diagnosis and the year of diagnosis.

POPULATION COMPONENT

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

SUMMARY

Describes the population component of the MISCAN–Fadia model.

OVERVIEW

The Population component simulates the demography of the simulated population.

DETAIL

The US population is simulated by 5–year birth cohorts starting from 1895–99 up to 1965–1969 and 1970 (the latter being a 1 year cohort which is necessary for simulating the year 2000), and all persons in the cohort are simulated from birth to death. Each cohort has its own lifetable (using 1 year age steps) for deaths from other causes which was derived directly from the Base Case data for other cause mortality, for the mid–year of each cohort (thus, 1892,1897,…). Death from other causes before age 30 is neglected in these lifetables because relevant model output is only produced for ages 30–79. The maximum lifetable age in the MISCAN is 100, at which all persons have died.

The relative size of each birth cohort (at birth) is calculated from the Base Case data for the size of the population in 1975, correcting for the probability of dying before 1975 (only for women who reached age 30 before 1975). The relative sizes of the cohorts are then translated into a proportion of the simulated population for each of the cohorts, see Table 11.

TABLE 11. MISCAN–Fadia. Proportion of the simulated population in each birth cohort

Birth cohort	Proportion
1895–99	4.1%
1900–04	4.6%
1905–09	5.2%
1910–14	5.3%
1915–19	5.6%
1920–24	6.0%
1925–29	5.7%
1930–34	5.2%
1935–39	5.4%
1940–44	6.5%
1945–49	7.1%
1950–54	8.5%
1955–59	9.5%
1960–64	10.1%
1965–69	9.4%
1970–71	1.8%

Relevant Assumptions

See Demography Assumptions

Relevant Parameters

See Demography Parameters

SCREENING COMPONENT

SUMMARY

Describes the Screening component of the MISCAN–Fadia model.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

OVERVIEW

The screening component simulates the dissemination and the effect of screening.

DETAIL

Screening usage

The common Cisnet screening dissemination model was used as an external program, and the MISCAN simulation procedure was adapted accordingly for runs that include screening. First, MISCAN–Fadia generates dates of birth for all simulated women and these are written to a file. Next, the dissemination model is run, using the dates of birth from this file to generate a second file with screening ages for all women. Then, MISCAN–Fadia is run again (using common random numbers and the same seed values for the random number generator), and for each woman the screening ages are read from the second file, and a complete life history is generated.

Characteristics of screen–detected and interval tumors.

The tumor diameter distribution of cancers is determined by the continuous tumor growth model. For cancers diagnosed in never screened women it is influenced by the (positive) correlation between the variables growth rate and diameter at clinical diagnosis because of the primary tumor (Table 1b). For screen–detected and interval cancers it is also determined by these two variables, and in addition by the variable threshold diameter for screen–detection. The probability to detect a dcis depends on the sensitivity of dcis.

RELEVANT ASSUMPTIONS

See Screening Assumptions

RELEVANT PARAMETERS

See Screening Parameters

TREATMENT COMPONENT

SUMMARY

Describes the treatment component of the MISCAN–Fadia model.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

OVERVIEW

The Treatment component is used to simulate the dissemination and the effect of adjuvant treatment.

DETAIL

Treatment dissemination is included in the MISCAN–Fadia model as a probability of being treated with a certain type of adjuvant treatment, i.e. chemotherapy or tamoxifen or both for two years, tamoxifen for 5 years, chemotherapy and tamoxifen for 5 years, or none. These probabilities depend on year, age, and stage, and are adopted from the common Base Case data.

The benefit of adjuvant treatment was modeled according to the results of the Cochrane meta–analyses[2], which reported proportional reductions in all cause mortality hazard for the different adjuvant treatment regimes (Base Case Treatment effect data). For chemotherapy, we used the age specific proportional reductions as reported in the meta–analysis directly. For tamoxifen, we calculated the age specific proportional reductions by multiplying the proportional reductions for women with ER+ tumors as reported in the meta–analysis with the proportion of ER+ tumors by age group as reported in the SEER for the period 1988–1993. Furthermore, the effects of chemotherapy and tamoxifen are assumed to be independent.

In MISCAN, a hazard reduction as reported in the meta–analysis cannot be applied directly because survival in absence of adjuvant treatment is described by lognormal survival distribution for women in whom the diameter at diagnosis exceeds the fatal diameter, and cure for the of women in whom the diameter at diagnosis is smaller than the fatal diameter. The effect of adjuvant treatment is modeled as a shift in the fatal disease diameter depending on the adjuvant treatment given, analogous to the way the time dependency of treatment prior to 1975 is modeled, with an extra correction for death from other causes – this correction was done in order to model the effect on breast cancer mortality, since hazard ratios were reported for all cause mortality[2]–. We approximated the new cure proportions for each adjuvant treatment cadjth, using the hazard ratio r as reported by Peto, the 1975 cure proportion c_{1975}, the probability distribution function $F(t)$ for the survival time since the moment at which the tumor reached its fatal diameter and the probability of dying from other causes $F_{oc}(t)$.

$$c_{adjth} = \frac{\{[1 - F_{oc}(t)]^{r-1}[1 - (1 - c_{1975})]F(t)\}^r + F(t) - 1}{F(t)}$$

We used *t=10* years, corresponding to the average follow–up in Peto's meta–analysis[2]. The probability of dying from other causes was approximated using Base Case data. For each adjuvant treatment, the new cure proportion c_{adjth} was then translated into a shift in fatal diameter.

For each type of treatment and age group, a value of the scale parameter of the fatal diameter that corresponds to adjuvant treatment was calculated, see Table 12. This approach will lead to under–estimation of the short–term effect of adjuvant treatment and to over–estimation of the long–term effect. The shift in the fatal diameter leads to an additional delay in the moment of death from breast cancer because the moment of death from breast cancer is described by a distribution that starts at the moment at which the fatal diameter is reached. This will lead to an additional beneficial effect of adjuvant treatment.

TABLE 12. MISCAN–Fadia. Median value of fatal diameter corresponding to adjuvant treatment, by age and by type of treatment. Note that the mode is 0 for all adjuvant treatments and all ages, since the fatal diameter is governed by a Weibull distribution with shape parameter

TYPE AND DURATION	AGE			
		50–59	60–69	70+
Chemotherapy 2yr	4.34	3.60	3.23	3.73
Tamoxifen 2yr	3.23	3.60	3.88	5.11
Tamoxifen 5yr	3.60	4.18	4.70	8.64
Both 2yr	5.11	4.51	4.70	7.75
Both 5yr	5.56	5.32	5.81	14.73

TABLE 13. MISCAN–Fadia. Comparison of simulated and observed results for 1975: (rates per 100000)

CLINICAL INCIDENCE			MORTALITY(1973–1975)			PREVALENCE OF BREAST CANCER PATIENTS		
Age	**MISCAN**	**APC**	**pct diff**	**MISCAN**	**BaseCase**	**pct diff**	**MISCAN**	**BaseCase**	**pct diff**
30–34	28	28	–1%	6	6	10%	84	75	12%
35–39	61	62	–2%	14	13	6%	235	268	–12%
40–44	114	116	–2%	28	24	16%	507	636	–20%
45–49	176	180	–2%	49	43	12%	1023	1096	–7%
50–54	200	201	–1%	64	59	8%	1611	1527	6%
55–59	224	224	0%	74	74	0%	2097	1993	5%
60–64	259	264	–2%	87	84	3%	2643	2289	15%
65–69	293	294	0%	100	93	8%	3241	2556	27%
70–74	322	322	0%	117	104	12%	3856	2904	33%
75–79	329	331	–1%	119	118	1%	4363	3058	43%

Relevant Assumptions

See Treatment Component

Relevant Parameters

See Treatment Parameters

References:

1 EBCTCG "Polychemotherapy for early breast cancer: an overview of the randomised trials" in Lancet 1998; 352: : 930-42

2 EBCTCG "Tamoxifen for early breast cancer: an overview of the randomised trials" 1998; 351: : 1451-67

DEMOGRAPHY PARAMETERS

Parameters for the MISCAN–Fadia Population Component :

a. number of birth cohorts
b. parameters for the distribution of the population among the birth cohorts
c. for each birth cohort parameters for its birth table.. Each birth cohort is defined by its first and last date of birth. A birth table gives the distribution of dates of birth within the birth cohort.
d. for each birth cohort the parameters of its life table
e. for each birth cohort the life time breast cancer risk

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

NATURAL HISTORY PARAMETERS

Parameters for the MISCAN–Fadia Natural History Component :

a. parameters for the age specific distribution of onset of the first screen detectable disease stage (dCIS or invasive)
b. parameters for the duration, regression and progression of dCIS
c. parameters for the distribution of the tumor growth rate
d. parameters for the distribution of the fatal diameter, scale parameter depends on year of diagnosis
e. parameters for the distribution of the survival duration after reaching the fatal diameter
f. parameters for distribution of tumor diameter at clinical diagnosis because of the primary tumor
g. parameter for the moment at which distant metastases lead to clinical diagnosis of the tumor, modeled as a constant fraction of the survival duration after reaching the fatal diameter
h. parameters for the distribution of the tumor diameter at inception of detectable lymph node metastases N1
i. difference between the tumor diameters at which N1 and N2 lymph node involvement occur
j. parameter for the moment at which distant metastases occur, modeled as a constant fraction of the survival duration after reaching the fatal diameter
k. correlation between tumor growth rate and survival duration after reaching the fatal diameter
l. correlation between tumor growth rate and tumor diameter at clinical diagnosis because of the primary tumor
m. correlation between survival duration after reaching the fatal diameter and tumor diameter at clinical diagnosis because of the primary tumor

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

SCREENING PARAMETERS

Parameters of the MISCAN–Fadia Screening Component :

a. parameters for the dissemination of mammography screening
b. parameters for distribution of threshold diameter for screen detection, scale parameter by age and year of diagnosis
c. sensitivity of the screening test for dCIS

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

TREATMENT PARAMETERS

Parameters for the Treatment Component

a. parameters for the dissemination of adjuvant treatment by age at diagnosis, year of diagnosis and treatment
b. for each specified adjuvant treatment the corresponding effects by age group, modeled as treatment dependent fatal diameter

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

TWO COUNTY STUDY RESULT

Overview

The analysis of fitting the Fadia Natural History Component to Two County Study data, as described in Model Calibration Procedures , resulted in parameter estimates presented in Table 1. The observed and simulated detection rates, interval cancer rates and stage distribution of screen–detected cancers, interval cancers, and cancers diagnosed in the control group (before screening started in this group) are presented in Table 2. The Fadia Natural History Component gives a reasonably good fit of TCS data. Note that the model predicts too few small tumors for the control group and too few the screen–detected cancers during the first round, and too many in interval cancers and cancers found at repeat screening. This corresponds with the finding that the observed stage distribution of breast cancers detected at a first screening round is often not more favorable than the distribution at repeat screenings[1]. Figure 2 shows the comparison between observed and simulated survival by tumor diameter.

The difference in mortality between study and control group was simulated including a screening in the control group at the end of the study period. The simulated mortality reduction after 11 years was 27% which is somewhat lower than the observed 30% reduction.[2]

The biological model structure makes quantification of MISCAN–Fadia less straightforward than we expected. For example, survival time is measured from the moment of reaching the fatal diameter, which means that survival parameters have to be estimated or calibrated, with the complication that survival since diagnosis depends on several model variables: tumor growth rate, clinical diagnosis diameter, survival time since moment of reaching the fatal diameter.

Details

TABLE 1. Maximum likelihood estimates for the parameters of the natural history module based on the data from the Two County Study. "Survival" refers to survival time since the moment of reaching the fatal diameter.

A. PARAMETERS OF THE DISTRIBUTION FUNCTIONS

Variable	Distribution	par1	par2	mean	St.dev.
growth rate (1/year)	Lognormal (μ, σ)	0.062	0.87	1.55	1.65
fatal diameter (cm)	Weibull (scale, shape)	2.93	1.42	2.66	1.90
survival (duration, years)	Lognormal (μ, σ)	2.43	1.13	21.5	34.6
clinical diagnosis (diameter, cm)	Lognormal (μ, σ)	0.84	0.59	2.76	1.78
screening threshold (diameter, cm)	Weibull (scale, shape)	1.02	2.95	0.91	0.34

B. CORRELATION BETWEEN VARIABLES

Variables	ρ
growth rate – survival (ρ_1)	–0.90
growth rate – clinical diagnosis diameter (ρ_2)	+0.41
clinical diagnosis diameter – survival (ρ_3)	–0.43

C. TIME SINCE START OF FATAL DISEASE AT WHICH METASTASES LEAD TO CLINICAL DIAGNOSIS OF THE TUMOR
(fraction of the total survival time after reaching the fatal diameter):

0.9

TABLE 2. Comparison of the Two County Study data with the number of cancers as predicted by the Cohort model

A. Screen detected cancers by round (study group)

	Observed	Simulated
1	286	286
2+3	303	265

B. Interval cancers, by round (study group)

	Observed	Simulated
1	76	77
2+3	107	124

C. Size distribution for screen detected cancers, first round (study group)

Tumor diameter	Observed	Simulated
	9%	8%
6–10 mm	32%	27%
11–20 mm	39%	44%
>20 mm	20%	21%

D. Size distribution for screen detected cancers, subsequent rounds (study group)

Tumor diameter	Observed	Simulated
	8%	10%
6–10 mm	31%	34%
11–20 mm	49%	44%
>20 mm	13%	12%

E. Size distribution for interval cancers (study group)

Tumor diameter	Observed	Simulated
	1%	4%
6–10 mm	17%	22%
11–20 mm	41%	40%
>20 mm	41%	34%

F. SIZE DISTRIBUTION FOR CLINICALLY DIAGNOSED CANCERS (CONTROL GROUP)

Tumor diameter	Observed	Simulated
	3%	1%
6–10 mm	11%	10%
11–20 mm	36%	34%
>20 mm	50%	56%

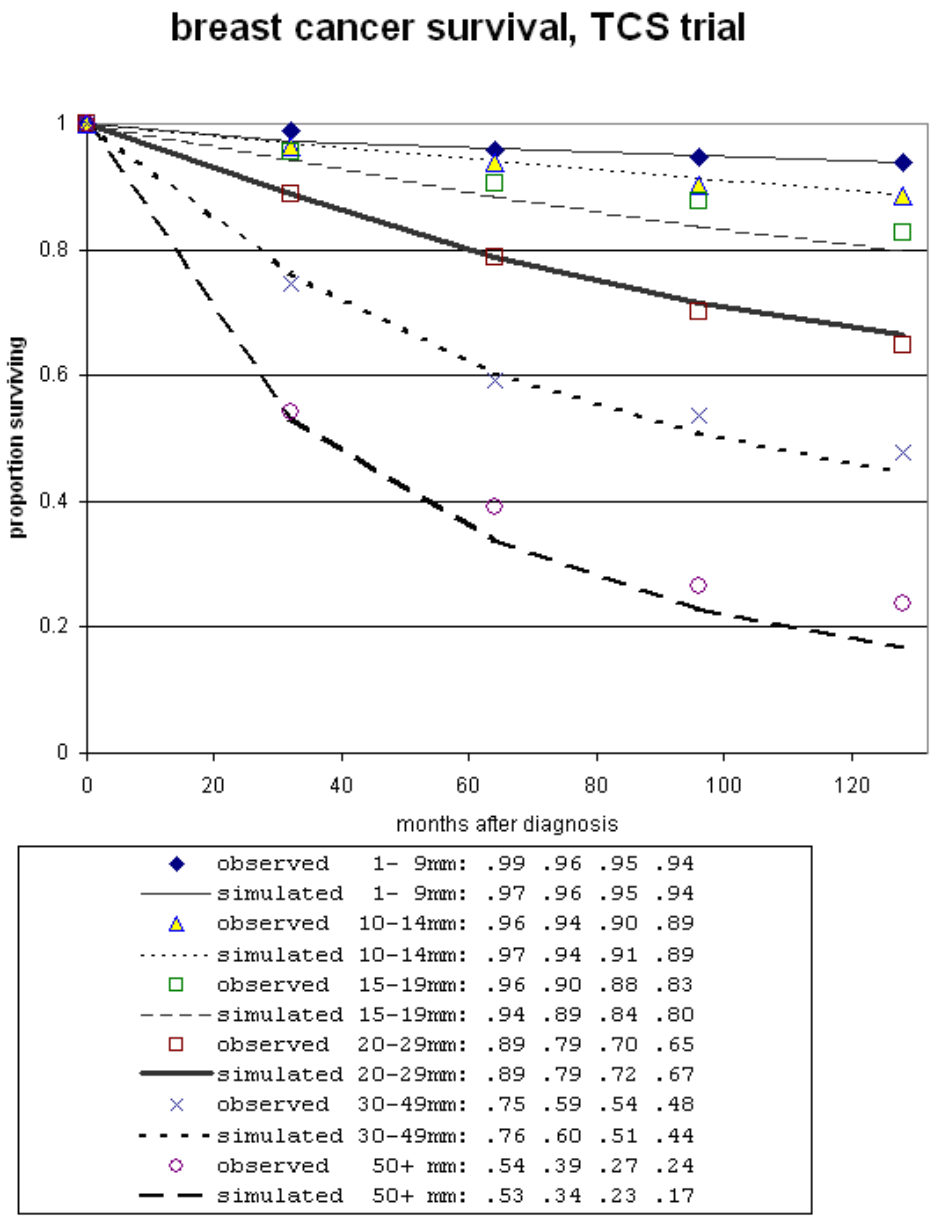


FIGURE 2: Cohort model. Comparison of simulated and observed survival by tumor diameter, Two County study. The legend displays observed and expected survival at 32, 64, 96 and 126 months since diagnosis.

REFERENCES:

[1] Boer, R, de Koning, H, Warmerdam, P, van Oortmarssen, G, van der Maas, P "Stage distribution at first and repeat examinations in breast cancer screening" in J Med Screen 1999b; 6: : 132-138

[2] Tabar, L, Fagerberg, G, Duffy, SW, Day, NE, Gad, A, Grontoft, O "Update of the Swedish two-county program of mammographic screening for breast cancer" in Radiol Clin North Am 1992; 30: : 187-210

CANCER INCIDENCE COMPONENT

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

SUMMARY

This document describes how cancer incidence is generated in the model.

OVERVIEW

In the MISCAN–Fadia model, incidence is modeled as a probability distribution for the onset of pre–clinical disease by age, which refers to the first possible preclinical disease state in the model: preclinical dcis. In the model, many women will not have a detectable dcis prior to the invasive cancer which is modeled as a zero dwelling time in this stage and will thus start with preclinical cancer. Only one cancer per woman can occur in the model.

DETAIL

The model input on the incidence of the onset is specified in two steps: the cumulative probability at age 85 which differs between birth cohorts, and the age distribution of the onset given that the woman will develop breast cancer before age 85 which is equal for all birth cohorts (Age–Cohort model). The cumulative onset of preclinical disease is calculated from the cumulative incidence of clinical breast cancer (up to age 84) (Base Case APC Incidence) by applying correction factors for the proportion of non–progressive preclinical dcis. The cumulative incidences are converted into cumulative probabilities.

Calculation of the age–distribution of the incidence of the onset of preclinical disease starts from the age specific clinical incidence rates for 1975 (Base Case APC Incidence). For each single–year age group from age 20–84, this clinical incidence is first adjusted for differences in the cumulative incidence between the birth cohorts and for differences in proportion of regressive dcis between ages. Next, the age–specific cumulative hazards are converted into age–specific cumulative probabilities. From these cumulative probabilities of being diagnosed with cancer for ages 20–84, the conditional probabilities for ages 20–84 were calculated of being diagnosed with cancer, given she will be diagnosed between age 20 and 84. These conditional probabilities were then averaged into 5 years age groups. Using the probability distribution of the duration of the preclinical stage (time between onset of dcis and clinical diagnosis), the proportion of onset cases that would become diagnosed in the same or in each of the subsequent five–year age categories was calculated. In a calibration procedure, these proportions were used to derive (non decreasing) onset rates (by five–year age groups) of dcis that yield the adjusted 1975 age–specific clinical incidence of breast cancer. The resulting onset distributions by birth cohort and by age are presented in Tables 9 and 10.

TABLE 9. MISCAN–Fadia. Cumulative probability (up to age 85) of the onset of preclinical breast cancer by birth cohort

Birth cohort	Cumulative incidence
1895–99	0.112
1900–04	0.122
1905–09	0.132
1910–14	0.141
1915–19	0.154
1920–24	0.169
1925–29	0.176
1930–34	0.182
1935–39	0.200
1940–44	0.220
1945–49	0.223
1950–54	0.204
1955–59	0.198
1960–64	0.193
1965–69	0.189
1970–71	0.187

TABLE 10. MISCAN–Fadia. Age–distribution of the incidence of the onset of pre–clinical breast cancer (incl. dcis).

Age	Cumulative probability
20	0.000
25	0.002
30	0.005
35	0.021
40	0.046
45	0.105
50	0.169
55	0.233
60	0.328
65	0.436
70	0.563
75	0.707
80	0.852
85	1.000

Survival And Mortality Component

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

Overview

The survival and mortality benefits of early detection

The survival and mortality benefits of early detection follow from the fatal disease concept (which is a special case of the "cure" type of screening model): for each woman there is a moment at which the disease can not be cured anymore, i.e. the moment at which the fatal tumor diameter is reached – this moment depends on the (adjuvant) treatment given at the moment of diagnosis. The screening benefit (cure) only occurs if the tumor is detected by screening before it has become fatal and would otherwise have been diagnosed after it had become fatal.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

BASE CASE RESULTS

Simulation results and Base Case data for 1975 are compared in Table 13 (see Treatment Component). Simulated clinical incidence matches APC incidence quite well. Simulated mortality is too high, compared to Base Case data (1973–1975 SEER mortality). Simulated prevalence is too low at younger ages and increasingly too high at older ages, compared to Base Case prevalence data.

The cancer incidence between 1975 and 2000 as simulated by the MISCAN–Fadia model for the situation without screening or adjuvant treatment is very close to the age–adjusted incidence as provided in the Base Case data. When the Base Case screening dissemination and treatment dissemination data are used, MISCAN simulates a too high age adjusted incidence of invasive cancers for almost all years in the actual screening run, compared to SEER data, especially for tumors
Without screening and adjuvant treatment the age–adjusted mortality rate was predicted to increase from 52.4 to 67.5 per 10^5 women; with actual screening and adjuvant treatment the rate decreases to 46.6 in the year 2000 (see figure 5). For the actual screening and adjuvant treatment run, the simulated age adjusted mortality rates are higher than SEER data, and the difference increases over time to a constant difference of around 12% for the period 1979–1997 and a 25% difference in 1999–2000 (see figure 4). According to the MISCAN–Fadia model, actual screening and treatment (according to the Base Case dissemination data for screening and adjuvant treatment) have similar effects on mortality; screening leads to a 15% mortality reduction and adjuvant treatment to a 21% mortality reduction, see table 14. Annual screening of all women between 1975 and 2000 would have resulted in 36% reduction in mortality.

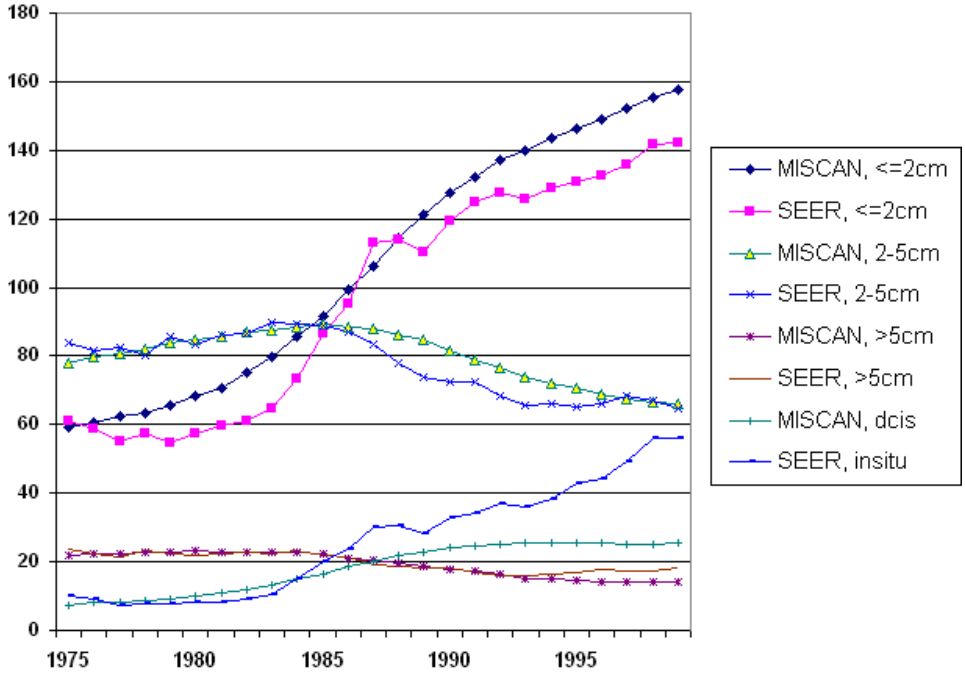


FIGURE 3: MISCAN–Fadia model. Simulated age adjusted incidence rates by tumor size (per 100,000) compared to SEER data (age adjusted to US 2000 standard population age 30–79)

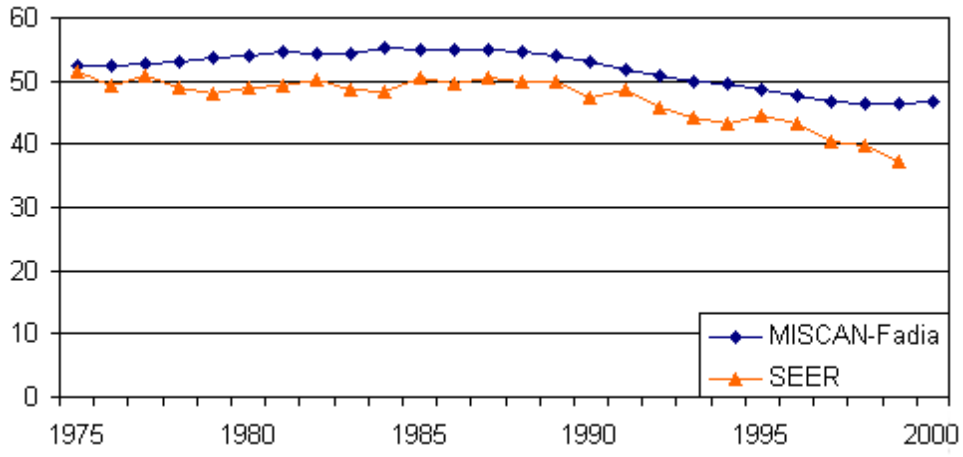


FIGURE 4: MISCAN–Fadia model. Simulated age adjusted mortality rates (per 100,000) compared to SEER data (age adjusted to US 2000 standard population age 30–79)

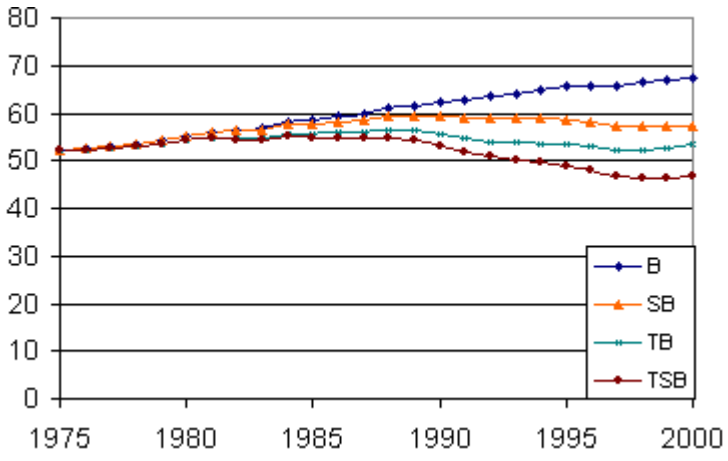


FIGURE 5: MISCAN–Fadia model. Simulated age adjusted mortality rates (per 100,000) for Base Case runs (age adjusted to US 2000 standard population age 30–79): B = Background risk only, SB = Mammography screening and background risk, TB = Adjuvant treatment and background risk, TSB = Adjuvant treatment, mammography screening and background risk

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

KEY REFERENCES

Boer, R, de Koning, H, Warmerdam, P, van Oortmarssen, G, van der Maas, P (1999b) Stage distribution at first and repeat examinations in breast cancer screening in *J Med Screen* 6, p 132–138

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

Chen, HH, Duffy, SW, Tabar, L, Day, NE (1997) Markov chain models for progression of breast cancer. Part 1: tumour attributes and the preclinical screen–detectable phase in *J Epidemioloy and Biostatistics* 2, p 9–23

Draisma, G, Boer, R, Otto, SJ, van der Cruijsen, IW, Damhuis, RA, Schroder, FH, de Koning, HJ Lead times and overdetection due to prostate–specific antigen screening: estimates from the European Randomized Study of Screening for Prostate Cancer in *J Natl Cancer Inst* 95, p 868–78

EBCTCG (1998) Polychemotherapy for early breast cancer: an overview of the randomised trials in *Lancet* 352, p 930–42

EBCTCG (1998) Tamoxifen for early breast cancer: an overview of the randomised trials 351, p 1451–67

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

Loeve, F, Boer, R, van Oortmarssen, GJ, van Ballegooijen, M, Habbema, JDF (1999) The MISCAN–COLON simulation model for the evaluation of colorectal cancer screening in *Comput Biomed Res* 32, p 13–33

Loeve, F, Brown, ML, Boer, R, van Ballegooijen, M, van Oortmarssen, GJ, Habbema, JDF (2000) Endoscopic colorectal cancer screening: a cost–saving analysis in *J Natl Cancer Inst* 92, p 557–63

Nyström, L, Andersson, I, Bjurstam, N, Frisell, J, Nordenskjold, B, Rutqvist, LE. (2002) Long–term effects of mammography screening: updated overview of the Swedish randomised trials 359, p 909–19

Otto, SJ, Fracheboud, J, Looman, CW, Broeders, MJ, Boer, R, Hendriks, JH, Verbeek, AL, de Koning, HJ (2003) Initiation of population–based mammography screening in Dutch municipalities and effect on breast–cancer in *Lancet* 361, p 1411–7

Rubinstein, R.Y., Shapiro, A. (1993) Discrete event systems. Sensitivity analysis and stochastic optimization by the score function method.,

Shwartz, M (1978) An analysis of the benefits of serial screening for breast cancer based upon a mathematical model of the disease in *Cancer* 41, p 1550–1564

Tabar, L, Fagerberg, G, Duffy, SW, Day, NE, Gad, A, Grontoft, O (1992) Update of the Swedish two–county program of mammographic screening for breast cancer in *Radiol Clin North Am* 30, p 187–210

Tabar, L, Vitak, B, Chen, HH, Duffy, SW, Yen, MF, Chiang, CF, Krusemo, UB, Tot, T, Smith, RA (2000) The Swedish Two–County Trial twenty years later. Updated mortality results and new insights from long–term follow–up in *Radiol Clin North Am* 38, p 625–51

Tan, SYGL, van Oortmarssen, GJ, Piersma, N (2003) Estimating parameters of a microsimulation model for breast cancer screening using the Score Function method in *Ann Oper Res* 119, p 43–61

de Koning, HJ, Boer, R, Warmerdam, PG, Beemsterboer, PM, van der Maas, PJ (1995) Quantitative interpretation of age–specific mortality reductions from the Swedish breast cancer–screening trials in *J Natl Cancer Inst* 87, p 1217–23

van den Akker–van Marle, ME, van Ballegooijen, M, van Oortmarssen, GJ, Boer, R, Habbema, JDF (2002) Cost–effectiveness of cervical cancer screening: comparison of screening policies in *J Natl Cancer Inst* 94, p 193–204

FLEXKB DOCUMENT
Version: HI.001.07242013.3521
Document generated: 07/24/2013

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

UNIVERSITY OF TEXAS MDACC

Important note: This document will remain archived as a technical appendix for publications. New versions will be added periodically as model refinements and updates are completed. The most current version is available at http://cisnet.cancer.gov/profiles. Note that unlike most PDF documents, the CISNET model profiles are not suitable for printing as they are not typically written or read in sequential fashion.

We recommend you let your interests guide you through this document, using the navigation tree as a general guide to the content available.

The intent of this document is to provide the interested reader with insight into ongoing research. Model parameters, structure, and results contained herein should be considered representative but preliminary in nature.

We encourage interested readers to contact the contributors for further information.

Go directly to the: Reader's Guide.

READERS GUIDE

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

Core Profile Documentation

These topics will provide an overview of the model without the burden of detail. Each can be read in about 5–10 minutes. Each contains links to more detailed information if required.

Model Purpose

This document describes the primary purpose of the model.

Model Overview

This document describes the primary aims and general purposes of this modeling effort.

Assumption Overview

An overview of the basic assumptions inherent in this model.

Parameter Overview

Describes the basic parameter set used to inform the model, more detailed information is available for each specific parameter.

Component Overview

A description of the basic computational building blocks (components) of the model.

Output Overview

Definitions and methodologies for the basic model outputs.

Results Overview

A guide to the results obtained from the model.

Key References

A list of references used in the development of the model.

Further Reading

These topics will provide a intermediate level view of the model. Consider these documents if you are interested gaining in a working knowledge of the model, its inputs and outputs.

JNCIMonograph Outline

This topic provides links to profile content organized according to the JNCI Monograph Outline for Model Description Chapters. Use this outline for comparisons focused on the CISNET Base Case simulations.

Advanced Reading

These topics denote more detailed documentation about specific and important aspects of the model structure

MODEL PURPOSE

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

SUMMARY

This document provides a description of the problems our model was designed to address.

PURPOSE

The decade from 1990 to 2000 has seen an over–all decrease in breast cancer mortality within the United States[1]. This encouraging trend has also been observed in a number of other countries including Canada and the United Kingdom[2]. While there are a variety of possible explanations for this decline in mortality, two of the most likely reasons are earlier detection and improved treatment.

The principal goal of our model is to provide estimates (and their associated uncertainties) of the relative contributions of screening mammography, tamoxifen use, and improvements in chemotherapy to the observed decrease in U.S. breast cancer mortality since 1990. We will also address the potential impact on future U.S. breast cancer mortality of changes in screening mammography schedules, increased use of tamoxifen, and improvements in chemotherapy.

REFERENCES:

[1] Cancer Surveillance Research Program, National Cancer Institute. "The Surveillance, Epidemiology, and End Results (SEER) Program" 1998;

[2] IARC "The CANCER-Mondial website" in International Agency for Research on Cancer 1999;

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

Model Overview

Summary

This document describes the methods we use to simulate the US population of women from 1975 through 2000 and estimate the breast cancer mortality for these years.

Purpose

Our principal goal is to provide estimates (and their associated uncertainties) of the relative contributions of screening mammography, tamoxifen use, and improvements in and greater use of chemotherapy to the observed decrease in U.S. breast cancer mortality since 1990. We also address the potential impact on future U.S. breast cancer mortality of changes in screening mammography prevalence, increased use of tamoxifen, and further improvements in chemotherapy.

Background

The decade from 1990 to 2000 has seen an overall decrease in breast cancer mortality within the United States[1]. This encouraging trend has also been observed in a number of other Western countries including Canada and the United Kingdom[2]. While there are a variety of possible explanations for such a decline, two of the most likely are earlier detection and improved treatment.

Model Description

Using innovative modeling and simulation techniques and available information we assess the impact that breast cancer interventions have had in the U.S. We use Bayesian updating[4] to estimate the contributions of mammography, chemotherapy, and tamoxifen use to the observed decline in breast cancer mortality in the United States since 1990. Computations of posterior distributions are effected using the "rejection method"[6]: an observation from the prior distribution is included in the posterior distribution depending on the value of its likelihood. In our application the likelihood function is very complicated and cannot be exhibited in closed form.

We begin with a cohort of women in 1975. We then follow this cohort until 2000, simulating the various breast cancer events on an annual basis. Our cohort is dynamic in that we allow women to enter (births, immigration) and leave (deaths, emigration) the population each year.

Breast cancer events depend on each woman's age, mammography use, and treatment (for those detected with breast cancer), all of which change over time. Each year each woman is assigned to be screened or not, depending on the patterns of screening by age in that year. Whether a woman is screened in any given year also depends on her screening history. Breast cancer is diagnosed (or not) depending on the woman's age, mode of detection, the time since her last mammogram, and the calendar year. If she is diagnosed with breast cancer, then her cancer is assigned a stage, nodal status, and estrogen–receptor status with frequencies appropriate for her age, mode of detection, and time since her last mammogram. Therapy is assigned according to the standards of the day, depending on the woman's and the cancer's characteristics. The effects of therapy are based on the observations sampled from the prior distributions for these effects.

We determine which women die depending on actuarial survival data, and we observe breast cancer mortality for the cohort of women to estimate breast cancer mortality from 1975 to 2000. This estimate is compared to the observed breast cancer mortality in the U.S. for each year from 1975 through 2000.

Further details of the model are described in Component Overview .

CONTRIBUTORS

Donald A. Berry, Ph.D.
Prinicipal Investigator
University of Texas M.D. Anderson Cancer Center
Department of Biostatistics & Applied Mathematics
1515 Holcombe Boulevard, Box 447
Houston, TX 77030
Telephone: (713) 794–4141
Fax: (713) 745–4940
E–mail: dberry@odin.mdacc.tmc.edu

Melissa Bondy, Ph.D.
Co–Principal Investigator
University of Texas M. D. Anderson Cancer Center
Telephone: (713) 794–5264
E–mail: mbondy@mdanderson.org

Lurdes Inoue, Ph.D.
Assistant Professor
University of Washington
Seattle, Washington
Telephone: (206) 616–6398
E–mail: linoue@u.washington.edu

Mark Munsell, M.S.
Research Statistician
University of Texas M. D. Anderson Cancer Center
Telephone: (713) 794–4179
E–mail: mfmunsell@mdanderson.org

Yu Shen. Ph.D.
Associate Professor
University of Texas M. D. Anderson Cancer Center
Telephone: (713) 794–4159
E–mail: yushen@odin.mdacc.tmc.edu

Richard L. Theriault, D.O., M.B.A.
Co–Principal Investigator
University of Texas M. D. Anderson Cancer Center
Telephone: (713) 745–3315
E–mail: rtheriau@mdanderson.org

John Venier, M.S.
Systems Anaylst III
University of Texas M. D. Anderson Cancer Center
Telephone: (713) 794–5985
E–mail: venier@mdanderson.org

Acknowledgements

We gratefuly acknowledge the National Cancer Institute's Breast Cancer Surveillance Consortium for working with us to provide invaluable data for this project.
We acknowledge the Centers for Disease Control and Prevention (CDC) and the state, territorial, and tribal organizations that are part of the National Breast and Cervical Cancer Early Detection Program (NBCCEDP) for their commitment to providing high–quality breast and cervical cancer screening, follow–up and treatment to women in medically underserved populations throughout the United States.
The NBCCEDP was created by Congress with the passage of the Breast and Cervical Cancer Mortality Prevention Act in 1990 (Public Law 101–354), authorizing funds for a national screening program for breast and cervical cancer for medically underserved women. The NBCCEDP is implemented through cooperative agreements with qualifying health agencies that provide free or low–cost screening to uninsured or underinsured low–income women, develop and dissminate public and professional education strategies, establish quality assurance systems, engage in surveillance and evaluation activities, and develop coalitions and partnerships

REFERENCES:

[1] Cancer Surveillance Research Program, National Cancer Institute. “The Surveillance, Epidemiology, and End Results (SEER) Program” 1998;
[2] IARC “The CANCER-Mondial website” in International Agency for Research on Cancer 1999;
[3] Berry, DA. “Statistics: A Bayesian Perspective” 1996;
[4] Berry, DA & Stangl, DK. “Bayesian Biostatistics” 1996;
[5] Spiegelhalter, DJ, Abrams KR, Myles JP. “Bayesian Approaches to Clinical Trials and Health-Care Evaluation.” 2004;
[6] Tanner, MA. “Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions” 1996;

ASSUMPTION OVERVIEW

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

SUMMARY

This document describes the key assumptions behind our model.

BACKGROUND

Population Dynamics

Our model allows for women to be born into our population or migrate into and out of our population.

Intervention Effects

Because we do not know the impact of adjuvant tamoxifen or adjuvant chemotherapy on the reduction in risk of breast cancer mortality, we impose separate and independent prior distributions on the reduction in the risk of breast cancer mortality due to these two interventions. We allow for the possibility of an additional survival benefit (beyond stage shift) due to mammography screening. More information on the assumptions regarding these intervention effects can be found in the Parameter Overview .

Tumor Characteristics

If breast cancer is detected in a woman, we base the tumor's characteristics on data from the Breast Cancer Surveillance Consortium[1], the National Breast and Cervical Cancer Early Detection Program[2], the Canadian National Breast Screening Studies[4], and the Health Insurance Plan Project[5], depending on the mode of detection. However, there are a few assumptions that we make regarding tumor characteristics.

ASSUMPTION LISTING

In our model we assume:

Population Dynamics

1. Women in the population are born on January 1 of their birth year.
2. Women age in discrete increments of 1 year.
3. Immigrants have had no screening mammograms before entering the population.
4. Emigrants are lost to follow–up as of the year they leave the population.

Intervention Effects

1. Observed decrease in mortality is caused by screening and treatment.

2. A priori, screening and treatment have independent effects.
3. All adjuvant chemotherapy regimens have the same effect.
4. A prior distribution of the reduction in risk of breast cancer mortality due to adjuvant tamoxifen.
5. A prior distribution of the reduction in risk of breast cancer mortality due to adjuvant chemotherapy.
6. A prior distribution on the reduction in risk of breast cancer mortality beyond stage shift.
7. Women with stage IV disease receive no survival benefit from chemotherapy or hormonal therapy.
8. Women aged 50 or younger receive an additional 10% reduction in the hazard of breast cancer mortality due to chemotherapy. (Based on the Overview results[6].)
9. Women with stage IV disease do receive no survival benefit from treatment with chemotherapy or hormonal therapy.
10. Women who are treated with taxanes receive an additional 14% survival benefit[7].

Tumor Characteristics

1. Given tumor characteristics, there is no race effect.
2. ER status is dependent on mode of detection.
3. Tumors detected more than 3 years after a screening mammogram have the same characteristics as clinically detected tumors.

REFERENCES:

[1] National Cancer Institute. "Breast Cancer Surveillance Consortium (BCSC)" 2003;
[2] Centers for Disease Control and Prevention. "National Breast and Cervical Cancer Early Detection Program (NBCCEDP)" 2002;
[3] Miller, AB, Baines, CJ, To, T, et al. "Canadian National Breast Screening Study: 1. Breast cancer detection and death rates among women aged 40 to 49 years." in Canadian Medical Association Journal 1992; 147: 1459-1476
[4] Miller, AB, Baines, CJ, To, T, et al. "Canadian National Breast Screening Study: 2. Breast cancer detection and death rates among women aged 50 to 59 years." in Canadian Medical Association Journal 1992; 147: 1477-1488
[5] Shapiro, S, Venet, W, Strax, P, et al. "Periodic Screening for Breast Cancer: The Health Insurance Plan Project and its Sequelae, 1963-1986" 1988;
[6] Early Breast Cancer Trialists' Collaborative Group. "Polychemotherapy for early breast cancer: an overview of the randomized trials." in The Lancet 1998; 352: 930-942
[7] Theriault R, Carlson R, Stockdale F. "(Personal communication)" 2003;

PARAMETER OVERVIEW

SUMMARY

This document describes the 6 parameters included in our model.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

BACKGROUND

Because we do not know the impact of adjuvant tamoxifen or adjuvant chemotherapy on the reduction in risk of breast cancer mortality, we impose separate and independent prior distributions on the reduction in the risk of breast cancer mortality due to these two interventions. Because of the lead time and the stage shift associated with screening mammography, women whose cancers are detected mammographically tend to have longer survival than those with cancers detected otherwise. We allow for the possibility of an additional survival benefit (beyond stage shift) due to mammography screening. We include separate prior distributions for the reduction in the risk of breast cancer mortality beyond stage shift for AJCC stages I–II and for AJCC stages III–IV.

We also allow for uncertainty in the underlying survival distributions by AJCC stage and age group by placing a prior distribution on the baseline hazard. An age–period–cohort (APC) model is used to estimate breast cancer incidence over time. We also impose a prior distribution on a parameter used to allow for uncertainty in the APC model.

PARAMETER LISTING OVERVIEW

1. Effect of Adjuvant Tamoxifen

The prior distribution of the reduction in risk of breast cancer mortality due to adjuvant tamoxifen follows a beta distribution with mean 0.28 and standard deviation 0.15 (i.e., beta(2.23, 5.73)). The mean for this prior distribution is from the 1998 report by the Early Breast Cancer Trialists' Collaborative Group[1]. Because this report is based on the results of randomized clinical trials, we incorporated additional uncertainty into our prior distribution and used a standard deviation that was three times the standard error in this report.

2. Effect of Adjuvant Chemotherapy

The prior distribution of the reduction in risk of breast cancer mortality due to adjuvant chemotherapy follows a beta distribution with mean 0.14 and standard deviation 0.16 (i.e., beta(0.52, 3.18)). The mean for this prior distribution is from the 1998 report by the Early Breast Cancer Trialists' Collaborative Group[2]. Again, we inflated teh standard error of the effect of chemotherapy in determining the standard deviation of our prior distribution.

3–4. Effect of Screening Mammography Beyond Stage Shift

In addition to any stage shift, we allow for an effect on survival beyond stage shift. We estimate the effects beyond stage shift from the Health Insurance Plan Project (HIP)[3] and the Canadian National Breast Screening Study (CNBSS)[5], and we use these data to derive the means and standard deviations of our prior distributions for our two

beyond–stage–shift parameters.

3. AJCC Stages I–II

We assume the prior distribution for the parameter estimating the effect beyond–stage–shift for AJCC stages I–II to be uniform(0, 0.80), having mean 0.40 and standard deviation 0.23.

4. AJCC Stages III–IV

We assume the prior distribution for the parameter estimating the effect beyond–stage–shift for AJCC stages III–IV to be uniform(0, 0.50), having mean 0.25 and standard deviation 0.14.

5. Underlying Breast Cancer Survival

We have an underlying survival distribution for non–screen–detected breast cancer for each AJCC stage I–IV and age group[6] that is not treated with either chemotherapy or tamoxifen. Because of the uncertainty in these underlying survival distributions, we allow for the data to modify them. We do this by imposing a uniform(0.8,1) prior distribution on the baseline hazard function of these survival distributions.

6. Age–Period Cohort Model

Women who have never had a screening mammogram have breast cancer detected with a probability that depends on her age and year of birth. The probabilities incorporate the secular trend in incidence from the age–period–cohort (APC) model develped by Holford[7]. However, this model is an estimate, and like all estimates is subject to uncertainty. To reflect this uncertainty we impose a uniform(0,1) prior distribution on the impact of the APC model. The alternative we consider to the APC model is constant background incidence over time. This method allows for the possibility tha the APC is not correct, and lets the actual observed mortality determine the weight attributed to the APC model.

REFERENCES:

[1] Early Breast Cancer Trialists' Collaborative Group. “Tamoxifen for early breast cancer: An Overview of the randomised trials” in The Lancet 1998; 351: 1451-1467

[2] Early Breast Cancer Trialists' Collaborative Group. “Polychemotherapy for early breast cancer: an overview of the randomized trials.” in The Lancet 1998; 352: 930-942

[3] Shapiro, S, Venet, W, Strax, P, et al. “Periodic Screening for Breast Cancer: The Health Insurance Plan Project and its Sequelae, 1963-1986” 1988;

[4] Miller, AB, Baines, CJ, To, T, et al. “Canadian National Breast Screening Study: 1. Breast cancer detection and death rates among women aged 40 to 49 years.” in Canadian Medical Association Journal 1992; 147: 1459-1476

[5] Miller, AB, Baines, CJ, To, T, et al. “Canadian National Breast Screening Study: 2. Breast cancer detection and death rates among women aged 50 to 59 years.” in Canadian Medical Association Journal 1992; 147: 1477-1488

[6] CISNET. “Modeling impact of mammography and adjuvant treatment on U.S. breast cancer mortality rates: collective results from the Cancer Intervention and Surveillance Modeling Network.” in Journal of the National Cancer Institute Monograph 2004;

[7] Holford, T. “Cancer Intervention and Surveillance Modeling Network; Base Case” 2003;

COMPONENT OVERVIEW

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

SUMMARY

This document provides an overview of the major components in the model.

OVERVIEW

Population Component

We recognize that in 1975, the intial year of our simulation, that there are women living with breast cancer. We must first identify these prevalent cases.
We then simulate a cohort of 2,000,000 women with an age distribution appropriate for 1975, allowing for prevalent cases. We then follow this cohort to 2000, allowing for births, deaths, and migration. Each year we identify which women are diagnosed with breast cancer. See the Cancer Incidence Component for details on how we diagnose breast cancer.

Screening Component

We assign each woman a screening schedule that she follows throughout her life. We assume that immigrants have had no screening mammograms before entering our cohort. As we follow the cohort from 1975 to 2000, each year we determine whether the woman had a screening mammogram based on her screening schedule.

Cancer Incidence Component

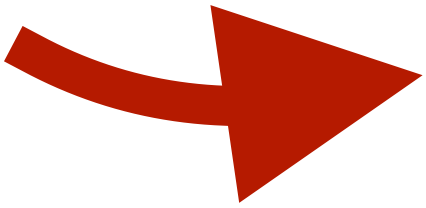
Each year we determine whether a woman is diagnosed with breast cancer. The probability of breast cancer detection depends on whether or not the woman has had a screening mammogram. If the woman has had a screening mammogram the probability of breast cancer detection depends on how long it has been since her last screening mammogram. We also allow for interval cases, which occur between screening mammograms. Tumor characteristics depend on how the breast cancer was detected, and our model recognizes this dependency.

Treatment Component

Treatment depends on a woman's and the tumor's characteristics. Treatment also depends on the calendar year, as there have been changes in treatment over time. Refer to the Survival And Mortality Component to see how treatment impact survival.

Survival And Mortality Component

Shorter of two lifetimes



Each woman who is diagnosed with breast cancer is assigned a lifetime with cause of death from breast cancer. Each woman also has a "natural" lifetime assigned to her when she enters the cohort. A woman's survival is defined as the shorter of these two lifetimes.

Results Component

Our model parameters are selected from prior distributions, which are based on available information from the literature and other sources (see Parameter Overview). We use Bayesian updating to populate the posterior distributions of these parameters. We are also able to obtain the joint posterior distributions of parameters, and from these parameters we estimate the impact of treatment and screening mammography on breast cancer mortality.

Component Listing

Population Component
Screening Component
Cancer Incidence Component
Treatment Component
Survival And Mortality Component
Results Component

For a more detailed listing of the steps in the simulation see Component Listing .

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

OUTPUT OVERVIEW

SUMMARY

This document describes the outputs generated by the model. Our model generates intermediate outputs that can be used to assess the operation of the model, as well as the primary outputs that are used to meet our principle goal (see Model Purpose).

OVERVIEW

Intermediate Outputs:

1. age distribution of women in the U.S. for each year 1975–2000
2. prevalence of breast cancer in 1975
3. tumor characterstics of breast cancer detected in each year 1975–2000
4. survival distribution for women diagnosed with breast cancer
5. survival distribution for women not diagnosed with breast cancer
6. screening mammography schedules
7. proportion of women who have ever had a screening mammogram
8. incidence of breast cancer by stage and by mode of detection, by age and year, and age–adjusted by year
9. breast cancer mortality by year of detection, prevalent in 1975, or incident in 1975 or later

Primary Outputs:

1. age–adjusted breast cancer mortality for each year 1975–2000
2. age–adjusted total mortality for each year 1975–2000
3. posterior distributions for parameters drawn from prior distributions such as the benefits of adjuvant tamoxifen and adjuvant chemotherapy

OUTPUT LISTING

All of the outputs are used in some form of testing and validation at one time or another, but the "intermediate outputs" listed above are primarily used for testing and validation.

Breast Cancer Mortality:

Our model yields estimates of breast cancer mortality for each year 1975–2000. Our modeling approach validates these estimates by comparing them to the known breast cancer mortality for each year 1975–2000.

Breast cancer mortality is also be used as the basis for the acceptance/rejection method[2] for determining the posterior distributions of the parameters which were drawn from the prior distributions. See the Parameter Overview for more details of these prior distributions.

REFERENCES:

[1] Spiegelhalter, DJ, Abrams KR, Myles JP. "Bayesian Approaches to Clinical Trials and Health-Care Evaluation." 2004;

[2] Tanner, MA. "Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions" 1996;

Results Overview

Summary

This document describes the results obtained from our model to address our principle goal (see Model Purpose).

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

Overview

We simulated approximately 80,000 populations in 1975 and followed them through the year 2000. For each of these 80,000 populations we simulated one set of parameters from our posterior distributions, as described above. Of these simulations we accepted 176 for our posterior distributions by the criteria illustrated in *Figure 1* (see Figures) and described above, using an acceptance window on each year of ±2.5 and a window on the slope of ± 0.17. The average of the breast cancer mortality estimates from these 176 accepted simulations is shown in *Figure 2* (see Figures).

Results List

Posterior Distributions of Model Parameters

The prior and posterior distributions for the 4 intervention parameters are shown in *Figure 3* (see Figures). The means and standard deviations of the posterior distributions of these 4 intervention parameters are summarized in *Table 2*. Also included in *Table 2* are the means and standard deviations of the other 2 parameters that we sample from prior distributions and discussed above. Recall that we place a prior distribution on the underlying survival distribution in the absence of treatment and on the impact of the age–period–cohort (APC) model for determining incidence of disease (see Parameter Overview).

From *Table 2* we see that the posterior mean effect of tamoxifen is 0.37, suggesting a 37% decrease in the hazard of breast cancer mortality due to treatment with tamoxifen. The posterior mean effect of screening mammography beyond stage shift for stages I–II is 0.28. This implies that screening mammography provides an additional reduction in hazard of 28% for those women who are diagnosed with stage I–II disease through screening. This reduction is in addition to any benefit that would be achieved due to the cancer being detected at an earlier stage than it might have been if detected clinically.

TABLE 2. Posterior Estimates of Model Parameters

	Mean	Std Dev
Tamoxifen	0.37	0.14
Chemotherapy	0.15	0.14
Beyond Stage Shift I–II	0.28	0.19
Beyond Stage Shift III–IV	0.23	0.14
Underlying Survival Dist	0.87	0.04
APC Incidence	0.61	0.29

The posterior means of the effect of chemotherapy and the benefit of screening

mammography beyond stage shift for stage III–IV disease are similar to the prior means. That is, we estimate that chemotherapy provides a 15% reduction in the hazard of breast cancer mortality, and the reduction in the hazard due to mammography beyond stage shift for stage III–IV disease is 23%.

Our model also estimates the adjustment to the hazard of the underlying survival distribution for women with breast cancer in the absence of treatment has a mean of 0.87 with a standard deviation of 0.04. That is, each underlying survival distribution Sij(t), for non–screen detected breast cancer of stage i, for i = 1, 2, 3, 4, and age group j in the absence of treatment is adjusted as $S^*_{ij}(t) = S_{ij}(t)^\lambda$, where λ has a distribution with mean 0.87 and standard deviation 0.04. The estimates of the effect of screening and treatment are in addition to this initial adjustment.

We discount the impact of the age–period–cohort (APC) model on estimating incidence of breast cancer by an average of 0.61 (sd=0.29). Recall that we placed a uniform(0, 1) distribution on the impact of the APC model. So, on average our model includes only 61% of the incidence estimated by the APC model.

Posterior Estimates of Intervention Effects

From each of our 176 accepted simulations we estimate the percent reduction in breast cancer mortality since 1990, and we estimate the contribution of treatment and screening to this reduction. By ignoring the effect of treatment in our model we estimate the impact of screening mammography on breast cancer mortality. Similarly, by ignoring the effect of screening we estimate the impact of treatment (both chemotherapy and tamoxifen) on breast cancer mortality.

The joint distribution of the contribution of screening and treatment is illustrated in Figures 4a and 4b (see Figures). It is clear from these figures that there is a negative correlation between the percent reduction in breast cancer mortality due to screening and due to treatment. Our model estimates this correlation to be –0.40.

Using our model we estimate a 0.90 posterior probability of a benefit of screening mammography. We estimate a 0.90 posterior probability of a benefit of treatment of at least 9.5%.

POPULATION COMPONENT

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

SUMMARY

This document describes how our model builds the initial cohort of women and follows this cohort over time.

OVERVIEW

We must first determine which women are living with breast cancer in 1975 (prevalent cases). Once we've identified these women, we simulate a cohort of women in 1975 with and age distribution appropriate for that year, including the prevalent cases. We then age our cohort in discrete yearly intervals, allowing for births, deaths, and migration.

DETAIL

Determining Prevalent Cases

To determine which women are the prevalent cases in 1975 we begin by simulating an initial cohort of 2,000,000 women in 1940. We follow this cohort to 1975, diagnosing women with breast cancer each year based on the incidence by age and stage for each year from 1940 to 1974. We assign each woman in this initial cohort a lifetime where cause of death is anything other than breast cancer[1]. Call this her "natural lifetime". We also simulate a lifetime with breast cancer as the cause of death[2], and we determine the cause of death from the shorter of these 2 lifetimes.

We do not allow women to enter this cohort, and women may exit this initial cohort only by dying (of any cause). The women in this cohort who have breast cancer in 1975 are the prevalent cases. We construct a new population of women in 1975 having the corresponding distribution of prevalent cases. We repeat this procedure for each simulation of the model.

Simulating Population of Women

Once we have identified the prevalent cases we simulate a population of 2,000,000 women with the age distribution appropriate for 1975 based on data from the 2001 Regional Database, Woods & Poole Economics, Inc.[3], including the prevalent cases. For each woman we simulate a natural lifetime[1], where cause of death is anything other than breast cancer. As we follow the population in discrete yearly intervals, each woman gets one year older and we determine whether she is diagnosed with breast cancer depending on the incidence of the disease for women her age in that year, and also depending on whether she had a screening mammogram in that year.

Each year we allow for births, deaths, and migration. Those women born into the population from 1975 on are not likely to develop breast cancer, but they do contribute to the size and age distribution of the population. We use the data from Woods & Poole Economics, Inc.[3] to define migration patterns by comparing the U.S. female age distributions in consecutive years. We assign a natural lifetime to each woman who immigrates into our population. We also assign her breast cancer events following the same procedure as for women who were initially in the population in 1975, as described in the Cancer Incidence Component .

Determining Cause of Death

For each woman who is diagnosed with breast cancer, her survival depends on her tumor's characteristics, the mode of detection of the tumor, and the treatment she received, as described in the Survival And Mortality Component . We compare this survival time to her natural lifetime simulated when she entered the population. If the survival time from breast cancer is shorter than her natural lifetime, then the woman is considered to have died from breast cancer and contributes to the breast cancer mortality. If the survival time from breast cancer is longer than her natural lifetime, then the woman is considered to have died from causes other than breast cancer. If a woman dies of other causes or emigrates she is censored as of that time.

Relevant Components

Cancer Incidence Component
Survival And Mortality Component

References:

[1] Rosenberg, M. "Annual probabilities of death from causes other than breast cancer; Base Case" 2002;
[2] CISNET. "Modeling impact of mammography and adjuvant treatment on U.S. breast cancer mortality rates: collective results from the Cancer Intervention and Surveillance Modeling Network." in Journal of the National Cancer Institute Monograph 2004;
[3] 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." 2001;

CANCER INCIDENCE COMPONENT

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

SUMMARY

This document describes how our model determines whether a woman is detected with breast cancer in a given year.

OVERVIEW

This component serves to determine if a woman has a breast cancer detected in the current year being simulated. The probability of detection (clinical or by screening) depends on an age–period–cohort model as well as the woman's screening status.

DETAIL

Breast Cancer Detected Clinically

In each year starting in 1975, we consider every woman who is at least 20 years old and determine whether or not she has a breast cancer detected. If she has not yet had a screening mammogram, she is detected with breast cancer with a probability that depends on her age and year of birth. These probabilities incorporate the secular trend in incidence estimated from the age–period–cohort model[1]. However, we impose a uniform(0, 1) prior distribution on the impact of the age–period–cohort model, and sample one value from this prior distribution for each population we simulate (see Parameter Overview).

Characteristics of tumors that are clinically detected are determined from the 1975 data in SEER[2] as adjusted and described in the Chapter 4 of CISNET[3]. These data provide a mechanism for assigning AJCC disease stage. We determine whether or not there were positive nodes based on data from HIP[4], and we determine ER status based on data from SEER[2].

Breast Cancer Detected by Screening Mammogram

For a woman who has a screening mammogram in the current year, the probability of breast cancer detection, depending on her age, is based on data from the National Breast and Cervical Cancer Early Detection Program (NBCCEDP)[5]. This probability also depends on whether it was her first mammogram, and if it was not, then it depends on the amount of time since her last screening mammogram. If it has been more than 3 years since her last screening mammogram, the probability of detecting breast cancer is the same as for a first screening mammogram.

Breast Cancer Detected in an Interval Between Screening Mammograms

We also simulate breast cancer incidence during intervals between screening mammograms (interval cases) by time since last screening mammogram, age and the current year. The tumor stage for these interval cases is assigned using data from a variety of sources. We used a hierarchical model based on data from the BCSC, the NBCCEDP, the Breast Cancer Detection Demonstration Project (BCDDP)[6], HIP[4], CNBSS[8], and data from 2 Scandinavian studies[10] to estimate the probability of an interval cancer being a given stage. Nodal status and estrogen receptor status was assigned based on data from the BCSC[11]. For those tumors detected more than 3 years after a screening mammogram, we assign tumor characteristics as if they were clinically detected tumors.

Relevant Assumptions

1. Immigrants have had no screening mammograms before entering the population.
2. Tumors detected more than 3 years after a screening mammogram have the same characteristics as clinically detected tumors.

See Assumption Overview .

Relevant Parameters

Age–Period–Cohort (APC) Model Parameter

Because the APC model is an estimate it is subject to uncertainty. To reflect this uncertainty we impose a uniform(0,1) prior distribution on the impact of the APC model. The alternative we consider to the APC model is constant background incidence over time. This method allows for the possibility tha the APC is not correct, and lets the actual observed mortality determine the weight attributed to the APC model. See Parameter Overview .

Relevant Components

Screening Component
Survival And Mortality Component

References:

1 Holford, T. “Cancer Intervention and Surveillance Modeling Network; Base Case” 2003;
2 Cancer Surveillance Research Program, National Cancer Institute. “The Surveillance, Epidemiology, and End Results (SEER) Program” 1998;

[3] CISNET. "Modeling impact of mammography and adjuvant treatment on U.S. breast cancer mortality rates: collective results from the Cancer Intervention and Surveillance Modeling Network." in Journal of the National Cancer Institute Monograph 2004;

[4] Shapiro, S, Venet, W, Strax, P, et al. "Periodic Screening for Breast Cancer: The Health Insurance Plan Project and its Sequelae, 1963-1986" 1988;

[5] Centers for Disease Control and Prevention. "National Breast and Cervical Cancer Early Detection Program (NBCCEDP)" 2002;

[6] Smart, CR, Byrne, C, Smith, RA, et al. "Twenty-year follow-up of the breast cancers diagnosed during the breast cancer detection demonstration project" in A Cancer Journal for Clinicians 1997; 47: 134-149

[7] Miller, AB, Baines, CJ, To, T, et al. "Canadian National Breast Screening Study: 1. Breast cancer detection and death rates among women aged 40 to 49 years." in Canadian Medical Association Journal 1992; 147: 1459-1476

[8] Miller, AB, Baines, CJ, To, T, et al. "Canadian National Breast Screening Study: 2. Breast cancer detection and death rates among women aged 50 to 59 years." in Canadian Medical Association Journal 1992; 147: 1477-1488

[9] Hakkama, M, Holli, K, Isola, J, et al. "Aggressiveness of screen-detected breast cancers" in The Lancet 1995; 345: 221-224

[10] Frisell, J Eklund, G, Hellstrom, L, Somell, A "Analysis of interval breast carcinomas in a randomized screening trial in Stockholm" in Breast Cancer Research and Treatment 1987; 9: 3: 219-225

[11] National Cancer Institute. "Breast Cancer Surveillance Consortium (BCSC)" 2003;

SCREENING COMPONENT

SUMMARY

This document describes how screening is modeled.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

OVERVIEW

Tumor characterstics, and thus survival, depend on the mode of detection of breast cancer. We determine each year whether a woman has a screening mammogram, and if she does, we determine whether breast cancer was detected.

DETAIL

Screening Dissemination

We use the screening mammogram dissemination model[1] to determine whether a woman will have screening mammograms. If so then we use the screening mammography dissemination model to determine her screening schedule.

Our model allows for immigration into our population. For a woman who is an immigrant, it is possible that the screening dissemination model would assign screening mammograms for her before she entered our population. Any such mammograms are ignored.

RELEVANT ASSUMPTIONS

Immigrants had no screening mammograms before entering our population.

RELEVANT COMPONENTS

Cancer Incidence Component

REFERENCES:

[1] Cronin, K, Krapcho, M. "Cancer Intervention and Surveillance Modeling Network; Base Case; unpublished data. " 2003;

TREATMENT COMPONENT

SUMMARY

This document describes how treatment is assigned in our model.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

OVERVIEW

We assign chemotherapy and tamoxifen to women who are detected with breast cancer, depending on the characteristics of the tumor. These treatment assignments will have an impact on survival, as described in the Survival And Mortality Component .

DETAIL

We use the treatment dissemination model developed by[1] to determine treatment for women who are diagnosed with breast cancer. The treatment depends on the tumor characteristics, as well as the woman's age and the year of detection.

In addition to polychemotherapy and tamoxifen, we consider the use of taxanes that were introduced into standard clinical practice in the late 1990s. Taxanes are not represented in the treatment dissemination model. Beginning in 1998 we allow any woman receiving chemotherapy to also receive a taxane. The proportion of women who receive a taxane depends on the stage of disease, and is based on expert opinion[2]. We assign an additional 14% survival benefit for women receiving taxanes.

RELEVANT ASSUMPTIONS

Women who are treated with taxanes receive an additional 14% survival benefit.

Relevant Parameters

1. Effect of Adjuvant Tamoxifen

The prior distribution of the reduction in risk of breast cancer mortality due to adjuvant tamoxifen follows a beta distribution with mean 0.28 and standard deviation 0.15 (i.e., beta(2.23, 5.73)). The mean for this prior distribution is from the 1998 report by the Early Breast Cancer Trialists' Collaborative Group[3]. Because this report is based on the results of randomized clinical trials, we incorporated additional uncertainty into our prior distribution and used a standard deviation that was three times the standard error in this report.

2. Effect of Adjuvant Chemotherapy

The prior distribution of the reduction in risk of breast cancer mortality due to adjuvant chemotherapy follows a beta distribution with mean 0.14 and standard deviation 0.16 (i.e., beta(0.52, 3.18)). The mean for this prior distribution is from the 1998 report by the Early Breast Cancer Trialists' Collaborative Group[4]. Again, we inflated teh standard error of the effect of chemotherapy in determining the standard deviation of our prior distribution.

See Parameter Overview .

Relevant Components

Survival And Mortality Component

References:

[1] 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” in Journal of Epidemology and Community Health 2002; 57: 7: 525-6

[2] Theriault R, Carlson R, Stockdale F. “(Personal communication)” 2003;

[3] Early Breast Cancer Trialists' Collaborative Group. “Tamoxifen for early breast cancer: An Overview of the randomised trials” in The Lancet 1998; 351: 1451-1467

[4] Early Breast Cancer Trialists' Collaborative Group. “Polychemotherapy for early breast cancer: an overview of the randomized trials.” in The Lancet 1998; 352: 930-942

Survival And Mortality Component

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

Summary

This document describes how survival and mortality from cancer are determined in the model.

Overview

This critical component of the model determines survival from cancer after both clinical and screen detection. Survival depends on several factors including mode of detection, stage, age, treatments used, and ER status.

For each woman who is diagnosed with breast cancer, her survival depends on her tumor's characteristics, the mode of detection of the tumor, and the treatment she received. We compare this survival time to her natural lifetime simulated when she entered the population. If the survival time from breast cancer is shorter than her natural lifetime, then the woman is considered to have died from breast cancer and her death contributes to breast cancer mortality. If the survival time from breast cancer is longer than her natural lifetime, then the woman is considered to have died from causes other than breast cancer. If a woman dies of other causes or emigrates she is removed from the at–risk population as of that time.

Detail

Baseline Survival

We have an underlying survival distribution, $S_{ij}(t)$, for non–screen detected breast cancer of stage *i*, for *i* = 1, 2, 3, 4, and age group *j* (Chapter 4 of CISNET, 2004[1]) that is not treated with either chemotherapy or tamoxifen. Because of the uncertainty in these underlying survival distributions, we allow for the data to modify them. We do this by imposing a uniform(0.80, 1) prior distribution on the hazard function of $S_{ij}(t)$. That is, for the simulation we sample a value, say λ, from a uniform(0.80, 1) distribution and adjust each of these underlying survival distributions as $S^*_{ij}(t) = S_{ij}(t)^{\lambda}$. This parameter λ is handled just like other uknown model parameters: it will be accepted as part of the posterior distribution if the resulting simulated breast cancer mortality is sufficiently close to the observed breast cancer mortality.

Impact of Interventions on Survival

We impose separate and independent prior distributions on the reduction in the risk of breast cancer mortality due to (1) adjuvant tamoxifen use, (2) adjuvant chemotherapy, and survival benefit beyond stage shift due to screening mammography. We have separate prior distributions for the reduction in the risk of breast cancer mortality beyond stage shift for (3) AJCC stages I–II and for (4) AJCC stages III–IV. The prior distributions are summarized in *Table 1*.

TABLE 1. Prior Distributions for Intervention Effects

		Mean	Std Dev
Tamoxifen	Beta(2.23, 5.73)	0.28	0.15
Chemotherapy	Beta(0.52, 3.18)	0.14	0.16
Beyond Stage Shift I–II	Uniform(0, 0.80)	0.40	0.23
Beyond Stage Shift III–IV	Uniform(0, 0.50)	0.25	0.14

Refer to the Parameter Overview for details on how these prior distributions were determined.

We sample once from each prior distribution to determine the reduction in risk of dying of breast cancer for each woman who is detected with the disease, depending on the tumor characteristics, whether the tumor was detected by a screening mammogram, and the treatment received. This parameter set is used in the simulation of the population from 1975 through 2000. Each time the population is simulated, we sample again from each prior distribution to obtain a parameter set to use for that population.

Relevant Assumptions

1. Observed decrease in mortality is caused by screening and treatment.
2. A priori, screening and treatment have independent effects.
3. All adjuvant chemotherapy regimens have the same effect.
4. A prior distribution of the reduction in risk of breast cancer mortality due to adjuvant tamoxifen.
5. A prior distribution of the reduction in risk of breast cancer mortality due to adjuvant chemotherapy.
6. A prior distribution on the reduction in risk of breast cancer mortality beyond stage shift.
7. Women with stage IV disease receive no survival benefit from chemotherapy or hormonal therapy.
8. Women aged 50 or younger receive an additional 10% reduction in the hazard of breast cancer mortality due to chemotherapy. (Based on the Overview results[2].)
9. Women with stage IV disease do receive no survival benefit from treatment with chemotherapy or hormonal therapy.

Relevant Parameters

1. Effect of Adjuvant Tamoxifen

The prior distribution of the reduction in risk of breast cancer mortality due to adjuvant tamoxifen follows a beta distribution with mean 0.28 and standard deviation 0.15 (i.e., beta(2.23, 5.73)). The mean for this prior distribution is from the 1998 report by the Early Breast Cancer Trialists' Collaborative Group[3]. Because this report is based on the results of randomized clinical trials, we incorporated additional uncertainty into our prior distribution and used a standard deviation that was three times the standard error in this report.

2. Effect of Adjuvant Chemotherapy

The prior distribution of the reduction in risk of breast cancer mortality due to adjuvant chemotherapy follows a beta distribution with mean 0.14 and standard deviation 0.16 (i.e., beta(0.52, 3.18)). The mean for this prior distribution is from the 1998 report by the Early Breast Cancer Trialists' Collaborative Group[2]. Again, we inflated teh standard error of the effect of chemotherapy in determining the standard deviation of our prior distribution.

3–4. Effect of Screening Mammography Beyond Stage Shift

In addition to any stage shift, we allow for an effect on survival beyond stage shift. We estimate the effects beyond stage shift from the Health Insurance Plan Project (HIP)[4] and the Canadian National Breast Screening Study (CNBSS)[6], and we use these data to derive the means and standard deviations of our prior distributions for our two beyond–stage–shift parameters.

3. AJCC Stages I–II

We assume the prior distribution for the parameter estimating the effect beyond–stage–shift for AJCC stages I–II to be uniform(0, 0.80), having mean 0.40 and standard deviation 0.23.

4. AJCC Stages III–IV

We assume the prior distribution for the parameter estimating the effect beyond–stage–shift for AJCC stages III–IV to be uniform(0, 0.50), having mean 0.25 and standard deviation 0.14.

5. Underlying Breast Cancer Survival

We have an underlying survival distribution for non–screen–detected breast cancer for each AJCC stage I–IV and age group[7] that is not treated with either chemotherapy or tamoxifen. Because of the uncertainty in these underlying survival distributions, we allow for the data to modify them. We do this by imposing a uniform(0.8,1) prior distribution on the baseline hazard function of these survival distributions.

Relevant Components

Cancer Incidence Component
Screening Component
Treatment Component
Results Component

References:

[1] CISNET. “Modeling impact of mammography and adjuvant treatment on U.S. breast cancer mortality rates: collective results from the Cancer Intervention and Surveillance Modeling Network.” in Journal of the National Cancer Institute Monograph 2004;

[2] Early Breast Cancer Trialists' Collaborative Group. “Polychemotherapy for early breast cancer: an overview of the randomized trials.” in The Lancet 1998; 352: 930-942

[3] Early Breast Cancer Trialists' Collaborative Group. “Tamoxifen for early breast cancer: An Overview of the randomised trials” in The Lancet 1998; 351: 1451-1467

[4] Shapiro, S, Venet, W, Strax, P, et al. "Periodic Screening for Breast Cancer: The Health Insurance Plan Project and its Sequelae, 1963-1986" 1988;

[5] Miller, AB, Baines, CJ, To, T, et al. "Canadian National Breast Screening Study: 1. Breast cancer detection and death rates among women aged 40 to 49 years." in Canadian Medical Association Journal 1992; 147: 1459-1476

[6] Miller, AB, Baines, CJ, To, T, et al. "Canadian National Breast Screening Study: 2. Breast cancer detection and death rates among women aged 50 to 59 years." in Canadian Medical Association Journal 1992; 147: 1477-1488

[7] CISNET. "Modeling impact of mammography and adjuvant treatment on U.S. breast cancer mortality rates: collective results from the Cancer Intervention and Surveillance Modeling Network." in Journal of the National Cancer Institute Monograph 2004;

RESULTS COMPONENT

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

SUMMARY

This document describes how we estimate the parameters for our model. We also describe here how we estimate the benefits of treatment and screening mammography.

OVERVIEW

We simulate a population of women and follow them from 1975 to 2000, assigning breast cancer events, screening, and treatment as appropriate for each year. We then compare the simulated breast cancer mortality to the observed breast cancer mortality for these years.

DETAIL

Updating the Posterior Distributions of Intervention Effects

To compare our simulated breast cancer mortality to the observed breast cancer mortality from 1975–2000 we implement the following strategy. We place an "acceptance window" on each year from 1975–2000. If the simulated mortality falls within this acceptance window for each year, then the parameters from the parameter set that we used in that simulation are candidates for acceptance into the respective posterior distributions.

We also divide the interval 1985–2000 into three five–year intervals (1985–1990, 1990–1995, 1995–2000). Then we calculate the slope of the observed mortality curve in each of these three intervals. For each of these slopes we define tolerance limits. For our simulated mortality curve we calculate the slope in these same three intervals. If the slope in each of the three intervals calculated from the simulated mortality curve falls within the tolerance limits of the slopes of the observed mortality curve, then the parameter values that were used to simulate the particular mortality curve are candidates for acceptance into the respective posterior distributions.

Only parameter sets that pass both tests described above are accepted into the respective posterior distributions. By simulating and following the population thousands of times, we will populate the posterior distributions with parameters accepted jointly in this fashion. *Figure 1* (see Figures) illustrates the acceptance algorithm, and *Figure 2* (see Figures) shows the average of our accepted simulations. The prior and posterior distributions of the 4 intervention parameters are illustrated in *Figure 3* (see Figures).

Estimating Impact of Interventions on Breast Cancer Mortality

Through simulation we can create populations of women where every woman aged 40 or older receives screening mammograms beginning in 1975. We can also simulate populations of women with the actual screening behavior that occurred from 1975 to 2000. Some women from each of these two groups will have developed breast cancer and some will have been treated with tamoxifen or adjuvant chemotherapy. By comparing the breast cancer mortality between these two populations of women we can obtain a posterior estimate of the effectiveness of screening mammography in reducing breast cancer mortality. Similarly, we can obtain posterior estimates of the effectiveness of tamoxifen and of chemotherapy.

We can also estimate the effectiveness of combinations of the various interventions as well as the effectiveness of each intervention in the presence of the others. By changing the proportion of women in each age cohort which use screening mammography in our model, we can estimate the potential impact on breast cancer mortality of future changes in the prevalence of screening mammography for each age cohort. Similarly, we can assess the potential impact of changes in the use of tamoxifen and chemotherapy. And we can estimate the effectiveness of combinations of these three interventions for specific age groups. Refer to the Results Overview for some results of our modeling.

Relevant Assumptions

See Assumption Overview .

Relevant Parameters

See Parameter Overview .

Relevant Components

Cancer Incidence Component
Treatment Component
Survival And Mortality Component

COMPONENT LISTING

A more detailed listing of the steps in the simulation follows.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

A. Select Parameters

- Sample parameters from their prior distributions or use fixed values for parameters
  - Relevant Inputs:
    - prior distribution on the reduction in risk of breast cancer mortality due to tamoxifen use or fixed value
    - prior distribution on the reduction in risk of breast cancer mortality due to improvements in chemotherapy or fixed value
    - prior distribution for reduction in risk of breast cancer due to mammography screening (beyond stage shift) or fixed values
    - prior distribution on hazard for underlying survival distributions by stage and age or fixed value
    - prior distribution on impact of age–period–cohort model or fixed value

B. Simulate Cohort in 1975

- simulate year of birth
  - Relevant Inputs:
    - age distribution of women in the U.S. in 1975
- simulate prevalent breast cancer cases in 1975 and their survival
  - Relevant Inputs:
    - incidence of breast cancer in the U.S. from 1940–1974
    - distribution of stage in clinically detected breast cancer
    - underlying breast cancer survival

C. Follow Cohort Through 2000

- Validation Step:
  - check number of women in the U.S. for each year 1975–2000

D. Allow Migration, Births, and Deaths

- births
  - input:
    - number of female births in the U.S. for each year 1975–2000
- deaths
  - Relevant Inputs:

- number of female deaths in the U.S., by age, due to causes other than breast cancer for each year 1975–2000

E. Simulate Breast Cancer Incidence

- simulate screening mammography dissemination
  - inputs:
    - screening mammogram dissemination generation software provided by NCI[1]
- simulate breast cancer incidence
  - inputs:
    - breast cancer incidence for women who have never had a screening mammogram, by age and year[2]
    - breast cancer incidence at screening mammograms, by age, adjusted for year
    - breast cancer incidence during intervals between screening mammograms, by characteristics of last screening mammogram, years since last screening mammogram, age, and year
  - validation:
    - breast cancer incidence in the U.S., by age, for each year 1975–2000

F. Simulate Breast Cancer Survival

- simulate tumor characteristics
  - input:
    - stage distribution for breast cancer in women who have never had a screening mammogram, by age[3]
    - stage distribution for breast cancer detected by a screening mammogram, by age and screening mammogram characteristics[4]
    - stage distribution for breast cancer detected during intervals between screening mammograms[6]
    - node status distribution by mode of detection of breast cancer[6]
    - estrogen receptor status distribution by age[6]
- simulate treatment dissemination[7]
  - input:
    - distribution of treatment type by age, stage, node status, year, and estrogen receptor status
    - distribution of tamoxifen duration by year
- simulate breast cancer survival[8]
  - input:

- modified baseline breast cancer survival by age and stage, modified by effect of treatment type, effect of tamoxifen duration, effect of ER status, etc., interactions, etc.

- validation:
  - number of female deaths in the U.S., by age, due to breast cancer for each year 1975–2000

G. Derive Posterior Distributions for the Parameters in Component 0 which Were Drawn from Prior Distributions, as Both a Validation and Inferential/Output Step

REFERENCES:

[1] Cronin, K., Kraphcho, M. "Cancer Intervention and Surveillance Modeling Network; Base Case" 2002;

[2] Holford, T. "Cancer Intervention and Surveillance Modeling Network; Base Case" 2003;

[3] Centers for Disease Control and Prevention. "National Breast and Cervical Cancer Early Detection Program (NBCCEDP)" 2002;

[4] National Cancer Institute. "Breast Cancer Surveillance Consortium (BCSC)" 2003;

[5] Shapiro, S, Venet, W, Strax, P, et al. "Periodic Screening for Breast Cancer: The Health Insurance Plan Project and its Sequelae, 1963-1986" 1988;

[6] Fracheboud, J, Groenewoud, JH, Boer, R, Broeders, MJM, Baan, CA, Verbeek, ALM, et al. "Landelijke evaluatie van bevolkingsonderzoek naar borstkanker in Nederland 2000 (VIII)" in Instituut Maatschappelijke Gezondheidszorg. : 35-41

[7] 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" in Journal of Epidemology and Community Health 2002; 57: 7: 525-6

[8] CISNET. "Modeling impact of mammography and adjuvant treatment on U.S. breast cancer mortality rates: collective results from the Cancer Intervention and Surveillance Modeling Network." in Journal of the National Cancer Institute Monograph 2004;

FIGURES

THE UNIVERSITY OF TEXAS MD ANDERSON CANCER CENTER Making Cancer History™
Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

FIGURE 1. Acceptance Criteria for Simulated Mortality

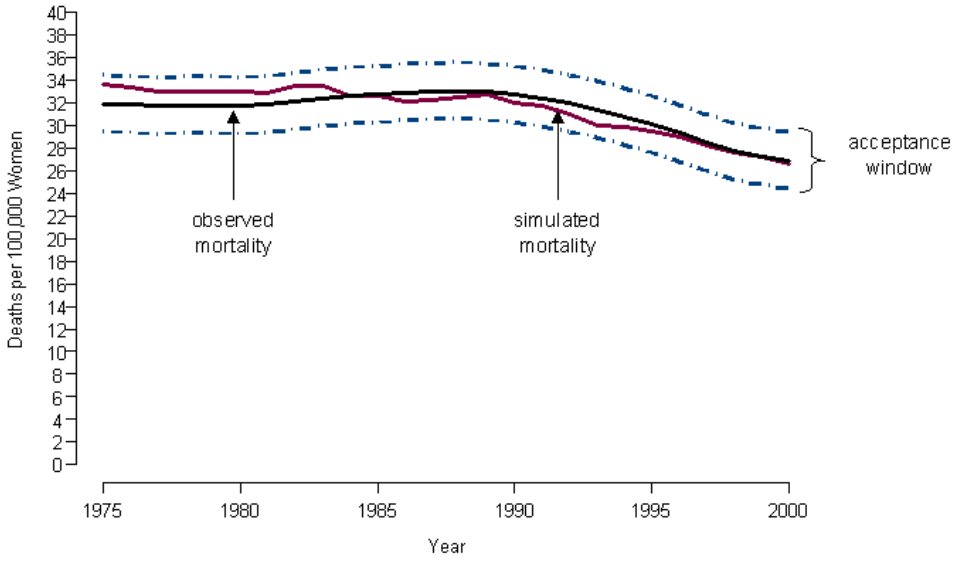


FIGURE 2. Simulated Mortality (Average)

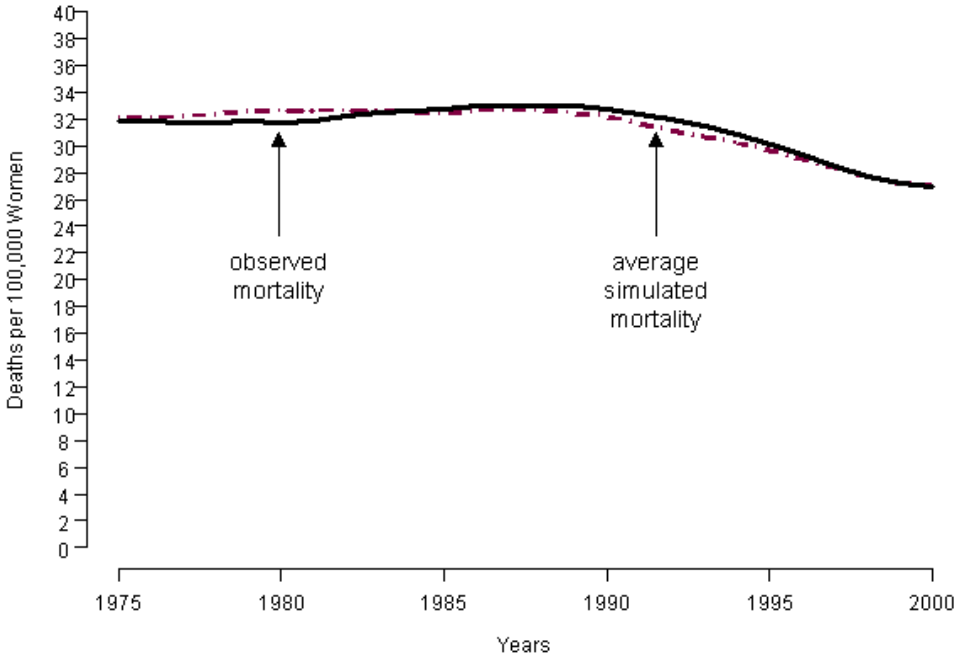


FIGURE 3. Prior and Posterior Distributions of Intervention Effects

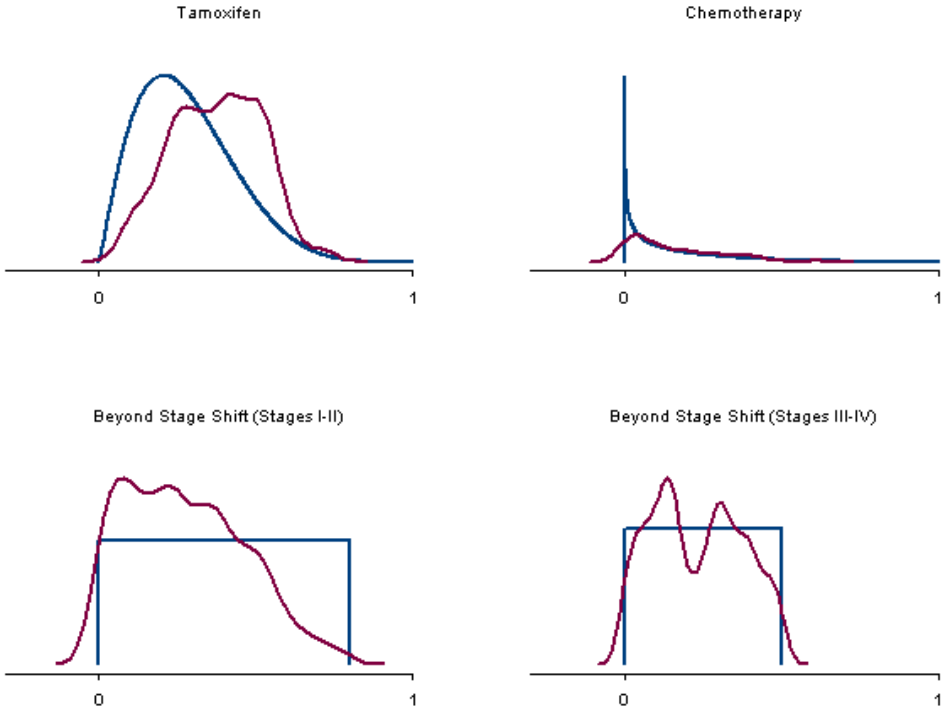


FIGURE 4A. % Reduction in Breast Cancer Mortality

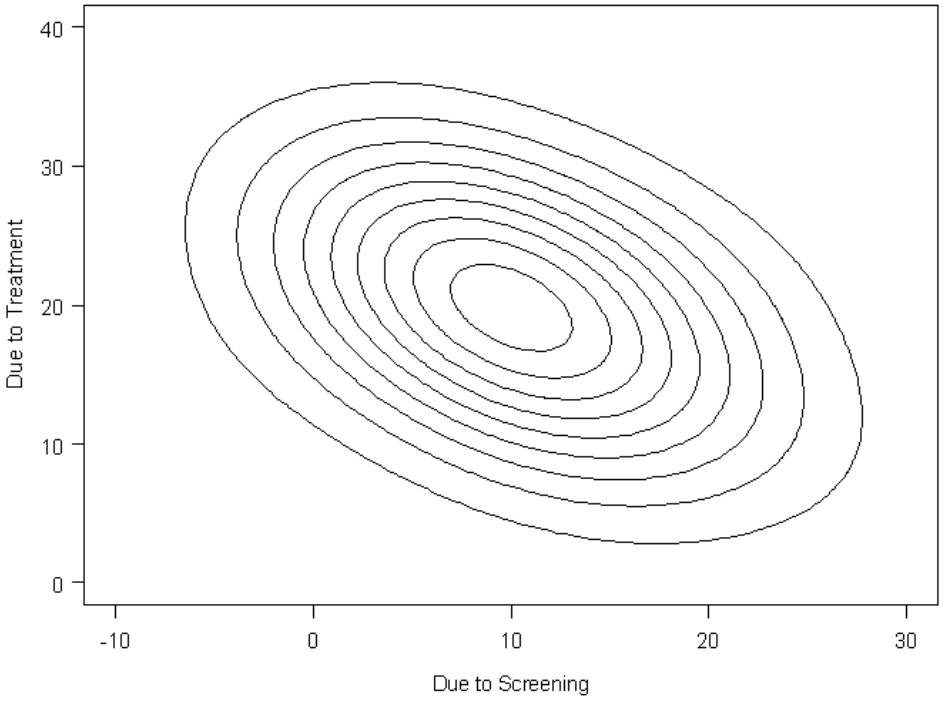
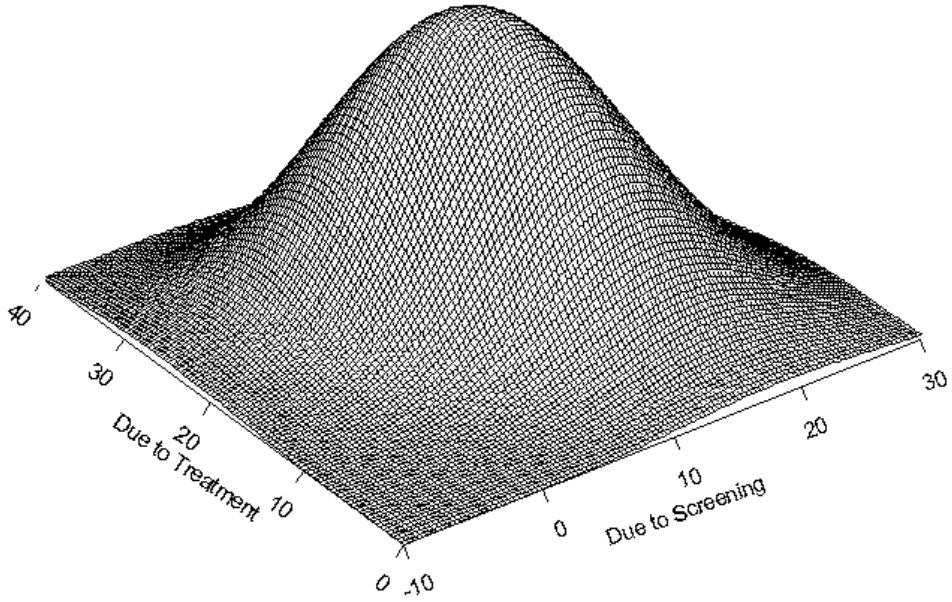


FIGURE 4B. % Reduction in Breast Cancer Mortality



Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

KEY REFERENCES

(1996) Trends in Cancer Screening – United States, 1987 and 1992 in *Morbidity and Mortality Weekly* 45:3, p 57–61

BCSC (2003) Breast Cancer Surveillance Consortium (BCSC) in *National Cancer Institute*,

Berry, DA & Parmigianni, G (1998) Assessing the Benefits of Testing for Breast Cancer Susceptibility Genes: A Decision Analysis in *Breast Disease* 10:(1,2), p 115–125

Berry, DA & Stangl, DK. (1996) Bayesian Biostatistics,

Berry, DA. (1996) Statistics: A Bayesian Perspective,

CISNET. (2004) Modeling impact of mammography and adjuvant treatment on U.S. breast cancer mortality rates: collective results from the Cancer Intervention and Surveillance Modeling Network. in *Journal of the National Cancer Institute Monograph*,

Cancer Surveillance Research Program, National Cancer Institute. (1998) The Surveillance, Epidemiology, and End Results (SEER) Program,

Centers for Disease Control and Prevention. (2002) National Breast and Cervical Cancer Early Detection Program (NBCCEDP),

Cronin, K, Krapcho, M. (2003) Cancer Intervention and Surveillance Modeling Network; Base Case; unpublished data.
,

Cronin, K., Kraphcho, M. (2002) Cancer Intervention and Surveillance Modeling Network; Base Case,

Early Breast Cancer Trialists' Collaborative Group. (1992) Systemic Treatment of Early Breast Cancer by Hormonal, Cytotoxic, or Immune Therapy: 133 Randomised Trial Involving 31,000 Recurrences and 24,000 Deaths Among 75,000 Women – Part 2 in *The Lancet* 339:8785,

Early Breast Cancer Trialists' Collaborative Group. (1992) Systemic treatment of early breast cancer by hormonal, cytotoxic, or immune therapy: 133 randomised trials involving 31,000 recurrences and 24,000 deaths among 75,000 women in *The Lancet*:339, p 1–15,71–85

Early Breast Cancer Trialists' Collaborative Group. (1998) Tamoxifen for early breast cancer: An Overview of the randomised trials in *The Lancet*:351, p 1451–1467

Early Breast Cancer Trialists' Collaborative Group. (1998) Polychemotherapy for early breast cancer: an overview of the randomized trials. in *The Lancet*:352, p 930–942

Fletcher, S. Black W, Harris, R, et al. (1993) Report of the international workshop on screening for breast cancer in *Journal of the National Cancer Institute* 85:20, p 1643–165

Fracheboud, J, Groenewoud, JH, Boer, R, Broeders, MJM, Baan, CA, Verbeek, ALM, et al. Landelijke evaluatie van bevolkingsonderzoek naar borstkanker in Nederland 2000 (VIII) in *Instituut Maatschappelijke Gezondheidszorg.*, p 35–41

Frisell, J Eklund, G, Hellstrom, L, Somell, A (1987) Analysis of interval breast carcinomas in a randomized screening trial in Stockholm in *Breast Cancer Research and Treatment* 9:3, p 219–225

Hakkama, M, Holli, K, Isola, J, et al. (1995) Aggressiveness of screen–detected breast cancers in *The Lancet*:345, p 221–224

Holford, T. (2003) Cancer Intervention and Surveillance Modeling Network; Base Case,

Hortobagyi GN, Hung M, and Buzdar AU (1999) Recent Developments in Breast Cancer Therapy in *Seminars in Oncology* 26(4):12, p 11–20

Hortobagyi, GN, Hung, M, Buzdar, AU (1999) Recent developments in breast cancer therapy in *Seminars in Oncology* 26:4, p Suppl 12:1

Hortobagyi, GN, Hung, M, Buzdar, AU (1999) Recent developments in breast cancer therapy
in *Seminars in Oncology* 26:4:Suppl 12, p 11–20

IARC (1999) The CANCER–Mondial website in *International Agency for Research on Cancer*,

International Agency for Research on Cancer, IARC (1999) CANCER–Mondial,

Jatoi, I (1999) Breast Cancer Screening in *The American Journal of Surgery* 177, p 518–524

Kerlikowske K, et al. (1995) Efficacy of Screening Mammography in *Journal of the American Medical Association* 273:2,

Kerlikowske, K, Grady, D, Rubin, SM, Sandrock, C, Ernster, VL (1995) Efficacy of screening mammorgraphy. A meta–analysis in *Journal of the American Medical Association* 273:2, p 149–54

Kricker, A, Farac, K, Smith, D, Sweeny, A, McCredie, M, Armstrong, BK (1999) Breast cancer in New South Wales in 1972–1995: Tumor size and the impact of mammographic screening. in *International Journal of Cancer* 81:6, p 877–80

Mariotto, A, Feuer, EJ, Harlan, LC, Wun, LM, Johnson, KA, Abrams, J. (2002) Trends in use of adjuvant multi–agent chemotherapy and tamoxifen for breast cancer in the United States: 1975–1999 in *Journal of Epidemology and Community Health* 57:7, p 525–6

Miller, AB, Baines, CJ, To, T, et al. (1992) Canadian National Breast Screening Study: 2. Breast cancer detection and death rates among women aged 50 to 59 years. in *Canadian Medical Association Journal*:147, p 1477–1488

Miller, AB, Baines, CJ, To, T, et al. (1992) Canadian National Breast Screening Study: 1. Breast cancer detection and death rates among women aged 40 to 49 years. in *Canadian Medical Association Journal*:147, p 1459–1476

NBCCEDP (2002) National Breast and Cervical Cancer Early Detection Program (NBCCEDP) in *Centers for Disease Control and Prevention*,

National Cancer Institute. (2003) Breast Cancer Surveillance Consortium (BCSC),

National Center for Health Statistics (1999) Health, United States, 1999 with Health and Aging Chartbook,

Rosenberg, M. (2002) Annual probabilities of death from causes other than breast cancer; Base Case,

Shapiro, S, Venet, W, Strax, P, et al. (1988) Periodic Screening for Breast Cancer: The Health Insurance Plan Project and its Sequelae, 1963–1986,

Smart, CR, Byrne, C, Smith, RA, et al. (1997) Twenty–year follow–up of the breast cancers diagnosed during the breast cancer detection demonstration project in *A Cancer Journal for Clinicians*:47, p 134–149

Spiegelhalter, DJ, Abrams KR, Myles JP. (2004) Bayesian Approaches to Clinical Trials and Health–Care Evaluation.,

Tanner, MA. (1996) Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions,

Tengs, TO, et al. (1999) Testing for BRCA1 Breast–Ovarian Cancer Susceptibility Gene: A Decision Analysis in *Medical Decision Making*,

Theriault R, Carlson R, Stockdale F. (2003) (Personal communication),

U.S. Department of Health and Human Services, National Center for Vital Statistics (1992) Vital Statistics of the United States,

Woods & Poole Economics Inc. (2001) 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.,

FLEXKB DOCUMENT
Version: HI.001.07242013.3522
Document generated: 07/24/2013

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

GEORGETOWN UNIVERSITY

Important note: This document will remain archived as a technical appendix for publications. New versions will be added periodically as model refinements and updates are completed. The most current version is available at http://cisnet.cancer.gov/profiles. Note that unlike most PDF documents, the CISNET model profiles are not suitable for printing as they are not typically written or read in sequential fashion.

We recommend you let your interests guide you through this document, using the navigation tree as a general guide to the content available.

The intent of this document is to provide the interested reader with insight into ongoing research. Model parameters, structure, and results contained herein should be considered representative but preliminary in nature.

We encourage interested readers to contact the contributors for further information.

Go directly to the: Reader's Guide.

READERS GUIDE

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

Core Profile Documentation

These topics will provide an overview of the model without the burden of detail. Each can be read in about 5–10 minutes. Each contains links to more detailed information if required.

Model Purpose

This document describes the primary purpose of the model.

Model Overview

This document describes the primary aims and general purposes of this modeling effort.

Assumption Overview

An overview of the basic assumptions inherent in this model.

Parameter Overview

Describes the basic parameter set used to inform the model, more detailed information is available for each specific parameter.

Component Overview

A description of the basic computational building blocks (components) of the model.

Output Overview

Definitions and methodologies for the basic model outputs.

Results Overview

A guide to the results obtained from the model.

Key References

A list of references used in the development of the model.

Further Reading

These topics will provide a intermediate level view of the model. Consider these documents if you are interested gaining in a working knowledge of the model, its inputs and outputs.

JNCIMonograph Outline

This topic provides links to profile content organized according to the JNCI Monograph Outline for Model Description Chapters. Use this outline for comparisons focused on the CISNET Base Case simulations.

Advanced Reading

These topics denote more detailed documentation about specific and important aspects of the model structure

MODEL PURPOSE

SUMMARY

Here we describe the historical and current purposes of our model of breast cancer.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

PURPOSE

Historically, this model was developed to carry out cost–effectiveness analysis of various approaches to breast cancer screening. Among the questions originally addressed were: what are the costs and benefits associated with various population–based interventions aimed at increasing the utilization of mammography in minority populations; what are the costs and benefits of continuing to screen elderly and very elderly women?

The original development of the model focused on simulating the experience of a hypothetical cohort of women of a certain age. For the CISNET project, the inner core of the simulation logic is "re–packaged" to simulate the entire female population of the United States between 1975 and 2000.

The logic of the model is based on the distributions of times to various events in the life of a simulated subject. There are no particular assumptions made about the mechanism by which breast cancer progresses or kills people. As such this model cannot be effectively used to test hypotheses about breast cancer biology, nor to calibrate the parameters of mechanistic models of breast cancer.

Model Overview

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

Summary

This document describes, at the broadest level, the method by which we simulate the US population incidence and mortality from breast cancer between 1975 and 2000.

Purpose

This model, originally developed to perform cost–effectiveness analysis of breast cancer screening programs, has been adapted to simulate the incidence and mortality of breast cancer in the US population between 1975 and 2000. In particular, it is designed to estimate the effects of screening and treatment improvements during that era.

Background

Interesting changes in the morbidity and mortality associated with breast cancer have occurred over the past 25 years, but little is known about the causes of this. While at one time mammography was almost universally agreed to be effective in reducing breast cancer mortality, more recent reviews of the original clinical trials have raised serious questions about this. As further large scale trials of mammography are unlikely to be conducted in the near future, it would be helpful if careful analytic approaches could disentangle the effects of increased mammography utilization, improvements in the efficacy of treatment, and other changes in the population.

This model builds on the basic breast–cancer simulation developed and used by this group for cost–effectiveness analysis of screening programs.

Model Description

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 pre–specified time–interval distributions. We simulate 55 million women to obtain reasonably smooth estimates of the mortality curves. Using US Census data, we begin with women born in or after 1890 to simulate the population distribution of adult women alive in 1975. Women who are destined to develop breast cancer may either be screen detected, present with clinical symptoms, or die of other causes before breast cancer is diagnosed. At presentation, the cancer has a stage assigned, based on whether the tumor is screen or clinically detected. The stage for screen–detected cancers is calculated from what the stage would have been had the tumor presented with symptoms and the lead time gained from screening using a formula derived from Bayes' theorem. Cancers are designated as being estrogen–receptor (ER) positive or negative. Survival is conditional on age and stage at diagnosis, ER status, and treatment.

Model inputs include:

- age distribution of US women in 1975, age and year–specific projections of breast cancer incidence in the absence of screening (from an age–period cohort model generated by NCI),

- birth–year specific annual US female mortality from all causes other than breast cancer (from the Berkeley mortality database, as modified by M. Rosenberg),
- age–specific distributions of stages of cancers diagnosed clinically (taken from SEER data in 1975),
- age–specific distributions of stages of cancers diagnosed through screening (taken from SEER data in the 1990's),
- sensitivity of mammography screening by age,
- mean tumor sojourn time by age,
- mean tumor dwell time in each clinical stage (DCIS, local, regional, distant),
- age and calendar–year estimates of the pattern of mammography utilization (provided by NCI),
- age–stage–ER specific distributions of treatment choices in different calendar years (provided by NCI),
- age–stage–ER specific breast cancer survival curves
- estimates of the odds–ratios of survival associated with use of adjuvant tamoxifen and adjuvant chemotherapy.

For each woman, the model produces a life history that identifies whether or not a diagnosis of breast cancer is made, and if so, in what stage it presents, what treatment was chosen, as well as a date of death and an indication of whether death is from breast cancer or other causes. The total number of mammography screenings is provided, as is a count of how many such screenings were positive. These life histories are then summarized to produce annual estimates of breast cancer incidence and mortality grouped by decade of age.

Key limitations of the model are that it does not allow for any effect of early detection unless a stage shift results, and that it assumes that all breast cancers (including ductal carcinoma in situ) are progressive. The latter limitation is of particular importance.

CONTRIBUTORS

Jeanne Mandelblatt, MD, MPH
Aimee M. Near, MPH
Clyde B. Schechter, MA, MD
Michael A. Stoto, PhD

All from Georgetown University, except Dr. Schechter who is at Albert Einstein College of Medicine

ASSUMPTION OVERVIEW

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

SUMMARY

This document give a broad overview of the assumptions inherent in the model.

BACKGROUND

This model makes no assumptions about the biological mechanisms of breast cancer progression and mortality. Indeed, the model could be used without further modification to simulate any multi–stage failure process for which the appropriate inputs (in particular process step time distributions) were available.

A number of assumptions about the mechanism by which mammography affects the natural history of breast cancer are detailed below.

ASSUMPTION LISTING

This model relies on the following assumptions:

1. The benefit of mammography screening is exactly represented by the effects of shifting diagnosis to an earlier stage of disease. Early detection that is not early enough to result in detection at an earlier stage does not, on average, alter survival. Furthermore, earliness of detection which does result in an earlier stage at diagnosis is "rewarded" with the full difference in stage–specific survivals.
2. In the absence of screening, the distribution of stages of clinically detected tumors would resemble the distribution of stages of clinically detected tumors from the early part of the 1975–2000 era.
3. Breast cancer progresses from a pre–clinical stage to a clinical presence, and then through stages of local, regional, and distant spread. The dwell times in each stage are assumed to have an exponential distribution. All tumors, including all ductal carcinomas in situ, have the potential to progress to metastatic disease and cause death.
4. Dwell times in each successive stage are independent of each other and of the sojourn time.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

PARAMETER OVERVIEW

SUMMARY

Focusing on those input parameters which are not common to all of the CISNET models, we describe the sources of our parameters concerning the stage–progression and sojourn time of breast cancer, as well as mammography operating characteristics.

BACKGROUND

Publications which report estimates for the dwell time in any stage of breast cancer, the sojourn time, or the sensitivity and specificity of mammography were reviewed. Studies not carried out in the industrialized world were excluded, as were studies carried out in highly restrictive populations. We did not distinguish studies which fit statistical models to screening data from studies which, in one way or another, directly observed the particular parameter. The median value of reported estimates was initially used as our base case parameter value. Because simulation results with these estimates showed a shortfall in predicted incidence which increased as screening disseminated, we experimented with other values to try to match the observed US population incidence curves. This resulted in selecting values of sensitivity and sojourn time which are about 2/3 of the way between the lowest and highest values in our literature reviews.

PARAMETER LISTING OVERVIEW

The parameters discussed here pertain only to the natural history of breast cancer component of the model. All of these parameters are used to calculate lead time, and to simulate stage at diagnosis with screening.

The parameters are:

Sojourn time (mean of exponential distribution)—this is taken to be age dependent.

Age	Mean
	1.7y
50–54	2.1y
55–60	3.3y
60–69	3.9y
70+	5.2y

Dwell time as DCIS or in local and regional stages (mean of exponential distributions).

Stage	Time
DCIS	2.97y
LOCAL	5.30y
REGIONAL	11.40y

Sensitivity of initial mammography:

Age	Sens
	0.77
40–49	0.87
50–59	0.94
60–69	0.94
70	0.91

Sensitivity of subsequent mammography:

0.85 at all ages.

Note that in the studies used to derive our base case values, different definitions of stages, or sensitivity have been used. We have ignored these differences and pretended that all studies were estimating the same parameter.

Component Overview

Summary

This document gives a broad view of the key components in the model.

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

Overview

This model overall includes 4 processes:

1. Subject generation. This part of the model generates simulated women to simulate the age distribution of US women in 1975, their age–specific breast–cancer incidence, and their overall mortality experience.
2. Screening simulation. This part of the model generates a screening schedule for each woman and determines whether it results in the early detection of a breast cancer, and if so, in what stage.
3. Course of Disease. This part of the model identifies what treatment approach the simulated woman will undergo and projects subsequent survival.
4. Bookkeeping. This part of the model doesn't simulate anything—it tallies the results as successive women are simulated.

All of these components operate under the "orchestration" of a general simulation engine. The simulation engine maintains a chronologically ordered event queue. This queue is re–initialized at the start of each simulated subject's processing. The queue contains events such as "gets clinically evident breast cancer," "has next mammogram," "dies of cause other than breast cancer," etc.

Component Listing

Population Component

The demographic component generates a population of simulated women having the age distribution of the female population of the United States in 1975. Using SEER

data, 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.

Natural History Component

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 conditional on the lead–time realized by the screening test that detects it.

Screening Component

The screening utilization component determines when simulated women undergo breast cancer screening based on a model of the observed diffusion through the population between 1975 and 1999.

Treatment Component

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 SEER data for age, stage, estrogen receptor status, and treatment–specific survival. Competing mortality is estimated using actuarial methods.

OUTPUT OVERVIEW

SUMMARY

This document describes the general types and forms of output from the simulator.

OVERVIEW

There are two major components to the simulation output.

The first component is tallies of incident breast cancers, breast cancer deaths, and total simulated female midyear population by age (single years) and calendar year between 1975 and 2000. (The simulation generates breast cancers and deaths outside the 1975–2000 window of interest, but these are suppressed and excluded from this component of the output.) The incident breast cancer figures are further disaggregated by stage and ER status. These figures are used to calculate agegroup specific incidence and mortality rates.

The second component is a simulated "cancer registry." In this segment of the output, a record is created for each simulated woman with breast cancer which shows are date of birth, date of diagnosis, stage at presentation, ER status, treatment, date of death, and indication of whether death is from breast cancer or not. A unique feature of this "registry" not available in real life is an entry for the date at which the tumor would have presented clinically (= actual date of diagnosis if tumor was not screen detected.) The registry also summarizes the woman's screening history by including the total number of true positive, true negative, positive, and negative mammograms she underwent during her lifetime. These data are used to calculate survival curves for the simulated cancers, and also to estimate mammogram program sensitivity and lead–time distribution. Note that a simulated woman who dies prior to 1975 is not recorded in this "cancer registry," but events occurring outside the 1975–2000 window which take place during the lifetime of a woman who remains alive at any time in this window are recorded in the "cancer registry."

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

POPULATION COMPONENT

DETAIL

Demographic Component of the Model

Each birth year is selected in our model with a frequency proportional to its prevalence among the US female population in 1975 and inversely proportional to the probability of survival to 1975.[1] This ensures that the 1975 age–distribution we simulate matches that of the given 1975 US 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.

REFERENCES:

[1] Woods & Poole Economic 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" 2001;

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

NATURAL HISTORY COMPONENT

DETAIL

Natural History of Disease

The model makes no explicit assumptions about the biological nature of breast cancer. Rather, 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. This implies, 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.[1]
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.[2] These incidence rates were estimated using an age–period–cohort model which is described elsewhere.[3] These incidence rates are used as the hazard in a survival process, where failure consists of incident breast cancer. The corresponding 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, the majority of 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.[7] 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. Dwell times (e.g., from DCIS to invasive cancer, from local to regional disease, or from regional to distant) were estimated based upon data from randomized clinical trials of breast cancer screening,[12] and simulating stage distributions in screened and unscreened settings (personal communication, William Lawrence, 2002).

When a tumor is diagnosed by screening, the lead–time is calculated. The stage at which the tumor would have presented clinically is "known" within the simulation. The conditional probability that a tumor in any given stage would have progressed to that known stage in the obtained lead–time is therefore calculated by convoluting the exponential stage dwell time distributions. A "prior" distribution for stage at screening is taken from the observed distribution of stages among tumors diagnosed recently (personal communication, Breast Cancer Surveillance Consortium, Diane Miglioretti and William Barlow, 2002).[13] Bayes' theorem is then applied to calculate a "posterior" distribution of stage at screening conditional on the stage at clinical presentation and lead–time. This posterior distribution is then sampled to identify the simulated stage at screen–detection.

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 multi–drug chemotherapy regimens, were 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. While 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 only model the incidence of 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. The underestimate of incidence never exceeds 1% for women under age 50, and only reaches 5.4% for women ages 75 to 79 after 1994.

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

This calibration was performed using a small number of simulations (5 million per woman). Sojourn time was calibrated together with test sensitivity (see below). Results were inspected for face validity and the final combination selected based on the most reasonable combination of values for each parameter that estimated the observed incidence and stage–distribution as closely as possible.

REFERENCES:

[1] Ashba, J, Traish, AM “Estrogen and progesteron receptor concentrations and prevalence of tumor hormonal phenotypes in older breast cancer patients” in Cancer Detect Prev 1999; 23: 3: 238-244

[2] CISNET “Female breast cancer incidence rates SEER 9 registries” 2004;

[3] CISNET “Breast base case age-specific secular trend parameter” 2004;

[4] Brekelmans, CTM, Westers, P, Faber, JAJ, Peeters, PHM, Collette, HJA “Age specific sensitivity and sojourn time in a breast cancer screening programme (DOM) in the netherlands: a comparison of different methods” in J Epidem Commun Hlth 1996; 50: : 68-71

[5] Paci, E, Duffy, SW “Modelling the analysis of breast cancer screening programmes: sensitivity lead time and predictive valus in the florence district programme (1975-1986)” in Intern'l J Epidemiol 1991; 20: 4: 852-858

[6] 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” in BMJ 1998; 317: 7155: 376-379

[7] Boer, R, de Koning, HJ, van der Maas, PJ “A longer breast carcinoma screening interval for women age older than 65 years?” in Cancer 1999; 86: 8: 1506-1510

[8] Fletcher, SW, Harris, RP, Gonzalez, JJ, Degnan, D, Lannin, DR, Strecher, VJ, et al "Increasing mammography utilization: A controlled study" in J Natl Cancer Inst 1993; 85: 2: 112-120

[9] 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" 1988; : 3-15

[10] Chamberlain, J, Coleman, D, Moss, et al "Sensitivity and specificity of screening in the UK trial of early detection of breast cancer" 1991; : 3-17

[11] 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" in CMAJ 1992; 147: 10: 1459-1476

[12] 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" in CMAJ 1992; 147: 10: 1477-1488

[13] Ballard-Barbash, R, Taplin, SH, Yankaskas, BC, Ernster, VL, Rosenberg, RD, Carney,PA, et al "Breast Cancer Surveillance Consortium: a national mammography screening and outcomes database " in AJR Am J Roentgenol 1997; 169: 4: 1001-1008

SCREENING COMPONENT

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

OVERVIEW

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 US population.[1] These data were based on fitting parametric frailty models to national screening data. However, these data only covered 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. Women born after 1970 were assumed to obtain screen at the same ages as women born in 1970.

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 over–estimated.

DETAIL

Mammogram Sensitivity

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.[5] 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 less data on the results for subsequent screens over time by age. We also model test sensitivity as a constant over the period of simulation. While test performance is likely to have improved over time between 1975 and 2000, there was sufficient variability in published estimates from large screening trials by time period[5] that no reasonable time–period–dependent curve could be fit to the observed data.

REFERENCES:

[1] CISNET "Breast base case mammography dissemination parameter" 2004;

[2] 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" in JAMA 1996; 276: 1: 39-43

[3] Fletcher, SW, Harris, RP, Gonzalez, JJ, Degnan, D, Lannin, DR, Strecher, VJ, et al "Increasing mammography utilization: A controlled study" in J Natl Cancer Inst 1993; 85: 2: 112-120

[4] 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" in Radiology 1986; 160: 2: 295-298

[5] Lawrence, WF, Liang, W, Mandelblatt, J, Gold, K, Freedman, M, Ascher, S, et al "Serendipity in diagnostic imaging: Magnetic resonance imaging of the breast" in J Nat'l Cancer Inst 1998; 90: 23: 1792-1800

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

TREATMENT COMPONENT

Treatment and Survival Component

Adjuvant treatments gradually disseminated into practice after 1975. We used data from Mariotto and colleagues to estimate the dissemination of non–hormonal chemotherapy and tamoxifen.[2] Since the two main surgical options – mastectomy and breast conservation – have equivalent survival[4]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 multi–agent chemotherapy or tamoxifen.[5] Women receiving tamoxifen or chemotherapy were assigned a survival time based on a modification of the 1975 survival curve using data from large meta–analyses.[8] 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. Only women with ER positive tumors are assumed to have survival benefits associated with tamoxifen. For ER positive women receiving both tamoxifen and adjuvant chemotherapy, the two odds ratios are multiplied. This, in effect, assumes that the two treatments are neither synergistic nor interfering.

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 time period of interest.

Competing Mortality Component

Death from causes other than breast cancer was estimated using birth cohort–specific annual mortality data.[9]

REFERENCES:

[1] CISNET "Breast base case treatment dissemination parameter" 2004;

[2] 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" in Trends in use of adjuvant multi-agent chemotherapy and tamoxifen for breast cancer in the United States: 1975-1999 2002; 94: 21: 1626-1634

[3] 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" in N Engl J Med 1995; 333: 22: 1456-1461

[4] NIH Consensus Conference "Treatment of early stage breast cancer" in JAMA 1991; 265: 391-395

[5] CISNET "Breast base case 1975 cause-specific survival parameter"

6 EBCTCG “Polychemotherapy for early breast cancer: an overview of the randomised trials” in Lancet 1998; 352: 9132: 930-942

7 EBCTCG “Tamoxifen for early breast cancer:an overview of the randomised trials” in Lancet 1998; 315: : 1451-1467

8 CISNET “Breast base case treatment effect parameter”

9 CISNET “Competing risks 2004” 2004;

KEY REFERENCES

Readers Guide
Model Overview
Assumption Overview
Parameter Overview
Component Overview
Output Overview
Results Overview
Key References

(2004) Breast cancer facts and figure in *American Cancer Society*,

(2003) Statistical Abstracts of the United States,

Allen, C, Cox, EB, Manton, KG, Cohen, HJ (1986) Breast cancer in the elderly -- current patterns of care in *J AM Geriatric Soc* 34, p 637-642

Andersson, I, Aspegren, K, Janzon, L, Landberg, T, Lindholm, K, Linell, F, et al (1988) Mammographic screening and mortality from breast cancer: The Malmo mammographic screening trial in *BMJ* 297:6654, p 943-948

Ashba, J, Traish, AM (1999) Estrogen and progesteron receptor concentrations and prevalence of tumor hormonal phenotypes in older breast cancer patients in *Cancer Detect Prev* 23:3, p 238-244

Baines, CJ, Miller, AB, Wall, C, McFarlane, DV, Simor, IS, Jong, R, et al (1986) Sensitivity and specificity of first screen mammography in the Canadian National Breast Screening Study: A preliminary report from five centers in *Radiology* 160:2, p 295-298

Ballard-Barbash, R, Taplin, SH, Yankaskas, BC, Ernster, VL, Rosenberg, RD, Carney,PA, et al (1997) Breast Cancer Surveillance Consortium: a national mammography screening and outcomes database in *AJR Am J Roentgenol* 169:4, p 1001-1008

Baquet, C, Ringen, K (1986) Cancer control in Blacks: epidemiology and NCI program plans in *Prog Clin Biol Res* 216, p 215-227

Boer, R, de Koning, H, Threlfall, A, Warmerdam, P, Street, A, Friedman, E, et al (1998) Cost effectiveness of shortening screening interval or extending age range of NHS breast screening programme: computer simulation study in *BMJ* 317:7155, p 376-379

Boer, R, de Koning, HJ, van der Maas, PJ (1999) A longer breast carcinoma screening interval for women age older than 65 years? in *Cancer* 86:8, p 1506-1510

Brekelmans, CTM, Westers, P, Faber, JAJ, Peeters, PHM, Collette, HJA (1996) Age specific sensitivity and sojourn time in a breast cancer screening programme (DOM) in the netherlands: a comparison of different methods in *J Epidem Commun Hlth* 50, p 68-71

CISNET Breast base case 1975 cause-specific survival parameter,

CISNET (2004) Breast base case age specific incidence in the absence of screening,

CISNET (2004) Breast base case age-specific secular trend parameter,

CISNET (2004) Breast base case treatment dissemination parameter,

CISNET (2004) Breast base case mammography dissemination parameter,

CISNET Breast base case treatment effect parameter,

CISNET (2004) Competing risks 2004,

CISNET (2004) Female breast cancer incidence rates SEER 9 registries,

CISNET (2004) Proportion of women receiving breast cancer adjuvant therapy by age, SEER stage and year,

CISNET (2004) Stage Distribution for Female Breast Cancer Cases Diagnosed in SEER 1975-1979, by Age,

Centers for Disease Control and Prevention (2003) US obesity trends 1985-2002,

Chamberlain, J, Coleman, D, Moss, et al (1991) Sensitivity and specificity of screening in the UK trial of early detection of breast cancer, p 3-17

Chu, J, Diehr, P, Feigle, P, Glaefke, G, Begg, C, Glicksman, A, et al (1987) The effect of age on the care of women with breast cancer in community hospitals in *J Gerontology* 42:2, p 185-190

Chu, KC, Tarone, RE, Brawley, OW (1999) Breast cancer trends of black women compared with white women in *Arch Fam Med* 8:6, p 521-528

Colditz, G, Hankinson, SE, Hunter, DJ, Willett, WC, Manson, JE, Stampfer, MJ, et al (1995) The use of estrogen and progestins and the risk of breast cancer in postmenopausal women in *New Engl J Med* 352:24, p 1589-1593

Costanza, ME (1994) The extent of breast cancer screening in older women in *Cancer* 74:7(suppl), p 2046-2050

Crowe, JP Jr, Gordon, NH, Hubay, CA, Pearson, OH, Marshall, JS, McGuire, WL (1986) The interaction of estrogen receptor status and race in predicting prognosis for stage II breast cancer patients in *Surgery* 100, p 599-605

EBCTCG (1998) Polychemotherapy for early breast cancer: an overview of the randomised trials in *Lancet* 352:9132, p 930-942

EBCTCG (1998) Tamoxifen for early breast cancer:an overview of the randomised trials in *Lancet* 315, p 1451-1467

Elledge, RM, Clark, GM, Chamness, GC, Osborne, CK (1994) Tumor biologic factors and breast cancer prognosis among white
Hispanic and black women in the United States in *J Natl Cancer Inst* 86, p 705-712

Feurer, EJ, Wun, LM (1992) How much of the recent rise in breast cancer incidence can be explained by increase in mammography utilization: A dynamic population model approach in *Am J Epidemiol* 136, p 1423-1436

Fisher B, Dignam J, Wolmark N, et al (1997) Tamoxifen and chemotherapy for lymph node-negative, estrogen receptor-positive breast cancer in *J Natl Cancer Inst*,

Fisher, B, Anderson, S, Redmond, C, Wolmark, N, Wickerham, D, Cronin, W (1995) 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 in *N Engl J Med* 333:22, p 1456-1461

Fisher, B, Anderson, S, Tan-Chiu, E, Wolmark, N, et al (2001) Tamoxifen and chemotherapy for axillary node-negative, estrogen receptor-negative breast cancer: findings from National Surgical Adjuvant Breast and Bowel Project B-23 in *J Clin Oncol* 19, p 931-942

Fisher, B, Costantino, JP, Wickerham, DL, Redmond, CK, Kavanah, M, Cronin, WM, et al (1998) Tamoxifen for prevention of breast cancer: report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study in *J Nat'l Cancer Inst* 90:18, p 1371-1388

Fisher, B, Redmond, C, Legault-Poisson, S, Dimitrov, NV, Brown, AM, Wickerham, DL, et al (1990) 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 fro in *J Clin Oncol* 8:6, p 1005-1018

Fletcher, SW, Harris, RP, Gonzalez, JJ, Degnan, D, Lannin, DR, Strecher, VJ, et al (1993) Increasing mammography utilization: A controlled study in *J Natl Cancer Inst* 85:2, p 112-120

Gold, MR, Siegel, JE, Russel, LB, Weinstein, MC (1996) Cost-effectiveness in health and medicine,

Kerlikowske, K, Grady, D, Barclay, J, Sickles, EA, Ernster, V (1996) Likelihood ratios for modern screening mammography, Risk of breast cancer based on age and mammographic interpretation in *JAMA* 276:1, p 39-43

Lawrence, WF, Liang, W, Mandelblatt, J, Gold, K, Freedman, M, Ascher, S, et al (1998) Serendipity in diagnostic imaging: Magnetic resonance imaging of the breast in *J Nat'l Cancer Inst* 90:23, p 1792-1800

Mandelblatt, J, Saha, S, Teutsch, S, Hoerger, T, Siu, A, Atkins, D, et al (2003) The cost-effectiveness of screening mammography beyond age 65: A systematic review for the US preventive services task force in *Ann of Intern Med* 139:10,

Mandelblatt, JS, Fryback, DG, Weinstein, MC, Russell, LB, Gold, MR, Hadorn, D (1996) Assessing the effectiveness of health interventions in *Cost - Effectiveness in Health and Medicine,*

Mandelblatt, JS, Fryback, DG, Weinstein, MC, Russell, LB, Gold, MR (1997) Assessing the effectiveness of health interventions for cost-effectiveness analysis. Panel on Cost-Effectiveness in Health and Medicine in *J Gen Intern Med* 12:9, p 551-558

Mariotto, A, Feuer, EJ, Harlan, LC, Wun, LM, Johnson, KA, Abrams, J (2002) Trends in use of adjuvant multi-agent chemotherapy and tamoxifen for breast cancer in the United States: 1975-1999 in *Trends in use of adjuvant multi-agent chemotherapy and tamoxifen for breast cancer in the United States: 1975-1999* 94:21, p 1626-1634

McWhorter, WP, Mayer, WJ (1987) Black/white differences in type of initial breast cancer treatment and implications for survival in *Am J Public Health* 77:12, p 1515-1517

Michaelson, JS, Halpern, E, Kopans, DB (1999) Breast cancer: computer simulation method for estimating optimal intervals for screening in *Radiology* 212:2, p 551-560

Miller, AB, Baines, CJ, To, T, Wall, C (1992) Canadian National Breast Screening Study: 1. Breast cancer detection and death rates among women aged 40 to 49 years in *CMAJ* 147:10, p 1459-1476

Miller, AB, Baines, CJ, To, T, Wall, C (1992) Canadian National Breast Screening Study: 2. Breast cancer detection and death rates among women aged 50 to 59 years in *CMAJ* 147:10, p 1477-1488

Mohla, S, Sampson, CC, Khan, T, et al (1982) Estrogen and progesterone receptors in breast cancer in Black Americans: Correlation of receptor data with tumor differentiation in *Cancer* 50, p 552-559

NIH Consensus Conference (1991) Treatment of early stage breast cancer in *JAMA*:265, p 391-395

National Cancer Institute, SEER Program (2003) Trends of US mortality age-adjusted rates: Total, US 1990-2000,

National Cancer Institute, SEER Program (2003) Trends of US mortality age-adjusted rates: Total, US 1990-2000,

National Center for Health Statistics, National Cancer Institute (2004) Groups at high risk for cancer deaths,

Newcomb, PA, Carbone, PP (1993) Cancer treatment and age: Patient perspectives in *J Natl Cancer Inst* 85, p 1580-1584

Oakar, MR (1992) Legislative effect of the 102nd Congress, Cancer prevention detection treatment and research in *Cancer* 69:7(suppl), p 1954-1956

Ownby, HE, Frederick, J, Russo, J, Brooks, SC, Swanson, GM, Heppner, GH, et al (1985) Racial differences in breast cancer patients in *J Natl Cancer Inst* 75, p 55-60

Paci, E, Duffy, SW (1991) Modelling the analysis of breast cancer screening programmes: sensitivity lead time and predictive valus in the florence district programme (1975-1986) in *Intern'l J Epidemiol* 20:4, p 852-858

Rimer, BK, Keintz, MK, Kessler, HB, Engstrom, PF, Rosan, JR (1989) Why women resist screening mammography: patient-related barriers in *Radiology* 172, p 243-246

Roberts, MM, Alexander, FE, Anderson, TJ, Chetty, U, Donnan, PT, Forrest, P, et al (1990) Edinburgh trial of screening for breast cancer: Mortality at seven years in *Lancet* 335:8684, p 241-246

Russell, LB, Siegel, JE, Daniels, N, Gold, MR, Luce, BR, Mandelblatt, JS (1996) Cost-effectiveness analysis as a guide to resource allocation in health: Roles and limitations in *Cost-effectiveness in health and medicine*, p 3-24

Shapiro, S, Venet, W, Strax, P, Venet, L (1988) Current results of the breast cancer screening randomized trial: The Health Insurance Plan (HIP) of Greater New York study, p 3-15

Shapiro, S, Venet, W, Strax, P, Venet, L Periodic screening for breast cancer: The Health Insurance Plan project and its sequelae, 1963-86,

Silverberg, E, Boring, CC, Squires, TS (1990) Cancer Statistics in *CA Cancer J Clin* 40:1, p 9-26

Soldo, BJ, Agree, EM (1988) America's elderly,

Surveillance Research Program, National Cancer Institute (2004) SEER*Stat sofware www.seer.cancer.gov/seerstat version 5.2.2,

US Census Bureau, Projections of the resident population by race (2003) Hispanic origin and nativity: Middle series 2025 to 2045,

Woods & Poole Economic Inc (2001) 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,