JOINT ESTIMATION OF MAMMOGRAPHIC SENSITIVITY AND TUMOR GROWTH

by

Heidi Wen Ashih

Institute of Statistics and Decision Sciences
Duke University

Date: _____

Approved:

_____
Dr. Don Berry, Supervisor

_____
Dr. Giovanni Parmigiani, Co-Supervisor

_____
Dr. Peter Mueller

_____
Dr. Gary Rosner

Dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy
in the Institute of Statistics and Decision Sciences
in the Graduate School of
Duke University

2000

ABSTRACT

(SUBJECT:Bayesian Statistics, Biostatistics)

JOINT ESTIMATION OF MAMMOGRAPHIC SENSITIVITY AND TUMOR GROWTH

by

Heidi Wen Ashih

Institute of Statistics and Decision Sciences
Duke University

Date: _____

Approved:

_____
Dr. Don Berry, Supervisor

_____
Dr. Giovanni Parmigiani, Co-Supervisor

_____
Dr. Peter Mueller

_____
Dr. Gary Rosner

An abstract of a dissertation submitted in partial
fulfillment of the requirements for the degree
of Doctor of Philosophy in the
Institute of Statistics and Decision Sciences in the Graduate School of
Duke University

2000

Abstract

This dissertation develops two Bayesian models for jointly estimating mammographic sensitivity and tumor growth rates. First, we fit a two-measurement design, where each patient has exactly two tumor measurements, and provide a basic model for such data. Using a Metropolis-Hastings step to update values of the parameters, we utilized a Markov Chain Monte Carlo algorithm to sample iteratively from our posterior distribution.

We then present our complete-measurement design. Its data would include all past negative mammograms and ages, symptomatic status at detection, and number of looks at each mammogram. We propose a joint likelihood to fit such an ideal dataset. In the cases of missing information or only one tumor measurement, we provide a marginal likelihood as well.

Within our complete-measurement design, we require the probability of a woman being symptomatic given her age and tumor size. We present a non-parametric method to interpolate symptomatic counts of bivariate categorical data into finer bins of equal size. By assigning each symptomatic count a growth rate, we multiply-impute associated asymptomatic counts. Using these counts, we create a table of symptomatic probabilities over age and tumor size.

We then generate twelve sets of 1000 patients, using pre-set values for our five parameters. Using the same MCMC approach with a Metropolis-Hastings step, we fit our complete-measurement design to each of the twelve sets of data. We present a comparison of our posterior MCMC means and pre-set values for the parameters of interest, and discuss possible bias issues.

Acknowledgements

My thanks go to my advisors Don Berry & Giovanni Parmigiani, who provided both statistical and editorial support for the completion of this thesis.

I also wish to thank my other committee members, Gary Rosner & Peter Mueller, for their efforts.

My deepest gratitude falls on Aaron C. W. Ashih. Throughout these years and particularly during this last month, your emotional support held me together though I felt on the verge of a nervous breakdown. My last-minute reprogramming into C would not have been possible without your knowledge and patience (special thanks go also to Herbie Lee in this endeavor).

Finally, I thank my parents, who successfully managed the impossible – to be supportive and yet not ever mention the words "thesis" or "dissertation" or "research" in the final months towards the impending deadline. To my dad, who died exactly one month before my defense, I dedicate my PhD and following professorship at Duke. I hope to make you proud.

Contents

List of Figures

Chapter 1

Introduction

1.1 Overview

Screening mammography detects some breast cancers before clinical symptoms appear (Morrison, 1989). To understand the benefit of such screening, it helps to know how many asymptomatic cancers would be found. This depends on the duration of the pre-clinical (asymptomatic) but detectable state, or pre-clinical sojourn time. The sojourn time depends on the tumor growth rate, which varies highly among patients and possibly within patients over time.

We present models for two designs, described below. In both, we assume exponential growth for the change in tumor volume. Some researchers consider logistic or Gompertzian growth to be better. But, with only two-data points, it is difficult to estimate such decelerating growth curves unless more assumptions are made.

Also, we model mammographic sensitivity as a logistic regression on the woman's age and her tumor radius. This is different than in current literature, where sensitivity generally depends only on age. We capture the idea that the larger tumors are easier to detect.

Another innovation of our model is the joint estimation of mammographic sensi-

tivity and tumor growth rate. This is important because previous negative mammograms could be due to one of two cases: 1) Tumor did not yet exist (or was below threshold size), and then grew very fast, or 2) Tumor was there (and above threshold size), but was missed by the mammogram.

In Chapter 2, we present the two-measurement design and its associated model. We use data from the Breast Cancer Detection and Demonstration Project (BCDDP). Each subject's tumor was detected mammographically from screening. After detection, the previous mammogram was viewed retrospectively, revealing the tumor at an earlier state. That is, after a woman's tumor is detected by a scheduled screening, her previous negative mammogram is re-read by the radiologist. Sometimes, the radiologist finds the tumor upon looking back at her previous mammogram. The data we use for the two-measurement design consist of such tumors. Thus, each tumor has two measurements. The two-measurement model jointly estimates tumor growth and mammographic sensitivity from these data. But, this design does not include single-measurement (or interval) cases, which tend to have faster growth rates. Thus, results from such analyses tend to be biased towards slower growth rates.

In Chapter 3, we present the complete-measurement design, which addresses previous negative readings and includes women with only one measurement in addition to those with two or more measurements. We do not have data for the complete-measurement design, but we define a collection strategy that would give us such data. This data contains the full mammographic history of each patient, including symptomatic status and ages at past negative mammograms.

Knowing a woman's symptomatic status is important because if she is symptomatic at the time of a mammogram then her tumor is sure to be found. Such a data point should not tell us about the sensitivity of mammography, but it can tell us about growth rate. Knowing the woman's ages at previous negative mammograms

helps better estimate mammographic sensitivity. If we have two measurements already, then we know the growth rate and can estimate the tumor size at each of these previous mammograms. If we have only one measurement, then we can properly weight different possible mammographic misses, given her ages at past negative readings.

In Chapter 4, we simulate a patient population based on given values for the parameters of interest (3 logistic parameters, and 2 for the log-normal distribution on growth rates). Then, we use Monte Carlo Markov Chains (MCMC) to find the posterior distribution of model parameters. We compare these estimated values to those used in the simulation, and discuss the results.

Within the complete-measurement model, we require the probability of becoming symptomatic given a woman's age and tumor size. In Chapter 5, we present a nonparametric method of interpolating and smoothing bivariate categorical data into finer bins of equal size. We then discuss converting these counts into symptomatic probabilities for use in our complete-measurement design analysis.

Our proposed design and associated analysis apply to any solid tumor and not just breast cancer. One can determine the effect of a screening program from tumor growth rates and assessment sensitivity. Growth rates can be converted to preclinical sojourn times by setting appropriate tumor size thresholds on detectability and symptoms. Given these sojourn times and the jointly estimated test sensitivity, the effectiveness of different screening intervals could be assessed. That is, one could estimate the number and size of cancers found due to different screening protocols. Unlike our work, previous studies have not jointly estimated both radiological sensitivity and tumor growth rates. Our proposed design jointly estimates these two, and captures a more representative sample of the population than most current designs.

1.2 Literature

In the following chapters, we discuss two experimental designs and their associated modeling and analysis. These models jointly estimate tumor growth and mammographic sensitivity. We base our designs on current models of tumor growth; extend current estimations of mammographic sensitivity to depend on both patient age and tumor size; and expand the types of patients included from screening programs. Much work has been done to estimate tumor growth and mammographic sensitivity, but separately. Here we review recent contributions to the literature which inspired our work.

Controlled trials have established that early detection reduces breast cancer mortality for women aged 50-69 (Nyström *et al.*, 1993; Chu *et al.*, 1988; Morrison *et al.*, 1988). Hence, several governmental, professional and scientific organizations recommend regular mammographic screening for these women. Controversy remains over the beneficial effects of screening women aged 40-49 (Bjurstam *et al.*, 1997; Fletcher, 1997; Taubes, 1997; Peer *et al.*, 1996; Claus *et al.*, 1998).

Better data and models are needed to understand how the benefits of screening depend on age and other risk factors. Such models require evaluating the sensitivity of mammography as it depends on the age of the patient and the size of the tumor. Most researchers consider the sensitivity of mammography across broad age ranges (Peer *et al.*, 1996; Schmitt and Threatt, 1982), but not across size.

Some studies (Kerlikowske *et al.*, 1998; Mushlin *et al.*, 1998; Duffy *et al.*, 1997; Straatman *et al.*, 1997; Peer *et al.*, 1996) estimate sensitivity of a screening program rather than sensitivity of a single mammogram (Launoy *et al.*, 1998). These researchers use cancers arising within 12-13 months after a screening exam as a way to count "false-negative" readings. Such counting, while useful, does not allow for the possibility that a fast-growing tumor that surfaces after a screening may not have

existed at the previous screen. From this perspective, such counting underestimates sensitivity. Their counting also does not include slow-growing tumors that may not become symptomatic within a year of being missed. These missed tumors would not be attributed to mammography, whose sensitivity would be overestimated in these cases. However, the researchers' estimations and assumptions are reasonable, given their available data.

Duffy *et al.* (1997) propose a Markov model to address the reduced effectiveness of screening mammography in younger women. The authors estimate the sensitivity of the screening program, and not the test (mammography) itself. They conclude that tumors found in women aged 40-49 progress more quickly than those in older women. While that may be true, their assumptions may not be realistic. First, they assume sojourn times of all women come from the same exponential distribution. Second, within their 3-state model (1 is no detectable disease, 2 is pre-clinical detectable disease, 3 is clinical symptomatic), they assume that while a woman is in state 2, her tumor would be found 100% of the time, if screened. For women with very dense breast tissue this may never occur. Indeed, it is a different definition of pre-clinical than commonly used, where it would be unreasonable to assume the sensitivity of mammography to be 100% for pre-clinical detectable disease (Kerlikowske *et al.*, 1998; Peer *et al.*, 1996).

Finally, the model of Duffy *et al.* assumes that knowing an individual's state at prior times provides no additional information in assessing her likely future. But, the longer a tumor is in a pre-clinical state, the slower its growth rate and the better the patient's prognosis. Unfortunately, Markov models do not allow such a dependence. We attempt to address all of the above issues in our design and analysis.

While Duffy *et al.* addressed the efficacy of mammographic screening through a Markov model of breast tumor progression, Mushlin *et al.* (1998) do so by combining

different datasets through a meta-analysis. Mushlin *et al.* estimate the mammographic sensitivity of the very first screen, and not subsequent screens. As they point out, tumors at first screen tend to be larger. "Prevalent cases are more likely to have larger tumors which are easier to detect; the accuracy in subsequent rounds would differ." But these subsequent rounds are important. The benefit of a routine screening program is enhanced by finding tumors at screens after the first.

The focus of their investigation is the "true positive rate" (TPR) defined as $TPR = \frac{\#\text{ found}}{\#\text{ found} + \#\text{ missed}}$. As mentioned earlier, (Mushlin *et al.*, 1998) is one the papers that defines the number missed (false-negatives) to be the number of cancers discovered within one year of the first screening, before the second screen.

One can more accurately assess the sensitivity of mammography by incorporating retrospective looks at negatively-read mammograms. When a tumor is detected, its previous negative mammograms can be reread to see if any were originally misread (false-negative). Retrospective looks, however, are not perfect, and may be misread as well. Counting symptomatic cancers within one year of screens excludes both fast and slow-growing tumors. Assessing sensitivity using retrospective looks does not depend as strongly on the growth rates of the tumors.

Like mammographic sensitivity, growth rate estimation has received considerable research attention. Peer *et al.* (1993) estimate growth rates based on two measurements of each tumor. Their model assumes an exponential tumor growth rate, and uses tumors whose earlier measurements were found retrospectively. Peer *et al.* acknowledge the possibility of a Gompertzian fit being better (von Fournier *et al.*, 1980). However, they point out that Gompertzian parameter estimation requires at least three measurements of each tumor, and they had only two. In addition, even if they had several measurements, exponential and Gompertzian growth are similar during early tumor growth.

Peer *et al.* analyzed the growth rates of 236 women from the Nijmegen program (Peeters *et al.*, 1989). They concluded that the tumors detected in younger women grew faster. In our work we will assume that growth rate has no dependence on age of onset. That their conclusion is not incompatible with our assumption can be seen as follows. Consider a cohort of women who have breast tumors that begin growing at age 35. The tumor growth rates vary, some slow and others fast. Those that are detected (either by screen or symptoms) before age 40 are faster-growing. Thus, tumors *detected* when the women are older tend to be slower growing.

Slower-growing tumors permit longer screening intervals. Thus, for our examples, the recommendation by Peer *et al.* for longer screening intervals in these older women has some justification. The conclusion that growth rate depends on age, however, needs to be qualified to be age at detection. The younger women with slower growth rates are not detected when they are younger, and thus present themselves only when they are older. One possible extension of the work of Peer *et al.* would be to extend all the growth rates back through time until the age of onset; and then look at the distribution of growth rates by age of onset.

Spratt *et al.* (1986) estimate shorter sojourn times in younger women. Their estimate is subject to the same bias as mentioned for Mushlin *et al.*. Their results agree with Peer *et al.* and indicate that growth rates depend inversely on age at detection.

In our design, we model each tumor's growth as beginning at a particular patient age, with an associated growth rate. Like Peer *et al.*, we assume exponential growth. Spratt *et al.* (1993) discuss the data we use for the two-measurement design. Women underwent routine mammography as part of breast cancer screening programs at the University of Heidelberg (Heuser *et al.*, 1979) and as part of the Breast Cancer Detection and Demonstration Project at the University of Louisville (von Fournier

et al., 1980). Paired serial mammographic tumor measurements (1 pair per patient) and the patient's associated ages were obtained for 338 patients: 306 from Heidelberg and 32 from Louisville. The earlier measurement was seen retrospectively, after the cancer was detected at the second mammogram. Thus for each patient, we have a measured tumor size at two consecutive screening times, representing untreated tumor growth (Spratt *et al.*, 1995; Spratt *et al.*, 1996). No information was available on the number of previous negative mammograms.

Spratt *et al.* (1993, 1995, 1996) and von Fournier *et al.* (1980) use this data to estimate Gompertzian and logistic growth rates. These decelerating tumor growth models are supported by various studies (Laird, 1965; Lala and Patt, 1966). When trying to fit such models with only two measurements, additional assumptions are required, such as for maximal size. As Peer *et al.* noted, these rate laws are well-approximated by exponential growth during early tumor development. Rather than require threshold assumptions, we use the data from Spratt *et al.* (1993) to fit exponential growth.

Brookmeyer *et al.* (1986) estimated pre-clinical sojourn time and test sensitivity jointly. Their work set the precedence for the combined evaluation of these two. As they point out, the quantities are enmeshed and interdependent: one cannot estimate sojourn time without test sensitivity, and vice versa. Their case-controlled method is quite different than ours; and their application area is pap smear testing for cervical cancer. Their test is a cellular scraping, not radiological. However, their ground-breaking work set the stage for our extension towards joint-estimation of tumor growth rate and mammographic sensitivity of breast cancer.

Indeed, all the mentioned works provide the various ingredients we bring together to create our models. We now summarize their contributions and associated issues. Most reports of mammographic sensitivity refer to the screening program rather than

to a single screen. As such, they focus on age dependence and not size dependence. While the two sensitivities are related, modeling of individual decision making could require explicit consideration of test (not program) sensitivity.

With only two measurements, exponential tumor growth is reasonable and approximates the preferred decelerating rates during early stages of growth. Current models of growth, however, do not include single-measurement cases. Their exclusion may lead to biased estimates. Also, current models do not jointly estimate both growth rate and mammographic sensitivity, an idea introduced in a different setting by Brookmeyer *et al.*. In the next section, we discuss how our model and complete-measurement design merges these ideas to address these issues.

1.3 The Model

Our design and analysis are novel in several ways. First, we estimate mammographic sensitivity based on both age and tumor size. We focus on the sensitivity of a single mammogram rather than of a particular screening regimen. Second, we jointly estimate tumor growth and mammographic sensitivity. Third, we incorporate data from women's previous negative mammograms, both with and without retrospective looks. Fourth, in our complete-measurement design we also include women whose tumors have only one measurement. These latter women are not included in the above referenced studies. Our work expands upon existing models, attempting to address the biases mentioned in the previous section.

1.4 Notation

Within the two-measurement and complete-measurement models, we use a common notation to discuss the data and parameters, as follows.

1.4.1 Observables

i patient index.

k mammogram index.

a_{ik} age of the i^{th} patient at her k^{th} mammogram. These can differ from patient to patient (e.g. due to lack of compliance); while a_{ik} will typically be random, we assume that their distribution does not depend on the model parameters.

t_{ik} total number of readings of the i^{th} patient's k^{th} mammogram (can be 1 or 2).

r^1_{ik} indicator of positive 1^{st} reading of the k^{th} mammogram: r^1_{ik} is 1 for the tumors measured at first reading, and is 0 for the tumors missed at first reading.

r^2_{ik} indicator of positive 2^{nd} reading of the k^{th} mammogram. It is only defined when a mammogram has been read twice, or $t_{ik} = 2$.

v_{ik} volume of the i^{th} patient's tumor at her k^{th} mammogram. It is observable at $k = n_i + 1$ (i.e., $r^1_{ik} = 1$) or when $r^2_{ik} = 1$.

$n_i + 1$ index of the mammogram at which the tumor was first detected. Patient i has n_i negative mammograms before her first positive mammogram. $v_{i(n_i+1)}$ and $a_{i(n_i+1)}$ are the tumor volume and her age when her tumor is first detected. n_i is a random variable and could be greater than the number of mammograms, although then its value would not be observed.

s_{ik} binary indicator of a symptomatic (i.e., diagnostic) mammogram. It equals zero if the k^{th} mammogram is a screening mammogram (i.e., the

patient has no symptoms at the time of the exam).

$\boldsymbol{D_i}$ all the observables whose distribution depends on the model's parameters (everything above except for the ages).

1.4.2 Unobservables

g_i exponential growth rate of tumor i. It is equal to the slope between successive log volumes $(\log(mm^3)/yr)$.

a_i^* age of patient i when her tumor volume is 1 mm^3; the associated volume is $v_i^* \equiv 1\ mm^3$.

m_i index of the first mammogram taken after the tumor volume exceeds 1 mm^3, or, in other words, the index of the mammogram immediately following a_i^*.

$\mu_1,\ \sigma_1^2$ mean and variance of the log-normal distribution of growth rates.

$\alpha_1,\ \beta_1,\ \beta_2$ intercept, $\sqrt[3]{vol}$ coefficient, and age coefficient in the logistic regression of the probability of a positive first reading, $p(r_{ik}^1 = 1)$.

1.5 Data for Two-Measurement Design

The data consist of asymptomatic women whose tumors were detected by screening mammography, and whose previous negative mammogram reveals the tumor upon second reading. For each of these women, the data consist of her age and estimated tumor volume at each of the two mammograms.

For example, a woman enters the screening program at age 50. She receives routine screening mammograms every two years, at ages 50, 52, and 54. At age 54, the woman's mammogram is positive, with a tumor measured to be 10 mm^3. The

radiologist looks back at her previous mammogram and decides that the tumor was in fact present at the time, and its volume was 5 mm^3. The data from this subject for the two-measurement design are presented in Table 1.1.

Table 1.1: Hypothetical two-measurement design data for patient i

k	n_i	$n_i + 1$
r_{ik}^1	0	1
r_{ik}^2	1	-
t_{ik}	2	1
v_{ik}	5	10
a_{ik}	52	54

The data we examine for the two-measurement design come from Spratt *et al.* (1993), as discussed in the literature review. Spratt *et al.* provide diameter measurements of the tumors along two axes. Using these measurements, we calculated each tumor's associated oblong spheroid volume. Figures 1.1 and 1.2 are graphs of the transformed data. In Figure 1.1, each line represents a patient; the x-axis is the woman's age, while the y-axis is her tumor size (log tumor volume, in $ln(mm^3)$). Some tumors have more than two measurements. We used only those women with exactly two measurements, corresponding to the lines in Figure 1.1 with 2 points. We take the slopes of the lines in Figure 1.1 to be the growth rates of the tumors (change in log volume/change in time). Figure 1.2 shows the patient growth rates across ages. These are log of the log growth rates, $ln(ln(mm^3)/yr)$. There is little evidence for an age trend in these data.

We assumed that growth rate is independent of age-of-onset and tumor size. Many researchers (Hart *et al.*, 1998; Heitjan, 1991; Moskowitz, 1986) indicate that growth should depend on one or both of these factors. Graphs of the data from Spratt

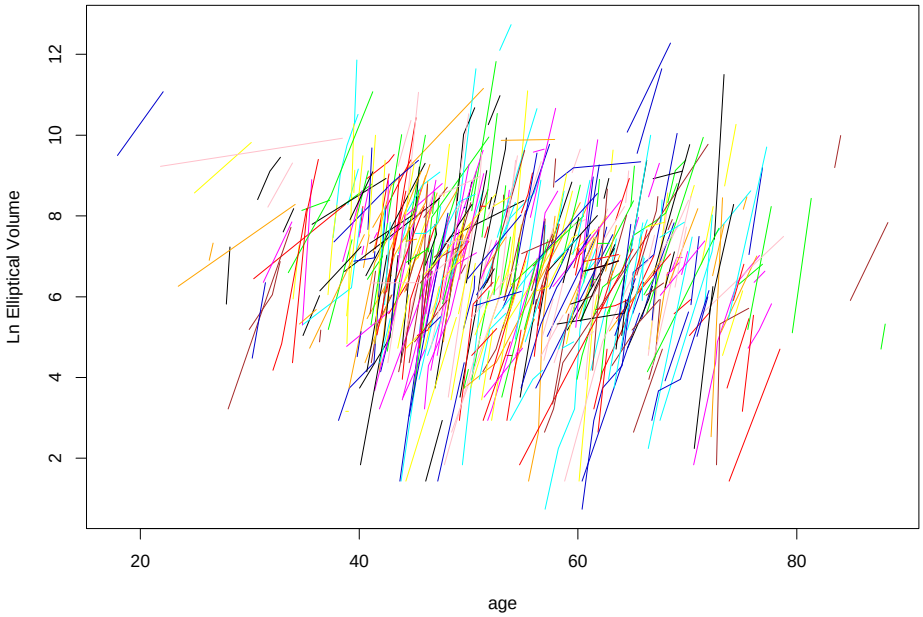

Figure 1.1: Log tumor volumes, by patient age.

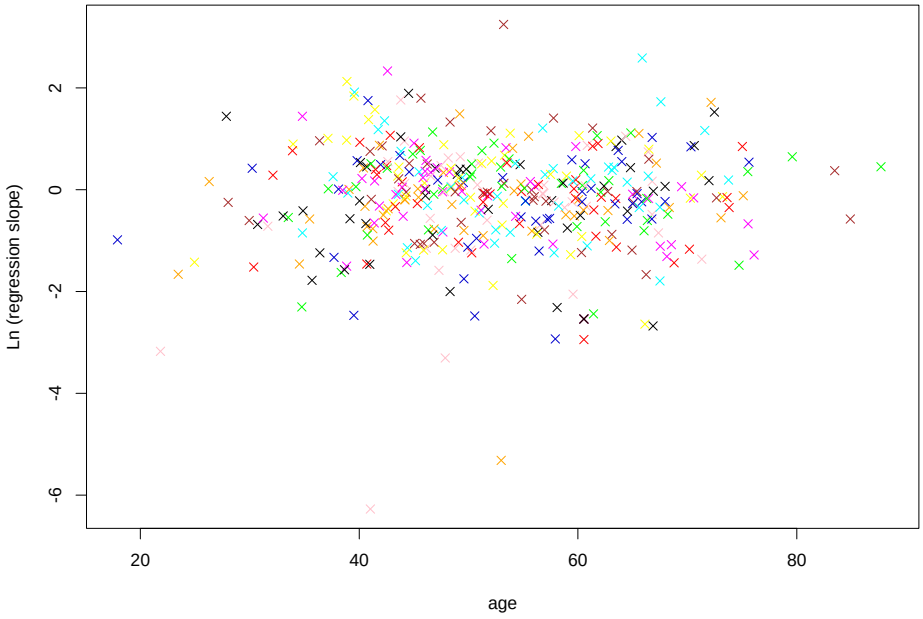

Figure 1.2: Log tumor growth rates, by patient age.

et al. (see Figures 1.2 and 1.3) show no such dependence and support our use of independent growth rates.

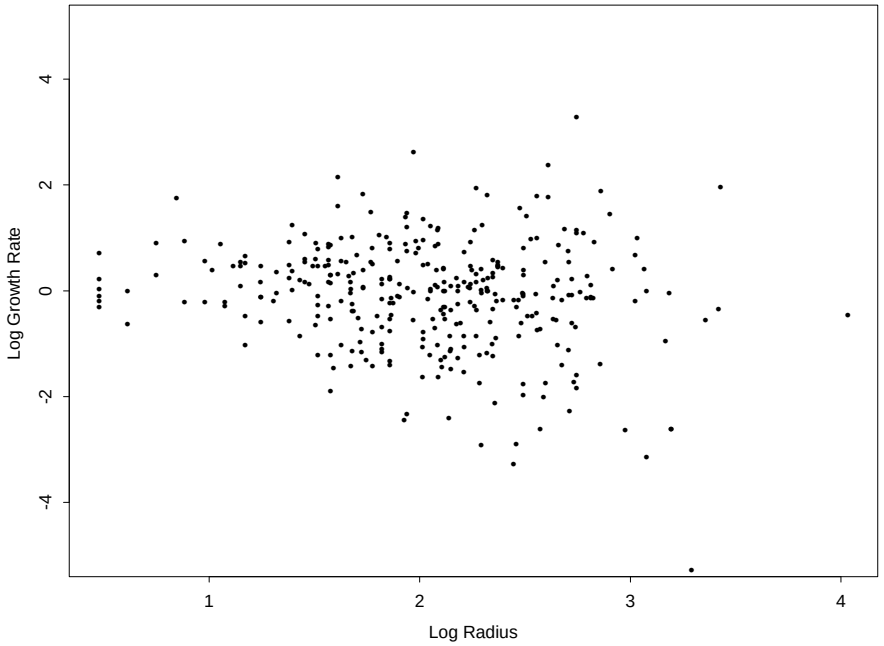

Figure 1.3: Two-Measurement Design Data: Log(Rate) vs. Log(Radius)

1.6 Data for Complete Measurement Design

Data for the complete-measurement design includes the following:

- Woman's age at each mammogram
- Reason for each mammogram: screening or diagnostic (symptomatic).
- Number of readings (1 or 2) for each mammogram
- Whether the tumor was detected, at each reading
- Measurement of tumor size, if detected

k	1	2	3	4	5	6	7
r_{ik}^1	0	0	0	0	0	0	1
r_{ik}^2	-	-	-	0	1	1	-
t_{ik}	1	1	1	2	2	2	1
v_{ik}	-	-	-	-	3.5	6.1	10
s_{ik}	0	0	0	0	0	0	0
a_{ik}	57	58	60	62	63	65	66

Table 1.2: Hypothetical full data, D_i, for patient i

In Table 1.2, we give example data for the complete-measurement design. In this example, patient i has a total of 7 screening mammograms. At each screening, we know her age (a_{ik}) and symptomatic status (s_{ik}). The first six mammograms received negative readings when they were first read by the radiologist ($r_{i1}^1 = 0, \ldots, r_{i6}^1 = 0$). When the 7^{th} mammogram is read as positive upon first look ($r_{i7}^1 = 1$), the radiologist then looks back at the mammograms before the 7^{th} one, starting with the 6^{th} and working back until no longer seeing a tumor (retrospectively). These mammograms (4 through 6, in our example) receive a second reading ($t_{ik}^1 = 2$). Of these, mammograms 5 and 6 are positive on second reading ($r_{i5}^2 = 1, r_{i6}^2 = 1$). We have tumor volumes (v_{ik}) for the mammograms that received a positive first or second reading (and therefore the volumes could be measured).

1.7 Biases

There is a bias towards slower growth rates inherent in any screening data (Zelen, 1976). In particular, because at least two measurements of a tumor are needed for any direct measure of growth rate, tumors must grow slowly enough to be measured on at least two screens. Single measurement tumors may constitute a large part of

the results from a screening program. A fast-growing tumor that is missed on a mammogram would likely become symptomatic before the woman's next mammogram. Therefore, this tumor would not be included in the two-measurement design. Fast-growing tumors are seldom seen; and slow-growing tumors are therefore over-represented. Thus, growth rate estimations based on the two-measurement design are biased.

We model the distribution of growth rates as a log-normal with parameters μ_1 and σ_1. When based on the two-measurement design, estimates of μ_1 and σ_1 are affected by the bias towards slower growth rates. An implication is that the estimate of μ_1 will have a downward bias. In addition, the estimate of σ_1 will also tend to be lower than that of the true population.

The two-measurement design also imposes a bias on estimates of mammographic sensitivity. The sensitivity of mammography is the probability that a radiologist detects a tumor upon reading a woman's mammogram when a tumor is present. We model this probability of a positive reading as a logistic regression on the tumor's $\sqrt[3]{vol}$ and the woman's age:

$$\log\left(\frac{p(r_{ik}^1 = 1|a_{ik}, v_{ik})}{1 - p(r_{ik}^1 = 1|a_{ik}, v_{ik})}\right) = \alpha_1 + \beta_1 \sqrt[3]{v_{ik}} + \beta_2 a_{ik}.$$

This bias towards slower growth rates gives rise to biased estimations of the three logistic parameters. In the two-measurement design, mammographic sensitivity is modeled using readings that were negative on first look, but positive retrospectively. These negative readings are the only ones used in the model. However, there are women whose tumors were missed even upon retrospective reading. Presumably, these tumors tend to be smaller. They are not included in the two-measurement design. We miss these single-measurement cases, which have a higher fraction of fast growing cancers. Therefore, based on the two-measurement design, estimates of mammographic sensitivity are too high. The parameter that determines overall

mammographic sensitivity is α_1. Thus, estimates of α_1 are too high. Because estimates of α_1, β_1 and β_2 depend strongly on the distribution of patient ages and tumor volumes, it is difficult to assess the direction of the bias in their estimates.

The complete-measurement design misses fewer cases than the two-measurement design. Knowing the patient ages at past negative mammograms allows us to estimate growth rate parameters (μ_1 and σ_1) and sensitivity parameters (α_1, β_1, and β_2) for the single-measurement cases, which tend to have faster growth rates in comparison to tumors with two or more measurements. Having these faster growth rates in the data expands the range of the observed growth rates, improving our estimates of σ_1. Having these higher rates also improves our estimates of the mean, μ_1, of their log-values. Because our growth rate parameter estimates are improved, we can more accurately estimate tumor sizes. For single measurement cases, we can now better estimate the size of the tumor not seen retrospectively. These tumors tend to be smaller and were subsequently missed. Thus, including single-measurement cases also aids our estimation of the mammographic sensitivity parameters, α_1, β_1, and β_2. Because it captures more of the population and more information per patient, the complete-measurement design resolves some of the bias issues in the two-measurement design.

While better than the two-measurement design, the complete-measurement design has limitations. First, retrospective readings are not perfect: some of the tumors missed on first reading are missed again on retrospective look. Second, there is still a selection bias of slower growth rates. While the complete-measurement design incorporates the single-measurement cases, which tend to be faster-growing, we cannot directly estimate a tumor's growth rate from only one data point. Later, we will discuss how one might incorporate single-measurement cases.

Chapter 2

Two-Measurement Design

2.1 Introduction

We begin by considering data from the two-measurement design, as described in Section 1.5. The data consist of 338 women who took part in a regular mammographic screening program. Each of these women had a tumor whose size could be determined at two different times from her screening mammograms. After detecting the larger volume at its screen, radiologists found the smaller measurement retrospectively. We assume no measurement error once a tumor is detected, even though we model the uncertainty of actual mammographic detection. In this chapter, we propose a model and analysis for data from the two-measurement design.

2.2 Model

Based on the two-measurement design, we propose a model to jointly estimate tumor growth and mammographic sensitivity. Recall from Chapter 1 that we assume that pre-clinical tumor growth is exponential. We model the distribution of these growth rates as log-normal with parameters μ_1 and σ_1; and mammographic sensitivity as a logistic regression with parameters α_1, β_1, and β_2, where β_1 is the effect of $\sqrt[3]{vol}$ and

β_2 is the effect of age.

We assume a non-informative normal prior on α_1, β_1 and β_2; a non-informative truncated normal prior on σ_1; and a proper though highly variable prior on μ_1.

$$\alpha_1, \beta_1, \beta_2 \overset{iid}{\sim} Normal(0, 1000000),$$

$$\sigma_1 \sim HalfNormal(0, 1000000),$$

$$\mu_1 \sim Normal(\mu_2 = -0.5, \sigma_2 = 3).$$

The mean of the log growth rates is μ_1. We base our informed, though vague, μ_1 prior on published estimates of tumor volume doubling time (see Table 2.1). Since we assume exponential tumor growth, the tumor doubling time is the ratio of the growth rate to ln(2). Our prior on μ_1 centered at -0.5 corresponds to a prior on growth rates being centered around 0.6 $\frac{log(mm^3)}{yr}$.

Table 2.1: Reported estimates of tumor volume doubling times. The variable n is the number of data points; DT is the reported mean doubling time; g is the corresponding exponential growth rate

Study	(von Fournier *et al.*, 1980)	(Heuser *et al.*, 1979)	(Spratt *et al.*, 1996)
n	147	32	352
DT	212 days = 0.581 yr	325 days = 0.890 yr	352 days = 0.964 yr
g	0.404	0.617	0.668

The contribution of a patient to the likelihood for $\theta = (\alpha_1, \beta_1, \beta_2, \sigma_1, \mu_1)$ is the probability of the patient's history (D) conditional on the parameters, the cancer existing during the screening program (c_e), and the cancer being found (c_f). We omit the subscript i for ease of reading. By laws of conditional probability:

$$p(D|\theta, c_e, c_f) = \frac{p(D, c_f|\theta, c_e)}{p(c_f|\theta, c_e)} = \frac{p(D|\theta, c_e)}{p(c_f|\theta, c_e)},$$

as the event c_f is contained in the history.

We assume that the denominator $p(c_f|\theta, c_e) \approx 1$. That is, given that a tumor exists, the probability of it being found during the screening program is approximately one. It is possible for a tumor to be missed repeatedly. But, as the tumor grows, the probability of detection or symptoms increases. Given that the tumor exists during the screening program, it is likely to be detected eventually.

In a long screening program, i.e., one in which each woman could be screened for several years, it seems reasonable that most cancers would be found, given they exist during the screens. Patients with very slow-growing tumors and whose age of onset is near the end of the screening program will be missed. This subset is assumed small relative to the rest of the subjects. We do not know exactly how long the 300,000 women in the dataset were screened. Data from (Spratt *et al.*, 1993; Spratt and Spratt, 1985) suggest that each woman could be screened up to 12 or more years. Thus, we argue it is reasonable to suggest $p(c_f|\theta, c_e) \approx 1$. Our likelihood is then approximated by the numerator

$$p(D|\theta, c_f, c_e) \approx p(D|\theta, c_e).$$

The data from Spratt *et al.* include both screen-detected and symptomatic cases, but we do not know which is which. Thus, in the two-measurement model, we do not incorporate symptomatic status. We model mammographic sensitivity as a logistic regression on the woman's age and tumor size, regardless of her symptomatic status. This is not ideal, but required due to the unavailability of further information. In Chapter 3, we discuss improvements in this area based on our complete-measurement design and associated model.

We now present our model of $p(D|\theta, c_e)$. Recall that the data from Spratt *et al.* consist of 338 women. Each woman (patient i) received routine screens until a tumor was found at her $(n_i + 1)^{th}$ mammogram ($r^1_{i(n_i+1)} = 1$). Retrospective reading of the n_i^{th} mammogram revealed a tumor that had originally been missed ($r^1_{in_i} = 0$ and

$r^2_{in_i} = 1$). Thus we have:

$$p(D_i|\theta, c_e) \propto \quad p(r^1_{in_i} = 0|a_{in_i}, v_{in_i}) \cdot p(r^1_{i(n_i+1)} = 1|a_{i(n_i+1)}, v_{i(n_i+1)})$$
$$\cdot p(r^2_{in_i} = 1|a_{in_i}, v_{in_i}) \cdot p(v_{in_i}, v_{i(n_i+1)}).$$

The available data contain no information about the value of n_i. For example, we do not know whether woman i was detected at her second screen or her seventh.

We now transform $p(v_{in_i}, v_{i(n_i+1)})$ to $p(a^*_i) \cdot p(g_i)$. This change of variables does not depend on parameters. Thus, its associated Jacobian can be ignored from the likelihood analysis.

$$p(D_i|\theta, c_e) \propto \quad p(r^1_{in_i} = 0|a_{in_i}, v_{in_i}) \cdot p(r^1_{i(n_i+1)} = 1|a_{i(n_i+1)}, v_{i(n_i+1)})$$
$$\cdot p(r^2_{in_i} = 1|a_{in_i}, v_{in_i}) \cdot p(a^*_i) \cdot p(g_i).$$

We model the probability of finding either tumor at first look as a logistic regression on the patient's age and tumor volume. That is:

$$\log\left(\frac{p(r^1_{ik} = 1 \mid a_{ik}, v_{ik})}{1 - p(r^1_{ik} = 1 \mid a_{ik}, v_{ik})}\right) = \alpha_1 + \beta_1 \sqrt[3]{v_{ik}} + \beta_2 a_{ik},$$

or equivalently:

$$p(r^1_{ik} = 1 \mid a_{ik}, v_{ik}) = \frac{e^{\alpha_1 + \beta_1 \sqrt[3]{v_{ik}} + \beta_2 a_{ik}}}{1 + e^{\alpha_1 + \beta_1 \sqrt[3]{v_{ik}} + \beta_2 a_{ik}}},$$
$$p(r^1_{ik} = 0 \mid a_{ik}, v_{ik}) = 1 - p(r^1_{ik} = 1 \mid a_{ik}, v_{ik})$$
$$= \frac{1}{1 + e^{\alpha_1 + \beta_1 \sqrt[3]{v_{ik}} + \beta_2 a_{ik}}}.$$

Given the tumor is detected at first look for the larger of the two measurements, we assume that the probability of finding the first volume at second look is a constant:

$$p(r^2_{ik} = 1|v_{ik} > 1mm^3) = p_2 = .95.$$

Our model contains no parameters of interest related to age of onset. Age of onset is assumed to follow a known gamma distribution.

$$a_i^* \sim Gamma(A = 30, B = .5).$$

This gamma distribution has a mean of 60 and standard deviation of 9.5. We base this probability density on results by Parmigiani and Skates (1999).

As noted earlier, the tumor growth rates are modeled using a log-normal distribution. That is:

$$log(g_i) \sim Normal(\mu_1, {\sigma_1}^2).$$

2.3 Computing

We are interested in marginal inference. Unfortunately, direct analytical integration of the posterior distribution is not feasible. Thus, we sample the posterior distribution via a Markov Chain. The mean of this sample ergodically approximates the true mean of the marginal posterior distributions (Roberts, 1995; Tierney, 1995).

We utilized a Markov Chain Monte Carlo algorithm (Gelfand and Smith, 1990; Tierney, 1994), iteratively sampling $\theta = (\alpha_1, \beta_1, \beta_2, \mu_1, \sigma_1)$ jointly to obtain a sample from the posterior distribution of θ. That is, we update current values of the parameters, denoted by θ^c, by using a Metropolis-Hastings step (Metropolis *et al.*, 1953; Hastings, 1970). Specifically, we generated proposal values, θ^p, for the parameters from normal distributions with mean θ^c and variance equal to twice the variance of their maximum likelihood estimates. These MLE's were used as initial values for θ^c; although after an appropriate "burn-in" period, the chain should no longer 'remember' its starting values. θ^p is accepted with the probability

$$\alpha(\theta^c, \theta^p) = min\left(1, \frac{p(\theta^p)}{p(\theta^c)}\right),$$

where $p(\theta)$ is the posterior $p(\theta|a_{i1}, v_{i1}, a_{i2}, v_{i2})$ evaluated at θ. Otherwise, the candidate θ^p is rejected and the last sampled value, θ^c, is kept. The acceptance rate was 30.34%. That is, 30.34% of the candidate values were accepted. (Gelman *et al.*, 1996) show that acceptance rates of 44% are optimal in one-dimensional problems and 24% in high-dimensional problems. Our acceptance rate fits well within these guidelines.

2.4 Results

We ran the MCMC chain for 20,000 iterations. Visual inspections of the output determined the length of the burn-in. We discarded the first 10,000 iterations. We decided to stop at 20,000 iterations because, with different starting values and after the burn-in, several runs agreed on estimates of posterior means within 2 or 3 significant digits. The second 10,000 iterations, which we kept, resulted in the following posterior averages: $\bar{\alpha_1^*} = -0.4048$, where α_1^* is defined below, $\bar{\beta_1} = 0.158$, $\bar{\beta_2} = 0.0217$, $\bar{\mu_1} = -0.337$, and $\bar{\sigma_1} = 2.85$. Figure 2.1 depicts the histograms of the posterior values. The volume and age values were re-centered for stability in convergence of the MCMC chain. Recall that

$$\log\left(\frac{p(r_{ik}^1 = 1|a_{ik}, v_{ik})}{1 - p(r_{ik}^1 = 1|a_{ik}, v_{ik})}\right) = \alpha_1 + \beta_1 \sqrt[3]{v_{ik}} + \beta_2 a_{ik}.$$

We have calculated

$$\log\left(\frac{p(r_{ik}^1 = 1|a_{ik}, v_{ik})}{1 - p(r_{ik}^1 = 1|a_{ik}, v_{ik})}\right) = \alpha_1^* + \beta_1(\sqrt[3]{v_{ik}} - 9) + \beta_2(a_{ik} - 53).$$

Our posterior value for α_1^* corresponds to the following original parameterization: $\alpha_1 = \alpha_1^* - \beta_1 \cdot 9 - \beta_2 \cdot 53$.

The resulting probability of mammographic detection given a woman's tumor size is shown in Figure 2.2 across three different patient ages. The variability of this

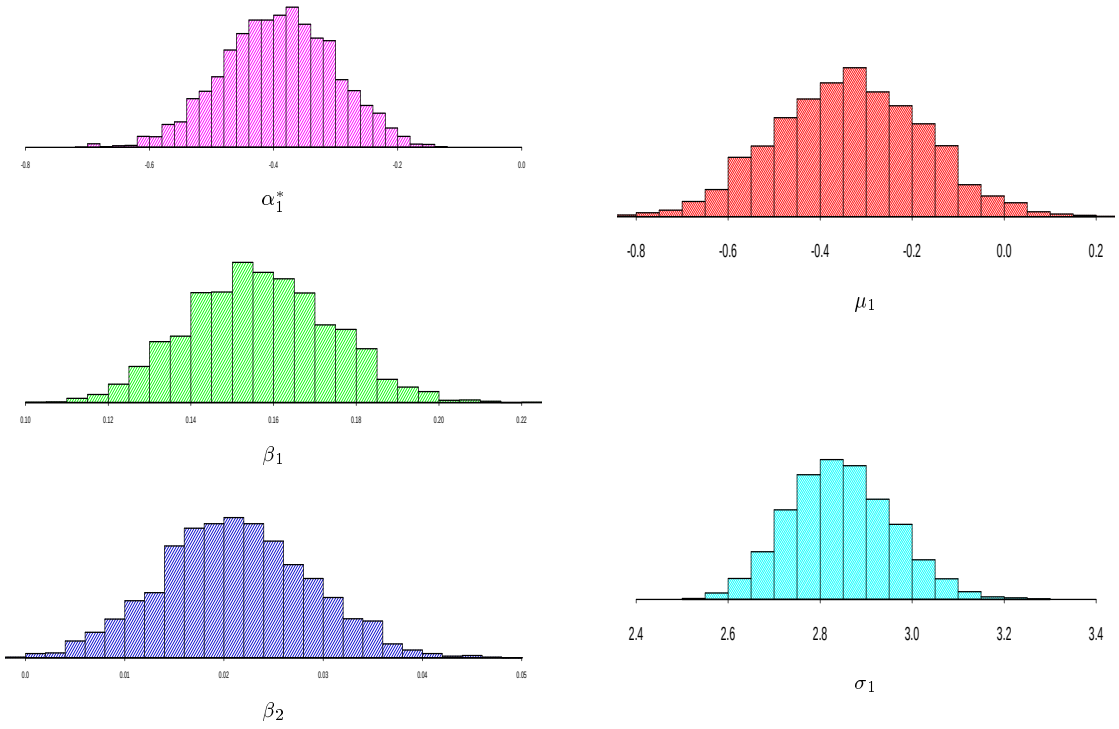

Figure 2.1: MCMC Output.

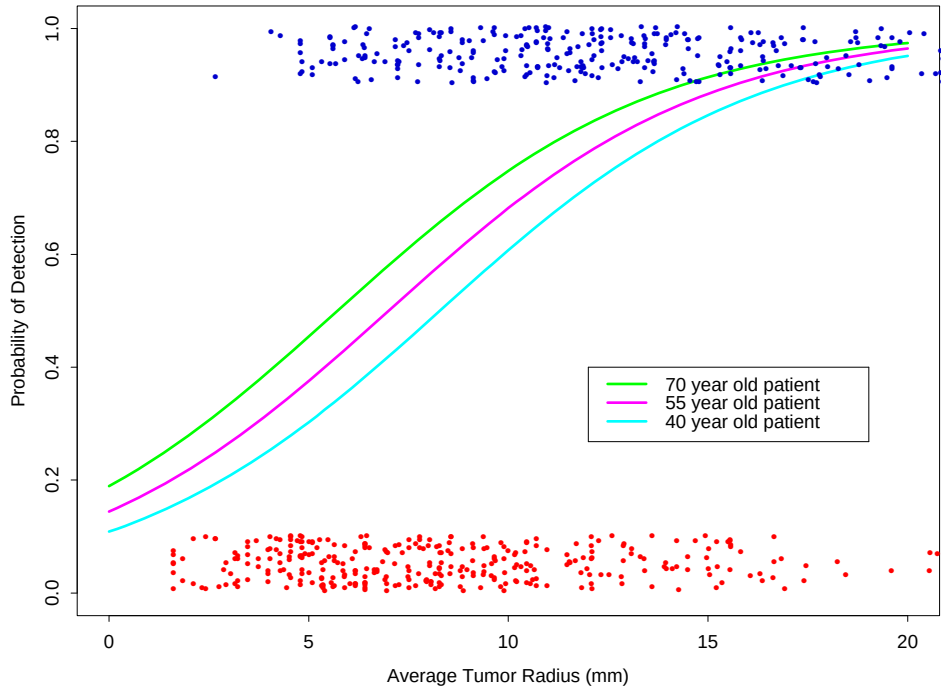

Figure 2.2: Probability of detection overlaid with points showing tumors detected (top of figure) and not detected (bottom).

estimate can be seen in Figure 2.3, which shows 20 different probability of detection curves for a patient who is 55 years old, where each line is drawn for a different triple of posterior values for (α_1, β_1, β_2) from the chain. The histogram in Figure 2.4 shows the actual patient growth rates. The curve in Figure 2.4 is the estimated distribution of growth rates, based on the posterior MCMC values. Its long tail is made possible by the log-normal assumption on growth rates and driven by a few extremely high growth rates (up to 25 $log(mm^3)/year$).

2.5 Discussion

We know of no measure of mammographic sensitivity by tumor size (van Dijck *et al.*, 1993). In recent literature, the 3 curves for mammographic sensitivity seen in Figure 2.2 are 3 horizontal lines, varying for age, but not for tumor size (Perloff *et al.*, 1996). It is important to jointly estimate, as we do, the patient's growth rate and the sensitivity of mammographic detection. Recall that if a patient has a tumor found at first reading, but upon looking back to the previous mammogram, there is no tumor (of size v_{in_i}) found upon retrospective second reading, then one of two cases occurred. Either there was no tumor and the growth rate was very fast, or the tumor was there but missed by mammography. Therefore, our model's joint estimation of growth rate and mammographic sensitivity address their inter-dependence for a given patient history.

The present data include only a subset of possible patient histories, namely those who have tumors viewable at 2 mammograms: the latter at time of screen, and the former upon retrospective reading. The exclusion of other cases, especially single-measurement cases, indicates a likely length bias in the growth rates (Zelen, 1976). That is, the data tend to include only slower growing tumors viewable on 2 screened mammograms. While we try to address this issue through directly modeling the

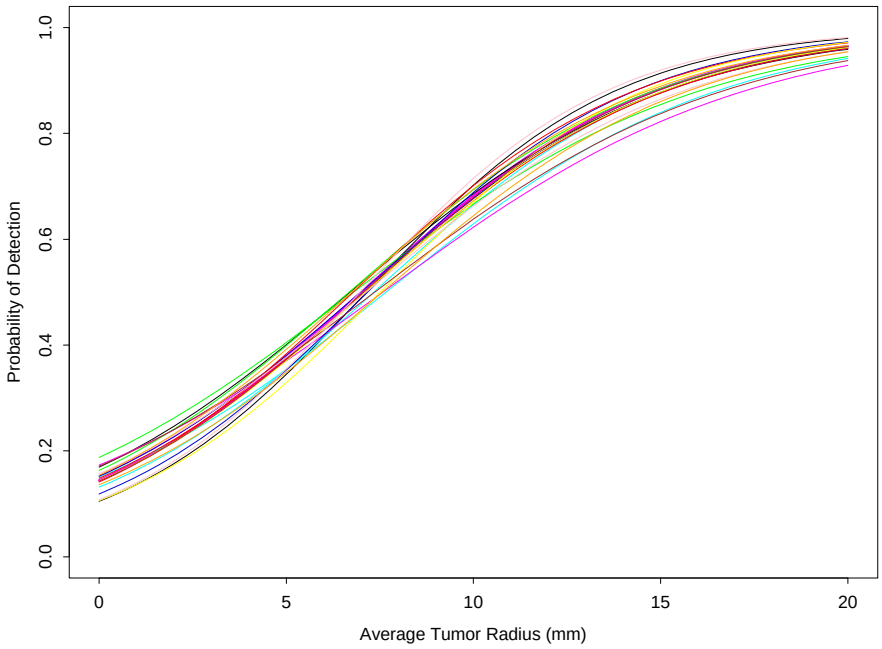

Figure 2.3: Sample of curves representing the probability of detection for a 55 year-old woman, using 20 different MCMC triples of α_1, β_1, and β_2.

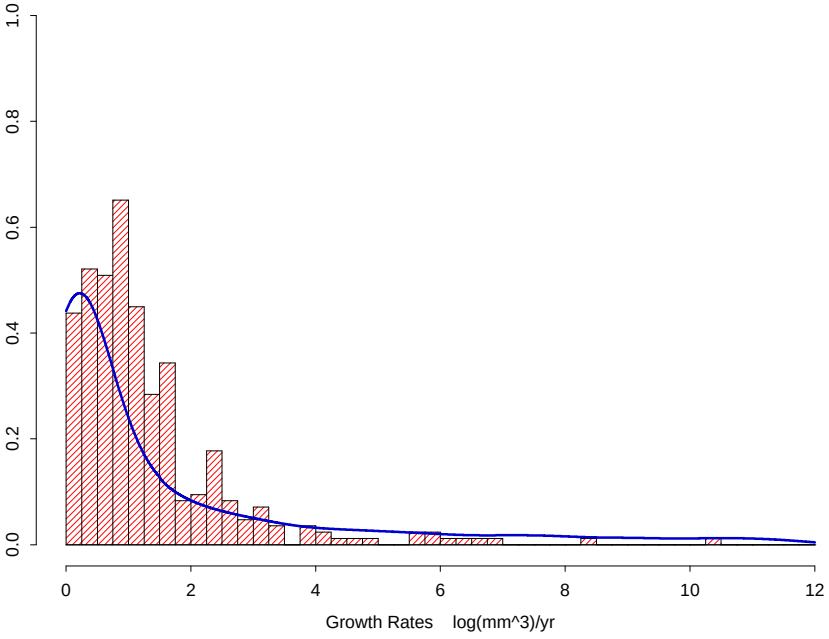

Figure 2.4: Distribution of predicted (curve) and observed (histogram) growth rates.

selection process, the absence of any other types of patients still presents a concrete bias concern. To further understand the bias of the data, patient histories were simulated based on the assumptions of our model (See Chapter 4). The patients whose tumors are found during a screening program fall into one of 4 cases: (1) The tumor is found at first screen ($k = 1$), *i.e.*, prevalent cases. (2) The tumor is not seen until $(n_i + 1)^{th}$ mammogram ($v_{i(n_i+1)}$), and retrospective reading of the previous mammogram is negative (this includes tumors that were missed upon both viewings, and those that started growing between the two mammograms). (3) The tumor is missed until first seen on $(n_i + 1)^{th}$ mammogram ($v_{i(n_i+1)}$), and is seen at previous mammogram (v_{in_i}) retrospectively. (4) No tumor is mammographically detected, but the woman becomes symptomatic (leaves pre-clinical phase and enters clinical phase). The data we used contained only those patients who are in group (3). From the simulation, around 50% fell into this group. The following chapters expand the model to accommodate data from all 4 groups with full information and explore bias inherent in both the model and the data.

Chapter 3

Complete-Measurement Design

We begin by considering data from the complete-measurement design, as described in Section 1.6. Recall that the data consist of women whose tumors were found during a screening program, either by a mammographic screen or due to symptoms. For each woman, we know her complete mammographic history, including ages at past negative mammograms, symptomatic status, and how many times each mammogram was read. In this chapter, we propose a model and present examples of its application to data from the complete-measurement design.

3.1 Full Likelihood

3.1.1 Likelihood Overview

Our model for the complete-measurement design builds upon our model for the two-measurement design. As described in Chapter 2, there are 5 parameters of interest. We retain the non-informative normal prior on α_1 (intercept), β_1 ($\sqrt[3]{v_{ik}}$ effect), and β_2 (a_{ik} effect); the non-informative truncated normal prior on σ_1; and the informed,

though dispersed, prior on μ_1:

$$\alpha_1, \beta_1, \beta_2, \overset{iid}{\sim} Normal(0, 1000000),$$

$$\sigma_1 \sim HalfNormal(0, 1000000),$$

$$\mu_1 \sim Normal(\mu_2 = -0.5, \sigma_2 = 3).$$

The probability of the observed data for woman i, $\boldsymbol{D_i}$, is equal to the product of the conditional probability of missing the tumor at each mammogram before it was finally detected, the conditional probability of finding it at the patient's age and tumor size when it was detected, and the conditional probability that she was asymptomatic during all of the intervals. Recall from Chapter 2 that we are conditioning on the probability that the cancer exists (c_e) and that the cancer is found (c_f); but, we assume $p(\boldsymbol{D} \mid \theta, c_f, c_e) \approx p(\boldsymbol{D} \mid \theta, c_e)$ (see Section 2.2). Assuming exponential tumor growth, and given a_i^* and g_i, one can calculate all volumes $\boldsymbol{v_i}$ for given ages $\boldsymbol{a_i}$. In symbols, the likelihood for data $\boldsymbol{D_i}$ and its associated growth rate g_i, given the parameter values and the ages at each mammogram, is

$$\begin{aligned}
&p(\boldsymbol{D_i}, g_i | \alpha_1, \beta_1, \beta_2, \mu_1, \sigma_1, \boldsymbol{a_i}, c_e) \\
&\quad \propto p(\text{all observed negative first readings } |\boldsymbol{a_i}, \boldsymbol{v_i}, \boldsymbol{s_i}) \\
&\qquad \cdot p(\text{symptomatic status } |\boldsymbol{a_i}, \boldsymbol{v_i}) \\
&\qquad \cdot p(\text{observed positive first reading at } v_{i(n_i+1)}, a_{i(n_i+1)} | a_{i(n_i+1)}, v_{i(n_i+1)}) \\
&\qquad \cdot p(\text{all observed positive } 2^{nd} \text{ readings } |\boldsymbol{v_i}) \\
&\qquad \cdot p(\text{all negative } 2^{nd} \text{ readings } |\boldsymbol{v_i}) \\
&\qquad \cdot p(a_i^*) \\
&\qquad \cdot p(g_i)
\end{aligned}$$

Rewriting this using notation developed earlier:

$$
\begin{aligned}
p(\boldsymbol{D_i}, g_i \mid \alpha_1, \beta_1, \beta_2, \mu_1, \sigma_1, c_e) \propto\ & p(r^1_{i1} = 0, \ldots, r^1_{in_i} = 0 \mid \boldsymbol{a_i}, \boldsymbol{v_i}, \boldsymbol{s_i}) && \textbf{(a)} \\
& \cdot p(r^1_{i(n_i+1)} = 1 \mid a_{i(n_i+1)}, v_{i(n_i+1)}, s_{i(n_i+1)}) && \textbf{(b)} \\
& \cdot p(\boldsymbol{s_i} | \boldsymbol{a_i}, \boldsymbol{v_i}) && \textbf{(c)} \\
& \cdot p(\text{all observed } r^2_{ik} = 1 \mid \boldsymbol{v_i}) && \textbf{(d)} \\
& \cdot p(\text{all observed } r^2_{ik} = 0 \mid \boldsymbol{v_i}) && \textbf{(e)} \\
& \cdot p(a^*_i) && \textbf{(f)} \\
& \cdot p(g_i). && \textbf{(g)}
\end{aligned}
\tag{3.1}
$$

3.1.2 Likelihood Decomposition

We continue to model the probability of mammographic detection of an asymptomatic tumor on its first reading as a logistic regression. The covariates are the patient's age and tumor volume, which is assumed to be greater than 1 mm^3.

$$
\log\left(\frac{p(r^1_{ik} = 1 | a_{ik}, v_{ik}, s_{ik} = 0)}{1 - p(r^1_{ik} = 1 | a_{ik}, v_{ik}, s_{ik} = 0)}\right) = \alpha_1 + \beta_1 \sqrt[3]{v_{ik}} + \beta_2 a_{ik}, \tag{3.2}
$$

$$
p(r^1_{ik} = 1 | a_{ik}, v_{ik}, s_{ik} = 0) = \frac{e^{\alpha_1 + \beta_1 \sqrt[3]{v_{ik}} + \beta_2 a_{ik}}}{1 + e^{\alpha_1 + \beta_1 \sqrt[3]{v_{ik}} + \beta_2 a_{ik}}}, \tag{3.3}
$$

$$
\begin{aligned}
p(r^1_{ik} = 0 | a_{ik}, v_{ik}, s_{ik} = 0) &= 1 - p(r^1_{ik} = 1 | a_{ik}, v_{ik}, s_{ik} = 0) \\
&= \frac{1}{1 + e^{\alpha_1 + \beta_1 \sqrt[3]{v_{ik}} + \beta_2 a_{ik}}}.
\end{aligned}
$$

In the two-measurement design for the data from Spratt, we could not tell if a woman was symptomatic. In the complete-measurement design, however, we take her symptomatic status as known and condition on it. We assume that if the patient is symptomatic ($s_{ik} = 1$) and has breast cancer, then the probability of a positive mammogram is 1,

$$
p(r^1_{ik} = 1 | a_{ik}, v_{ik}, s_{ik} = 1) = 1. \tag{3.4}
$$

Based on the above modeling assumptions, we develop each term, **(a)** through **(g)**, of the likelihood (Eq 3.1).

We now address factor **(a)**, which becomes

$$p(r^1_{i1} = 0, \ldots, r^1_{in_i} = 0 \mid \boldsymbol{a_i}, \boldsymbol{v_i}, \boldsymbol{s_i}) = \prod_{k=1}^{n_i} p(r^1_{ik} = 0 | a_{ik}, v_{ik}, s_{ik}).$$

The tumor's volume is less than 1 mm^3 for the first $(m_i - 1)$ mammograms[1]. We assume that the probability of missing a tumor whose volume is less than 1 mm^3 is 1. Therefore,

$$p(r^1_{i0} = 0, \ldots, r^1_{i(m_i-1)} = 0) = 1.$$

Thus,

$$p(r^1_{i0} = 0, \ldots, r^1_{i(m_i-1)} = 0, r^1_{im_i} = 0, \ldots, r^1_{in_i} = 0 | \boldsymbol{v_i}, \boldsymbol{a_i}, \boldsymbol{s_i}) =$$

$$\left(\prod_{k=1}^{m_i-1} 1 \right) \cdot \left(\prod_{k=m_i}^{n_i} \left[1 - p(r^1_{ik} = 1 \mid a_{ik}, v_{ik} = e^{log(v_{i(n_i+1)}) - g_i \cdot (a_{i(n_i+1)} - a_{ik})}, s_{ik}) \right] \right)$$

We next address factor **(b)**.

In view of Equations (3.3) and (3.4):

$$p(r^1_{i(n+1)} = 1 \mid, a_{i(n_i+1)}, v_{i(n_i+1)}, s_{i(n_i+1)} = 0) = \frac{e^{\alpha_1 + \beta_1 \sqrt[3]{v_{ik}} + \beta_2 a_{ik}}}{1 + e^{\alpha_1 + \beta_1 \sqrt[3]{v_{ik}} + \beta_2 a_{ik}}}$$

$$p(r^1_{i(n_i+1)} = 1 \mid a_{i(n_i+1)}, v_{i(n_i+1)}, s_{i(n_i+1)} = 1) = 1$$

We next address factor **(c)**, $p(\boldsymbol{s_i} \mid \boldsymbol{a_i}, \boldsymbol{v_i})$. The probability that a woman becomes symptomatic given her age and tumor size is $p(s_{ik} = 1 \mid a_{ik},\ v_{ik})$. We estimated these probabilities based on the SEER (Surveillance, Epidemiology, and End Results) database of 96,821 cases of breast cancer reported from 1973 to 1982. The interpolations and smoothing of these probabilities are described in Chapter 5.

[1]The age a_{im_i} (and thus m_i) is calculated based on $v_{i(n_i+1)}$; the ages at all mammograms, $a_{i1}, ..., a_{i(n_i+1)}$; and the growth rate, g_i. a_{im_i} is the age of the patient at the mammogram that immediately follows the age-of-onset $a^*_i = a_{i(n_i+1)} - \frac{log(v_{i(n_i+1)})}{g_i}$.

We next address factor **(d)**, $p(\text{all observed } r_{ik}^2 = 1 \mid \boldsymbol{v_i})$. Given a positive first reading for $v_{i(n_i+1)}$, we assume that the probability of a positive second (retrospective) reading for any previous v_{ik} is 0.95,

$$p(r_{ik}^2 = 1 | v_{ik} > 1\, mm^3) = p_2 = 0.95, \forall k \leq n_i.$$

By definition, m_i is the mammogram at which the tumor first exceeds $1\, mm^3$. Thus,

$$p(r_{ik}^2 = 1, k \geq m_i) = p_2 = .95.$$

We next address factor **(e)**, $p(\text{all observed } r_{ik}^2 = 0 \mid \boldsymbol{v_i})$. The probability of a negative retrospective reading of a mammogram is

$$p(r_{ik}^2 = 0 | v_{ik} > 1\, mm^3) = 1 - p(r_{ik}^2 = 1 | v_{ik} > 1\, mm^3) = 1 - p_2 = .05.$$

We next address factor **(f)**, $p(a_i^*)$. Given $a_{i(n_i+1)}$ and $v_{i(n_i+1)}$, age-of-onset is determined by the growth rate. From our assumption of exponential growth and the definition of age-of-onset, a_i^*, we have

$$g_i = \frac{\log(v_{i(n_i+1)}) - 0}{a_{i(n_i+1)} - a_i^*}.$$

Solving for a_i^*,

$$a_i^* = a_{i(n_i+1)} - \frac{\log(v_{i(n_i+1)})}{g_i}.$$

As in the two-measurement model, age-of-onset is assumed to follow a gamma distribution. There is still no learning about age-of-onset. We base this distribution on results published by Parmigiani & Skates (1999).

$$a_i^* \sim Gamma(A = 30, B = .5).$$

Finally, we address factor **(g)**, $p(g_i)$. As in the two-measurement model, tumor growth rates have a log-normal distribution:

$$\log(g_i) \sim Normal(\mu_1, {\sigma_1}^2).$$

Once a tumor is detected, the previous mammograms may be read a second time. There are two possibilities: First, all retrospective second readings are negative, which includes the possibility that the tumor may be present at an earlier time but was not detected, either prospectively or retrospectively. From these single-measurement cases, we may gather information on mammographic sensitivity from the single detection. Having only one measurement, we cannot discern the tumor's growth rate, g_i. The tumor may have been present on previous mammograms but was missed, even retrospectively. Thus, we cannot set a strict lower bound on the growth rate. The second possibility is that upon second look, some earlier mammograms are positive. Based on two or more tumor measurements, we can calculate the tumor's growth rate, g_i, and jointly estimate the mammographic sensitivity, as in the two-measurement design.

3.1.3 Joint Likelihood of $(\boldsymbol{D}, g)$

Suppose we can estimate g_i. Putting together the factors **(a)** through **(g)** from Eq. (3.1), the probability of the observed data, $\boldsymbol{D_i}$, and the associated growth rate,

g_i, of woman i's tumor is

$$
\begin{aligned}
& p(\boldsymbol{D_i}, g_i \mid \alpha_1, \beta_1, \beta_2, \mu_1, \sigma_1, c_e) \\
& \quad = \left(\prod_{k=m_i}^{n_i} \left(1 - p(r_{ik}^1 = 1 \mid a_{ik}, v_{ik} = e^{log(v_{i(n_i+1)}) - g_i \cdot (a_{i(n_i+1)} - a_{ik})}, s_{ik} = 0)\right) \right. \\
& \qquad \left. \cdot \left\{ \begin{array}{ll} 1, & t_{ik} = 1 \\ p_2, & r_{ik}^2 = 1 \\ 1 - p_2, & r_{ik}^2 = 0 \end{array} \right\} \cdot p(s_{ik} = 0 \mid a_{ik}, v_{ik}) \right) \\
& \cdot \left\{ \begin{array}{ll} \Big[p(r_{i(n_i+1)}^1 = 1 \mid a_{i(n_i+1)}, v_{i(n_i+1)}, s_{i(n_i+1)} = 0) & \\ \quad \cdot p(s_{i(n_i+1)} = 0 \mid a_{i(n_i+1)}, v_{i(n_i+1)})\Big] & s_{n_i+1} = 0 \\ 1 \ \cdot p(s_{i(n_i+1)} = 1 \mid a_{i(n_i+1)}, v_{i(n_i+1)}) & s_{n_i+1} = 1 \end{array} \right\} \\
& \cdot p(a_i^*) \\
& \cdot p(g_i)
\end{aligned}
\tag{3.5}
$$

The application of the above likelihood to an example case is found in Section 3.2.

3.1.4 Marginal Likelihood

We drop the patient subscript i in the following discussion. When the growth rate is not known, we integrate the likelihood over all possible growth rates with respect to g. Recall that m is the index of the mammogram immediately following the age-of-onset. Because age-of-onset depends on g, m changes with g. Thus, there is a range of growth rate that translates to age-of-onset falling between the ages of 2 consecutive mammograms. For each interval between successive mammograms, we integrate over this range of growth rates (see Fig. 3.1). We obtain the marginal of the above product (Eq. 3.5) by summing over all possible m's the integral over the corresponding range of g.

The limits of growth rate for a given m are the slopes (see Fig. 3.1) where the first

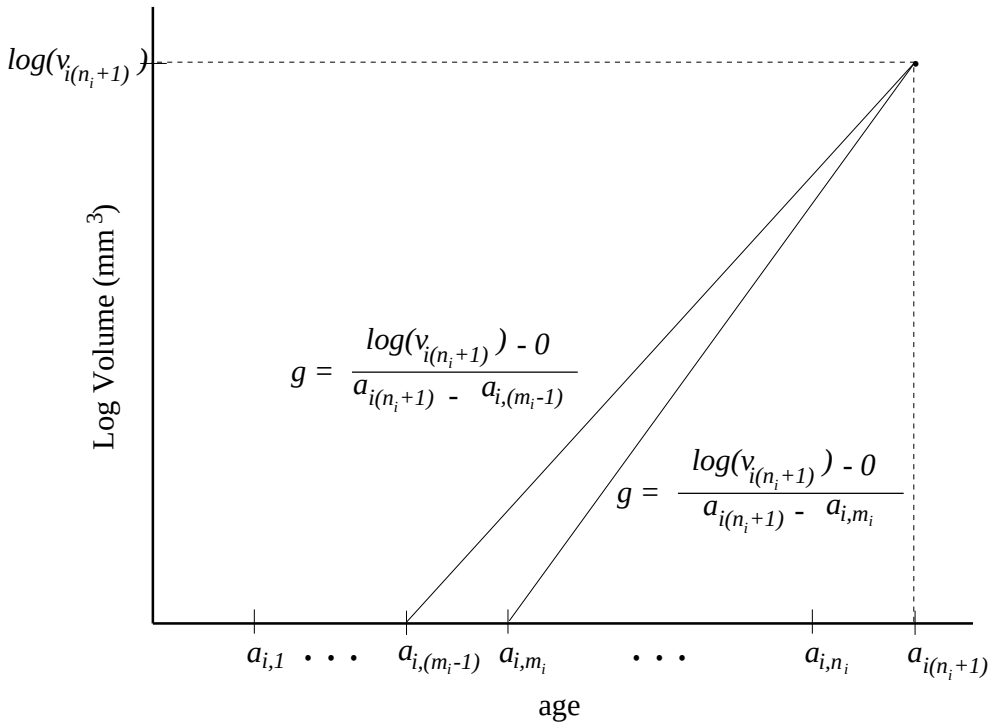

Figure 3.1: Range of Possible Growth Rates if tumor was first detectable between age $a_{i(m_i-1)}$ and a_{im_i}.

point is on the age-axis between a_{m-1} and a_m, with log-volume $\log(1mm^3) = 0$; and the second point is $(a_{i(n_i+1)}, \log(v_{i(n_i+1)}))$. Actually, the growth rate is the minimum of these values and the maximum possible growth rate, g_{max}. We set the maximum possible growth rate to that of a tumor with 1 cm diameter (5 mm radius) doubling to 2 cm diameter (10 mm radius) in 1 month:

$$g_{max} = \frac{\log(\frac{4}{3}\pi(5mm)^3) - \log(\frac{4}{3}\pi(10mm)^3)}{\frac{1}{12}year} = 24.95\,\frac{\log(mm^3)}{yr}$$

Putting this together we have,

$$
\begin{aligned}
&p(\boldsymbol{D_i} \mid \alpha_1, \beta_1, \beta_2, \mu_1, \sigma_1, c_e) \\
&\quad = \sum_{m_i=1}^{n_i} \int_{g=\min\left(\frac{\log(v_{i(n_i+1)})-0}{a_{i(n_i+1)}-a_{im_i}}, g_{max}\right)}^{g=\min\left(\frac{\log(v_{i(n_i+1)})-0}{a_{i(n_i+1)}-a_{i(m_i-1)}}, g_{max}\right)} p(\boldsymbol{D_i}, g_i \mid \alpha_1, \beta_1, \beta_2, \mu_1, \sigma_1) dg \\
&\quad = \sum_{m_i=1}^{n_i} \int_{g=\min\left(\frac{\log(v_{i(n_i+1)})-0}{a_{i(n_i+1)}-a_{im_i}}, g_{max}\right)}^{g=\min\left(\frac{\log(v_{i(n_i+1)})-0}{a_{i(n_i+1)}-a_{i(m_i-1)}}, g_{max}\right)} \left(\prod_{k=1}^{m_i-1} 1\right) \\
&\qquad \cdot \left(\prod_{k=m_i}^{n_i} \left(1 - p(r_{ik}^1 = 1 \mid a_{ik}, v_{ik} = e^{log(v_{i(n_i+1)}) - g\cdot(a_{i(n_i+1)} - a_{ik})}, s_{ik} = 0)\right)\right. \\
&\qquad\qquad \left. \cdot \left\{ \begin{array}{ll} 1 & t_{ik} = 1 \\ p_2 & r_{ik}^2 = 1 \\ 1 - p_2 & r_{ik}^2 = 0 \end{array} \right\} \cdot p(s_{ik} = 0 \mid a_{ik}, v_{ik}) \right) \\
&\qquad \cdot \left\{ \begin{array}{ll} \Big[p(r_{i(n_i+1)}^1 = 1 \mid a_{i(n_i+1)}, v_{i(n_i+1)}, s_{i(n_i+1)} = 0) & \\ \quad \cdot p(s_{i(n_i+1)} = 0 \mid a_{i(n_i+1)}, v_{i(n_i+1)})\Big] & s_{n_i+1} = 0 \\ 1 \ \cdot p(s_{i(n_i+1)} = 1 \mid a_{i(n_i+1)}, v_{i(n_i+1)}) & s_{n_i+1} = 1 \end{array} \right\} \\
&\qquad \cdot p(a_i^*) \\
&\qquad \cdot p(g)\ dg
\end{aligned}
\tag{3.6}
$$

We demonstrate two different hypothetical single-measurement cases and the application of the above marginal likelihood in the following section (see type A patients).

3.2 Examples

We now give examples of different types of women with breast cancer.

Type A - single-measurement cases.

Type B - two or more measurements.

Type C - no measurements.

We describe the different cases of women in each category, and present their contribution to the likelihood.

Type A: Recall that single-measurement cases occur in 2 ways: (1) The radiologist prospectively detects the tumor at the first screen ($r_{i1}^1 = 1$). (2) The first retrospective reading is read as negative ($r_{in_i}^2 = 0$). We present hypothetical data for both cases in Table 3.1.

Table 3.1: Hypothetical data for Type A patients.

(1)

k	1
r_{ik}^1	1
r_{ik}^2	-
t_{ik}	1
v_{ik}	5
s_{ik}	0
a_{ik}	50

(2)

k	1	2	3
r_{ik}^1	0	0	1
r_{ik}^2	-	0	-
t_{ik}	1	2	1
v_{ik}	-	-	10
s_{ik}	0	0	0
a_{ik}	50	51	52

For the example data in Table 3.1(1),

$$\begin{aligned} p(\boldsymbol{D_i} \mid \boldsymbol{a_i}, \boldsymbol{v_i}, \boldsymbol{s_i}, c_e) &\propto p(r_{i1}^1 = 1 \mid a_{i1} = 50, v_{i1} = 5, s_{i1} = 0) \\ &\quad \cdot p(s_{i1} = 0 \mid a_{i1} = 50, v_{i1} = 5) \\ &= \frac{e^{\alpha_1 + \beta_1 \sqrt[3]{5} + \beta_2 \cdot 50}}{1 + e^{\alpha_1 + \beta_1 \sqrt[3]{5} + \beta_2 \cdot 50}} \cdot p(s_{i1} = 0 \mid a_{i1} = 50, v_{i1} = 5) \end{aligned}$$

The likelihood for the type A example data given in Table 3.1(2) can be separated

into 3 disjoint subcases.

$$p(\boldsymbol{D_i} \mid \boldsymbol{a_i}, \boldsymbol{v_i}, \boldsymbol{s_i}, c_e) = \int p(\boldsymbol{D_i}, g \mid \boldsymbol{a_i}, \boldsymbol{v_i}, \boldsymbol{s_i}, c_e)\, dg$$

$$= \int_{I_1} p(\boldsymbol{D_i}, g \mid \boldsymbol{a_i}, \boldsymbol{v_i}, \boldsymbol{s_i}, c_e)\, dg + \int_{I_2} p(\boldsymbol{D_i}, g \mid \boldsymbol{a_i}, \boldsymbol{v_i}, \boldsymbol{s_i}, c_e)\, dg$$

$$+ \int_{I_3} p(\boldsymbol{D_i}, g \mid \boldsymbol{a_i}, \boldsymbol{v_i}, \boldsymbol{s_i}, c_e)\, dg$$

Each disjoint interval is a range of g corresponding to a range of ages for the age-of-onset, a_i^*. Interval I_1 corresponds to the tumor existing before the first measurement, $a_i^* < a_{i1}$. Interval I_2 corresponds to $a_{i1} \leq a_i^* < a_{i2}$, and I_3 corresponds to $a_{i2} \leq a_i^* < a_{i3}$. Specifically,

$$I_1 = \left[\frac{\log(v_{i3}) - 0}{a_{i3} - 0}, \frac{\log(v_{i3}) - 0}{a_{i3} - a_{i1}}\right) = [\log(10)/52, \log(10)/2),$$

$$I_2 = \left[\frac{\log(v_{i3}) - 0}{a_{i3} - a_{i1}}, \frac{\log(v_{i3}) - 0}{a_{i3} - a_{i2}}\right) = [\log(10)/2, \log(10)),$$

$$I_3 = \left[\frac{\log(v_{i3}) - 0}{a_{i3} - a_{i2}}, g_{max}\right) = [\log(10), g_{max}).$$

The likelihood for the subcases will differ due to the different probabilities for the readings in each subcase. Over the first interval, the tumor exists before the first screen and was missed twice before being detected. From Equation (3.6), we have

the probability of the first subcase.

$$\begin{aligned}\int_{I_1} p(\boldsymbol{D_i}, g \mid \boldsymbol{a_i}, \boldsymbol{v_i}, \boldsymbol{s_i}, c_e)\, dg = &\int_{I_1} p(r^1_{i1} = 0 \mid a_{i1} = 50, v_{i1} = e^{\log(v_{i3}) - g\cdot(a_{i3} - a_{i1})}, s_{i1} = 0) \\ &\cdot p(s_{i1} = 0 \mid a_{i1}, v_{i1}) \\ &\cdot p(r^1_{i2} = 0 \mid a_{i2} = 51, v_{i2} = e^{\log(v_{i3}) - g\cdot(a_{i3} - a_{i2})}, s_{i2} = 0) \\ &\cdot p(r^2_{i2} = 0 \mid v_{i2} > 1) \\ &\cdot p(s_{i2} = 0 \mid a_{i2}, v_{i2}) \\ &\cdot p(r^1_{i3} = 1 \mid a_{i3} = 52, v_{i3} = 5, s_{i3} = 0) \\ &\cdot p(s_{i3} = 0 \mid a_{i3}, v_{i3}) \\ &\cdot p\left(a^*_i = a_{i3} - \frac{\log(v_{i3})}{g}\right) \\ &\cdot p(g)\, dg.\end{aligned}$$

Applying the data in Table 3.1(2), we have

$$\begin{aligned}\int_{I_1} p(\boldsymbol{D_i}, g \mid \boldsymbol{a_i}, \boldsymbol{v_i}, \boldsymbol{s_i}, c_e)\, dg = &\int_{I_1} \frac{1}{1 + e^{\alpha_1 + \beta_1 \sqrt[3]{e^{\log(10) - g\cdot 2}} + \beta_2 \cdot 50}} \\ &\cdot p(s_{i1} = 0 \mid a_{i1} = 50, v_{i1} = e^{\log(10) - g\cdot 2}) \\ &\cdot \frac{1}{1 + e^{\alpha_1 + \beta_1 \sqrt[3]{e^{\log(10) - g\cdot 1}} + \beta_2 \cdot 51}} \\ &\cdot (1 - p_2) \\ &\cdot p(s_{i2} = 0 \mid a_{i2} = 51, v_{i2} = e^{\log(10) - g\cdot 1}) \\ &\cdot \frac{e^{\alpha_1 + \beta_1 \sqrt[3]{10} + \beta_2 \cdot 52}}{1 + e^{\alpha_1 + \beta_1 \sqrt[3]{10} + \beta_2 \cdot 52}} \\ &\cdot p(s_{i3} = 0 \mid a_{i3} = 52, v_{i3} = 10) \\ &\cdot p\left(a^*_i = 52 - \frac{\log(10)}{g}\right) \\ &\cdot p(g)\, dg.\end{aligned}$$

Over the second interval, the tumor started growing between a_{i1} and a_{i2}. Thus, it was only missed once before detection. Again from Equation (3.6), the probability of the second subcase is

$$\begin{aligned}\int_{I_2} p(\boldsymbol{D_i}, g \mid \boldsymbol{a_i}, \boldsymbol{v_i}, \boldsymbol{s_i}, c_e)\, dg = &\int_{I_2} p(r_{i2}^1 = 0 \mid a_{i2} = 51, v_{i2} = e^{\log(v_{i3}) - g\cdot(a_{i3} - a_{i2})}, s_{i2} = 0) \\ &\cdot p(r_{i2}^2 = 0 \mid v_{i2} > 1) \\ &\cdot p(s_{i2} = 0 \mid a_{i2}, v_{i2}) \\ &\cdot p(r_{i3}^1 = 1 \mid a_{i3} = 52, v_{i3} = 5, s_{i3} = 0) \\ &\cdot p(s_{i3} = 0 \mid a_{i3}, v_{i3}) \\ &\cdot p\left(a_i^* = a_{i3} - \frac{\log(v_{i3})}{g}\right) \\ &\cdot p(g)\, dg.\end{aligned}$$

Applying the data in Table 3.1(2), we have

$$\begin{aligned}\int_{I_2} p(\boldsymbol{D_i}, g \mid \boldsymbol{a_i}, \boldsymbol{v_i}, \boldsymbol{s_i}, c_e)\, dg = &\int_{I_2} \frac{1}{1 + e^{\alpha_1 + \beta_1 \sqrt[3]{e^{\log(10) - g\cdot 1}} + \beta_2 \cdot 51}} \\ &\cdot (1 - p_2) \\ &\cdot p(s_{i2} = 0 \mid a_{i2} = 51, v_{i2} = e^{\log(10) - g\cdot 1}) \\ &\cdot \frac{e^{\alpha_1 + \beta_1 \sqrt[3]{10} + \beta_2 \cdot 52}}{1 + e^{\alpha_1 + \beta_1 \sqrt[3]{10} + \beta_2 \cdot 52}} \\ &\cdot p(s_{i3} = 0 \mid a_{i3} = 52, v_{i3} = 10) \\ &\cdot p\left(a_i^* = 52 - \frac{\log(10)}{g}\right) \\ &\cdot p(g)\, dg.\end{aligned}$$

And over the last interval, the tumor started growing between a_{i2} and a_{i3} and was

detected at the first possible mammogram. The probability of the third subcase is

$$\begin{aligned}\int_{I_3} p(\boldsymbol{D_i}, g \mid \boldsymbol{a_i}, \boldsymbol{v_i}, \boldsymbol{s_i}, c_e)\, dg = &\int_{I_3} p(r^1_{i3} = 1 \mid a_{i3} = 52, v_{i3} = 5, s_{i3} = 0) \\ &\cdot p(s_{i3} = 0 \mid a_{i3}, v_{i3}) \\ &\cdot p\left(a^*_i = a_{i3} - \frac{\log(v_{i3})}{g}\right) \\ &\cdot p(g)\, dg \\ = &\int_{I_3} \frac{e^{\alpha_1 + \beta_1 \sqrt[3]{10} + \beta_2 \cdot 52}}{1 + e^{\alpha_1 + \beta_1 \sqrt[3]{10} + \beta_2 \cdot 52}} \\ &\cdot p(s_{i3} = 0 \mid a_{i3} = 52, v_{i3} = 10) \\ &\cdot p\left(a^*_i = 52 - \frac{\log(10)}{g}\right) \\ &\cdot p(g)\, dg.\end{aligned}$$

Type B: Patients whose tumors have two or more measurements (at least one $r^2_{ik} = 1$).

Table 3.2: Hypothetical data for Case B patient i.

k	1	2	3
r^1_{ik}	0	0	1
r^2_{ik}	0	1	-
t_{ik}	2	2	1
v_{ik}	-	8.5	10
s_{ik}	0	0	0
a_{ik}	50	51	52

For the case B example data in Table 3.2,

$$\begin{aligned} p(\boldsymbol{D_i}, g_i \mid \theta, c_e) \propto\ & p(r_{i1}^1 = 0 \mid a_{i1} = 50, v_{i1} = e^{\log(10) - g_i \cdot (52-50)}, s_{i1} = 0) \\ & \cdot p(r_{i1}^2 = 0) \\ & \cdot p(s_{i1} = 0 | a_{i1} = 50, v_{i1}) \\ & \cdot p(r_{i2}^1 = 0 \mid a_{i2} = 51, v_{i2} = 8.5, s_{i2} = 0) \\ & \cdot p(r_{i2}^2 = 1) \\ & \cdot p(s_{i2} = 0 | a_{i2} = 51, v_{i2} = 8.5) \\ & \cdot p(r_{i3}^1 = 1 | a_{i3} = 52, v_{i3} = 10) \\ & \cdot p(a_i^*) \\ & \cdot p(g_i). \end{aligned}$$

From the 2 measurements and their associated ages, we calculate the tumors growth rate as

$$g_i = \frac{\log(v_{i3}) - \log(v_{i2})}{a_{i3} - a_{i2}} = \frac{\log(10) - \log(8.5)}{1},$$

and the age-of-onset as

$$a_i^* = a_{i3} - \frac{\log(v_{i3})}{g_i} = 52 - \frac{\log(10)}{\log(10) - \log(8.5)}.$$

Thus we have

$$
\begin{aligned}
p(\boldsymbol{D_i}, g_i \mid \theta, c_e) \propto & \frac{1}{1+e^{\alpha_1+\beta_1\sqrt[3]{e^{\log(10)-2(\log(10)-\log(8.5))}}+\beta_2\cdot 50}} \\
& \cdot (1-p_2) \\
& \cdot p(s_{i1}=0 \mid a_{i1}=50, v_{i1}=e^{\log(10)-2(\log(10)-\log(8.5))}) \\
& \cdot \frac{1}{1+e^{\alpha_1+\beta_1\sqrt[3]{8.5}+\beta_2\cdot 51}} \\
& \cdot p_2 \\
& \cdot p(s_{i2}=0 | a_{i2}=51, v_{i2}=8.5) \\
& \cdot \frac{e^{\alpha_1+\beta_1\sqrt[3]{10}+\beta_2\cdot 52}}{1+e^{\alpha_1+\beta_1\sqrt[3]{10}+\beta_2\cdot 52}} \\
& \cdot p(a_i^* = 52 - \frac{\log(10)}{\log(10)-\log(8.5)}) \\
& \cdot p(g_i = \log(10)-\log(8.5)).
\end{aligned}
$$

Type C: There are 3 cases of women whose tumors are not captured in the complete-measurement design. (1) Patients who are symptomatic prior to age a_{i1} (and are thus not in the screening program); (2) patients whose age-of-onset a_i^* is greater than their age at the end of the screening program; and (3) patients who had a tumor during the program, but whose readings were all negative ($\forall k, r_{ik}^1 = 0$) and never became symptomatic. Type C patients have no recorded tumor measurements during the screening program. Thus, they do not contribute to our likelihood. Our complete-measurement design cannot include these tumors. But such women are not available from a screening program. Our model and proposed design address only the data directly attainable from a screening program.

In the following chapters, we simulate women in a screening program following the complete-measurement design. We simulate data because we do not know of any available data with which to fit our complete-measurement model. We present

our simulation scheme and the breakdown of the women into each of the 3 types. We then check our model by comparing its parameter estimates to those used in the simulation. Also, we present our derivation, interpolation, and smoothing of symptomatic probabilities used in the complete-measurement model.

Chapter 4

Complete-Measurement Simulation & Model Validation

4.1 Simulation of Data

To validate our model and because we know of no complete-measurement design data available, we simulated data based on our assumptions. In creating this complete data set and in the associated analysis, we hope to demonstrate the utility of our methodologies, as outlined in Chapter 3. We simulate the data based on fixed preset values for the five parameters: α_1, β_1, β_2, μ_1, and σ_1. Figure 4.1 depicts the simulation process.

Our simulation is in three main steps: Set-up, Screening Loop, and Retrospective Readings. In Figure 4.1, the Set-up is the top two lines, leading into the first simulation of s_{ik}; the Screening Loop starts where Set-up leaves off and ends at the decision of $r_{ik} \neq 1$; and the Retrospective Readings begin with the decision of $r_{ik} = 1$ and goes until the end of the flowchart.

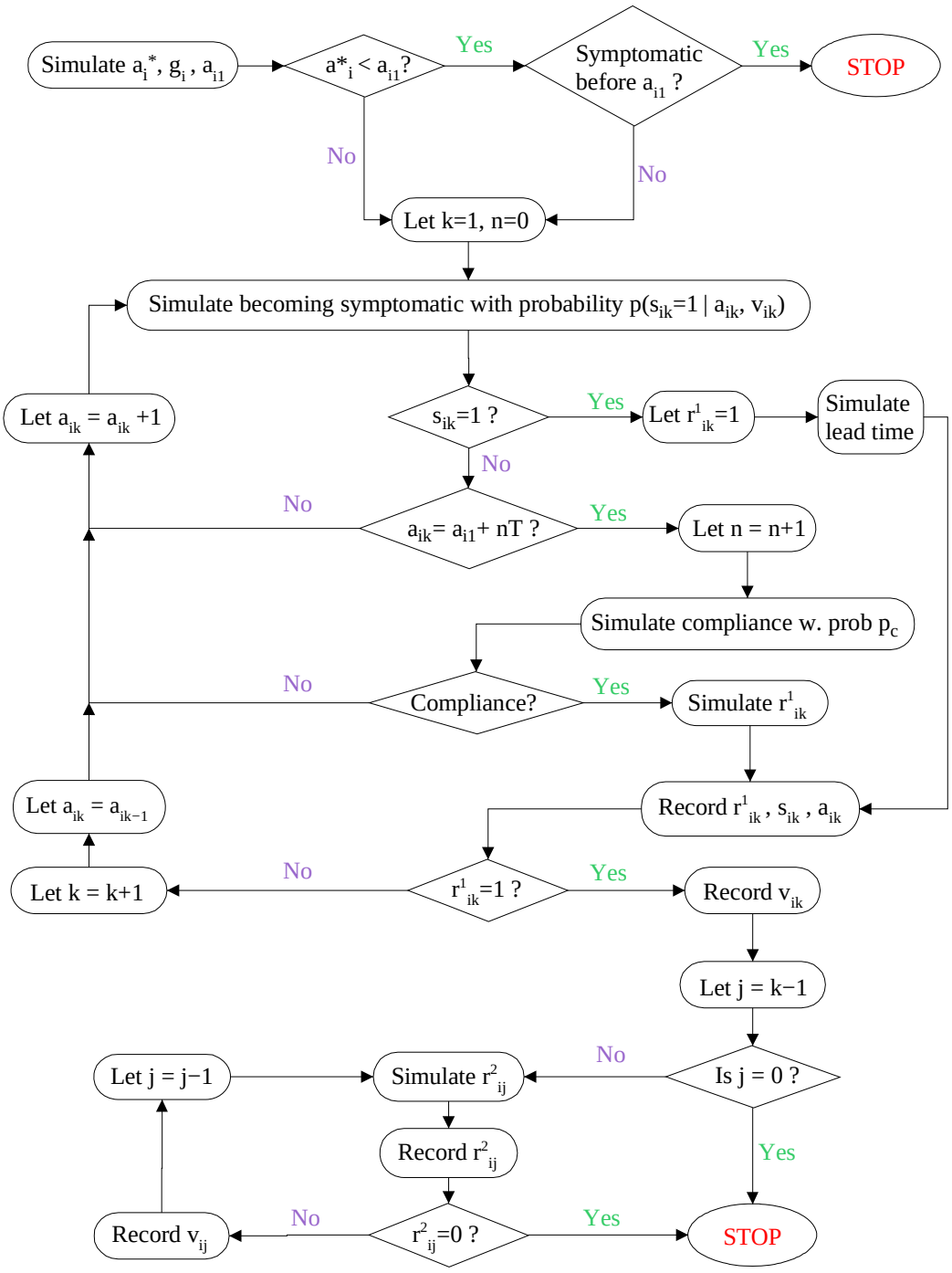

Figure 4.1: Simulation Flowchart for Full Data

4.1.1 Set-up

We only simulate cancer cases and we assume that each woman has at most one tumor. We begin by simulating age of onset ($a_i^* \sim Gamma(30, 0.5)$), growth rate ($g_i \sim lognormal(\mu_1, \sigma_1)$), and age of first mammogram (a_{i1}, which we set equal to 50, but could be drawn from any distribution). Simulated growth rates larger than $g_{max} = 25$ are reduced to g_{max}. If the age of onset is before the age of first mammogram ($a_i^* < a_{i1}$), then we check at every birthday (and associated tumor size) from a_i^* to a_{i1} for symptoms. If the patient becomes symptomatic before her first screen, then she is not recorded in our "reported" data, as she would not make it into the screening program. If she is not symptomatic before her first screen (including cases where $a_i^* > a_{i1}$), then we set $k = 1$ and $n = 0$.

4.1.2 Screening Loop

There are four parts to the Screening Loop: simulation of symptoms, check if in the screening schedule, compliance, and first readings (r_{ik}^1).

The value of k is incremented only when $r_{ik}^1 = 0$ (i.e., when the woman has a negative mammogram).

Once a woman enters the Screening Loop with initialized values for k and n, we check for symptoms at every birthday after a_{i1}, regardless of whether she is screened. In a given year, a woman may not be screened either because it is not part of her screening schedule (every T years) or due to non-compliance.

If a woman is symptomatic, then she has a positive first reading ($r_{ik}^1 = 1$). After we record her age (a_{ik}), positive reading (r_{ik}^1), and symptomatic status ($s_{ik} = 1$), she exits the Screening Loop.

If a woman is not symptomatic, then we check to see if the current year is part of her screening schedule (every T years, incremented by nT, where $n = 0, 1, \ldots$).

For years not in the screening schedule, we increment age and restart the Screening Loop at the next age. On screening years in which the woman is asymptomatic, we simulate her compliance with probability p_c. If the woman is non-compliant, then we increment age and restart the Screening Loop at the next age.

If an asymptomatic woman is compliant with her scheduled examination, then we simulate r_{ik}^1 based on the fixed pre-set sensitivity parameters $(\alpha_1, \beta_1, \beta_2)$, her age and volume. We record the value of r_{ik}^1, symptomatic status (s_{ik}), and age (a_{ik}). If $r_{ik}^1 = 1$, then we also record the volume (v_{ik}) and exit the Screening Loop to enter the Retrospective Readings loop.

If $r_{ik}^1 = 0$, then we do not record v_{ik}, but increment k and a_{ik}, and re-enter the Screening Loop. Suppose an asymptomatic 50 year-old woman complies at $k = 1$ and has a negative first mammogram. Then, $a_{i1} = 50, r_{i1}^1 = 0, s_{i1} = 0$. We increment k to 2 and set $a_{i2} = a_{i1} + 1 = 51$. At $k = 2, a_{i2} = 51$, she is then checked by the simulation algorithm for symptoms as she re-enters the Screening Loop.

4.1.3 Retrospective Readings

Once $r_{ik}^1 = 1$, we record volume v_{ik} and begin retrospective looks backwards. If the tumor was detectable at mammogram $k-1$ $(v_{i(k-1)} > 1mm^3)$, then we simulate a positive second reading with probability p_2. If the tumor volume is less than 1 mm^3, then the second reading is negative. If $r_{ij}^2 = 1$, then we record v_{ij}. We begin at the $(k-1)^{th}$ mammogram and continue looking backwards until $k = 1$ or $r_{ik}^2 = 0$.

4.1.4 Simulated Women

Our simulated population falls into three groups of women, as described in Chapter 3:

- Type A - Women with only one tumor measurement:
  - Subjects found on the first screen $(r_{i1}^1 = 1)$

  - Subjects whose first retrospective look back was negative ($r^2_{in_i} = 0$).

- Type B - Women with two or more measurements (subjects with at least one $r^2_{ik} = 1$).

- Type C - Women who give no information:

  - Subjects who were symptomatic prior to a_{i1}
  - Subjects whose $a^*_i >$ age at the end of the screening program
  - Subjects whose $r^1_{ik} = 0$ for all k. (NB: if a subject becomes symptomatic during the screening program, her $r^1_{ik} = 1$.)

Example data for each type and its associated likelihood contribution have been presented in Section 3.2. We now discuss, and later illustrate by example, how each type contributes differently to the likelihood. Because type C individuals are not recorded, they do not contribute any data to our likelihood. Type A women have only one recorded measurement. Two measurements are needed to determine a tumor's growth rate. Thus growth rates cannot be exactly determined from type A data, though some information is gained in knowing that their previous mammograms were negative. In contrast, type B patients have 2 or more tumor measurements. Their growth rates are known. Hence, type A provide less information about growth rate, and its associated parameters (μ_1, σ_1), than type B.

For type A subjects, we integrate the likelihood over all possible growth rates. A woman can have exactly one measurement in 2 ways: either her tumor had a slow growth rate and it was missed on past mammograms, or her tumor grew quickly and her past mammograms were truly negative. The faster growth rates are more probable because it is unlikely to have continually mis-read the previous mammograms. We have no firm lower bound on these tumor's growth rates, though they do contribute information about faster rates to our likelihood.

On the other hand, both types A and B could have cases detected due to symptoms ($s_{ik} = 1$) and not by mammography. Their data at detection ($k = n_i + 1$) do not contribute information towards the mammographic sensitivity parameters $(\alpha_1, \beta_1, \beta_2)$. However, for each symptomatic tumor, data of mammograms prior to symptoms ($k \leq n_i$) could include retrospective measurements and history of past negative readings. These data would contribute, as they did for screen-detected cases, information towards mammographic sensitivity and growth rates.

We simulated twelve sets of 1000 women. In each set of 1000, the women generally split into 36% type A, 41% type B, and 23% type C, all $\pm 5\%$.

4.2 Validation of the Model

The twelve sets of simulated women, who fall into types A-C, were given identical pre-set values for parameters α_1, β_1, β_2, μ_1, and σ_1. With each set of 1000 women, we ran our full Monte Carlo Markov Chain (MCMC) model, using the Metropolis Hastings step, as described in Chapter 2, for 600,000 iterations after a burn-in of 200,000, keeping every 20th observation. The average and range of the 12 posterior means along side the pre-set values used for simulation are presented in Table 4.1.

Table 4.1: Comparison of Values used in Simulation and MCMC posterior means

	pre-set simulation values	MCMC posterior means	range
α_1	-0.4048	-0.516	(-0.65 , -0.40)
β_1	0.1576	0.139	(0.136 , 0.152)
β_2	0.0217	0.028	(0.018 , 0.032)
μ_1	-0.3374	-0.633	(-0.788 , -0.424)
σ_1	2.8463	2.395	(2.36 , 2.46)

4.3 Bias Check

The main difference between the posterior means and the pre-set simulation values is in μ_1. There are several possible reasons why the MCMC posterior mean for μ_1 (-0.633) may differ from the pre-set simulation value of -0.3374. The difference between the posterior (-0.633) and pre-set (-0.337) values on μ_1 could be caused by the prior, which was centered on -0.5. We test this hypothesis two ways. First, we reran the MCMC chains on the data with the original pre-set value for μ_1, but with a new prior centered on this preset value ($\mu_2 = -0.3374$). This resulted in an almost identical posterior mean, $\bar{\mu_1} = -0.621$. Second, we simulated a new set of data values using $\mu_1 = -0.5$ (same as μ_2). The resulting MCMC yielded $\bar{\mu_1} = -1.081$. This value is as different from our new pre-set μ_1 as before. Since neither of these attempts aided in a better estimation of μ_1, we conclude that the prior is not influencing μ_1 towards the wrong value. We must consider an alternative cause for the observed discrepancy.

Another possibility is that our likelihood does not correctly allow for the truncation of growth rates from above. Recall that we restrict growth rates to be below g_{max} for the single measurement likelihood and for the simulation of all data. Such exclusion of the fastest growth rates could result in lower estimates of μ_1 and σ_1.

We do, however, include type A subjects. These single-measurement cases tend to have faster growth rates (as compared to type B, see Figure 4.2). Type A patients give no direct information about their growth rates. Recall that slower growth rates tend to be type B tumors with 2 or more measurements, giving exact growth rate information to the likelihood. Thus, faster growth rates are not contributing as much information to the likelihood as slower growth rates. This bias would also result in underestimation of the mean growth rate parameter, μ_1, and its variance, σ_1.

Since we simulated the data, we know the actual growth rates for all of the women, including the ones where only one measurement could have resulted. To test our

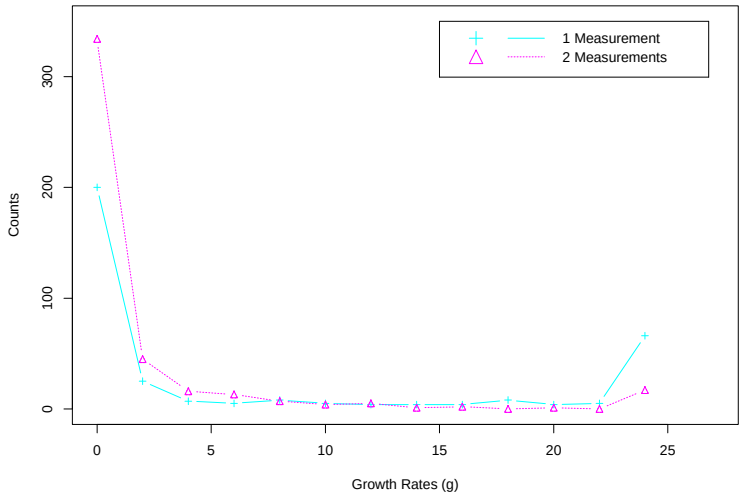

Figure 4.2: Histogram comparison of growth rates from simulated data

hypothesis that we were not gathering as much information about g from those with only one measurement, we reran our MCMC model using the actual growth rate[1], g, that was assigned to each woman. The values from the resulting MCMC runs were very close to the actual pre-set values. In fact, $\bar{\mu_1} = -0.36$ and $\bar{\sigma_1} = 2.81$, as compared to the values used in simulating the data of $\mu_1 = -0.3374$ and $\sigma_1 = 2.8463$. These values are much more accurate than the ones obtained by our model when g is not known. Thus, we have verified that the estimated μ_1 and σ_1 biases are due to not fully incorporating the faster growth rates.

Our model, however, does estimate μ_1 and σ_1 better when given both type A and B data than when given only type B. The latter case is similar to the two-measurement design proposed in Chapter 2, which had exactly two measurements, but no history of negative screens prior to the first measurement. We reran our model, using only data from those with two or more measurements, including their entire history. These values were more heavily skewed towards lower estimates of g, and hence they are more

[1]Recall that these growth rates are the minimum of the simulated values and g_{max}

Table 4.2: Comparison of simulation values to different MCMC posterior means

	pre-set values (simulation)	full model (2+ volumes)	full model (1 & 2+volumes)	full model + g (using assigned g)
α_1	-0.4048	-0.594	-0.516	-0.455
β_1	0.1576	0.130	0.139	0.142
β_2	0.0217	0.031	0.028	0.025
μ_1	-0.3374	-0.758	-0.633	-0.356
σ_1	2.8463	2.369	2.395	2.659

biased estimates of μ_1 and σ_1. Thus, considering only those with 2 measurements increased the bias, further underestimating growth rate. Additionally, the estimates of the logistic parameters (α_1, β_1, and β_2) were further from the pre-set values. Hence, incorporation of type A patients, as for our complete-measurement design, results in improved estimates of all five parameters. See Table 4.2 for a comparison of the different values. The best our model could do would be with omniscience regarding tumor growth. In terms of knowing g, including the single-measurement cases reduces the bias on the logistic parameters by a factor of at least 2. The β_1 bias (coefficient for $\sqrt[3]{vol}$) is reduced by a factor of $4 = (.130 - .142)/(.139 - .142)$.

While the complete-measurement design still leads to biased estimates, it performs better when taking into account all available measurements in contrast to using only those with two or more. Future work will concentrate on improved estimations of μ_1 and σ_1 via more appropriate growth rate bias corrections.

Chapter 5

Interpolation of Probabilities from SEER Data

5.1 Introduction to SEER Data, Our Goals and Methods

In our full model likelihood (Equation 3.5), we need the probability that a woman becomes symptomatic in one year, given her age and tumor size, $p(s_{ik}|a_{ik}, v_{ik})$. In this chapter, we present our interpolation of these probabilities from categorical counts of symptomatic women. We first redistribute the counts into finer, evenly-sized categories. Then, we assign each woman a growth rate and impute back her age and tumor size to create (age, size) pairs for asymptomatic women. Finally, we calculate and smooth the probabilities of a woman being symptomatic given her age and tumor size.

The data for symptomatic patients' age and tumor size (diameter) come from the SEER (Surveillance, Epidemiology, and End Results) database on 96,821 reported cases of breast cancer from 1973 to 1982 (Table A.1) (National Cancer Institute, 1997). Since screening programs were not commonplace prior to 1980's (Moskowitz, 1984), most breast cancers in this time period were detected because of symptoms.

The SEER data provides counts of symptomatic patients for each age (16 - 102 years of age) and tumor size category. Their diameter measurements are provided as counts in only eight variably sized categories (Table A.1 & Figure 5.1). The resolution is not adequate for our modeling. To estimate $p(s_{ik}|a_{ik}, v_{ik})$, we redistribute the counts into finer, uniformly sized bins and then convert the resulting counts into probabilities.

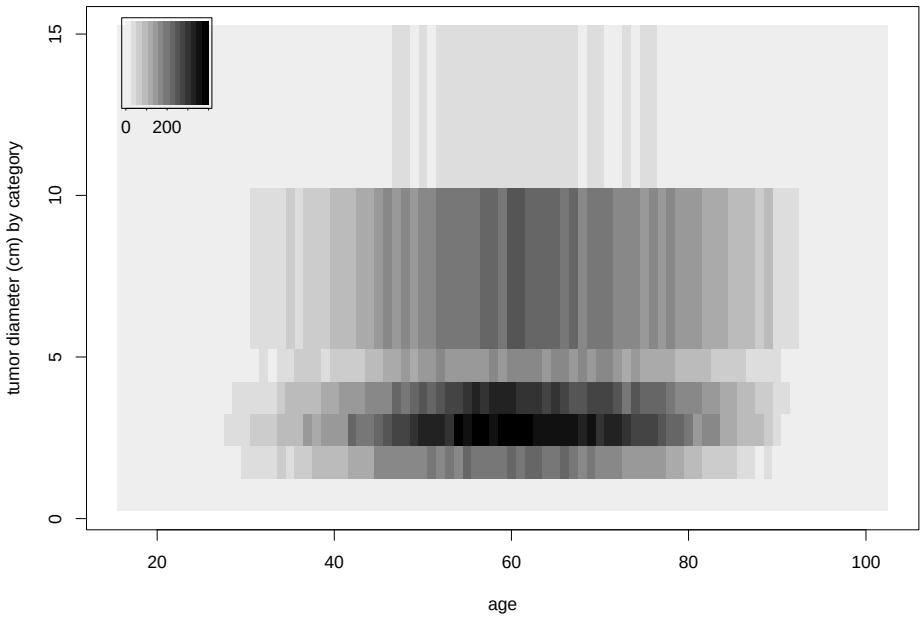

Figure 5.1: Counts for SEER Symptomatic Patients, 1973-1982

In the original data table (Table A.1 & Figure 5.1), the eight categories are[1]:

1	< 0.5 cm
2	0.5 cm - 1.0 cm
3	1.0 cm - 2.0 cm
4	2.0 cm - 3.0 cm
5	3.0 cm - 4.0 cm

[1]measurements falling directly on the interval breaks are placed in the larger category

```
6       4.0 cm - 5.0 cm
7       5.0 cm - 10.0 cm
8             > 10.0 cm
```

We create a table of probabilities over age (by year, as in SEER) and size by 31 half-centimeter increments from 0 to 15.5.

The goals of this chapter are to explain the methods used to construct this table of probabilities and to test the sensitivity of our calculations to the growth rate assumptions. Here we provide an overview of the methods used in this chapter. In the following sections, we discuss each item in detail.

5.2 Interpolate the SEER symptomatic counts into finer size categories.

5.2.1 Total across ages.

5.2.2 Redistribute original size totals across a finer grid.

5.2.3 Redistribute new finer size totals back across ages.

5.3 Create associated table of asymptomatic patients.

5.3.1 Assign each symptomatic patient a growth rate.

5.3.2 Back-calculate age and size for each patient, each year.

5.4 Create table of probabilities from the two tables of counts.

5.5 Test sensitivity of probability calculations to growth rates.

5.6 Smooth the table of probabilities.

5.2 Non-parametric interpolation of the SEER symptomatic counts into finer size categories

The first step in creating our desired table of probabilities over finer size categories is to interpolate the original SEER counts of symptomatic patients into a larger table of counts over finer categories of tumor size.

5.2.1 Total across ages

First, we compute the totals of the original table of SEER counts over ages (see column totals in Table A.1):

Original Categories (cm)	0–.5	.5–1	1–2	2–3	3–4	4–5	5–10	10+
Symptomatic Counts	282	749	8185	14607	11884	6659	8978	1389

5.2.2 Expand size totals

Next, we redistribute the above totals into 31 size categories (see Section 5.1). We smooth using density estimation. Figure 5.2 shows the process:

In **(a)**, the SEER count totals from Section 5.2.1 are plotted as boxed crosses. These are first divided by their bin size (thus, bin 7's total of 8978 is divided by 5 cm to arrive at 1795.6) and then plotted across the bins. These are the histogram bars, whose areas are equal to the total counts.

In **(b)**, the histogram from **(a)** is again plotted, and the midpoints (represented by solid pink circles) of each histogram bar are joined linearly (solid pink lines). The points along these lines that correspond to the midpoints of the desired finer bins (the hollow navy triangles) are then recorded.

In **(c)**, these recorded midpoint values are used to distribute each original SEER total out into the finer bins. The original counts are again shown as boxed crosses.

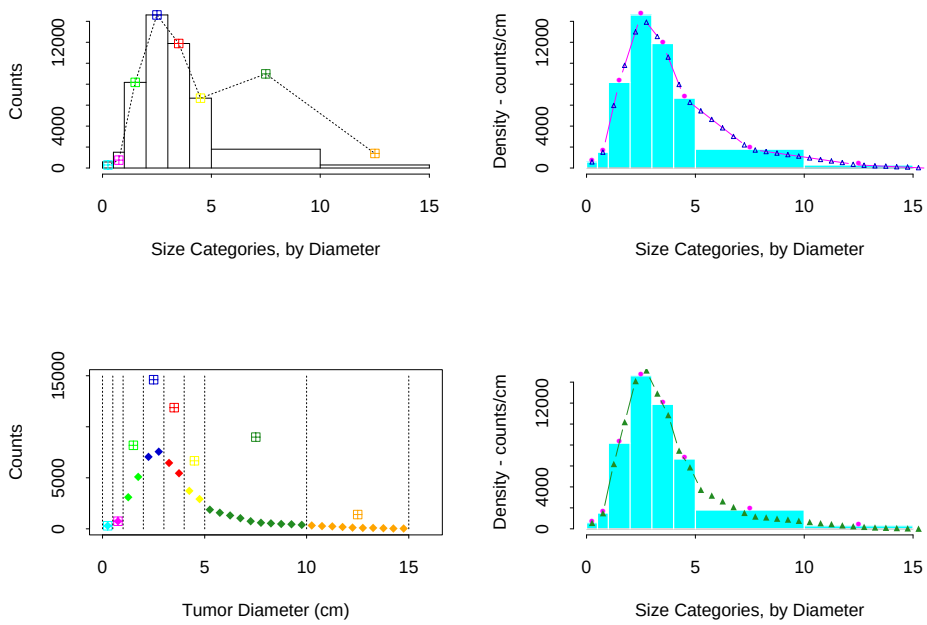

Figure 5.2: Smoothing of Uneven Bins. Along the left are graphs of the counts, while on the right are graphs of the density. In the upper left, **(a)**, are the original counts and then their associated histogram. In the upper right, **(b)**, are the midpoints of the histograms connected by lines, along which the midpoints for the finer bins are drawn. In the lower left, **(c)**, are the original counts within each bin (shown by vertical lines) then redistributed into the finer bins. In the lower right, **(d)**, are the two densities (the original histogram and the newly created smoothed version).

The original bins are depicted as black vertical lines. Within each bin, the original total is redistributed in the proportion of the midpoint heights found in **(b)**.

For example, size category 6 (4.0 - 5.0 cm) has an original total of 6659 patients. We divide this into two subcategories (4.0 - 4.5 cm and 4.5 - 5.0 cm). The midpoints (along the x-axis) of these new bins are 4.25 cm and 4.75 cm. In **(b)**, the two hollow navy triangles, that correspond to these diameters, fall on the pink lines connecting the midpoints (solid pink circles) of the original histogram. Midpoints 4.25 cm and 4.75 cm have density values of 7965.25 and 6253.72 count/cm in **(b)**.

To redistribute the SEER total of 6659 patients, we do so in proportion to these density values. Let $count_{4.25}$ be the newly assigned value to bin [4.0 - 4.5 cm), and

$density_{4.25}$ be its associated density value:

$$\begin{aligned} count_{4.25} &= \frac{density_{4.25}}{density_{4.25} + density_{4.75}} \cdot (\text{total count for original bin}) \\ &= \frac{7965.25}{7965.25 + 6253.72} \cdot 6659 \\ &= 3730. \end{aligned}$$

Likewise,

$$\begin{aligned} count_{4.75} &= \frac{density_{4.75}}{density_{4.25} + density_{4.75}} \cdot (\text{total count for original bin}) \\ &= \frac{6253.72}{7965.25 + 6253.72} \cdot 6659 \\ &= 2929. \end{aligned}$$

Thus, the new counts for bins 4 - 4.5 cm & 4.5 - 5 cm are 3730 & 2929, respectively. A generalization of this method is:

$$new.count_{new.bin} = \frac{density_{new.bin}}{\sum density_{all.bins.in.original.bin}} \cdot (\text{total count for original bin})$$

In this way, we have 31 expanded size totals. These are represented by the solid diamonds in **(c)**, where the colors match the counts that were originally in the same bin.

In **(d)**, the newly expanded totals are divided by their bin size of 0.5 cm (see the solid green triangles) to make a density curve (in terms of counts/cm) comparable to the original histogram.

5.2.3 Redistribute size totals across ages

These 31 size totals (see solid diamonds in Figure 5.2 **(c)**) are now extrapolated back into the age categories. We do this by redistributing each of the new totals back through the ages in the same proportion as the original bins.

We illustrate with a simplified example. Suppose the original data were the following:

age	bin 1
50	10
51	20
52	10
53	10
col total	50

Suppose when we split this bin into 2 new finer bins, the split totals were 30 and 20 each. We then redistribute the 30 and 20 back through the ages by the same proportion in each age group as was seen in the original bin:

age	new bin 1	new bin 2
50	6	4
51	12	8
52	6	4
53	6	4
col total	30	20

We carry out the above-demonstrated redistribution for all 31 size totals across all 87 age years (16 - 102), and arrive at our newly extrapolated table of symptomatic counts (see Figure 5.3).

5.3 Create associated table of asymptomatic patients

Using this newly interpolated table of symptomatic counts (see Figure 5.3), we multiply impute its corresponding table of asymptomatic counts (Figure 5.4).

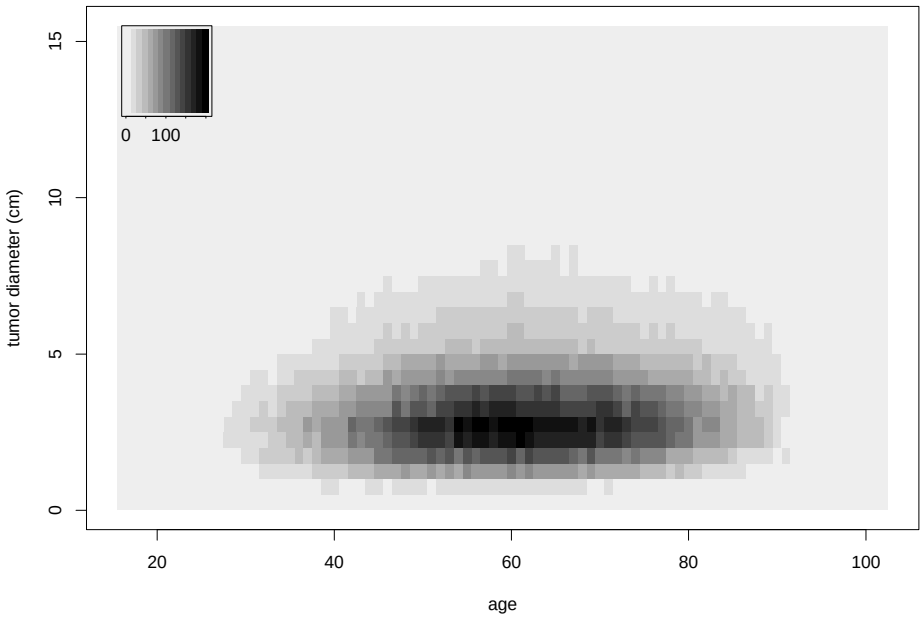

Figure 5.3: Extrapolated Counts for SEER Symptomatic Patients, 1973-1982

5.3.1 Assign each symptomatic patient a growth rate

Each of the 96,821 symptomatic patients is randomly assigned a growth rate, g_i, based on the predictive distribution of g_i from the two-measurement data analysis (see Figure 2.4).

5.3.2 Back-calculate age and size for each patient, each year

Patients are assumed to be asymptomatic prior to detection. To estimate the age and tumor size distribution of asymptomatic patients we back-calculate tumor size for each preceding year for each patient. The patient's age at onset is assumed to be when her tumor volume was 1 mm^3. The patient age and tumor size back-calculation continues until her age-of-onset or age 16 is reached, whichever comes later in life. Thus, we augment the (age, tumor size) distribution of symptomatic women from SEER (Figure 5.3) using information from the two-measurement data analysis

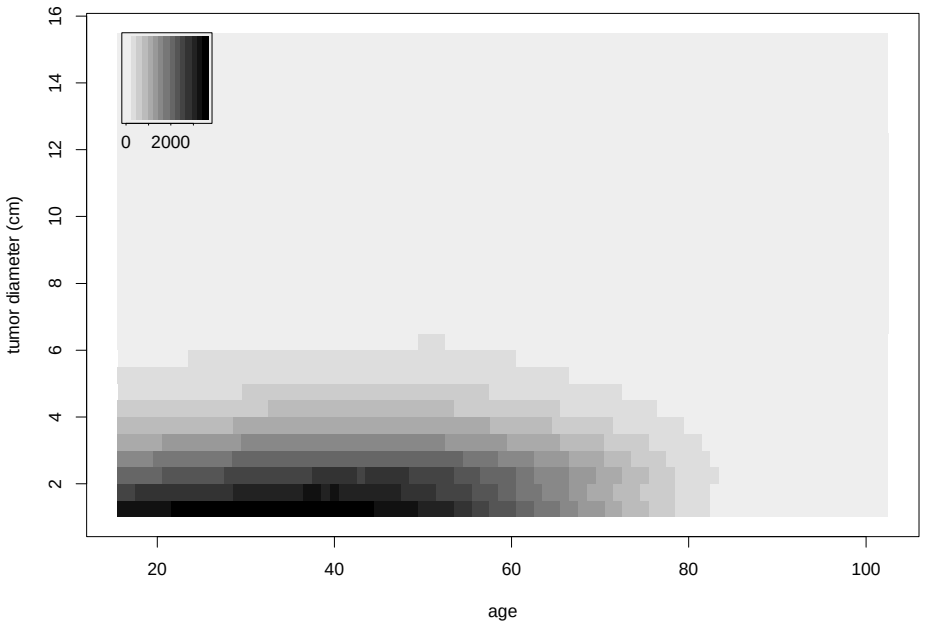

Figure 5.4: Back-Calculated Counts for Asymptomatic Patients, based on symptomatic SEER Patients, 1973-1982, assuming they were asymptomatic before their cancer was identified.

to calculate the corresponding (age, tumor size) joint distribution of asymptomatic patients (see Figure 5.4).

5.4 Create table of probabilities from table of counts

To find the probability that a patient becomes symptomatic given her tumor size and age, we need the extrapolated table of symptomatic counts and its associated table of asymptomatic counts, as we derived in Section 5.3.2. Once we have these two tables, we can easily find the table of probabilities:

$$P(s_{ik} \mid a_{ik}, v_{ik}) = \frac{\#\text{ symptomatic at } (a_{ik}, v_{ik})}{\#\text{ asymptomatic at } (a_{ik}, v_{ik}) \ + \ \#\text{ symptomatic at } (a_{ik}, v_{ik})}$$

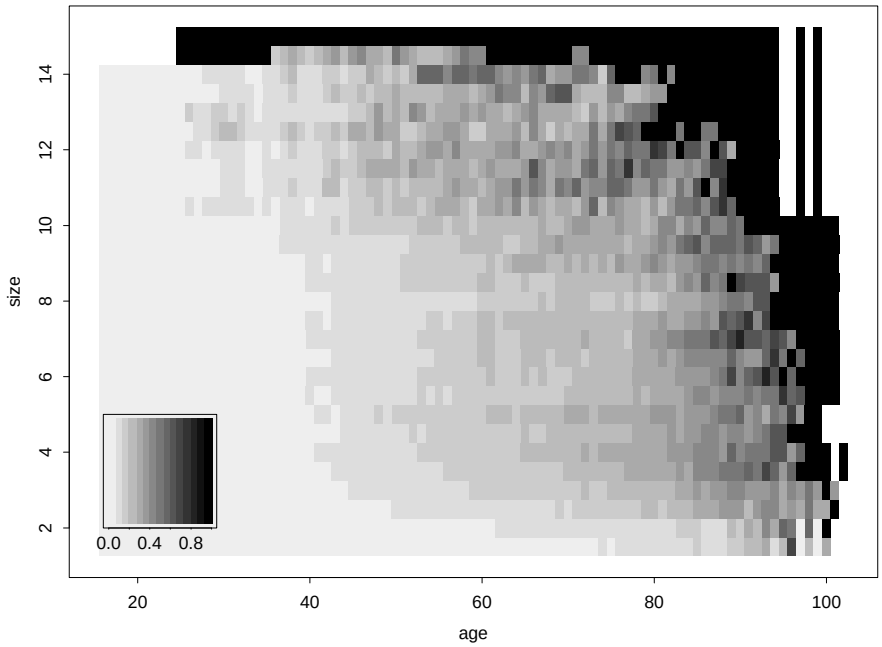

Figure 5.5: Probability of a Woman Becoming Symptomatic at a Given Age and Tumor Size Category, using g - Posterior Predictive Distribution of Growth Rates from Spratt Data

The probability that a patient becomes symptomatic at a given age and size can thus be seen in Figure 5.5.

In the next section, we show that our probabilities are not overly sensitive to the growth rate estimates we derived from Spratt *et al.* (1993).

5.5 Test sensitivity of probability calculations to growth rates

We tested the sensitivity of these symptomatic probabilities to growth rates by re-calculating the table of probabilities twice. First, we assigned each patient a growth rate that was half the growth rate she was assigned in Section 5.3.1, and reconstructed the table (see the original calculation in Figure 5.5 and this new one in Figure 5.6). These new growth rates increase the bias already present in the Spratt

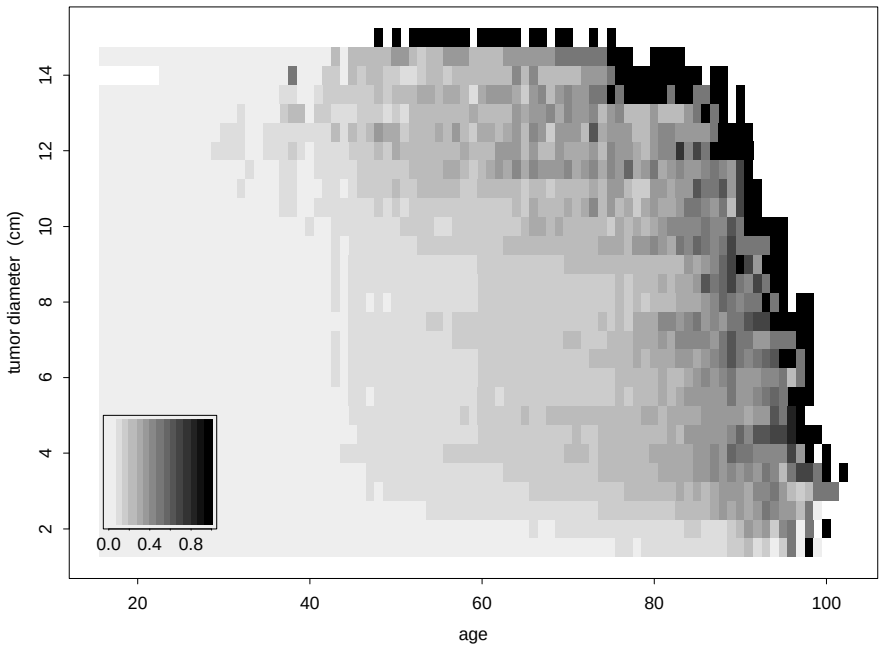

Figure 5.6: Probability of a Woman Becoming Symptomatic at a Given Age and Tumor Size Category, using $\frac{g}{2}$.

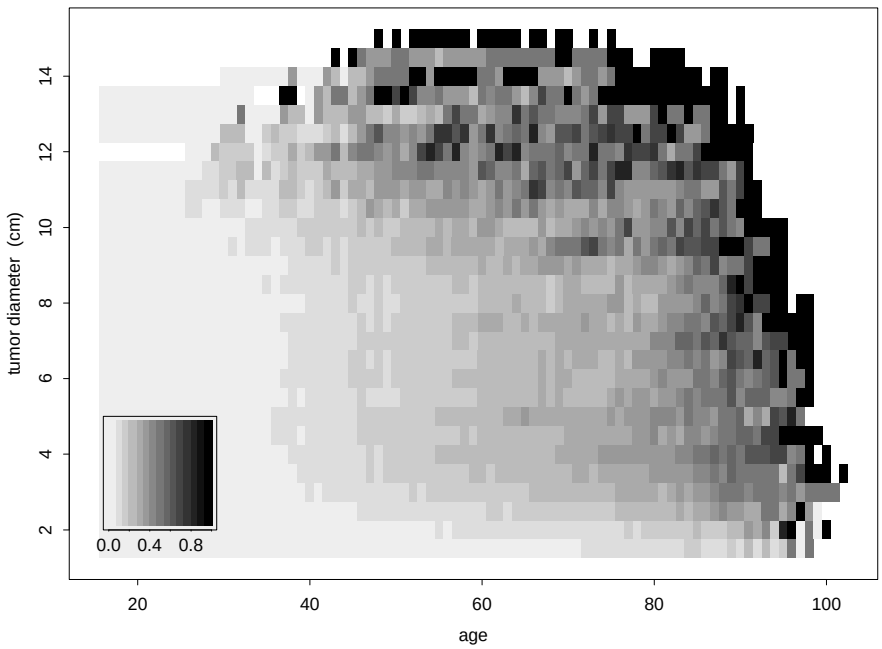

Figure 5.7: Probability of a Woman Becoming Symptomatic at a Given Age and Tumor Size Category, using $2g$.

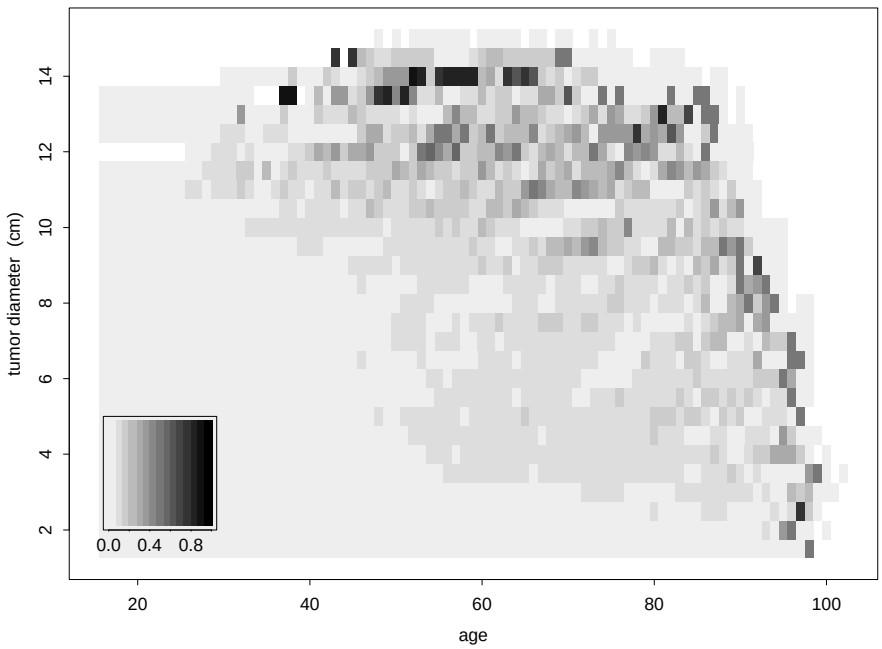

Figure 5.8: Difference Probability of a Woman Becoming Symptomatic at a Given Age and Tumor Size Category, between using $2g$ & $\frac{g}{2}$.

two-measurement data. Second, we repeated the calculation, now with each patient being assigned a growth rate twice the original value used in Section 5.3.1 (see these new probabilities in Figure 5.7). This increase may compensate for the estimated bias we observed in Section 4.3. Then, we considered the absolute differences of these two new tables (Figures 5.6 & 5.7), as seen in Figure 5.8. The absolute differences tend to be small, with a mean of 0.09. Thus our calculations are not overly sensitive to our choice of growth rates, g. Future work will attempt to quantify the growth rate bias from Spratt and accordingly adjust the assigned rates. Such a correction would yield a slightly more accurate probability table.

5.6 Smooth the table of probabilities

We used the generalized additive model as implemented in S-plus to smooth the above probabilities across both age and size bins. In particular, we used a Loess fit, with a span of 0.5 and weights equal to the number of observations (counts of

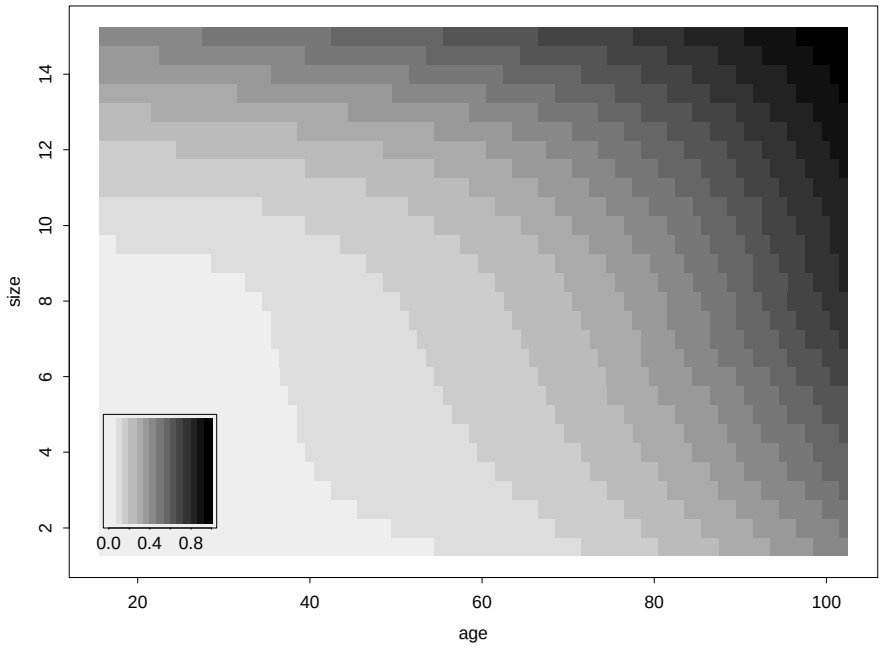

Figure 5.9: Smoothed Probability of a Woman Becoming Symptomatic at a Given Age and Tumor Size Category, using g.

both symptomatic and asymptomatic patients) in each bin. We use these smoothed probabilities (see Figure 5.9) in our likelihood for the complete-measurement design.

Chapter 6

Conclusion & Future Developments

Overall, we presented two designs and their associated likelihoods for the joint estimation of tumor growth and mammographic sensitivity. For both designs, our model of mammographic sensitivity depends on both patient age and tumor size. We are not aware of any other estimate of mammographic sensitivity by tumor size (van Dijck *et al.*, 1993). For example, the three curves seen in Figure 2.2 are usually 3 horizontal lines, varying for age, but not for tumor size (for example, Peer *et al.*, 1996; Schmitt and Threatt, 1982).

The two-measurement design utilizes only data from patients who have exactly two tumor measurements. We base our analysis on data from Spratt *et al.* (1993). Our complete-measurement design incorporates data from single-measurement cases, which are usually left out of tumor growth estimations. Another innovative feature of our complete-measurement design is the use of patient ages at past *negative* mammograms to better estimate tumor growth and mammographic sensitivity. Our novel joint estimation is appropriate because a negative mammogram one year before a positive mammogram can be due to either a rapid growth rate or poor mammographic sensitivity.

In Chapter 2 we introduced the two-measurement design that included a small

selected subset of a patient population, provided by Spratt *et al.* We presented the complete analysis of posteriors based on the MCMC sample. The results included the predictive distribution of growth rates and estimates of mammographic sensitivity based on age and tumor size.

In Chapter 3, we outlined the complete-measurement design, detailing the information that would be recorded. This new design includes many women who are excluded from the two-measurement design. In particular, the complete-measurement design includes single-measurement cases. We built a model tailored to this data. Application of this model requires a table of symptomatic probabilities over age and tumor size. In Chapter 5, we presented a non-parametric method of smoothing coarse non-uniform symptomatic counts into finer, uniformly sized bins. We then presented our conversion of counts into probabilities, for use in our complete-measurement model.

Compared to the two-measurement design, our complete-measurement model incorporates more information from a screening program. We provided both a joint and marginal likelihood to fit both single- and multiple-measurement cases. Since we knew of no data available to fit our model, we simulated our own patient population in Chapter 4. The flowchart (see Figure 4.1) demonstrated our assumptions and beliefs about the course of a patient's tumor and details the information to be recorded.

We generated twelve sets of 1000 patients each based on one set of predetermined parameter values, and tested our model's ability to reproduce these values. Using the complete-measurement likelihood, we compared runs based on patients with two or more measurements (and thus direct values of growth rate), on all recorded patients (including those with only one measurement), and on all recorded patients with omniscience (additionally using the growth rate assigned to those with only one

measurement). Of course the omniscient model performed best, but in real data there is no way to know an interval case's true growth rate.

Our complete-measurement model fits the data quite well, and better than if patients with only two or more measurements are used, as in the two-measurement model. However, an obvious bias in growth rates still exists within our complete-measurement model. Future work will concentrate on addressing this issue. We also plan to estimate sojourn times from our jointly obtained tumor growth rates and mammographic sensitivity. From these sojourn times, we will show the results of different screening protocols.

Appendix A

Original SEER Data

In the old table (Table A.1 & Figure 5.1) , the cutoffs were:

1	< 0.5 cm
2	0.5 cm - 0.9 cm
3	1.0 cm - 1.9 cm
4	2.0 cm - 2.9 cm
5	3.0 cm - 3.9 cm
6	4.0 cm - 4.9 cm
7	5.0 cm - 9.9 cm
8	10.0 cm -

Table A.1: Symptomatic breast cancer counts of women by age and tumor size (diameter) category - SEER 1973-1982

	Original Size Categories								row
Age	1	2	3	4	5	6	7	8	total
16	0	0	0	0	1	0	0	0	1
17	0	0	0	0	0	2	1	0	3
18	0	0	1	0	0	1	0	0	2
19	0	0	1	1	2	2	0	0	6
20	0	0	0	0	0	0	2	0	2
21	0	0	1	0	1	1	0	0	3
22	0	0	0	3	3	0	2	0	8
23	0	0	6	7	1	3	12	0	29
24	0	0	6	7	7	1	3	0	24
25	0	0	6	8	10	5	7	1	37
26	0	2	11	24	5	2	9	3	56
27	1	1	14	21	20	12	18	3	90
28	0	1	21	30	14	15	16	4	101
29	1	5	20	39	27	11	20	5	128
30	1	6	29	44	37	9	23	7	156
31	6	4	34	56	40	25	34	7	206
32	3	3	50	56	52	28	27	9	228
33	3	5	48	71	47	21	50	5	250
34	3	8	56	84	61	39	38	3	292
35	4	9	50	99	86	43	55	7	353
36	4	2	68	89	82	54	50	7	356
37	4	12	65	139	84	58	63	12	437
38	3	11	82	108	105	59	75	14	457
39	5	14	98	148	112	49	76	9	511

Table A.1: *continued*

	Original Size Categories								row
Age	1	2	3	4	5	6	7	8	total
40	4	14	104	158	117	59	101	8	565
41	6	13	105	151	133	74	96	11	589
42	3	12	119	212	144	74	93	17	674
43	8	12	121	190	154	67	130	19	701
44	9	17	129	211	170	87	113	12	748
45	6	17	167	219	176	93	149	18	845
46	9	13	171	242	176	122	168	25	926
47	3	14	179	282	237	121	138	28	1002
48	7	22	181	264	192	141	161	33	1001
49	8	17	180	295	229	126	140	23	1018
50	5	17	180	322	251	134	169	34	1112
51	3	11	201	327	226	146	178	23	1115
52	7	20	180	331	253	159	190	33	1173
53	7	15	201	298	271	142	204	34	1172
54	5	17	163	374	278	153	203	35	1228
55	6	15	219	349	294	153	207	30	1273
56	10	20	196	379	322	153	192	35	1307
57	10	20	206	374	295	157	222	38	1322
58	8	20	209	357	318	164	212	33	1321
59	3	21	202	387	325	148	206	28	1320
60	10	24	220	379	324	176	252	30	1415
61	7	21	197	396	293	177	252	35	1378
62	8	18	236	380	296	176	217	36	1367
63	7	18	212	345	299	178	219	31	1309
64	7	14	201	352	290	157	218	42	1281

Table A.1: *continued*

	Original Size Categories								row
Age	1	2	3	4	5	6	7	8	total
65	4	18	206	361	301	179	229	28	1326
66	8	20	216	357	266	159	206	46	1278
67	8	17	187	351	263	147	229	31	1233
68	7	20	171	341	260	166	180	25	1170
69	5	12	198	351	274	157	207	32	1236
70	7	12	168	308	279	160	196	31	1161
71	5	14	170	329	278	167	195	25	1183
72	4	11	165	330	262	139	177	26	1114
73	2	10	149	293	210	129	165	35	993
74	3	10	154	277	253	134	161	19	1011
75	1	12	133	276	236	123	152	32	965
76	2	12	139	277	230	118	164	27	969
77	3	6	132	260	214	112	134	22	883
78	6	8	106	234	192	111	165	15	837
79	1	8	109	223	180	103	142	17	783
80	3	11	93	201	175	103	141	26	753
81	2	7	86	151	165	83	146	20	660
82	3	3	72	162	141	96	110	18	605
83	1	8	69	167	139	64	119	23	590
84	1	7	66	124	112	78	121	16	525
85	0	3	60	130	124	68	102	16	503
86	1	6	38	101	99	55	100	12	412
87	0	2	28	96	99	46	81	13	365
88	0	2	22	92	58	39	75	15	303
89	1	2	28	65	66	34	82	6	284

Table A.1: *continued*

	Original Size Categories								row
Age	1	2	3	4	5	6	7	8	total
90	0	3	17	51	37	33	42	10	193
91	0	0	22	25	38	17	33	7	142
92	0	0	9	23	24	16	27	3	102
93	0	0	9	16	20	10	26	1	82
94	0	0	2	10	10	14	19	2	57
95	0	0	6	6	4	7	14	0	37
96	0	0	6	4	5	10	7	0	32
97	0	0	0	2	3	3	8	1	17
98	0	0	2	2	3	1	8	0	16
99	0	0	0	1	1	1	2	2	7
100	0	0	1	1	2	0	1	0	5
101	0	0	0	1	0	0	1	0	2
102	0	0	0	0	1	0	0	0	1
col total	282	749	8185	14607	11884	6659	8978	1389	52733

Bibliography

Bjurstam, N., Björneld, L., Duffy, S.W., Smith, T.C., Cahlin, E., Erikson, O., Hafström, L.-O., Lingaas, H., Mattsson, H., Persson, S., Rudenstam, C.-M. and Säve-Söderbergh, J. (1997) The gothemburg breast screening trial. *Cancer*, **80,** 2091–2099.

Brookmeyer, R., Day, N. E. and Moss, S. (1986) Case-control studies for estimation of the natural history of preclinical disease from screening data. *Statistics in Medicine*, **5,** 127–138.

Chen, H.H., Duffy, S.W. and Tabar, Laszlo (1996) A Markov chain method to estimate the tumour progression rate from preclinical to clinical phase, sensitivity and positive predictive value for mammography in breast cancer screening. *The Statistician*, **45,** 307–317.

Chu, K.C., Smart, C.R. and Tarone, R.E. (1988) Analysis of breast cancer mortality and stage distribution by age for the health insurance plan clinical trial. *Journal of National Cancer Institute*, **80,** 1125–1132.

Claus, E.B., Schildkraut, J., Iversen, E.S. Jr., Berry, D. and Parmigiani, G. (1998) Effect of BRCA1 and BRCA2 on the association between breast cancer risk and family history. *Journal of the National Cancer Institute*, **90(23),** 1824–1829.

van Dijck, J.A.A.M., Verbeek, A.L.M., Hendricks, J.H.C.L. and Holland, R. (1993) The current detectability of breast cancer in a mammographic screening program. *Cancer*, **72,** 1933–1938.

Duffy, S.W., Day, N.E., Tabár, L. and Chen, H. (1997) Markov models of breast tumor progression: some age-specific specific results. *Journal of the National Cancer Institute Monographs*, **22,** 93–97.

Fletcher, S.W. (1997) Breast cancer screening among women in their forties: An overview of the issues. *Monographs of the National Cancer Institute*, **22,** 5–9.

von Fournier, D., Weber, E., Hoeffken, W., Bauer, M., Kubli, F. and Barth, V. (1980) Growth rate of 147 mammary carcinomas. *Cancer*, **45,** 2198–2207.

Gelfand, Alan E. and Smith, Adrian F. M. (1990) Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, **85,** 398–409.

Gelman, A., Roberts, G.O. and Gilks, W.R. (1996) Efficient metropolis jumping rules. In *Bayesian Statistics 5* (ed. J. Bernardo et al.), pp. 599–607. Oxford University Press.

Hart, D., Shochat, E. and Agur, Z. (1998) The growth law of primary breast cancer as inferred from mammography screening trials data. *British Journal of Cancer*, **78(3),** 382–387.

Hastings, W.K. (1970) Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, **57,** 97–109.

Heitjan, Daniel F. (1991) Generalized norton-simon models of tumor growth. *Statistics in Medicine*, **10,** 1075–1088.

Heuser, L., Spratt, J.S. and Polk, H.C. (1979) Growth rates of primary breast cancers. *Cancer*, **43,** 1888–1894.

Kerlikowske, K., Grady, D., Barclay, J., Frankel, S.T., Ominsky, S.H., Sickles, E.A. and Erster, V. (1998) Variability and accuracy in mammographic interpretation using the american college of radiology breast imaging reporting and data system. *Journal of the National Cancer Institute*, **90,** 1801–1809.

Laird, A.K. (1965) Dynamics of tumor growth: comparison of growth rates and extrapolation of growth curve to one cell. *Journal of Breast Cancer*, **19,** 278–291.

Lala, P.K. and Patt, H.M. (1966) Cytokinetic analysis of tumor growth. *Proceedings of the National Academy of Science USA*, **56,** 1735–1742.

Launoy, G., Duffy, S.W., Prevost, T.C. and Bouvier, V. (1998) Dépistage des cancers, sensibilité du test et sensibilité du programme de dépistage (*Test sensitivity and program sensitivty for cancer screening*). *Revue d'Épidémiologie et de Santé Publique*, **46(5),** 420–426.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. and Teller, E. (1953) Equations of state calculations by fast computing machine. *Journal of Chemical Physics*, **21,** 1087–1091.

Morrison, A.S. (1989) Review of evidence on the early detection and treatment of breast cancer. *Cancer*, **64,** 2651–2656.

Morrison, A.S., Brisson, J. and Khalid, N. (1988) Breast cancer incidence and mortality in the breast cancer detection demonstration project. *Journal of the National Cancer Institute*, **80,** 1540–1547.

Moskowitz, M. (1984) Mammography to screen asymptomatic women for breast cancer. *American Journal of Roentgenology*, **143(3),** 457–459.

Moskowitz, M. (1986) Breast cancer: Age-specific growth rates and screening strategies. *Radiology*, **161,** 37–41.

Mushlin, A.I., Kouides, R.W. and Shapiro, D.E. (1998) Estimating the accuracy of screening mammography: a meta-analysis. *American Journal of Preventative Medicine*, **14,** 143–153.

National Cancer Institute (1997) Surveillance, Epidemiology, and End Results (SEER) Program. `http://www-seer.ims.nci.nih.gov`.

Nyström, L., Rutqvist, L.E., Wall, S., Lindgren, A., Lindqvist, M., Rydén, S., Andersson, I., Bjurstam, N., Fagerberg, G., Frisell, J., Tabár, L. and Larsson, L.-G. (1993) Breast cancer screening with mammography: overview of Swedish randomised trials. *The Lancet*, **341,** 973–978.

Parmigiani, G. and Skates, S. (1999) Estimating the age of onset of detectable asymptomatic cancer. Technical Report. Institute of Statistics & Decision Sciences, Duke University. http://ftp.isds.duke.edu/WorkingPapers/99-03.ps.

Peer, P.G., Verbeek, A.L., Mravunac, M., Hendriks, J.H. and Holland, R. (1996) Prognosis of younger and older patients with early breast cancer. *British Journal of Cancer*, **73,** 382–385.

Peer, P.G.M., van Dijck, J.A.A.M., Hendriks, J.H.C.L., Holland, R. and Verbeek, A.L.M. (1993) Age-dependent growth rate of primary breast cancer. *Cancer*, **71,** 3547–3551.

Peer, P.G.M., Verbeek, A.L.M., Straatman, H, Hendriks, J.H.C.L. and Holland, R. (1996) Age-specific sensitivities of mammographic screening for breast cancer. *Breast Cancer Research and Treatment*, **38,** 153–160.

Peeters, P.H., Verbeek, A.L., Straatman, H., Holland, R., Hendriks, J.H., Mravunac, M., Rothengatter, C., Van Dijk-Milatz, A. and Werre, J.M. (1989) Evaluation of overdiagnosis of breast cancer in screening with mammography: results of the nijmegen programme. *International Journal of Epidemiology*, **18,** 295–299.

Perloff, M., Norton, L., Korzun, A.H., Wood, W.C., Carey, R.W., Gottlieb, A. *et al.* (1996) Postsurgical adjuvant chemotherapy of stage ii breast carcinoma with or without crossover to a non-cross-resistant regimen: a cancer and leukemia group b study. *Journal of Clinical Oncology*, **14,** 1589–98.

Roberts, G.O. (1995) *Markov Chain Monte Carlo in Practice*, ch. Markov chain concepts related to sampling algorithms, pp. 45–57. Chapman & Hall. eds Gilks, WR and Richardson, S and Spiegelhalter, DJ.

Schmitt, E.L. and Threatt, B. (1982) Tumor location and detectability in mammographic screening. *American Journal of Radiology*, **139,** 761–765.

Spratt, J.A., von Fournier, D., Spratt, J.S. and Weber, E.E. (1993) Mammographic assessment of human breast cancer growth and duration. *Cancer*, **71,** 2020–2026.

Spratt, J.S., Greenberg, R.A. and Heuser, L.S. (1986) Geometry, growth rates, and duration of cancer and carcinoma in situ of the breast before detection by screening. *Cancer Research*, **46,** 970–974.

Spratt, J.S., Meyer, J.S. and Spratt, J.A. (1995) Rates of growth of human neoplasms: Part i. *Journal of Surgical Oncology*, **60,** 137–146.

Spratt, J.S., Meyer, J.S. and Spratt, J.A. (1996) Rates of growth of human neoplasms: Part ii. *Journal of Surgical Oncology*, **61,** 68–83.

Spratt, J.S. and Spratt, J.A. (1985) What is breast cancer doing before we can detect it? *Journal of Surgical Oncology*, **30,** 156–160.

Straatman, H., Peer, P.G. and Verbeek, A.L. (1997) Estimating lead time and sensitivity in a screening program without estimating the incidence in the screened group. *Biometrics*, **53,** 217–229.

Taubes, G. (1997) The breast–screening brawl. *Science*, **275,** 1056–1059.

Tierney, L (1995) *Markov Chain Monte Carlo in Practice*, ch. Introduction to general state-space Markov chain theory, pp. 59–74. Chapman & Hall. eds Gilks, WR and Richardson, S and Spiegelhalter, DJ.

Tierney, Luke (1994) Markov chains for exploring posterior distributions (disc: P1728-1762). *The Annals of Statistics*, **22,** 1701–1728.

Walter, Stephen D. and Day, Nicholas E. (1983) Estimation of the duration of a pre-clinical disease state using screening data. *American Journal of Epidemiology*, **118,** 865–886.

Zelen, M. (1976) The theory of early detection of breast cancer in the general population. In *Breast Cancer: Trends in Research and Treatment, New York* (eds J C Heuson, W H Mattheiem and M Rozencweig), pp. 287–300. Raven Press.

Biography

Heidi Wen Ashih (Maiden name: Shih) was born in West Palm Beach, FL several months early on the 10^{th} of June, 1974. Ever impatient, she skipped kindergarten and never learned to share. Being an only child only solidified her ego-centric belief in "Heidi First!" She met, through grace of God, Aaron Christopher Wesley Ashih (Formerly: Ashford) during her first week at MIT. He properly proposed just 3 weeks later. They live in fear of television, with their kitties, and enjoy running, singing, and sunshine. Now that this blasted thing is finished, they hope to spend some time with their friends.

In her spare time (ha!), she likes to be a teen counselor at UU camps. Heidi is particularly fond of her new SUUSI friends, especially Conor M. Curtis.

She graduated from the Massachusetts Institute of Technology (MIT) with a Bachelor of Science (S.B.) in mathematics in May 1995. She continued her studies at Duke University, where she earned her Master's Degree (M.S.) in May 1997 and now her Doctorate (Ph.D) in July 2000.

Her published articles include:

Heidi W. Ashih, M.S., Tara Gustilo-Ashby, M.D., Evan R. Myers, M.D., M.P.H., Jeffrey Andrews, M.D., Daniel L. Clarke-Pearson, M.D., Donald Berry, Ph.D. and Andrew Berchuck, M.D. (1999) Cost-Effectiveness of Treatment of Early Stage Endometrial Cancer. *Gynecologic Oncology*, **74(2)**, pp. 208–216.

Ashih, Heidi W., Berry, Don A. and Parmigiani, Giovanni (1998) Modeling Natural History of Breast Cancer Tumor Growth. *ASA Proceedings of the Biometrics Section*, pp. 182–185. American Statistical Association (Alexandria, VA).

Awards, honors, and fellowships are too embarrassing to list, but will be given in person under duress.