

## Use Authorization

In presenting this thesis in partial fulfillment of the requirements for an advanced degree at Idaho State University, I agree that the Library shall make it freely available for inspection. I further state that permission to download and/or print my thesis for scholarly purposes may be granted by the Dean of the Graduate School, Dean of my academic division, or by the University Librarian. It is understood that any copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Signature \_\_\_\_\_

Date \_\_\_\_\_

**A Stochastic Model of Cancer Progression:  
Mathematical Analysis and Biomedical  
Implications**

by

**Jason Derek Rose**

A thesis submitted in partial fulfillment of the  
requirements for the degree of

**DOCTOR OF ARTS  
IN  
MATHEMATICS**

**IDAHO STATE UNIVERSITY**

**Fall 2014**

Copyright 2014 Jason Derek Rose

To the Graduate Faculty:

The members of the committee appointed to examine the thesis of JASON ROSE find it satisfactory and recommend that it be accepted.

---

Dr. Leonid Hanin  
Major Advisor

---

Dr. Yu Chen  
Committee Member

---

Dr. DeWayne Derryberry  
Committee Member

---

Dr. Ann Gironella  
Committee Member

---

Dr. Ken Aho  
Graduate Faculty Representative

## Contents

List of Figures	viii
List of Tables	x
Abstract	xi
Chapter 1. Introduction and Background	1
1.1. The relation between primary tumor size and the rate of metastasis shedding	5
1.2. The timing of metastatic development relative to the development of the primary tumor	7
1.3. Metastatic dormancy	7
1.4. The interaction between the primary tumor and metastases	8
1.5. Summary	10
Chapter 2. Mathematical Model of Metastatic Cancer	12
2.1. A mathematical model of cancer progression	12
2.2. Distribution of the sizes of detectable metastases	16
2.3. The distribution of the sizes of detectable metastases for exponentially growing tumors with exponentially distributed latency times	23
2.4. A Note on Parameters and Parameter Recovery	27
2.5. Model Analysis	29
Chapter 3. Submodels and Limiting Cases	30
3.1. Submodels arising from alterations to the Poisson process	32
3.2. Submodels arising from alterations to the growth model	39

Chapter 4. Data Description and Conversion	47
4.1. MSKCC data and conversion	47
4.2. Autopsy data and conversion	48
4.3. Full Model and Submodels written in terms of diameters	51
4.4. Expected diameter method	54
4.5. Computing $F_Z$ and $f_Z$ for the True Diameter Methods	57
Chapter 5. Perils of Maximum Likelihood Estimation	70
5.1. Infinite Likelihood Using Volume Data and Expected Diameters	70
5.2. CSPT Using the Method of True Diameters	78
5.3. IISMG Using the Method of True Diameters	79
Chapter 6. Data Analysis and Parameter Estimation	80
6.1. Patient A	82
6.2. Protocol 10	88
6.3. Protocol 17	96
Chapter 7. Results and Conclusions	106
7.1. Comparison of the methods of parameter estimation.	106
7.2. Stem-like cancer cells	108
7.3. Suppression of metastatic growth by the primary tumor	109
7.4. Acceleration of metastatic growth by primary tumor resection	110
7.5. Duration of metastatic latency	111
7.6. Timing of cancer onset.	111
7.7. Primary Tumor Growth Rates and Progression Time	112
7.8. Was the primary tumor detectable at inception of the first metastasis?	112
7.9. Summary	115
Bibliography	117
Appendix A. Some Useful Results Concerning Poisson Processes	121

A.1.	Definition of a Poisson process	121
A.2.	Joint distribution of the time of occurrence of events in a Poisson process given their number	123
A.3.	Independent Classification of Events	130
A.4.	Intensity of a filtered Poisson process	132
A.5.	Intensity of a delayed Poisson process	134
Appendix B. Catalog of Models		136
B.1.	Full Model	136
B.2.	Instantaneous Seeding	138
B.3.	Homogeneous	140
B.4.	Heavy-Seeding/Long-Latency (HSLL)	142
B.5.	Complete Suppression by Primary Tumor (CSPT)	144
B.6.	Instantaneous Infinite Shedding and Metastasis Growth (IISMG)	146
Appendix C. Some Computations for the CSPT and IISMG Models		147
C.1.	Proof that $Q$ can be determined from $\sigma$ in the CSPT model	147
C.2.	Proof that $g(b_1)$ is decreasing	149

## List of Figures

Integral Simplification Case i	19
Integral Simplification Case ii	20
Relationship between spherical diameter and profile diameter	49
Maximum error between profile diameter and spherical diameter	50
$Q$ as a function of $\sigma$ , CSPT model, Patient A	84
Graphs of empirical and theoretical cdfs, likelihood maximizing parameters, CSPT and IISMG models, Patient A	84
Graphs of NNLL profile functions, likelihood maximizing parameters, Compromise Likelihood, Patient A	85
Graphs of $L^2$ norm profile functions, $L^2$ minimizing parameters, Full model, Patient A	86
Graphs of empirical and theoretical cdfs, likelihood maximizing and $L^2$ minimizing parameters, Full model, Patient A	86
Graphs of NNLL profile functions, likelihood maximizing parameters, Homogeneous model, Protocol 10 Transformed data	90
Graphs of empirical and theoretical cdfs, likelihood maximizing and $L^2$ minimizing parameters, Homogeneous model, Protocol 10, Transformed data	91
Graphs of NNLL profile functions, likelihood maximizing parameters, Full model, Protocol 10, True diameters without rounding	92
Graphs of empirical and theoretical cdfs, likelihood maximizing and $L^2$ minimizing parameters, Full Model, Protocol 10, True Diameters without rounding	93

Graphs of NNLL profile functions, likelihood maximizing parameters, Full model, Protocol 10, True diameters with rounding	94
Graphs of piecewise continuous linear empirical cdf and theoretical cdfs, likelihood maximizing parameters, Full model, Protocol 10, True Diameters with Rounding	95
$Q$ as a function of $\sigma$ , CSPT model, Protocol 17	98
Graphs of empirical and theoretical cdfs, likelihood maximizing parameters, CSPT and IISMG models, Protocol 17, Transformed data	98
Graphs of NNLL profile functions, maximum likelihood parameters, Full model, Protocol 17, Method of Expected Diameters	99
Graphs of empirical and theoretical cdfs, likelihood maximizing and $L^2$ minimizing parameters, Full model, Protocol 17, Transformed data	100
Graphs of $L^2$ norm profile functions, $L^2$ norm minimizing parameters, Full model, Protocol 17, Transformed data	101
Graphs of NNLL profile functions, maximum likelihood parameters, HSSL model, Protocol 17, True Diameters with Rounding	103
Graphs of piecewise continuous linear empirical cdfs and theoretical cdfs, likelihood maximizing and $L^2$ minimizing parameters, HSSL model, Protocol 17, True Diameters with Rounding	104
Graphs of $L^2$ norm profile functions, $L^2$ norm minimizing parameters, HSSL model, Protocol 17, True Diameters with Rounding	105

## List of Tables

Summary of Submodels and Limiting Cases	32
Summary of parameters for Patient A	88
Expected diameters of metastases, Protocol 10	89
Summary of parameters for Protocol 10	96
Expected diameters of metastases, Protocol 17	97
Summary of parameters for Protocol 17	104

A Stochastic Model of Cancer Progression: Mathematical Analysis and Biomedical  
Implications

Thesis Abstract – Idaho State University (2014)

A stochastic model of metastatic cancer that accounts for metastasis shedding by the primary tumor, growth of the primary tumor, metastatic latency, and growth of metastases before and after resection of the primary tumor is reviewed and analyzed. Parameters of this model were estimated by using data from a breast cancer patient who underwent surgery and whose metastases were measured at a later time. The model is adapted to be used for autopsy data obtained by surveying cross sections of metastases in thin slices of liver tissue. Model parameters are then estimated for two autopsy patients. Results corroborate the hypotheses that (1) the spread of metastases may begin very early, (2) metastases may remain latent for a long period of time, and (3) that removal of the primary tumor can cause an increase in the growth rate of metastases.

## CHAPTER 1

### Introduction and Background

Worldwide, cancer accounts for about 13% of all deaths [36]. This year, about 570,000 Americans are expected to die from cancer, which makes it responsible for nearly 1 in 4 deaths in this country and the second most common cause of death in the U.S. after heart disease [46].

At the cellular level, cancer begins with a mutation in a single cell that affects the regulation of cell proliferation, cycling, and apoptosis. Such mutations can occur in a cell in any part of the body and may occur spontaneously or be caused by either external agents or inherited genetic factors [36]. Cell division allows the mutation to be replicated. The resulting group of abnormal daughter cells can invade and destroy surrounding normal tissue and form a tumor. While this primary tumor can certainly be disruptive to the normal function of the body, in many solid cancers a much more serious condition known as metastasis can occur. In metastasis, the malignant cells spread through blood and lymph vessels to neighboring and even distant parts of the body to create secondary tumors referred to as metastases. Even if the primary tumor is located in a non-vital organ, its metastases may spread to vital organs. For this reason, metastasis is the major cause of mortality, accounting for about 90% of cancer-related deaths [36].

Throughout recorded history, efforts to find an effective treatment for cancer have generally ended in frustration. As early as 3000 B.C., the Egyptian physician Imhotep described how to distinguish cancer from a mere infection and pronounced “there is no treatment [32].” Writing around 400 BC, Hypocrites observed that “It is better not to apply any treatment in cases of occult cancer; for, if treated, the patients die quickly; but if not treated, they hold out for a long time [12].”

The Roman encyclopedist Aulus Cornelius Celsus (30 BC to 38 AD) observed that only the smallest tumors could be successfully removed. “Nor was any person ever relieved by medicine but after cauterizing the tumours have been quickened in their progress and increased till they proved mortal when they have been cut out and cicatrized they have notwithstanding returned and occasioned death [6].”

Modern efforts have also met with limited success. The discovery that cells propagate by division (by Robert Remak in the 1850’s [26]) lead to a deeper understanding of the nature of cancer. Building on Remak’s discovery and autopsy studies, the pathologist and doctor Rudolf Virchow (sometimes referred to as the “father of modern pathology”) proposed a model for breast cancer that became very influential in the search for a cure. According to cancer researcher Michael Retsky and his colleagues [39], Virchow “suggested that the disease started as a single focus within the breast, expanding with time and then migrating along lymphatic channels to the lymph glands in the axilla. These glands were said to act as a first line of defense filtering out the cancer cells. Once these filters became saturated the glands themselves acted as a nidus for [further] spread to a second and third line of defense like the curtain walls around a medieval citadel. Ultimately when all defenses were exhausted the disease spread along tissue planes to the skeleton and vital organs.” The logical conclusion of this “linear” model of cancer progression was that early detection and removal of the primary tumor and surrounding tissue might cure the disease.

In Virchow’s time, surgery was a dangerous proposition. However, by the end of the 19th century, antiseptics and anesthesia made surgery more practical and the surgeon William Halstead, guided by Virchow’s model, pioneered the treatment of breast cancer through mastectomy. Halstead’s early success and subsequent training programs helped to position mastectomy as the standard treatment for the next 75 years [11]. When patients who underwent this procedure still continued on to develop metastatic disease, the natural assumption was that not enough had been removed. The surgical procedures became more extensive by removing all of the breast, the

overlying skin, the underlying muscles, and the lymph nodes and, later, using follow-up radiation to scour the tissue that remained near the surgical site. Radical surgery and postoperative radiotherapy did reduce the chance of local recurrence, but it did not improve overall disease-free survival. An overview study published in 1995 concluded that “Some of the local therapies for breast cancer had substantially different effects on the rates of local recurrence – such as the reduced recurrence with the addition of radiotherapy to surgery – but there were no definite differences in overall survival at 10 years [18].”

As radical surgery was found to have limitations, focus shifted to earlier detection through more widespread screening. Unfortunately, results for breast and prostate cancer screening have been mixed. In the case of breast cancer, the advent of mammography raised hopes that early detection and removal of primary tumors could pre-empt the formation of metastases and lead to higher survival rates. As data from large-scale clinical trials has been analyzed and interpreted, it has been observed that for every life saved by early detection, 10 women were treated unnecessarily for cancer that was not life-threatening and may even have gone away on its own if left untreated. A more dramatic result of these trials was the observation of what has been called the mammography paradox: for pre-menopausal women aged 40-49 with node-positive breast cancer, early screening was associated with an increase in mortality rates [38], [39].

A similar scenario arose in the case of prostate cancer. In 1970, Richard Ablin identified a protein produced by the cells of the prostate gland known as prostate-specific antigen (PSA) [37]. High levels of PSA were associated with both localized and metastatic prostate cancer and, once a quantitative blood test was developed, could be used to screen for prostate cancer. PSA became widely used in screening for prostate cancer, but results have been disappointing. It has been estimated that for every one death prevented, about 50 men are needlessly treated for prostate cancer. As Hanin [19] has observed, “very much like mammography, PSA screening led to massive overdiagnosis, combined with failure to timely detect deadly tumors, decreased

survival in a subset of cases, and reduced quality of life for those unduly treated.” A study published in March 2011 concluded “After 20 years of follow-up the rate of death from prostate cancer did not differ significantly between men in the screening group and those in the control group.” [44]

The failure of early detection and treatment to prevent metastatic disease points to a more complex reality underlying the development of the primary tumor, the mechanism by which metastases are spread, and the interaction between the primary tumor and metastases. The process of development and progression of cancer is generally hidden from the observer and the initiating events (such as the development of the first malignant cell) are separated from the final outcomes by large temporal distances.

Mathematics is an indispensable tool for modeling unobservable events and processes and for linking the cancer-initiating event to clinical outcomes that become observable months, years, or even decades later. It can therefore play an important role in understanding both the progress of cancer and its response to treatment. In this work, we present a mathematical model for the progression of metastatic cancer. We then fit this model to data in order to investigate four aspects of the complex dynamics of cancer:

1. the relation between primary tumor size and the rate of shedding of metastases,
2. the timing of metastatic progression relative to the development of the primary tumor,
3. metastatic dormancy, and
4. the interaction between the primary tumor and metastases.

In the remainder of this chapter we discuss some of the current ideas and observations relevant to these four aspects.

## 1.1. The relation between primary tumor size and the rate of metastasis shedding

In order to understand the process of metastasis formation, we first examine how the intensity of metastasis shedding is related to the size of the primary tumor. In order to shed metastases, a tumor must have access to blood or lymph vessels. Beyond that, it would seem reasonable to suppose that the larger the tumor, the more metastatic cells can be shed. Cells near the surface of the tumor are known to be more actively proliferating than those in the center [4], so metastasis shedding may have more to do with the surface area of the tumor than the volume. It is also possible that the surface area of the vascular network that supplies the tumor with nutrients, oxygen, and growth factors might have an important effect.

**1.1.1. Cancer Stem Cells.** While our understanding of stem cells is still evolving, stem cells are generally defined as cells that have the ability to perpetuate themselves through self-renewal, the capacity for fast proliferation, and pluripotency, i.e. the ability to generate mature functional cells of a particular tissue through differentiation [41]. Stem cells show varying degrees of differentiation: cells of a fertilized egg (embryonic stem cells) can obviously differentiate into all cell types, but other cells within the body (such as hematopoietic stem cells) have been shown to have some degree of differentiation, but can still give rise to further differentiated cells within their particular lineage. Just as many normal tissues have a very small subpopulation of stem cells, it may be that a tumor possesses a very small sub-population of cancer stem cells (CSCs.) Because they would be less differentiated than the surrounding tumor cells, CSCs may be less susceptible to drugs targeting tumor cells and form a drug-resistant core that allows tumors to survive chemotherapy [7]. Given their hypothesized pluripotency, high proliferative potential, and ability to produce metastases, a small core of surviving CSCs would be able to perpetuate the disease even after extensive treatment.

The evidence for CSCs is strongest for blood cancers. For these cancers, tritium-labeling studies conducted in the 1960's suggested the presence of a sub-population of primitive-appearing cells within the larger group of cancer cells [7]. But it wasn't until the late 1990's that more conclusive evidence was produced by studies of acute myelogenous lymphoma (AML) which showed that when human AML cell populations are sorted according to certain cell-surface markers, some types can regenerate a human AML cell population within immunodeficient mice while other types cannot [3]. In the case of breast cancer, it has been found that cells with two particular surface markers could be used to induce tumors in recipient animals with doses as small as 100 cells while doses of tens of thousands of cells with a different marker did not induce tumors [7], [25]. Similar effects have been observed for certain brain tumors and for human prostate cancer [28], [7].

While cancer stem cells offer an explanation for resistance to chemotherapy, they also have implications for the effect of primary tumor size on the rate of metastasis shedding. If CSCs are a key component of the metastasis shedding mechanism, then metastasis formation would have more to do with the size of this subpopulation than the overall size of the tumor. As suggested by self-renewal, stem cells are able to divide in an asymmetric way. Rather than dividing symmetrically into two new stem cells (which stem cells can do to increase their population) they can also divide into one stem cell and one differentiated cell. Asymmetric division of CSCs would preserve the number of CSCs while giving rise to other differentiated cells within the tumor. Thus a tumor could shed many, many cells, and yet only those which are cancer stem cells or poorly differentiated cells would have the ability to create a new tumor in a metastatic site. In this case, the rate of metastasis shedding would be relatively constant and proportional to the size of the relatively stable population of CSCs.

## **1.2. The timing of metastatic development relative to the development of the primary tumor**

In the 1970's, the failure of radical mastectomy to increase breast cancer survival rates prompted a prominent researcher from the University of Pittsburgh, Bernard Fisher, to speculate that cancer spreads via the blood stream and lymphatics even before its clinical detection [13]. This suggests that the extent of surgery would have little effect on patients' survival. Numerous clinical trials showed that this is indeed the case [13]. Thus, at the time of diagnosis, cancer cells may already be circulating in the blood or lymph system and metastases may have already been seeded and begun to grow in a host site.

In fact, circulating tumor cells (CTCs) were first observed in 1869 [2]. Since then, CSCs have been confirmed by numerous studies. A recent study showed that about one-third of breast cancer patients 7-22 years after mastectomy and without any evidence of the disease had CTCs [30]. Levels of circulating tumor cells have been shown to fall off rapidly after removal of the primary tumor [39], so continued presence of CTCs years after removal of the primary tumor may suggest the presence of undetected primary or secondary tumors. If so, the presence of CTCs from undetected tumors gives credence to Fisher's idea that metastases may be spreading long before detection. Fisher posited that the early dissemination of metastases would mean that adjuvant systemic therapies (such as chemotherapy) that treat for metastatic disease would provide benefit even for tumors that appear to be very localized at the time of diagnosis. This has been shown to be the case, but, unfortunately, with only a modest effect [18].

## **1.3. Metastatic dormancy**

Regardless of the timing of dissemination, metastasis progression is a complex and highly selective process, and may involve a considerable period of time. A metastatic cell must be shed off the primary, migrate to and penetrate a blood or lymph vessel, evade detection by the immune system, extravasate at a suitable location and begin

to grow. Tumor cells in the lymph system are not simply caught in the lymph nodes, and tumor cells in the circulatory system do not simply lodge in the first capillary bed they encounter. They are much more mobile and able to traverse both systems as well as cross the interstitial space between organs [13].

In addition to the time a metastatic cell spends in travel, it has been shown that metastatic cells undergo a period of dormancy before they begin an irreversible proliferation leading to a detectable secondary metastasis [33]. In order to grow beyond about  $10^6$  cells (1-2mm in diameter), a solid tumor must develop its own blood supply to overcome resource limitations [14], thus solitary cells or pockets of cells may remain dormant in metastatic sites at undetectable levels for an extended period of time.

#### 1.4. The interaction between the primary tumor and metastases

Animal studies have repeatedly demonstrated that there is an interaction between the primary tumor and metastases. In 1905, Paul Ehrlich observed that mice that already had a primary tumor exhibited resistance to grafts of a second tumor [17]. In some experiments, mice with artificially established tumors became refractory to inoculation with a dose of tumor cells several orders of magnitude higher than the dose required to produce a primary tumor in a normal mouse. At first this phenomenon was thought to be caused by the animal's immune system, and it was dubbed *concomitant immunity* but later experiments with immunologically compromised mice showed that this protection could not be due to an immune response [16]. It was also observed that the level of protection increased with the size of the primary tumor and that when implanted tumors were partially excised, it caused metastases to proliferate.

Research gradually came to focus on the interplay between tumors, growth and angiogenesis promoters, and growth and angiogenesis inhibitors. In the 1970's, a prominent oncologist and cancer researcher, Judah Folkman, was able to show that tumors without a blood supply can grow to at most 1-2 mm in diameter [14]. He also used

Millipore filters to separate tumors from their environs and show that tumors can induce angiogenesis on the other side of the filter by a diffusible factor (an angiogenesis promoter) that can travel 3-5 mm. Tumors were also found to produce angiogenesis inhibitors, but these inhibitors are more stable than the promoters and could act over much larger ranges [35], [34]. Rather than a single angiogenesis promoter/inhibitor pair, there are actually multiple angiogenesis and tumor growth promotion/inhibition pathways that may have some degree of interaction. This gives an explanation as to how one tumor might suppress the growth of another: while promoting its own vascularization and/or growth through short-range angiogenesis/growth promoters, a tumor releases long-range angiogenesis/growth inhibitors that suppress the vascularization/growth of other tumors. In the presence of growth inhibitors and without the ability to obtain a blood supply, metastases could only grow to a limiting size and then remain effectively dormant, unable to sustain further growth.

Removal of the primary tumor also removes the source of growth and angiogenesis inhibitors which allows metastases to develop their own blood supply and accelerate their growth without bound. Certainly metastases develop naturally even when the primary tumor is not removed, but removal of the primary may throw the “angiogenic switch” and hasten their progression. In addition, the wound healing process that attends surgery would trigger the systemic production of growth and angiogenesis factors.

Although much of the research on concomitant immunity has been conducted in animals, there is evidence to suggest that it may also occur in humans. Studies in terminally ill patients have demonstrated the suppression of a secondary tumor by a primary tumor [15]. In certain cases of testicular cancer in humans, partial excision of the primary tumor has been observed to exacerbate the disease [29]. In the mid-90’s, Romano Demichelli et. al. [8] published a study of relapse data from 1173 early-stage breast cancer patients who had been treated with only surgical excision of the primary. When they examined the hazard rate for these patients, they found it to be

trimodal with a sharp peak at 18 months, another at 30 months, a nadir at 50 months, a broad peak at 60, and a long tail extending 15-20 years. Retsky et al. [39], [40] posit that the third peak represents the slow natural progression of metastases from single cells to avascular micrometastases and then to vascular metastases that grow large enough for detection. They see the first and second peaks as corroborative of a rapid development of metastases brought on by surgery and the attendant release of growth and angiogenesis promoters as well as removal of the growth and angiogenesis inhibitors that had been secreted by the primary tumor. They posit that avascular micrometastases "awakened" by surgery lead to the first peak and single metastatic cells awakened by surgery lead to the second peak.

### 1.5. Summary

The conceptual framework for understanding the progression of cancer has evolved in response to clinical/experimental observations and epidemiological studies. The main shift has been from a view of cancer at a genomic and cellular level (one mutated cell divides and slowly conquers regional territory until metastatic cells can be spread to produce independent colonies in other organs) to a view of cancer at a systemic level (primary and secondary tumors develop and interact between themselves and with their microenvironment in a way that may be similar to the cells that comprise organs.) In the next chapter, we develop a mathematical model of cancer progression that is flexible enough to capture some of these systemic effects. In particular, the results of our model-based analysis will shed light on

1. the relation between primary tumor size and the rate of shedding of metastases,
2. the existence of cancer stem cells,
3. the timing of metastatic dissemination relative to the development of the primary tumor,
4. the duration of metastatic dormancy, and
5. the effects of primary tumor resection on the rate of growth of metastases.

In the following chapters, we fit this model to data and then examine the resulting parameters for insight into the processes at work.

## Mathematical Model of Metastatic Cancer

### 2.1. A mathematical model of cancer progression

In our model, we divide the natural history of invasive cancer into three time periods which we call the disease-free period, primary tumor growth, and metastasis formation. Note that the last two periods may overlap.

**Disease-free period.** This period begins with the birth of the individual (or the start of the exposure to a carcinogen in the case of an induced tumor) and ends with the production of the first malignant cell whose clone will go on to create the primary tumor. The production of this first cell marks the *onset of the disease* and we use the variable  $T$  to denote the onset time (as measured from the time of birth or start of carcinogen exposure.)

**Primary tumor growth.** We assume that once the primary tumor emerges, its growth will be a deterministic function of time,  $t$ , as measured from the disease onset at time  $T$ . We denote this function by  $\Phi(t)$  and assume that it is strictly increasing and continuous until the time of resection, if any. We will measure  $\Phi$  in terms of the actual count of cells, and so we have the initial condition  $\Phi(0) = 1$ . The inverse function of  $\Phi$  will be denoted by  $\phi$ .

**Metastasis formation.** We assume that metastases will be shed off the primary tumor according to a non-homogeneous Poisson process with intensity  $\mu$ . In the case of some cancers, including breast and prostate cancer, there is evidence to suggest that only certain cells within the primary tumor are capable of producing metastases [25], [28], [7]. If we denote the size of this sub-population of metastasis-producing cells by  $N(t)$ , then the shedding intensity should be proportional to  $N$  so that we can find a constant  $\alpha_0 > 0$  such that  $\mu(t) = \alpha_0 N(t)$ . It is reasonable to

assume that the number of metastasis-producing cells should be related to the size,  $\Phi$ , of the primary tumor. Specifically, we assume that there is a constant  $\alpha_1$  such that  $N(t) = \alpha_1 \Phi^\theta(t)$  for some  $\theta \geq 0$ . Combining  $\alpha_0$  and  $\alpha_1$  into a single constant,  $\alpha$ , gives  $\mu(t) = \alpha_0 \alpha_1 \Phi^\theta(t) = \alpha \Phi^\theta(t)$ . Because actively proliferating clonogenic cells are typically located close to the tumor surface, one possible scenario is that the intensity of metastasis shedding is proportional to the surface area of the primary tumor, and because  $\Phi$  is proportional to the volume of the tumor, the intensity,  $\mu$ , would be proportional to  $\Phi^{2/3}$ . This would give  $\mu(t) = \alpha \Phi^\theta$ , where  $\theta = \frac{2}{3}$  and  $\alpha$  is some (positive) constant of proportionality. Another possibility is that every cell in the primary tumor has a chance to become metastatic; in this case  $\theta = 1$ . On the other hand, if only a small, persistent group of tumor cells are capable of producing metastases, then we would expect  $N(t)$  to be constant. This would mean that the intensity would be approximately constant in time and we would expect  $\theta$  to be close to zero. As discussed in the introduction, such would be the case for cancer stem cells, which can divide asymmetrically to produce one new stem cell and one differentiated cell, thus limiting the growth of their population. To allow our model more generality, we leave the power on  $\Phi$  unspecified and assume that  $\mu(t) = \alpha \Phi^\theta$  where  $\theta$  is some non-negative exponent.

Once a potential metastasis has been shed from the primary tumor, it still faces many obstacles that would prevent it from becoming a detectable metastasis in a secondary site. After it is separated from the primary tumor, it must migrate to and penetrate a blood or lymph vessel in order to travel to another site. It then must extravasate and invade the new site, survive through a period of dormancy, start to proliferate, and induce angiogenesis in order to support its growth. All of this must be accomplished while it evades the host's immune system. We will assume that each potential metastasis that is shed from the primary tumor has a fixed probability,  $q$ , of accomplishing these tasks and becoming a detectable metastasis in a given secondary site.

Thus the production of viable metastases in a given site (as opposed to shedding of potential metastases) is governed by what is known as a “filtered” Poisson process. In the Appendix, we show that a filtered Poisson process of intensity  $\mu(t)$  is a Poisson process with intensity  $\nu(t) = q\mu(t) = q\alpha\Phi^\theta(t)$ .

We use the term *inception* to refer to the start of irreversible proliferation leading to a detectable secondary tumor. As discussed in Chapter 1, there is experimental evidence to suggest that metastases spend a period of time between detachment from the primary tumor and inception. We assume that these *latency times* for different metastases bound for a given secondary site are independent and identically distributed with some probability density function (pdf)  $f$ . In the Appendix, we show that this results in a “delayed” Poisson process with intensity

$$(2.1) \quad \lambda(t) = \int_0^t \nu(s)f(t-s)ds.$$

**Observable events.** Although the time of disease onset,  $T$ , is unobservable, we can assume that at some time  $U \geq T$ , a patient will be diagnosed with a primary tumor and its size,  $S$ , will be measured. According to our growth law,  $S = \Phi(U - T)$  so we can use the inverse function  $\phi$  to infer that  $\phi(S) = U - T$  or

$$(2.2) \quad T = U - \phi(S).$$

If the primary tumor is resected, it will occur at some time,  $V$ , at or after diagnosis, i.e.  $V \geq U$ . We will assume that at some time  $W$  (at either a follow-up visit or at the time of death) the patient will be found to have developed observable metastases. We will denote the number of observable metastases by  $n$  and their volumes by  $X_1, X_2, \dots, X_n$ , where  $X_1 \leq X_2 \leq \dots < X_n$ . It is clear that  $W \geq U$ , and we can assume that  $W \geq V$  because resection of the tumor after time  $W$  makes no difference as far as our model or inference on its parameters are concerned. In cases where the

primary tumor was not resected, we set  $V = W$ . Thus we have the timeline

$$0 \text{ (birth)} \leq T \text{ (onset)} \leq U \text{ (diagnosis)} \\ \leq V \text{ (resection)} \leq W \text{ (measurement of metastases)}.$$

**Metastasis growth.** In Chapter 1, we observed that there is evidence in both animals and humans to suggest that the presence of the primary tumor has an effect on the growth rate of metastases. In order to investigate this effect, we will assume that the growth of metastases in a given host site follows one function before resection and another function after resection. Before resection, we assume that growth is described by the function  $\Psi_0$ . After resection, we assume that a new growth function,  $\Psi_1$ , acts multiplicatively on the size of the metastasis at the time of resection (as would, for example, be the case if growth were exponential before and after resection, but with different growth rates.) We assume that both  $\Psi_0$  and  $\Psi_1$  are strictly-increasing, differentiable functions of time with  $\Psi_i(0) = 1, i = 0, 1$ . Because metastases begin growth in their host sites at different times, each metastasis will grow for a different amount of time pre- and/or post-resection. We define  $Y$  to be the time of inception of a metastasis in a given secondary site as measured from the onset of the disease. Then the size of the metastasis at time  $t$  (measured from inception) is represented by

$$\Psi_Y(t) = \begin{cases} \Psi_0(t) & \text{if } t \leq V - T - Y \text{ and } Y < V - T, \\ \Psi_0(V - T - Y)\Psi_1(t - (V - T - Y)) & \text{if } t > V - T - Y \text{ and } Y < V - T, \\ \Psi_1(t) & \text{if } Y \geq V - T. \end{cases}$$

Notice that  $V - T$  is the amount of time from onset to resection,  $V - T - Y$  is the amount of time between inception of the metastasis and resection,  $\Psi_0(V - T - Y)$  is the size of a metastasis at the time of resection, and  $t - (V - T - Y)$  is the amount of time the metastasis has been growing beyond the time of resection.

**Secondary metastasis.** Although it is conceivable that the metastases themselves might shed metastases of their own, we will assume that this shedding is negligible.

**Metastasis detection.** The volume of a metastasis becomes measurable when it reaches some threshold value,  $m$ , determined by the sensitivity of the imaging technology or survey method used to detect the metastases.

## 2.2. Distribution of the sizes of detectable metastases

Assume that at age  $W$  (the time of our final survey) a metastasis of size  $X$  is detected. If we denote, as above, the inception time by  $Y$  (relative to the onset of the disease), then according to our metastasis growth law,

$$X = \Psi_Y(W - T - Y) = \begin{cases} \Psi_0(V - T - Y)\Psi_1(W - V) & \text{if } Y < V - T, \\ \Psi_1(W - T - Y) & \text{if } Y \geq V - T. \end{cases}$$

We now show that we can invert this relationship to discover the inception in terms of the size at time  $W$ . Notice that  $W - T$  is the time from onset to final survey, so  $W - T - Y$  is the total growth time for the metastasis from its inception to detection at time  $W$ . We call  $W - T - Y$  the *progression time* of the metastasis and label it by the variable  $y$ . Thus  $X = \Psi_Y(W - T - Y) = \Psi_Y(y)$ . By replacing  $W - T - Y$  with  $y$ , the expression  $V - T - Y$  can be written as  $V + y - W$  or  $y - (W - V)$ . Also, the inequality  $Y < V - T$  can be written as  $W - V < W - T - Y$  or  $W - V < y$ . Similarly, the inequality  $Y \geq V - T$  can be written as  $y \leq W - V$ . With these simplifications, the metastasis size,  $X$ , can be written in terms of the progression time,  $y$ , as

$$X = \Psi(y) = \begin{cases} \Psi_0(y - (W - V))\Psi_1(W - V) & \text{if } W - V < y \leq W - T, \\ \Psi_1(y) & \text{if } 0 \leq y \leq W - V. \end{cases}$$

Thus, the function  $\Psi$  that relates the size of a metastasis to its progression time is independent of  $Y$ ! So we can write  $X = \Psi(W - T - Y)$ . The progression time cannot

exceed the time from disease onset to the final survey, so the maximum value of  $\Psi$  is

$$M = \Psi(W - T) = \Psi_0(V - T)\Psi_1(W - V).$$

Because of our assumptions on  $\Psi_0$  and  $\Psi_1$ ,  $\Psi$  is strictly increasing, continuous, piecewise-differentiable, and satisfies  $\Psi(0) = 1$ . We find the inverse function  $\psi = \Psi^{-1}$  to be

$$(2.3) \quad \psi(x) = \begin{cases} \Psi_1^{-1}(x) & \text{if } 1 \leq x \leq \Psi_1(W - V), \\ \Psi_0^{-1}\left(\frac{x}{\Psi_1(W - V)}\right) + W - V & \text{if } \Psi_1(W - V) < x \leq M, \end{cases}$$

which gives the metastasis' progression time as a function of its size.

**Theorem 1.** The sizes  $X_1 < X_2 < \dots < X_n$  of metastases in a given host site that are detectable at age  $W$  are equidistributed, given their number  $n$ , with the vector of order statistics for a random sample of size  $n$  drawn from the distribution with pdf defined by

$$p(x) = \begin{cases} \omega(W - T - \psi(x))\psi'(x), & \text{if } m \leq x \leq M, \\ 0, & \text{if } x \notin [m, M], \end{cases}$$

where

$$(2.4) \quad \omega(t) = \frac{\int_0^{\min\{t, V-T\}} \Phi^\theta(s) f(t-s) ds}{\int_0^{\min\{W-T-\psi(m), V-T\}} \Phi^\theta(s) F(W-T-\psi(m)-s) ds}$$

if  $0 \leq t \leq W - T - \psi(m)$ ,

and  $F$  is the cumulative distribution function (cdf) of the metastasis latency time corresponding to the pdf  $f$ .

**Proof:** Given the sizes of detectable metastases,  $X_1 < X_2 < \dots < X_n$ , at the time of final survey,  $W$ , let  $\tau = (\tau_1, \tau_2, \dots, \tau_n)$  be the vector of corresponding inception times counted from the onset of the disease. (These times can be found using  $\tau_i = W - \psi(X_i) - T$ .) Because our growth function,  $\Psi$ , is increasing, the larger the tumor, the earlier its inception time, so  $\tau_1 > \tau_2 > \dots > \tau_n$ . With  $m$  as our smallest detectable

size, the latest possible inception time is  $W - \psi(m) - T$ , so a metastasis is detectable if and only if  $0 \leq \tau_i \leq W - \psi(m) - T$ . Recall that the intensity of shedding of viable metastases is given by  $\nu(t) = q\mu = q\alpha\Phi^\theta(t)$ . This holds until the primary tumor is removed at time  $V$ , after which we set the shedding intensity to zero. Because of the random latency time, the intensity of actual metastasis inception is (from (2.1)) given by

$$(2.5) \quad \lambda(t) = q\alpha \int_0^{\min\{t, V-T\}} \Phi^\theta(s) f(t-s) ds.$$

In the Appendix, we show that for a Poisson process  $X(t)$  with time-dependent locally integrable rate  $\lambda(t)$ , the joint pdf of the occurrence times given the number,  $n$ , of events is

$$f_{W_1 \dots W_n | X(t)=n}(w_1, \dots, w_n) = \begin{cases} n! \omega(w_1) \dots \omega(w_n), & \text{for } t \geq w_n > w_{n-1} > \dots > w_1 > 0, \\ 0 & \text{otherwise,} \end{cases}$$

where  $\omega(u) = \frac{\lambda(u)}{\int_0^t \lambda(s) ds}$ .

Thus, in our context, the pdf of inception times is

$$p_\tau(t_1, t_2, \dots, t_n | N = n) = \begin{cases} n! \omega(t_1) \omega(t_2) \dots \omega(t_n) & \text{for } W - T - \psi(m) \geq t_1 > t_2 > \dots > t_n \geq 0 \\ 0 & \text{otherwise,} \end{cases}$$

where

$$\omega(t) = \frac{\lambda(t)}{\int_0^{W-T-\psi(m)} \lambda(s) ds}, \quad 0 \leq t \leq W - T - \psi(m).$$

Replacing  $\lambda(t)$  with its expression in (2.5) gives

$$\begin{aligned} \omega(t) &= \frac{q\alpha \int_0^{\min\{t, V-T\}} \Phi^\theta(s) f(t-s) ds}{\int_0^{W-T-\psi(m)} q\alpha \int_0^{\min\{u, V-T\}} \Phi^\theta(s) f(u-s) ds du} \\ &= \frac{\int_0^{\min\{t, V-T\}} \Phi^\theta(s) f(t-s) ds}{\int_0^{W-T-\psi(m)} \int_0^{\min\{u, V-T\}} \Phi^\theta(s) f(u-s) ds du}. \end{aligned}$$

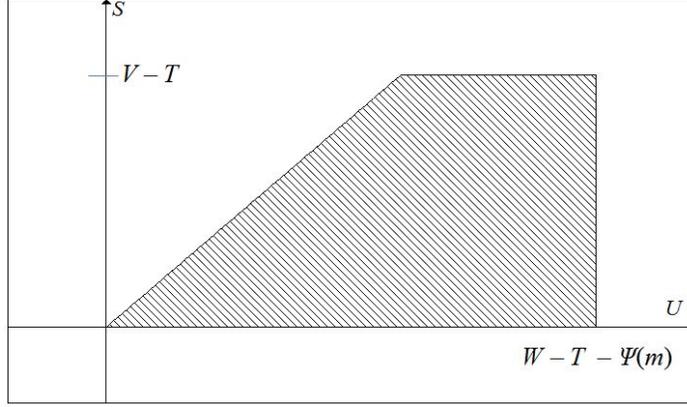


FIGURE 1. Case i)  $V - T < W - T - \Psi(m)$

In order to simplify the integral in the denominator, we notice that the domain of integration is given by one of two cases:

Case i)  $V - T < W - T - \psi(m)$ . In this case,  $s$  runs from 0 to  $\min\{u, V - T\}$  and  $u$  runs from 0 to  $W - T - \psi(m)$ , see Figure 1. Reversing the order of integration has  $u$  running from  $s$  to  $W - T - \psi(m)$  and  $s$  running from 0 to  $V - T$ . The integral becomes

$$\int_0^{V-T} \int_s^{W-T-\psi(m)} \Phi^\theta(s) f(u-s) du ds$$

Case ii)  $V - T \geq W - T - \psi(m)$ . In this case,  $s$  runs from 0 to  $u$  and  $u$  runs from 0 to  $W - T - \psi(m)$ , see Figure 2. Reversing the order of integration gives  $u$  running from  $s$  to  $W - T - \psi(m)$  while  $s$  runs from 0 to  $W - T - \psi(m)$ . The integral becomes

$$\int_0^{W-T-\psi(m)} \int_s^{W-T-\psi(m)} \Phi^\theta(s) f(u-s) du ds$$

The only difference in these cases is the upper bound on the outer integral, which is  $V - T$  if  $V - T \leq W - T - \psi(m)$  and  $W - T - \psi(m)$  if  $V - T > W - T - \psi(m)$ .

So the two cases can be represented together by

$$\int_0^{\min\{W-T-\psi(m), V-T\}} \int_s^{W-T-\psi(m)} \Phi^\theta(s) f(u-s) du ds.$$

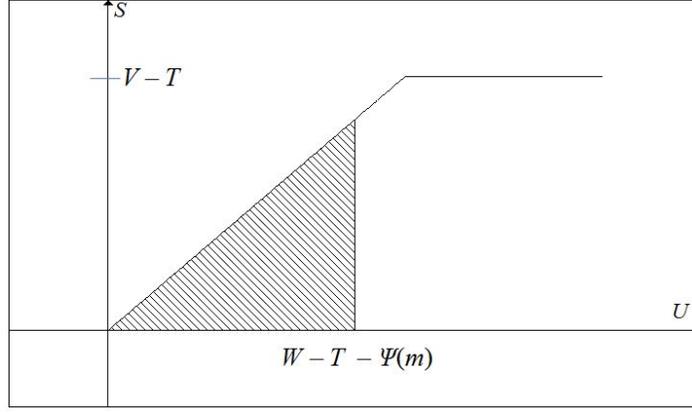


FIGURE 2. Case ii)  $V - T \geq W - T - \Psi(m)$

But this simplifies to

$$\begin{aligned} & \int_0^{\min\{W-T-\psi(m), V-T\}} \Phi^\theta(s) \int_s^{W-T-\psi(m)} f(u-s) du ds \\ &= \int_0^{\min\{W-T-\psi(m), V-T\}} \Phi^\theta(s) F(W-T-\psi(m)-s) ds. \end{aligned}$$

Thus we have

$$\omega(t) = \frac{\int_0^{\min\{t, V-T\}} \Phi^\theta(s) f(t-s) ds}{\int_0^{\min\{W-T-\psi(m), V-T\}} \Phi^\theta(s) F(W-T-\psi(m)-s) ds}, 0 \leq t \leq W-T-\psi(m)$$

As we have seen above, each inception time,  $t_i$ , can be linked to a corresponding metastasis size,  $x_i$ , by the transformation  $x_i = \Psi(W - T - t_i)$  and we can thus use the conditional distribution of the inception times to evaluate the conditional distribution of metastasis sizes  $X = (X_1, X_2, \dots, X_n)$  given that  $N = n$ .

We begin by choosing  $\delta > 0$  to be sufficiently small so that the intervals in the set  $\{[x_i - \delta, x_i + \delta]\}_{i=1}^n$  are disjoint. We then compute

$$\begin{aligned} & \Pr\{X_i \in [x_i - \delta, x_i + \delta], 1 \leq i \leq n | n = N\} \\ &= \Pr\{T_i \in [W - T - \psi(x_i + \delta), W - T - \psi(x_i - \delta)], 1 \leq i \leq n | N = n\} \\ &= n! \prod_{i=1}^n \int_{W-T-\psi(x_i+\delta)}^{W-T-\psi(x_i-\delta)} \omega(t) dt. \end{aligned}$$

As shown in (A.7) in Appendix A, dividing both sides by  $(2\delta)^n$  and taking the limit as  $\delta \rightarrow 0$  gives  $p_X(x_1, x_2, \dots, x_n | N = n)$  on the left. In order to evaluate the limit on the right, we first prove the following lemma.

**Lemma:** If  $\phi$  is an increasing differentiable function on an interval  $(c, d)$  containing  $x$ , and  $\omega$  is a continuous function on an interval  $[a, b]$  whose interior contains  $\phi(x)$ , then  $\lim_{\delta \rightarrow 0} \frac{1}{2\delta} \int_{\phi(x-\delta)}^{\phi(x+\delta)} \omega(t) dt = \omega(\phi(x))\phi'(x)$ .

**Proof:**

$$\begin{aligned} & \lim_{\delta \rightarrow 0} \frac{1}{2\delta} \int_{\phi(x-\delta)}^{\phi(x+\delta)} \omega(t) dt \\ &= \lim_{\delta \rightarrow 0} \frac{\int_a^{\phi(x+\delta)} \omega(t) dt - \int_a^{\phi(x)} \omega(t) dt + \int_a^{\phi(x)} \omega(t) dt + \int_a^{\phi(x-\delta)} \omega(t) dt}{2\delta} \\ &= \frac{1}{2} \lim_{\delta \rightarrow 0} \frac{\int_a^{\phi(x+\delta)} \omega(t) dt - \int_a^{\phi(x)} \omega(t) dt}{\delta} + \frac{1}{2} \lim_{\delta \rightarrow 0} \frac{\int_a^{\phi(x)} \omega(t) dt - \int_a^{\phi(x-\delta)} \omega(t) dt}{\delta} \\ &= \frac{1}{2} \frac{d}{dx} \int_a^{\phi(x)} \omega(t) dt + \frac{1}{2} \frac{d}{dx} \int_a^{\phi(x)} \omega(t) dt = \frac{d}{dx} \int_a^{\phi(x)} \omega(t) dt. \end{aligned}$$

By the Fundamental Theorem of Calculus,  $\frac{d}{dx} \int_a^x \omega(t) dt = \omega(x)$  for every  $x \in (a, b)$ . Notice that the function  $\int_a^{\phi(x)} \omega(t) dt$  is the composition of two differentiable functions, so that by the chain rule,  $\frac{d}{dx} \int_a^{\phi(x)} \omega(t) dt = \omega(\phi(x))\phi'(x)$  which completes the proof of the Lemma.

Applying the previous Lemma, we see that

$$\lim_{\delta \rightarrow 0} \frac{\int_{W-T-\psi(x_i+\delta)}^{W-T-\psi(x_i-\delta)} \omega(t) dt}{2\delta} = \omega(W-T-\psi(x_i))\psi'(x_i)$$

and therefore

$$\lim_{\delta \rightarrow 0} \frac{1}{2\delta} n! \prod_{i=1}^n \int_{W-T-\psi(x_i+\delta)}^{W-T-\psi(x_i-\delta)} \omega(t) dt = n! \prod_{i=1}^n \omega(W-T-\psi(x_i))\psi'(x_i)$$

Thus we have  $p_X(x_1, x_2, \dots, x_n | N = n) = n! \prod_{i=1}^n p(x_i)$  where

$p(x) = \omega(W-T-\psi(x))\psi'(x)$  and this concludes the proof of the theorem.

**Remark 1.** If the duration of the metastasis latency time is negligible, then all inceptions must occur before resection of the primary tumor (i.e.  $t \leq V - T$ ) and we have  $f(t - s) \approx \delta(t - s)$  and  $F(s) = 1$  for  $s > 0$ . In this case, the integral  $\int_0^{\min\{t, V-T\}} \Phi^\theta(s) f(t-s) ds$  becomes  $\int_0^{\min\{t, V-T\}} \Phi^\theta(s) \delta(t-s) ds = \Phi^\theta(t)$  and the integral  $\int_0^{\min\{W-T-\psi(m), V-T\}} \Phi^\theta(s) F(W-T-\psi(m)-s) ds$  becomes  $\int_0^{\min\{W-T-\psi(m), V-T\}} \Phi^\theta(s) ds$  so that

$$(2.6) \quad \omega(t) = \frac{\Phi^\theta(t)}{\int_0^{\min\{W-T-\psi(m), V-T\}} \Phi^\theta(s) ds}, \quad 0 \leq t \leq \min\{W - T - \psi(m), V - T\}$$

**Remark 2.** The parameter  $q\alpha$  plays an important role in determining how many metastases will be present at a given time. As a direct consequence of modeling the inception of tumors as a Poisson process, the distribution of the number of metastases in a given site that are detectable at time  $W$  is Poisson with parameter

$$\begin{aligned} \int_0^{W-T-\psi(m)} \lambda(u) du &= \int_0^{W-T-\psi(m)} \int_0^{\min\{u, V-T\}} q\alpha \Phi^\theta(s) f(u-s) ds du \\ &= q\alpha \int_0^{\min\{W-T-\psi(m), V-T\}} \Phi^\theta(s) F(W-T-\psi(m)-s) ds. \end{aligned}$$

However, the pdf  $p(x) = \omega(W - T - \psi(x)) \psi'(x)$  is independent of the number of metastases,  $n$ , as well as the intensity of metastasis seeding determined by the parameter  $q\alpha$ . This is due to the fact that our distribution was conditional on the number of metastases. Thus, the knowledge of the volumes of metastases detected in a given site does not enable inference on the distribution of their number.

**Remark 3.** Maximum likelihood techniques can be used to estimate parameters of a distribution when a random sample of size  $n$  is drawn from the distribution. The likelihood of a particular set of observations is the product of the pdf evaluated at each of the  $n$  observations. Although the observed volumes of detectable metastases  $x_1, x_2, \dots, x_n$  do not form a random sample from a probability distribution, the above Theorem makes it possible to apply the method of maximum likelihood for estimation

of model parameters. In fact, with the exception of the factor  $n!$ , the conditional pdf of secondary tumor sizes  $p_X(x_1, x_2, \dots, x_n | N = n) = n! \prod_{i=1}^n p(x_i)$ ,  $x_1 < x_2 < \dots < x_n$ , has the same form as the likelihood of the corresponding random sample from pdf  $p$ , which suggests using  $L(X_1, X_2, \dots, X_n) = n! \prod_{i=1}^n p(X_i)$  to estimate the identifiable parameters of the model. The same is true for any other rearrangement-invariant statistic.

### 2.3. The distribution of the sizes of detectable metastases for exponentially growing tumors with exponentially distributed latency times

In order to fit our model to data, we must assume some parametric form for the growth of the primary tumor and metastases as well as for the distribution of tumor latency times. The simplest model of tumor growth is exponential; so, we will assume that the growth laws  $\Phi$ ,  $\Psi_0$ , and  $\Psi_1$  are exponential, namely  $\Phi(t) = e^{\beta t}$ ,  $\Psi_0(t) = e^{\gamma_0 t}$  and  $\Psi_1(t) = e^{\gamma_1 t}$ . If we assume that the event of tumor inception is purely random with hazard rate  $1/\rho$ , then tumor latency times will follow the exponential distribution  $f(s) = \frac{1}{\rho} e^{-s/\rho}$  where  $\rho$  is the mean latency time. We compute

$$\begin{aligned} \int_0^{\min\{t, V-T\}} \Phi^\theta(s) f(t-s) ds &= \int_0^{\min\{t, V-T\}} \frac{1}{\rho} e^{\theta\beta s} e^{-(t-s)/\rho} ds \\ &= \frac{1}{\rho} e^{-t/\rho} \int_0^{\min\{t, V-T\}} e^{(\theta\beta+1/\rho)s} ds = \frac{1}{\theta\beta\rho+1} e^{-t/\rho} [e^{(\theta\beta+1/\rho)\min\{t, V-T\}} - 1] \end{aligned}$$

and thus

$$\omega(t) = \frac{\int_0^{\min\{t, V-T\}} \Phi^\theta(s) f(t-s) ds}{\int_0^{\min\{W-T-\psi(m), V-T\}} \Phi^\theta(s) F(W-T-\psi(m)-s) ds}, 0 \leq t \leq W-T-\psi(m)$$

becomes

$$(2.7) \quad \omega(t) = \frac{1}{C} e^{-t/\rho} [e^{(\theta\beta+1/\rho)\min\{t, V-T\}} - 1], 0 \leq t \leq W-T-\psi(m)$$

where

$$(2.8) \quad C = (\theta\beta\rho + 1) \int_0^{\min\{W-T-\psi(m), V-T\}} \Phi^\theta(s) F(W-T-\psi(m)-s) ds.$$

The inverse of  $\Phi(t) = e^{\beta t}$  is  $\phi(x) = \log(x)/\beta$ . From this, if the primary tumor is of size  $X$  at time  $U$ , then the onset of disease must have occurred  $\log(X)/\beta$  units of time before  $U$ . We can then determine that the onset of the disease occurred at time  $T = U - \log(X)/\beta$ , where clearly we must have  $T \geq 0$ .

The inverses of  $\Psi_0(t) = e^{\gamma_0 t}$  and  $\Psi_1(t) = e^{\gamma_1 t}$  are  $\Psi_0^{-1}(t) = \gamma_0^{-1} \log t$  and  $\Psi_1^{-1}(t) = \gamma_1^{-1} \log t$  and using these in (2.3) gives

$$\psi(x) = \begin{cases} \gamma_1^{-1} \log x & \text{if } 1 \leq x \leq e^{\gamma_1(W-V)}, \\ \gamma_0^{-1} \log [xe^{-\gamma_1(W-V)}] + W - V & \text{if } e^{\gamma_1(W-V)} < x \leq e^{\gamma_0(V-T)+\gamma_1(W-V)}. \end{cases}$$

Simplifying the second part of this function gives

$$(2.9) \quad \psi(x) = \begin{cases} \gamma_1^{-1} \log x & \text{if } 1 \leq x \leq e^{\gamma_1(W-V)}, \\ \gamma_0^{-1} \log x + \left(1 - \frac{\gamma_1}{\gamma_0}\right) (W - V) & \text{if } e^{\gamma_1(W-V)} < x \leq e^{\gamma_0(V-T)+\gamma_1(W-V)}. \end{cases}$$

Before proceeding further, we simplify by using the substitutions

$$(2.10) \quad \sigma = \theta\beta$$

$$(2.11) \quad Q = V - T \text{ and}$$

$$(2.12) \quad R = W - V.$$

Note that  $Q$  is the primary tumor progression time (of quiescence) from onset to resection, and  $R$  is the time from resection to metastasis survey. With these, (2.9) becomes

$$\psi(x) = \begin{cases} \gamma_1^{-1} \log x, & 1 \leq x \leq e^{\gamma_1 R}, \\ \gamma_0^{-1} \log x + \left(1 - \frac{\gamma_1}{\gamma_0}\right) R, & e^{\gamma_1 R} < x \leq e^{\gamma_0 Q + \gamma_1 R} \end{cases}$$

and using this expression for  $\psi$ , we compute

$$(2.13) \quad \psi'(x) = \begin{cases} (\gamma_1 x)^{-1}, & 1 \leq x \leq e^{\gamma_1 R}, \\ (\gamma_0 x)^{-1} & e^{\gamma_1 R} < x \leq e^{\gamma_0 Q + \gamma_1 R} \end{cases}$$

$$(2.14) \quad W - T - \psi(x) = Q + R - \psi(x) = \begin{cases} Q + R - \gamma_1^{-1} \log x, & 1 \leq x \leq e^{\gamma_1 R}, \\ Q - \gamma_0^{-1} \log x + \frac{\gamma_1}{\gamma_0} R, & e^{\gamma_1 R} < x \leq e^{\gamma_0 Q + \gamma_1 R}, \end{cases}$$

and

$$(2.15) \quad \min \{W - T - \psi(x), V - T\} \\ = \min \{Q + R - \psi(x), Q\} = \begin{cases} Q, & 1 \leq x \leq e^{\gamma_1 R}, \\ Q - \gamma_0^{-1} \log x + \frac{\gamma_1}{\gamma_0} R, & e^{\gamma_1 R} < x \leq e^{\gamma_0 Q + \gamma_1 R}. \end{cases}$$

Thus in the case when  $e^{\gamma_1 R} \leq m$ , we have

$$\omega(W - T - \psi(x)) \\ = \frac{1}{C} e^{-(Q - \gamma_0^{-1} \log x + \frac{\gamma_1}{\gamma_0} R)/\rho} \left[ e^{(\sigma+1/\rho)(Q - \gamma_0^{-1} \log x + \frac{\gamma_1}{\gamma_0} R)} - 1 \right], \quad m < x \leq e^{\gamma_0 Q + \gamma_1 R}$$

and in the case when  $e^{\gamma_1 R} > m$  we have

$$\omega(W - T - \psi(x)) \\ = \begin{cases} \frac{1}{C} e^{-(Q+R-\gamma_1^{-1} \log x)/\rho} \left[ e^{(\sigma+1/\rho)Q} - 1 \right], & m \leq x \leq e^{\gamma_1 R}, \\ \frac{1}{C} e^{-(Q-\gamma_0^{-1} \log x + \frac{\gamma_1}{\gamma_0} R)/\rho} \left[ e^{(\sigma+1/\rho)(Q-\gamma_0^{-1} \log x + \frac{\gamma_1}{\gamma_0} R)} - 1 \right], & e^{\gamma_1 Q} < x \leq e^{\gamma_0 Q + \gamma_1 R}. \end{cases}$$

We can now use  $p(x) = \omega(W - T - \psi(x))\psi'(x)$  to obtain

Case i)  $e^{\gamma_1 R} \leq m$

$$p(x) = (C_1 x)^{-1} e^{-(Q - \gamma_0^{-1} \log x + \frac{\gamma_1}{\gamma_0} R)/\rho} \left[ e^{(\sigma+1/\rho)(Q - \gamma_0^{-1} \log x + \frac{\gamma_1}{\gamma_0} R)} - 1 \right], \quad m < x \leq e^{\gamma_0 Q + \gamma_1 R}$$

which simplifies to

$$(2.16) \quad p(x) = (C_1 x)^{-1} \left[ \left( \frac{e^{\gamma_0 Q + \gamma_1 R}}{x} \right)^{\frac{\sigma}{\gamma_0}} - \left( \frac{x}{e^{\gamma_0 Q + \gamma_1 R}} \right)^{\frac{1}{\gamma_0 \rho}} \right], \quad m \leq x \leq e^{\gamma_0 Q + \gamma_1 R},$$

where the constant,  $C_1$ , is found from (2.8)

$$(2.17) \quad C_1 = \frac{\gamma_0}{\sigma} \left[ \left( \frac{e^{\gamma_0 Q + \gamma_1 R}}{m} \right)^{\frac{\sigma}{\gamma_0}} - 1 \right] + \gamma_0 \rho \left[ \left( \frac{m}{e^{\gamma_0 Q + \gamma_1 R}} \right)^{\frac{1}{\gamma_0 \rho}} - 1 \right]$$

Case ii)  $e^{\gamma_1 R} > m$

$$(2.18) \quad p(x) = \begin{cases} (C_2 x)^{-1} \frac{\gamma_0}{\gamma_1} \left( \frac{x}{e^{\gamma_1 R}} \right)^{\frac{1}{\gamma_1 \rho}} \left( e^{\sigma Q} - e^{-\frac{Q}{\rho}} \right), & m \leq x \leq e^{\gamma_1 R}, \\ (C_2 x)^{-1} \left[ \left( \frac{e^{\gamma_0 Q + \gamma_1 R}}{x} \right)^{\frac{\sigma}{\gamma_0}} - \left( \frac{x}{e^{\gamma_0 Q + \gamma_1 R}} \right)^{\frac{1}{\gamma_0 \rho}} \right], & e^{\gamma_1 R} < x \leq e^{\gamma_0 Q + \gamma_1 R} \end{cases}$$

where

$$(2.19) \quad C_2 = \gamma_0 \rho \left( e^{\sigma Q} - e^{-\frac{Q}{\rho}} \right) \left[ 1 - \left( \frac{m}{e^{\gamma_1 R}} \right)^{\frac{1}{\gamma_1 \rho}} \right] + \frac{\gamma_0}{\sigma} \left( e^{\sigma Q} - 1 \right) + \gamma_0 \rho \left( e^{-\frac{Q}{\rho}} - 1 \right)$$

We call the model in (2.16) and (2.18) the “full model.”

We can simplify further by making the substitutions

$$(2.20) \quad a = \frac{\sigma}{\gamma_0}, b_0 = \frac{1}{\gamma_0 \rho}, b_1 = \frac{1}{\gamma_1 \rho}, A = e^{\gamma_1 R}, \text{ and } M = e^{\gamma_0 Q + \gamma_1 R}.$$

Notice that  $A$  is the size that a metastasis incepted at the time of resection could grow to by the time of survey and that  $M$  is the size that a metastasis incepted at the time of disease onset could grow to by the time of survey, which is the maximum size of a metastasis allowed by the model. These allow us to write  $p(x)$  as

Case i)  $A \leq m$ ,

$$(2.21) \quad p(x) = (C_1 x)^{-1} \left[ \left( \frac{M}{x} \right)^a - \left( \frac{x}{M} \right)^{b_0} \right], \quad m \leq x \leq M,$$

where

$$C_1 = \frac{1}{a} \left[ \left( \frac{M}{m} \right)^a - 1 \right] + \frac{1}{b_0} \left[ \left( \frac{m}{M} \right)^{b_0} - 1 \right]$$

and

Case ii)  $A > m$ ,

$$(2.22) \quad p(x) = \begin{cases} (C_2 x)^{-1} \frac{b_1}{b_0} \left(\frac{x}{A}\right)^{b_1} \left[ \left(\frac{M}{A}\right)^a - \left(\frac{A}{M}\right)^{b_0} \right], & m \leq x \leq A, \\ (C_2 x)^{-1} \left[ \left(\frac{M}{x}\right)^a - \left(\frac{x}{M}\right)^{b_0} \right], & A < x \leq M \end{cases}$$

where

$$C_2 = \frac{1}{b_0} \left[ \left(\frac{M}{A}\right)^a - \left(\frac{A}{M}\right)^{b_0} \right] \left[ 1 - \left(\frac{m}{A}\right)^{b_1} \right] + \frac{1}{a} \left[ \left(\frac{M}{A}\right)^a - 1 \right] + \frac{1}{b_0} \left[ \left(\frac{A}{M}\right)^{b_0} - 1 \right].$$

#### 2.4. A Note on Parameters and Parameter Recovery

The original biological parameters of our model were  $\alpha_0$ ,  $\alpha_1$ ,  $\beta$ ,  $\gamma_0$ ,  $\gamma_1$ ,  $\rho$ ,  $\theta$ ,  $q$ , and  $T$ . Because of the conditional nature of the distribution of the sizes of metastases,  $\alpha_0$ ,  $\alpha_1$ , and  $q$  disappear (see Theorem 1 above.) This leaves the quantities  $\beta$ ,  $\gamma_0$ ,  $\gamma_1$ ,  $\rho$ ,  $\theta$ , and  $T$ . Further reduction of the set of model parameters depends on whether or not the size,  $S$ , of the primary tumor at resection is known. When  $S$  is not known,  $\theta$  and  $\beta$  always appear together in the model as the quantity  $\sigma = \theta\beta$  and cannot be separated. This reduces us to the five parameters  $\gamma_0$ ,  $\gamma_1$ ,  $\rho$ ,  $\sigma$ , and  $T$  (or  $\gamma_0$ ,  $\gamma_1$ ,  $\rho$ ,  $\sigma$ , and  $Q$  where  $Q = V - T$ .) When  $S$  is known, then  $\beta$  and  $T$  are related by

$$\beta = \frac{\log(S)}{V - T}$$

and this allows  $\theta$  to be separated from  $\beta$ . In this case we can still consider the parameters to be  $\sigma$  and  $T$  (or  $\sigma$  and  $Q$ ), but we can also use either  $\beta$  and  $\theta$  or  $\theta$  and  $T$  as the parameters. Thus in any case the model becomes five-parametric.

We call the parameters  $A$ ,  $M$ ,  $a$ ,  $b_0$ , and  $b_1$  as defined in (2.20) “simplifying parameters” because they tend to simplify the way the model is expressed. On the other hand, we call the parameters  $\gamma_0$ ,  $\gamma_1$ ,  $\rho$ ,  $\sigma$ , and  $Q$  “native parameters” because they are either original parameters (in the case of  $\gamma_0$ ,  $\gamma_1$ , and  $\rho$ ) or they retain a clear relationship to the original parameters (in the case of  $Q = V - T$  and  $\sigma = \theta\beta$ .) Writing the

model in terms of native parameters is helpful for relating changes in the parameters to their biomedical implications. For example, a decrease in a native parameter like  $\rho$  would indicate a decrease in the average latency of metastases whereas a decrease in a simplifying parameter like  $b_0$  might mean an increase in  $\gamma_0$  (metastasis growth rate prior to resection) or  $\rho$  (mean latency) or both. As shown in [22], parameters  $A$ ,  $M$ ,  $a$ ,  $b_0$ , and  $b_1$  are identifiable from the distribution  $p(x)$  of the sizes of metastases in a given secondary site. We have already mentioned that  $A$  and  $M$  do have a clear physical significance, but they are tied to  $a$ ,  $b_0$ , and  $b_1$  through  $\gamma_0$ ,  $\gamma_1$ , and  $Q$ .

The equations in (2.20) can be inverted to give the native parameters in terms of the simplifying parameters. First,

$$\gamma_1 = \frac{\log A}{R}$$

and this allows us to find

$$\rho = \frac{1}{b_1 \gamma_1} = \frac{R}{b_1 \log A}.$$

Once we have  $\gamma_1$  and  $\rho$ , we can obtain

$$\gamma_0 = \frac{1}{b_0 \rho} = \frac{b_1 \gamma_1}{b_0} = \frac{b_1 \log A}{b_0 R},$$

$$Q = \frac{1}{\gamma_0} \log \frac{M}{A} = b_0 \rho \log \frac{M}{A} = \frac{b_0 R \log(M/A)}{b_1 \log A},$$

and

$$\sigma = \theta \beta = a \gamma_0 = \frac{a b_1 \log A}{b_0 R}.$$

If the size of the primary tumor at resection is known, then  $S = e^{\beta Q}$  allows us to find

$$\beta = \frac{\log S}{Q} = \frac{\log S}{\frac{b_0 R \log(M/A)}{b_1 \log A}} = \frac{b_1 \log A \log S}{b_0 R \log(M/A)},$$

which, in turn, allows us to find

$$\theta = \frac{\sigma}{\beta} = \frac{a b_1 \log A}{b_0 R \beta} = \frac{a b_1 Q \log A}{b_0 R \log S} = \frac{a \log(M/A)}{\log S}.$$

## 2.5. Model Analysis

When  $b_0 \neq b_1$  (i.e. when the growth rates  $\gamma_0$  and  $\gamma_1$  are unequal) the equation for  $p$  in (2.22) has a discontinuity at  $x = A$  where the pdf jumps from

$(C_2)^{-1} \frac{b_1}{b_0} \left(\frac{A}{M}\right) \left[\left(\frac{M}{A}\right)^a - \left(\frac{A}{M}\right)^{b_0}\right]$  to  $(C_2)^{-1} \left(\frac{A}{M}\right) \left[\left(\frac{M}{A}\right)^a - \left(\frac{A}{M}\right)^{b_0}\right]$ . Because of this discontinuity, we can consider the value of  $p$  at  $x = A$  to be either

$p(A) = (C_2)^{-1} \frac{b_1}{b_0} \left(\frac{A}{M}\right) \left[\left(\frac{M}{A}\right)^a - \left(\frac{A}{M}\right)^{b_0}\right]$  (particularly if  $x$  is approaching  $A$  from the left) or to be

$p(A) = (C_2)^{-1} \left(\frac{A}{M}\right) \left[\left(\frac{M}{A}\right)^a - \left(\frac{A}{M}\right)^{b_0}\right]$  (particularly if  $x$  is approaching  $A$  from right.)

For  $m \leq x < A$ , we compute that

$$\frac{dp}{dx} = (C_2)^{-1} (b_1 - 1) x^{b_1-2} \frac{b_1}{b_0} \left(\frac{1}{A}\right)^{b_1} \left[\left(\frac{M}{A}\right)^a - \left(\frac{A}{M}\right)^{b_0}\right]$$

and therefore  $p$  is increasing when  $b_1 > 1$ , decreasing when  $b_1 < 1$ , and constant when  $b_1 = 1$ . This corresponds to an increasing function of  $x$  when  $\rho < 1/\gamma_1$  and decreasing when  $\rho > 1/\gamma_1$ .

Observe also that for  $A < x \leq M$ ,  $p(x) = (C_2 x)^{-1} \left[\left(\frac{M}{x}\right)^a - \left(\frac{x}{M}\right)^{b_0}\right]$  is a decreasing function of  $x$  with  $p(M) = 0$ .

## CHAPTER 3

### Submodels and Limiting Cases

In considering submodels of the Full model (2.16) and (2.18), it is helpful to consider the model as having come from two stages: a nonhomogeneous Poisson process whose events are the times of metastasis inception and a deterministic growth model that converts inception times into observed sizes at the time of metastasis survey. When written in native parameters, the Full model involves the five parameters  $Q$ ,  $\sigma$ ,  $\rho$ ,  $\gamma_0$ , and  $\gamma_1$ . Three of them,  $Q$ ,  $\rho$ , and  $\sigma$ , pertain to the Poisson process while the remaining two,  $\gamma_0$  and  $\gamma_1$ , pertain to the deterministic growth model.

We now consider the bounds on these parameters and the models that result as these parameters approach and/or attain their bounds.

Because  $Q$  is the time elapsed between onset and resection, we have  $0 \leq Q \leq V$ . Note that  $Q = 0$  indicates onset occurred at the time of resection and  $Q = V$  indicates onset occurred at birth.

Because  $\rho$  is the mean latency time, it can take any value zero or greater. We call the case when  $\rho = 0$  the “Instantaneous Seeding” case because in this case each metastasis begins to grow immediately after being shed off the primary tumor. As  $\rho$  tends to  $\infty$ , the model takes on another form which we describe below and call the “Heavy-Seeding/Long-Latency” model.

The parameter  $\sigma$  relates to the intensity of metastasis shedding which, under the assumption of exponential growth of the primary tumor, is given by  $\mu(t) = \alpha e^{\theta\beta t} = \alpha e^{\sigma t}$  with  $\alpha$  not appearing in the Full model because it is a conditional distribution. The parameter  $\beta$  is the growth rate of the primary tumor and is bounded above because there is a practical limit on the rate of cell division. (If  $t_d$  is the shortest time in which a cell could complete the cell cycle, then  $\beta \leq \log(2)/t_d$ .) The parameter  $\theta$  is intended to

model the relationship between the size of the primary tumor and metastasis shedding, e.g. if  $\theta = 1$ , then shedding is proportional to primary tumor volume (which means that every cancer cell has metastatic potential), if  $\theta = 2/3$  shedding would be proportional to primary tumor surface area (which typically contains actively proliferating cells), and if  $\theta = 0$ , shedding intensity would be constant (which suggests the existence of a stable subpopulation of cancer cells with high metastatic potential, perhaps “cancer stem cells”.) There is, however, no theoretical upper bound on the value of  $\theta$ . It is true that the number of metastases shed in a particular time could not exceed the number of cells that existed or were produced in the primary tumor during that same time period, but this number could be controlled by adjusting  $\alpha$ , which we have noted does not appear in the conditional distribution. Hence,  $\sigma$  is only bounded by 0 and  $\infty$ . Biologically, it would be possible for the growth rate of the primary tumor to be 0 at times, but in our model, the growth rate,  $\beta$  is constant and therefore must be nonzero or no primary tumor would emerge. Hence the case when  $\sigma = 0$  occurs if and only if  $\theta = 0$ . If  $\theta = 0$ , then the shedding intensity remains constant while the primary tumor is in place. We call the case when  $\sigma = 0$  the “Homogeneous” case.

Like  $\beta$ , the growth rates  $\gamma_0$  and  $\gamma_1$  are bounded between 0 and a value  $\gamma_{\max}$  which represents the biological upper limit on the rate of cell division. Notice that  $\gamma_0$  and  $\gamma_1$  cannot both be zero if metastases are observed. A tumor that was incepted at the time of disease onset would grow to size  $M = e^{\gamma_0 Q + \gamma_1 R}$  by the time of metastasis survey. Thus if  $x_n$  is the size of the largest observed metastasis, then we must have  $x_n \leq M$  or  $\gamma_0 Q + \gamma_1 R \geq \log x_n$ . If  $\gamma_0 = 0$ , all metastasis growth is suppressed by the primary tumor and  $\gamma_1$  must be at least  $\log(x_n)/R$ . If  $\gamma_1 = 0$ , all metastasis growth is suppressed by resection of the primary tumor or its biological effects and  $\gamma_0$  must be at least  $\log(x_n)/Q$ . Although  $\gamma_0$  is bounded above by  $\gamma_{\max}$ , when we search for optimal values of the parameters by the method of maximum likelihood, we find that  $\gamma_0$  and  $\sigma$  will become very large while  $\rho$  approaches 0. For this reason we consider this limiting

Parameter and bounds	Model at lower bound	Model at upper bound
$\rho \geq 0$	$\rho = 0$ : Instantaneous Seeding model	$\rho \rightarrow \infty$ : Heavy-Seeding/ Long-Latency model
$0 \leq \sigma < \infty$	$\sigma = 0$ : Homogeneous model	$\sigma \rightarrow \infty$ : infinite shedding intensity
$0 \leq Q \leq V$	$Q = 0$ : onset at resection	$Q = V$ : onset at birth.
$0 \leq \gamma_0 \leq \gamma_{\max},$ $\gamma_0 Q + \gamma_1 R \geq \log(x_n)$	$\gamma_0 = 0$ : complete suppression by primary. This requires $\gamma_1 R \geq \log(x_n)$	$\gamma_0 = \theta\beta = \infty, \rho = 0$ : IISMG
$0 \leq \gamma_1 \leq \gamma_{\max},$ $\gamma_0 Q + \gamma_1 R \geq \log(x_n)$	$\gamma_1 = 0$ : complete suppression after resection. This requires $\gamma_0 Q \geq \log(x_n)$	$\gamma_1 = \gamma_{\max}$ : maximum growth after resection

TABLE 1. Summary of Submodels and Limiting Cases. The value  $x_n$  denotes the largest observed metastasis size (in number of cells.)

case when  $\gamma_0, \sigma \rightarrow \infty, \rho \rightarrow 0$  and call this the “Instantaneous Infinite Shedding and Metastasis Growth” (IISMG) model.

Table 1 sums up the cases. In the remainder of this chapter we derive the submodels and limiting cases that were necessary for our data analysis. For convenience, these models are summarized in Appendix B.

### 3.1. Submodels arising from alterations to the Poisson process

We can develop expressions for the submodels in effect at the bounds of the Poisson parameters  $Q, \sigma$ , and  $\rho$  by considering the metastasis inception intensity function. In the previous analysis, we developed the Full model by combining the metastasis inception intensity function

$$\begin{aligned}
\lambda(t) &= q\alpha \int_0^{\min\{t, V-T\}} e^{\theta\beta s} \frac{1}{\rho} e^{-(t-s)/\rho} ds \\
&= \frac{q\alpha}{\rho} e^{-t/\rho} \left[ \frac{1}{(\theta\beta + 1/\rho)} e^{(\theta\beta + 1/\rho)s} \right] \Bigg|_0^{\min\{t, V-T\}} \\
&= \frac{q\alpha e^{-t/\rho}}{\rho\theta\beta + 1} (e^{(\theta\beta + 1/\rho)\min\{t, V-T\}} - 1)
\end{aligned}$$

$$= \frac{q\alpha}{\rho\theta\beta + 1} \begin{cases} e^{\theta\beta t} - e^{-t/\rho}, & 0 \leq t \leq V - T \\ e^{-t/\rho} [e^{\theta\beta(V-T) + (1/\rho)(V-T)} - 1], & t > V - T \end{cases}$$

with the metastasis growth function  $\Psi$ . Examining  $\lambda$ , we see that after onset and until the primary tumor is removed, the intensity of metastasis inception grows proportionally with the size of the primary tumor, but with a reduction in rate due to metastases that are delayed by the random latency time. After resection, the intensity decays exponentially as metastases that were shed from the primary tumor complete their latency times.

**3.1.1. Instantaneous Seeding Model.** If we had not included the latency time delay, i.e. if  $\rho = 0$ , we would have had an intensity of

$$\lambda(t) = q\alpha \begin{cases} e^{\theta\beta t}, & 0 \leq t \leq V - T \\ 0, & t > V - T. \end{cases}$$

We call the model arising in this case the ‘‘Instantaneous Seeding’’ model because it assumes that once a metastasis is shed, it instantaneously begins to develop in a host site. While the primary tumor is in place, the intensity is proportional to a power of the size of the tumor, but drops to zero once the primary is removed because there are no potential metastases awaiting inception. We develop the conditional size distribution  $p(x)$  by computing

$$\begin{aligned} \omega(t) &= \frac{\lambda_{ins}(t)}{\int_0^{W-T-\psi(m)} \lambda_{ins}(s) ds} = \frac{e^{\theta\beta t}}{\int_0^{\min\{V-T, W-T-\psi(m)\}} e^{\theta\beta s} ds} \\ &= \frac{\theta\beta e^{\theta\beta t}}{e^{\theta\beta \min\{V-T, W-T-\psi(m)\}} - 1}, \quad 0 \leq t \leq \min\{V-T, W-T-\psi(m)\} \end{aligned}$$

Thus

$$\begin{aligned} \omega(W-T-\psi(x)) &= \frac{\theta\beta e^{\theta\beta(W-T-\psi(x))}}{e^{\theta\beta \min\{V-T, W-T-\psi(m)\}} - 1}, \\ &0 \leq W-T-\psi(x) \leq \min\{V-T, W-T-\psi(m)\} \end{aligned}$$

and we consider two cases.

Case i)  $e^{\gamma_1(W-V)} \leq m$

$$\begin{aligned}
p(x) &= \omega(W - T - \psi(x)) \psi'(x) \\
&= (\gamma_0 x)^{-1} \frac{\theta \beta e^{\theta \beta (V-T - \gamma_0^{-1} \log x + \frac{\gamma_1}{\gamma_0} (W-V))}}{e^{\theta \beta (V-T - \gamma_0^{-1} \log m + \frac{\gamma_1}{\gamma_0} (W-V))} - 1}, \quad m \leq x \leq e^{\gamma_0(V-T) + \gamma_1(W-V)} \\
&= (\gamma_0 x)^{-1} \frac{\theta \beta \left( \frac{e^{\gamma_0(V-T) + \gamma_1(W-V)}}{x} \right)^{\frac{\theta \beta}{\gamma_0}}}{\left( \frac{e^{\gamma_0(V-T) + \gamma_1(W-V)}}{m} \right)^{\frac{\theta \beta}{\gamma_0}} - 1}, \quad m \leq x \leq e^{\gamma_0(V-T) + \gamma_1(W-V)}
\end{aligned}$$

This simplifies to

$$(3.1) \quad p(x) = C_3 x^{-\frac{\sigma}{\gamma_0} - 1}, \quad m \leq x \leq e^{\gamma_0 Q + \gamma_1 R}$$

where

$$C_3 = \frac{\sigma / \gamma_0}{m^{-\frac{\sigma}{\gamma_0}} - (e^{\gamma_0 Q + \gamma_1 R})^{-\frac{\sigma}{\gamma_0}}}$$

in native parameters or to

$$(3.2) \quad p(x) = C_3 x^{-a-1}, \quad m \leq x \leq M$$

where

$$C_3 = \frac{a}{m^{-a} - M^{-a}}.$$

in simplifying parameters.

The same result as above can be obtained by taking the pointwise limit of  $p(x)$  as  $\rho \rightarrow 0$  in (2.16).

Case ii)  $e^{\gamma_1(W-V)} > m$

$$\begin{aligned}
p(x) &= \omega(W - T - \psi(x)) \psi'(x) \\
&= (\gamma_0 x)^{-1} \frac{\theta \beta \left( \frac{e^{\gamma_0(V-T) + \gamma_1(W-V)}}{x} \right)^{\frac{\theta \beta}{\gamma_0}}}{(e^{\gamma_0(V-T)})^{\frac{\theta \beta}{\gamma_0}} - 1}, \quad e^{\gamma_1(W-V)} \leq x \leq e^{\gamma_0(V-T) + \gamma_1(W-V)}
\end{aligned}$$

This simplifies to

$$(3.3) \quad p(x) = C_4 x^{-\frac{\sigma}{\gamma_0}-1}, \quad e^{\gamma_1 R} \leq x \leq e^{\gamma_0 Q + \gamma_1 R}$$

where

$$C_4 = \frac{\sigma/\gamma_0}{(e^{\gamma_1 R})^{-\frac{\sigma}{\gamma_0}} - (e^{\gamma_0 Q + \gamma_1 R})^{-\frac{\sigma}{\gamma_0}}}$$

in native parameters or to

$$(3.4) \quad p(x) = C_4 x^{-a-1}, \quad A \leq x \leq M$$

where

$$C_4 = \frac{a}{A^{-a} - M^{-a}}$$

in simplifying parameters.

Taking the pointwise limit of  $p(x)$  as  $\rho \rightarrow 0$  in (2.18) leads to a similar result that only differs at  $x = A$  and  $x = M$ . For the Full model  $p(M) = 0$ , so  $\lim_{\rho \rightarrow 0} p(M) = 0$  and if we were to evaluate  $p(A)$  using the value on the left hand of the discontinuity, we would have

$$\lim_{\rho \rightarrow 0} p(A) = \frac{b_1}{b_0} \frac{aA^{-a-1}}{A^{-a} - M^{-a}}$$

instead of

$$\lim_{\rho \rightarrow 0} p(A) = \frac{aA^{-a-1}}{A^{-a} - M^{-a}}.$$

The difference at one point is insignificant to the probability, but when the likelihood is computed, it depends on the value of the pdf at the specific data values. Thus, the limit as  $\rho \rightarrow 0$  of the likelihood obtained from the Full model may not approach the value of the likelihood obtained from the Instantaneous Seeding model.

**3.1.2. Homogeneous Model.** If we remove the dependence of the inception intensity on the size of the primary tumor but still include the latency time delay, we

would have an inception intensity of

$$\lambda(t) = q\alpha \int_0^{\min\{t, V-T\}} \frac{1}{\rho} e^{-(t-s)/\rho} ds = q\alpha \begin{cases} 1 - e^{-t/\rho}, & 0 \leq t \leq V - T \\ e^{(V-T-t)/\rho} - e^{-t/\rho}, & t > V - T \end{cases}$$

We call the model arising in this case the ‘‘Homogeneous’’ model because the shedding intensity is independent of the size of the primary tumor making the Poisson process of metastasis shedding homogeneous. Because of the delay due to latency, the intensity of inception is initially zero but rapidly climbs toward a constant level as metastases that were delayed complete their latency times. After resection, the intensity fades as the remaining metastases that were shed prior to resection complete their latency times.

As was done in the Instantaneous Seeding case, we can develop the conditional size distribution  $p(x)$  from  $\lambda(t)$  by computing  $\omega(t)$  and setting  $p(x) = \omega(W - T - \psi(x))\psi'(x)$ , but we omit the details because, in this case, we obtain the same result by simply taking the pointwise limit of the Full model in (2.16) or (2.18) as  $\sigma \rightarrow 0$ .

Case i)  $e^{\gamma_1(W-V)} \leq m$  (i.e.  $A \leq m$ )

In this case

$$p(x) = (C_5 x)^{-1} \left[ 1 - \left( \frac{x}{e^{\gamma_0 Q + \gamma_1 R}} \right)^{\frac{1}{\gamma_0 \rho}} \right], m \leq x \leq e^{\gamma_0 Q + \gamma_1 R}$$

where

$$C_5 = \gamma_0 Q + \gamma_1 R - \log m + \gamma_0 \rho \left[ \left( \frac{m}{e^{\gamma_0 Q + \gamma_1 R}} \right)^{\frac{1}{\gamma_0 \rho}} - 1 \right]$$

in native parameters. In simplifying parameters, this is

$$(3.5) \quad p(x) = (C_5 x)^{-1} \left[ 1 - \left( \frac{x}{M} \right)^{b_0} \right], m \leq x \leq M$$

where

$$C_5 = \log \frac{M}{m} + \frac{1}{b_0} \left[ \left( \frac{m}{M} \right)^{b_0} - 1 \right].$$

Case ii)  $e^{\gamma_1(W-V)} > m$  (i.e.  $A > m$ )

In native parameters we have

$$p(x) = \begin{cases} (C_6 x)^{-1} \frac{\gamma_0}{\gamma_1} \left[ 1 - (e^{-\gamma_0 Q})^{\frac{1}{\gamma_0 \rho}} \right] \left( \frac{x}{e^{\gamma_1 R}} \right)^{\frac{1}{\gamma_1 \rho}}, & m \leq x < e^{\gamma_1 R} \\ (C_6 x)^{-1} \left[ 1 - \left( \frac{x}{e^{\gamma_0 Q + \gamma_1 R}} \right)^{\frac{1}{\gamma_0 \rho}} \right], & e^{\gamma_1 R} \leq x \leq e^{\gamma_0 Q + \gamma_1 R} \end{cases}$$

where

$$C_6 = \gamma_0 Q - \gamma_0 \rho \left( 1 - e^{-\frac{Q}{\rho}} \right) \left( \frac{m}{e^{\gamma_1 R}} \right)^{\frac{1}{\gamma_1 \rho}}.$$

In simplifying parameters, this is

$$(3.6) \quad p(x) = \begin{cases} (C_6 x)^{-1} \frac{b_1}{b_0} \left[ 1 - \left( \frac{A}{M} \right)^{b_0} \right] \left( \frac{x}{A} \right)^{b_1}, & m \leq x < A \\ (C_6 x)^{-1} \left[ 1 - \left( \frac{x}{M} \right)^{b_0} \right], & A \leq x \leq M \end{cases}$$

where

$$C_6 = \log \frac{M}{A} - \frac{1}{b_0} \left[ 1 - \left( \frac{A}{M} \right)^{b_0} \right] \left( \frac{m}{A} \right)^{b_1}.$$

**3.1.3. Heavy-Seeding/Long-Latency Model.** When fitting our model to data, we occasionally observe that the best fit occurs as  $\rho \rightarrow \infty$ . Because  $\rho$  is the mean latency time, this would indicate that metastases are seeded but remain dormant. However, if we start from the inception intensity of the Full model and compute the limit as  $\rho \rightarrow \infty$ , we have

$$\begin{aligned} \lim_{\rho \rightarrow \infty} \lambda(t) &= \lim_{\rho \rightarrow \infty} \frac{q\alpha}{\rho\theta\beta + 1} \begin{cases} e^{\theta\beta t} - e^{-t/\rho}, & 0 \leq t \leq V - T \\ e^{\theta\beta(V-T) + (1/\rho)(V-T-t)} - e^{-t/\rho}, & t > V - T \end{cases} \\ &= \frac{q\alpha}{\infty + 1} \begin{cases} e^{\theta\beta t} - 1, & 0 \leq t \leq V - T \\ e^{\theta\beta(V-T)} - 1, & t > V - T \end{cases} = 0, \end{aligned}$$

which is as we would suspect because if metastases, on average, spend an infinite amount of time in dormancy, we are not likely see many metastases develop. In order to compensate for the drop in inception intensity due to increasing the latency time,

we could increase  $q\alpha$  in such a way that the ratio  $q\alpha/\rho$  approaches a constant,  $c > 0$ . Then we obtain the intensity function

$$\lambda(t) = \frac{c}{\theta\beta} \begin{cases} e^{\theta\beta t} - 1, & 0 \leq t \leq V - T \\ e^{\theta\beta(V-T)} - 1, & t > V - T \end{cases}.$$

We call this the ‘Heavy-Seeding/Long-Latency’ model. One biological interpretation of the intensity function in this limiting case is that as the primary tumor grows, it sends out more and more metastases. Some may begin to develop but most remain dormant and thus a large pool of dormant metastases develops. Once the primary tumor is removed, this large pool of potential metastases remains. Assuming that the number of metastases that transition from dormancy to active development is very small in comparison to the size of the pool of dormant metastases, the intensity would remain approximately constant, even after the primary tumor was removed. This is consistent with large  $q\alpha$  (very many metastases shed from the primary) and large  $\rho$  (most metastases remain dormant and do not develop.)

Again, we could develop the conditional size distribution  $p(x)$  from  $\lambda(t)$  by computing  $\omega(t)$  and setting  $p(x) = \omega(W - T - \psi(x))\psi'(x)$ , but we omit the details because, in this case, we obtain the same result by simply taking the pointwise limit of the Full model in (2.16) or (2.18) as  $\rho \rightarrow \infty$ .

$$\text{Case i) } e^{\gamma_1(W-V)} \leq m$$

In native parameters we have

$$(3.7) \quad p(x) = (C_7 x)^{-1} \left[ \left( \frac{e^{\gamma_0 Q + \gamma_1 R}}{x} \right)^{\frac{\sigma}{\gamma_0}} - 1 \right], \quad m \leq x \leq e^{\gamma_0 Q + \gamma_1 R}$$

where

$$C_7 = \frac{\gamma_0}{\sigma} \left[ \left( \frac{e^{\gamma_0 Q + \gamma_1 R}}{m} \right)^{\frac{\sigma}{\gamma_0}} - 1 \right] + \log m - \gamma_0 Q - \gamma_1 R$$

which is

$$p(x) = (C_7 x)^{-1} \left[ \left( \frac{M}{x} \right)^a - 1 \right], \quad m \leq x \leq M$$

where

$$C_7 = \frac{1}{a} \left[ \left( \frac{M}{m} \right)^a - 1 \right] + \log \left( \frac{m}{M} \right)$$

in simplifying parameters.

Case ii)  $e^{\gamma_1(W-V)} > m$

In native parameters we have

$$(3.8) \quad p(x) = \begin{cases} \frac{\gamma_0}{\gamma_1} (C_8 x)^{-1} (e^{\sigma Q} - 1) & m \leq x < e^{\gamma_1 R} \\ (C_8 x)^{-1} \left[ \left( \frac{e^{\gamma_0 Q + \gamma_1 R}}{x} \right)^{\frac{\sigma}{\gamma_0}} - 1 \right] & e^{\gamma_1 R} \leq x \leq e^{\gamma_0 Q + \gamma_1 R} \end{cases}$$

where

$$C_8 = -\gamma_0 Q + (e^{\sigma Q} - 1) \left( \frac{\gamma_0}{\sigma} + \gamma_0 R - \frac{\gamma_0}{\gamma_1} \log m \right)$$

which is

$$p(x) = \begin{cases} \frac{b_1}{b_0} (C_8 x)^{-1} \left[ \left( \frac{M}{A} \right)^a - 1 \right] & m \leq x < A \\ (C_8 x)^{-1} \left[ \left( \frac{M}{x} \right)^a - 1 \right] & A \leq x \leq M \end{cases}$$

where

$$C_8 = -\log \left( \frac{M}{A} \right) + \left[ \left( \frac{M}{A} \right)^a - 1 \right] \left[ \frac{1}{a} + \frac{b_1}{b_0} \log \left( \frac{A}{m} \right) \right]$$

in simplifying parameters.

Notice that in this second case  $b_1$  and  $b_0$  always appear together as the fraction  $\frac{b_1}{b_0}$  and thus  $\frac{b_1}{b_0} = \frac{\gamma_0}{\gamma_1}$  is a single parameter.

### 3.2. Submodels arising from alterations to the growth model

While  $Q$ ,  $\sigma$  and  $\rho$  control the seeding and inception of metastases, the parameters  $\gamma_0$  and  $\gamma_1$  control the growth of those metastases after inception. We find that two cases are of particular interest when using the method of maximum likelihood to find parameters from data: the case when  $\gamma_0 \rightarrow 0$  and the case when  $\gamma_0 \rightarrow \infty$  while  $\frac{\sigma}{\gamma_0}$  and  $\frac{1}{\rho\gamma_0}$  are held constant.

**3.2.1. Complete Suppression by the Primary Tumor.** The parameter  $\gamma_0$  controls the growth of metastases prior to resection of the primary tumor. If we allow

$\gamma_0$  to approach 0 in the Full model so that the growth rate is essentially zero, then metastases are shed and complete their latency, but are unable to grow further and remain as either a collection of very few isolated cells or perhaps as small clumps of cells at the threshold of vascularization until the primary tumor is removed. After resection, all of the tumors that have lain dormant begin to grow at the same time, producing a concentrated group of tumors at or near the maximum size,  $e^{\gamma_1 R}$ . If  $e^{\gamma_1 R} < m$ , then no metastases will be observable and hence the pdf (and the likelihood) will be 0. On the other hand, if  $e^{\gamma_1 R} > m$  then growth post resection is sufficient to produce detectable metastases from at least some of the metastases incepted after resection. If we factor out  $\gamma_0$  from the normalizing constant, we can re-write the Full model as:

$$p(x) = \begin{cases} \frac{1}{\gamma_1} \left( \tilde{C}_2 x \right)^{-1} \left( \frac{x}{e^{\gamma_1 R}} \right)^{1/(\rho\gamma_1)} \left( e^{\sigma Q} - e^{-\frac{Q}{\rho}} \right), & m \leq x < e^{\gamma_1 R} \\ \frac{1}{\gamma_0} \left( \tilde{C}_2 x \right)^{-1} \left\{ \left[ \frac{e^{(\gamma_0 Q + \gamma_1 R)}}{x} \right]^{\frac{\sigma}{\gamma_0}} - \left[ \frac{x}{e^{(\gamma_0 Q + \gamma_1 R)}} \right]^{\frac{1}{\rho\gamma_0}} \right\}, & e^{\gamma_1 R} \leq x \leq e^{\gamma_0 Q + \gamma_1 R} \end{cases}$$

where

$$\tilde{C}_2 = \rho \left( e^{\sigma Q} - e^{-\frac{Q}{\rho}} \right) \left[ 1 - \left( \frac{m}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)} \right] + \frac{1}{\sigma} (e^{\sigma Q} - 1) + \rho \left( e^{-\frac{Q}{\rho}} - 1 \right).$$

We see that the first component of  $p$  (for  $m \leq x < e^{\gamma_1 R}$ ) accounts for metastases incepted after resection (conditional on the total number of metastases) while the second component (for  $e^{\gamma_1 R} \leq x \leq e^{\gamma_0 Q + \gamma_1 R}$ ) accounts for metastases that were incepted before resection. Thus the fraction of the total number of observed metastases that were incepted before resection is

$$r = \int_{e^{\gamma_1 R}}^{e^{\gamma_1 R + \gamma_0 Q}} p(x) dx = \frac{\frac{1}{\sigma} (e^{\sigma Q} - 1) + \rho \left( e^{-\frac{Q}{\rho}} - 1 \right)}{\rho \left( e^{\sigma Q} - e^{-\frac{Q}{\rho}} \right) \left( 1 - \left( \frac{m}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)} \right) + \frac{1}{\sigma} (e^{\sigma Q} - 1) + \rho \left( e^{-\frac{Q}{\rho}} - 1 \right)}$$

which we note is independent of  $\gamma_0$ . The growth rate  $\gamma_1$  appears in the denominator because it determines how many metastases incepted after resection will grow to detectable size.

As  $\gamma_0 \rightarrow 0$ , the Full model converges pointwise to

$$(3.9) \quad p(x) = \begin{cases} \frac{1}{\gamma_1} (C_9 x)^{-1} \left(\frac{x}{e^{\gamma_1 R}}\right)^{1/(\rho \gamma_1)} \left(e^{\sigma Q} - e^{-\frac{Q}{\rho}}\right), & m \leq x < e^{\gamma_1 R} \\ \infty, & x = e^{\gamma_1 R} \end{cases}$$

where

$$C_9 = \rho \left(e^{\sigma Q} - e^{-\frac{Q}{\rho}}\right) \left[1 - \left(\frac{m}{e^{\gamma_1 R}}\right)^{1/(\gamma_1 \rho)}\right] + \frac{1}{\sigma} (e^{\sigma Q} - 1) + \rho \left(e^{-\frac{Q}{\rho}} - 1\right).$$

If we use the maximum likelihood method with the Full model, then we see that we will have infinite likelihood if we set  $e^{\gamma_1 R} = x_n$  and let  $\gamma_0 \rightarrow 0$ . But as  $\gamma_0 \rightarrow 0$ , the fraction of tumors in the size range  $[e^{\gamma_1 R}, e^{\gamma_1 R + \gamma_0 Q}]$  remains the same even as the width of the interval shrinks to zero, so it is more instructive to think of the distribution of the sizes of detectable metastases as a combination of a continuous component for the metastases incepted after resection and a discrete component for the metastases incepted prior to resection:

$$(3.10) \quad p(x) = \frac{1}{\gamma_1} (C_9 x)^{-1} \left(\frac{x}{e^{\gamma_1 R}}\right)^{1/(\rho \gamma_1)} \left(e^{\sigma Q} - e^{-\frac{Q}{\rho}}\right), \quad m \leq x < e^{\gamma_1 R}$$

and

$$\Pr(x = e^{\gamma_1 R}) = (C_9)^{-1} \left[\frac{1}{\sigma} (e^{\sigma Q} - 1) + \rho \left(e^{-\frac{Q}{\rho}} - 1\right)\right]$$

In this case, the cdf would be smoothly increasing in  $x$  until it jumps up to 1 at  $x = e^{\gamma_1 R}$ . In (3.10), if we set  $v = \Pr(x = e^{\gamma_1 R})$  then we can solve for  $C_9$  to obtain

$$C_9 = \frac{1}{v} \left[\frac{1}{\sigma} (e^{\sigma Q} - 1) + \rho \left(e^{-\frac{Q}{\rho}} - 1\right)\right].$$

Equating this to the original expression for  $C_9$ , we have

$$\begin{aligned} & \frac{1}{v} \left[\frac{1}{\sigma} (e^{\sigma Q} - 1) + \rho \left(e^{-\frac{Q}{\rho}} - 1\right)\right] \\ &= \rho \left(e^{\sigma Q} - e^{-\frac{Q}{\rho}}\right) \left[1 - \left(\frac{m}{e^{\gamma_1 R}}\right)^{1/(\gamma_1 \rho)}\right] + \frac{1}{\sigma} (e^{\sigma Q} - 1) + \rho \left(e^{-\frac{Q}{\rho}} - 1\right) \end{aligned}$$

which we rearrange to obtain

$$\frac{1}{\sigma} (e^{\sigma Q} - 1) + \rho \left( e^{-\frac{Q}{\rho}} - 1 \right) = \frac{v}{1-v} \rho \left( e^{\sigma Q} - e^{-\frac{Q}{\rho}} \right) \left( 1 - \left( \frac{m}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)} \right).$$

This then leads to an alternative form of  $C_9$  as

$$C_9 = \frac{1}{1-v} \rho \left( e^{\sigma Q} - e^{-\frac{Q}{\rho}} \right) \left[ 1 - \left( \frac{m}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)} \right].$$

When we substitute this expression for  $C_9$  into the continuous part of the pdf we obtain

$$\begin{aligned} p(x) &= \frac{x^{-1} \left( \frac{x}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)} \left( e^{\sigma Q} - e^{-\frac{Q}{\rho}} \right)}{\gamma_1 \left( \frac{1}{1-v} \right) \rho \left( e^{\sigma Q} - e^{-\frac{Q}{\rho}} \right) \left[ 1 - \left( \frac{m}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)} \right]} \\ &= \frac{(1-v) x^{-1} \left( \frac{x}{e^{\gamma_1 R}} \right)^{1/(\rho \gamma_1)}}{\gamma_1 \rho \left[ 1 - \left( \frac{m}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)} \right]}, \quad m \leq x < e^{\gamma_1 R}. \end{aligned}$$

Thus the CSPT model can be written as

$$(3.11) \quad p(x) = \frac{(1-v) x^{-1} \left( \frac{x}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)}}{\gamma_1 \rho \left[ 1 - \left( \frac{m}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)} \right]}, \quad m \leq x < e^{\gamma_1 R},$$

$$\Pr(x = e^{\gamma_1 R}) = v$$

where

$$(3.12) \quad v = \frac{\frac{1}{\sigma} (e^{\sigma Q} - 1) + \rho \left( e^{-\frac{Q}{\rho}} - 1 \right)}{\rho \left( e^{\sigma Q} - e^{-\frac{Q}{\rho}} \right) \left[ 1 - \left( \frac{m}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)} \right] + \frac{1}{\sigma} (e^{\sigma Q} - 1) + \rho \left( e^{-\frac{Q}{\rho}} - 1 \right)}.$$

In Appendix C, we show that given  $\gamma_1$ ,  $\rho$ , and  $v$ , (3.12) can be solved for  $Q$  in terms of  $\sigma$  and has a unique positive solution in terms of  $\sigma$  if and only if

$$0 \leq \sigma < \frac{1-v}{\rho v \left[ 1 - \left( \frac{m}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)} \right]}.$$

When it exists, that solution satisfies

$$(3.13) \quad \frac{e^{\sigma Q} - 1}{e^{-\frac{Q}{\rho}} - 1} = \frac{\sigma \rho \left[ 1 - v \left( \frac{m}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)} \right]}{\sigma \rho v \left[ 1 - \left( \frac{m}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)} \right] + v - 1}.$$

The five-parameter Full model has degenerated into a three-parameter mixed model involving  $\gamma_1$ ,  $\rho$ , and  $v$ , where  $v$  can be used to specify a relationship between  $\sigma$ ,  $Q$ ,  $\gamma_1$  and  $\rho$ . We note that  $v$  represents the proportion of observed tumors that were incepted prior to resection; once values of  $\gamma_1$  and  $\rho$  are determined, if  $v$  is high, it may be that disease onset occurred very early (large  $Q$ ) or that metastasis shedding was particularly intense (large  $\sigma$ ) or some combination of the two.

We can write this model in simplifying parameters  $A$ ,  $b_1$ , and  $v$  as

$$p(x) = \frac{b_1 (1 - v) x^{-1} \left( \frac{x}{A} \right)^{b_1}}{1 - \left( \frac{m}{A} \right)^{b_1}}, m \leq x < A, \text{ and } \Pr(x = A) = v$$

Once  $b_1$ ,  $v$ , and  $A$  are found, then  $\rho$  and  $\gamma_1$  can be recovered and the relationship between  $\sigma$  and  $Q$  is specified by (3.13).

**3.2.2. Instantaneous Infinite Shedding and Metastasis Growth.** We also find that the likelihood becomes unbounded when we let  $\rho$  approach 0 while letting  $\sigma, \gamma_0 \rightarrow \infty$ . In this case, we have near instantaneous seeding with very large primary growth rate and very large metastasis growth rate prior to resection. A large value of  $\sigma$  means that the largest number of metastases (relative to the total) is shed just before resection and with  $\rho$  approaching 0, most metastases complete their latencies either just before or just after resection. Because the metastasis growth rate is so large prior to resection, metastases that were seeded before resection grow very large, whereas metastases that complete their latency after resection grow at a much slower rate. This concentration of metastases near the size  $e^{\gamma_1 R}$  is what causes the likelihood to be unbounded. Before finding the limiting pdf in the case  $A = e^{\gamma_1 R \geq m}$ , we adjust the inequality for  $x$  in the Full model so that the value of  $p(x)$  at  $x = e^{\gamma_1 R}$  is defined

by the left-hand equation:

$$p(x) = \begin{cases} \frac{1}{\gamma_1} (C_2 x)^{-1} \left(\frac{x}{e^{\gamma_1 R}}\right)^{1/(\rho \gamma_1)} \left(e^{\sigma Q} - e^{-\frac{Q}{\rho}}\right), & m \leq x \leq e^{\gamma_1 R} \\ \frac{1}{\gamma_0} (C_2 x)^{-1} \left[ \left(\frac{e^{(\gamma_0 Q + \gamma_1 R)}}{x}\right)^{\frac{\sigma}{\gamma_0}} - \left(\frac{x}{e^{(\gamma_0 Q + \gamma_1 R)}}\right)^{\frac{1}{\rho \gamma_0}} \right], & e^{\gamma_1 R} < x \leq e^{\gamma_0 Q + \gamma_1 R} \end{cases}$$

Next, we make the substitutions  $\sigma = a\gamma_0$  and  $\rho = 1/(b_0\gamma_0)$

$$p(x) = \begin{cases} \frac{(\gamma_1 x)^{-1} \left(\frac{x}{e^{\gamma_1 R}}\right)^{b_0 \gamma_0 / \gamma_1} (e^{a \gamma_0 Q} - e^{-b_0 \gamma_0 Q})}{\frac{1}{b_0 \gamma_0} (e^{a \gamma_0 Q} - e^{-b_0 \gamma_0 Q}) \left(1 - \left(\frac{m}{e^{\gamma_1 R}}\right)^{b_0 \gamma_0 / \gamma_1}\right) + \frac{1}{a \gamma_0} (e^{a \gamma_0 Q} - 1) + \frac{1}{b_0 \gamma_0} (e^{-b_0 \gamma_0 Q} - 1)}, & m \leq x \leq e^{\gamma_1 R} \\ \frac{(\gamma_0 x)^{-1} \left(\frac{e^{(\gamma_0 Q + \gamma_1 R)}}{x}\right)^a - \left(\frac{x}{e^{(\gamma_0 Q + \gamma_1 R)}}\right)^{b_0 \gamma_0 / \gamma_1}}{\frac{1}{b_0 \gamma_0} (e^{a \gamma_0 Q} - e^{-b_0 \gamma_0 Q}) \left(1 - \left(\frac{m}{e^{\gamma_1 R}}\right)^{b_0 \gamma_0 / \gamma_1}\right) + \frac{1}{a \gamma_0} (e^{a \gamma_0 Q} - 1) + \frac{1}{b_0 \gamma_0} (e^{-b_0 \gamma_0 Q} - 1)}, & e^{\gamma_1 R} < x \leq e^{\gamma_0 Q + \gamma_1 R} \end{cases}$$

Next we consider pointwise convergence of the model by first simplifying

$$p(x) = \begin{cases} \frac{(\gamma_1 x)^{-1} \gamma_0 \left(\frac{x}{e^{\gamma_1 R}}\right)^{b_0 \gamma_0 / \gamma_1} (1 - e^{-a \gamma_0 Q - b_0 \gamma_0 Q})}{\frac{1}{b_0} (1 - e^{-a \gamma_0 Q - b_0 \gamma_0 Q}) \left(1 - \left(\frac{m}{e^{\gamma_1 R}}\right)^{b_0 \gamma_0 / \gamma_1}\right) + \frac{1}{a} (1 - e^{-a \gamma_0 Q}) + \frac{1}{b_0} (e^{-a \gamma_0 Q - b_0 \gamma_0 Q} - e^{-a \gamma_0 Q})}, & m \leq x < e^{\gamma_1 R} \\ \frac{(\gamma_1 x)^{-1} \gamma_0 (1 - e^{-a \gamma_0 Q - b_0 \gamma_0 Q})}{\frac{1}{b_0} (1 - e^{-a \gamma_0 Q - b_0 \gamma_0 Q}) \left(1 - \left(\frac{m}{e^{\gamma_1 R}}\right)^{b_0 \gamma_0 / \gamma_1}\right) + \frac{1}{a} (1 - e^{-a \gamma_0 Q}) + \frac{1}{b_0} (e^{-a \gamma_0 Q - b_0 \gamma_0 Q} - e^{-a \gamma_0 Q})}, & x = e^{\gamma_1 R} \\ \frac{(x)^{-1} \left(\frac{e^{\gamma_1 R}}{x}\right)^a - e^{-\gamma_0 Q a} \left(\frac{x}{e^{(\gamma_0 Q + \gamma_1 R)}}\right)^{b_0 \gamma_0 / \gamma_1}}{\frac{1}{b_0} (1 - e^{-a \gamma_0 Q - b_0 \gamma_0 Q}) \left(1 - \left(\frac{m}{e^{\gamma_1 R}}\right)^{b_0 \gamma_0 / \gamma_1}\right) + \frac{1}{a} (1 - e^{-a \gamma_0 Q}) + \frac{1}{b_0} (e^{-a \gamma_0 Q - b_0 \gamma_0 Q} - e^{-a \gamma_0 Q})}, & e^{\gamma_1 R} < x \leq e^{\gamma_0 Q + \gamma_1 R} \\ 0, & x = e^{\gamma_0 Q + \gamma_1 R} \end{cases}$$

and then taking the limit as  $\gamma_0 \rightarrow \infty$  (assuming that  $a, \gamma_1$ , and  $b_0$  are fixed):

$$(3.14) \quad p(x) = \begin{cases} 0, & m \leq x < e^{\gamma_1 R} \\ \infty, & x = e^{\gamma_1 R} \\ x^{-1} \frac{ab_0 \left(\frac{e^{\gamma_1 R}}{x}\right)^a}{a + b_0}, & e^{\gamma_1 R} < x < \infty \end{cases}.$$

If we write this in simplifying parameters, we have

$$p(x) = \begin{cases} 0 & m \leq x < A \\ \infty & x = A \\ \frac{b_0}{a+b_0} \frac{a}{A^{-a}} x^{-a-1} & A < x < \infty \end{cases}$$

We recognize this, aside from the infinite value at  $x = A$  and the attendant change in the normalizing constant, as the Instantaneous Seeding model (3.1) with  $M \rightarrow \infty$ .

Note that this distribution is improper in that

$$\int_A^\infty p(x) dx = \frac{b_0}{a+b_0} < 1.$$

Although the pdf is converging pointwise to zero for  $x$  between  $m$  and  $e^{\gamma_1 R}$  and to infinity for  $x = e^{\gamma_1 R}$ , we can show that the (conditional) probability of an observable metastasis between  $m$  and  $e^{\gamma_1 R}$  approaches a constant value. We begin by noting that in the Full model, the fraction of the total number of observed metastases that were incepted after resection is

$$\int_m^{e^{\gamma_1 R}} p(x) dx = \frac{\rho \left( e^{\sigma Q} - e^{-\frac{Q}{\rho}} \right) \left( 1 - \left( \frac{m}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)} \right)}{\rho \left( e^{\sigma Q} - e^{-\frac{Q}{\rho}} \right) \left( 1 - \left( \frac{m}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)} \right) + \frac{1}{\sigma} (e^{\sigma Q} - 1) + \rho \left( e^{-\frac{Q}{\rho}} - 1 \right)}.$$

We again let  $\sigma = a\gamma_0$  and let  $\rho = 1/(b_0\gamma_0)$  and consider the limit as  $\gamma_0 \rightarrow \infty$ :

$$\begin{aligned} \lim_{\gamma_0 \rightarrow \infty} \int_m^{e^{\gamma_1 R}} p(x) dx &= \\ \lim_{\gamma_0 \rightarrow \infty} \frac{(1 - e^{-a\gamma_0 Q - b_0\gamma_0 Q}) \left( 1 - \left( \frac{m}{e^{\gamma_1 R}} \right)^{b_0\gamma_0/\gamma_1} \right)}{(1 - e^{-a\gamma_0 Q - b_0\gamma_0 Q}) \left( 1 - \left( \frac{m}{e^{\gamma_1 R}} \right)^{b_0\gamma_0/\gamma_1} \right) + \frac{b_0}{a} (1 - e^{-a\gamma_0 Q}) + (e^{-a\gamma_0 Q - b_0\gamma_0 Q} - e^{-a\gamma_0 Q})} &= \\ &= \frac{a}{a+b_0} \end{aligned}$$

Because as  $\gamma_0 \rightarrow \infty$ , the fraction of tumors in the size range  $[m, e^{\gamma_1 R}]$  remains the same even as the pdf converges pointwise to zero, we can think of this pdf as a combination of a continuous pdf for the metastases incepted before resection and a discrete pdf for

the metastases incepted after resection.

$$(3.15) \quad p(x) = x^{-1} \frac{ab_0}{a+b_0} \left( \frac{e^{\gamma_1 R}}{x} \right)^a, e^{\gamma_1 R} < x < \infty$$

and

$$\Pr(x = e^{\gamma_1 R}) = \frac{a}{a+b_0}.$$

In simplifying parameters this is

$$p(x) = x^{-1} \frac{ab_0}{a+b_0} \left( \frac{A}{x} \right)^a, A < x < \infty$$

and

$$\Pr(x = A) = \frac{a}{a+b_0}.$$

In this case, the cdf would be zero up until  $x = A$  at which point it would jump up to  $\frac{a}{a+b_0}$  and then increase to 1 as  $x \rightarrow \infty$ .

Finally, in the case when  $A < m$ , i.e.  $\gamma_1 \leq \log(m)/R$ , we obtain

$$(3.16) \quad p(x) = am^a x^{-1-a}, m \leq x < \infty$$

We recognize this as the Instantaneous Seeding model (3.1) with  $M \rightarrow \infty$  and this means that in the case when  $\gamma_1 \leq \log(m)/R$ , the IISMG model can do no better at fitting the data than the Instantaneous Seeding model.

## CHAPTER 4

### Data Description and Conversion

In order to estimate the parameters of our model, we obtained data from two sources. The first source was a database of cancer patients diagnosed and treated at the Memorial Sloan-Kettering Cancer Center (MSKCC) in New York. The second source was autopsy data collected and reported by the Australian pathologist J.R.S. Douglas [9].

#### 4.1. MSKCC data and conversion

Most of the necessary information reported in the MSKCC data is easily encoded for use in our model. For example, the age at diagnosis, the volume of the primary tumor at diagnosis, the age at the time of primary tumor resection, and the age at which metastases were surveyed are the observables  $U$ ,  $S$ ,  $V$ , and  $W$ , respectively. The data collected from patients also includes additional information such as the type of surgery performed, any adjuvant radiation-, chemo-, or hormonal therapies given after diagnosis, and the site, number, and volumes of metastases. These clinical variables can be used to determine whether a particular subject will be compatible with the assumptions of our model. Compatible subjects would have undergone surgery to remove a primary tumor of known size and at a later time, would have had a careful survey within at least one organ or tissue that discovered and measured a large number of metastases. Ideally, there would have been no adjuvant therapy after tumor excision, or failing that, no significant change in therapy between the time of primary tumor removal and the later survey of metastases. The most critical element of data collection is the survey of metastasis sizes which, in the case of the MSKCC data,

was accomplished by a painstaking analysis of PET/CT scans supervised by a nuclear medicine specialist.

The metastasis data is reported in voxels and must be converted to a volume measurement. To be clearly identified on the scan, a tumor must have a volume of at least  $0.5 \text{ cm}^3$ . For metastases meeting this condition, the precision in volume determination is one voxel which represents a volume of  $0.065 \text{ cm}^3$ . Assuming that a cancerous cell has a volume of  $10^{-9} \text{ cm}^3$ , a typical value for solid tumors [21], then a detectable metastasis would contain at least  $5 \times 10^8$  cells.

## 4.2. Autopsy data and conversion

Douglas obtained metastasis sizes post-mortem by hardening affected organs in formalin and then slicing them into sections of uniform thickness,  $l$ . The thickness was constant for a particular organ, but varied from organ to organ with  $l$  always between 5 mm and 7 mm. Slicing through a metastasis revealed a circular profile whose diameter could then be measured. Because the hardened organs were not completely opaque, Douglas was able to identify all sections in which a particular tumor appeared (whether one or many) and recorded the largest diameter measured in any one such section. All measurements were rounded to the nearest millimeter, so that only tumors with diameters that exceed the threshold profile diameter of  $d_m = 0.5 \text{ mm}$  were reported.

We assume (as Douglas did) that the metastases are spherical. In order to use this data in our model, we must convert diameters of circular profiles to actual spherical diameters which can then be used to represent metastasis volume.

Notice that the spherical diameter can be no less than the diameter of the observed circular profile. But if the center of the sphere does not lie in the sectioning plane, then the circular profile will have a diameter that is less than the spherical diameter. For slices of thickness  $l$ , and for a given metastasis of spherical diameter,  $y$ , with center located at a distance  $k$  from the plane of the section in which it appears, the profile diameter,  $z$ , will be given by  $z = \sqrt{y^2 - 4k^2}$  (see Figure 1). Because Douglas reported

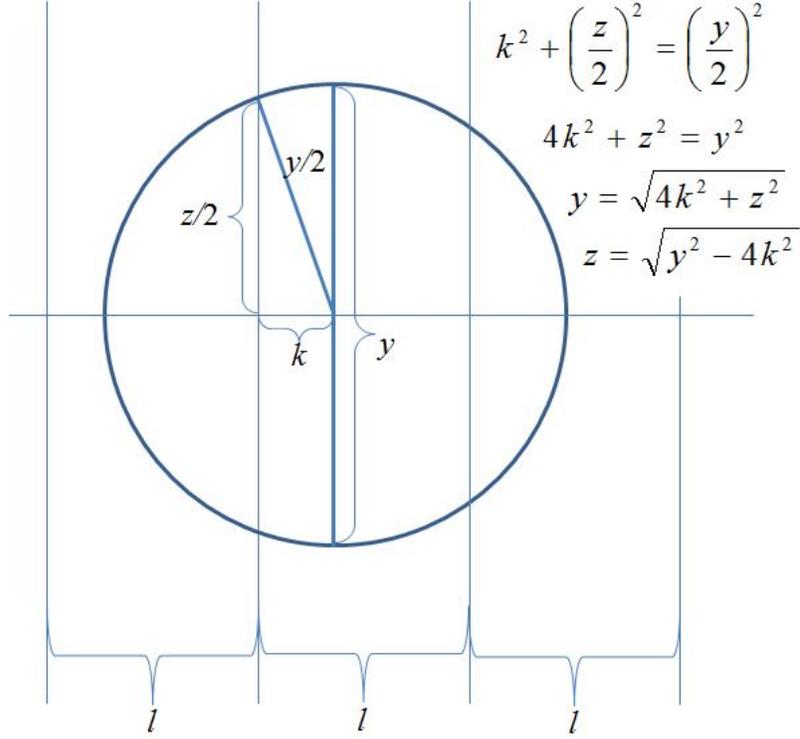


FIGURE 1. Relationship between spherical diameter,  $y$ , and profile diameter,  $z$ .

the largest profile diameter, we have  $0 \leq k \leq l/2$  (otherwise the center will be closer to a different section and  $z$  would have been reported from that closer section.) Note that for a given spherical diameter,  $y$ , we will observe the smallest profile diameter,  $z$ , when the distance from the nearest plane of section to the center of the sphere is  $k = l/2$ . This gives a minimum value for  $z$  of  $z_{\min} = \sqrt{y^2 - l^2}$  or, alternatively, when a profile of diameter  $z$  is observed, the maximum value of  $y$  is given by  $y_{\max} = \sqrt{z^2 + l^2}$ . Thus, for a given  $z$ , the spherical diameter,  $y$ , is bounded by  $z \leq y \leq \sqrt{z^2 + l^2}$ . Rewriting this inequality in terms of the error gives

$$(4.1) \quad 0 \leq y - z \leq \sqrt{z^2 + l^2} - z = \frac{l}{\sqrt{(z/l)^2 + 1} + z/l}$$

so that the larger  $z/l$  is, the less error can be incurred by estimating  $y$  with  $z$ .

As an example, Douglas' protocol 10 gives hepatic metastasis sizes for a 65-year-old male who died from an oat cell carcinoma of the lung that had metastasized to the

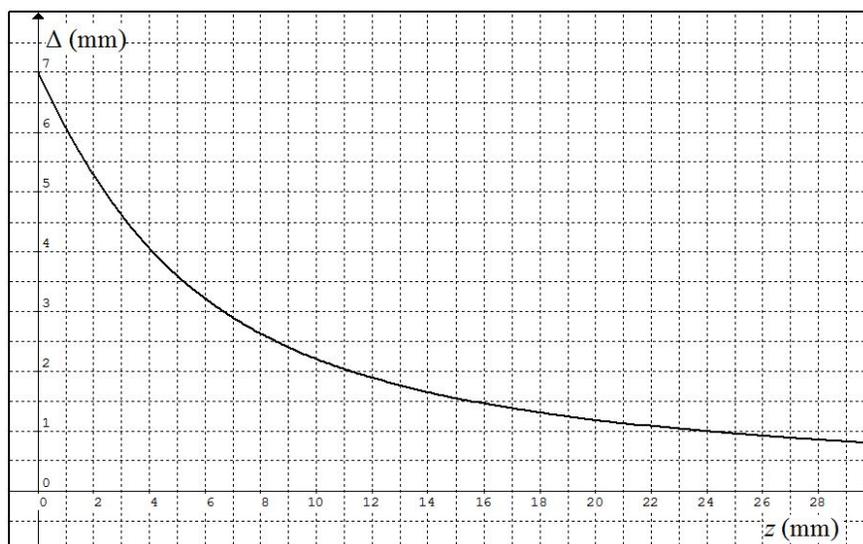


FIGURE 2. Graph of maximum error  $\Delta = \sqrt{z^2 + l^2} - z$  with  $l = 7$  mm.

liver. This protocol is of particular interest because it is the only one for which Douglas reports the size of the primary tumor which was *in situ* at the time of death with a diameter of 105 mm. Douglas microscopically determined that the mean cell diameter was 0.009 mm. The liver section width used was 7 mm. Figure 2 shows the maximum error in mm given by formula 4.1 as a function of  $z$  when  $l = 7$  mm. Given an observed diameter of a metastasis cross-section, the range of possible errors in actual diameter determination is about 7 mm wide for observed diameters near 0 mm and about 1 mm wide for the largest observed diameter in Protocol 10, which was 25 mm.

A further complication arises when we consider that metastases of small but appreciable size may lurk between the sections without being detected. It is clear that if the actual diameter,  $y$ , of a spherical tumor is greater than the section width,  $l$ , then a circular profile of the tumor will certainly be present on the surface of at least one section, but it will not be observed unless that circular profile has diameter greater than or equal to the minimum observable diameter,  $d_m$ . If a tumor is to be guaranteed of detection, then using the relationship  $z \geq \sqrt{y^2 - l^2}$ , we must have  $d_m \leq \sqrt{y^2 - l^2}$  which can be re-written as  $y \geq \sqrt{d_m^2 + l^2}$ . Thus metastases of spherical diameter greater than  $\sqrt{d_m^2 + l^2}$  will certainly be detected, metastases with spherical diameters

between  $d_m$  and  $\sqrt{d_m^2 + l^2}$  may or may not be detected, and metastases with  $y < d_m$  will go undetected.

In searching for a way to utilize Douglas' data, two lines of attack present themselves:

1. Obtain the conditional distribution of  $y$  given  $z$  and use the expected value of  $y$  given  $z$  to transform the data for direct use in our model. This may give a first estimate of model parameters, but will not account for metastases of spherical diameter between  $d_m$  and  $\sqrt{d_m^2 + l^2}$  that may escape detection because of their placement with regard to the slices. A more refined method is to
2. Compound the model-based distribution of the true secondary tumor diameters,  $y$ , with the conditional distribution of  $z$  given  $y$  to obtain the distribution of the tumor diameters  $z$  observed on autopsy slices directly and estimate model parameters from this compounded distribution.

We will refer to these methods as the Expected Diameter Method and the True Diameter Method, respectively. We now develop the mathematical formulae necessary for both approaches.

### 4.3. Full Model and Submodels written in terms of diameters

Regardless of which approach we take, we will need to convert the pdfs for the full model and our submodels from volumes,  $X$ , to diameters,  $Y$ , by using the transformation  $X = \frac{\pi Y^3}{6}$ . In each of the following cases, we apply this transformation and use the fact that  $\frac{dX}{dY} = \frac{\pi Y^2}{2}$  with the density transformation principle (see, for example [42]) to obtain the pdf of observable diameters. In the following,  $d_m = \sqrt[3]{\frac{6m}{\pi}}$  is the smallest observable diameter,  $d_M = \sqrt[3]{\frac{6M}{\pi}}$  is the largest possible diameter, and  $d_A = \sqrt[3]{\frac{6A}{\pi}}$  is the diameter that would be attained by a metastasis whose inception occurred at the time of primary tumor resection.

**4.3.1. Density  $f_Y$  for the Full Model.** Using (2.21), we have, for the case when  $d_A \leq d_m$ ,

$$(4.2) \quad f_Y(y) = \left(C_1 \frac{y}{3}\right)^{-1} \left[ \left(\frac{d_M}{y}\right)^{3a} - \left(\frac{y}{d_M}\right)^{3b_0} \right], \quad d_m \leq y \leq d_M,$$

where

$$C_1 = \frac{1}{a} \left[ \left(\frac{d_M}{d_m}\right)^{3a} - 1 \right] - \frac{1}{b_0} \left[ 1 - \left(\frac{d_m}{d_M}\right)^{3b_0} \right]$$

For the case when  $d_A > d_m$ , we transform (2.22) to obtain

$$(4.3) \quad f_Y(y) = \begin{cases} \frac{b_1}{b_0} \left(C_2 \frac{y}{3}\right)^{-1} \left[ \left(\frac{d_M}{d_A}\right)^{3a} - \left(\frac{d_A}{d_M}\right)^{3b_0} \right] \left(\frac{y}{d_A}\right)^{3b_1}, & d_m \leq y < d_A \\ \left(C_2 \frac{y}{3}\right)^{-1} \left[ \left(\frac{d_M}{y}\right)^{3a} - \left(\frac{y}{d_M}\right)^{3b_0} \right], & d_A \leq y \leq d_M. \end{cases}$$

The constant  $C_2$  is given by

$$C_2 = \frac{1}{b_0} \left[ \left(\frac{d_M}{d_A}\right)^{3a} - \left(\frac{d_A}{d_M}\right)^{3b_0} \right] \left[ 1 - \left(\frac{d_m}{d_A}\right)^{3b_1} \right] + \frac{1}{a} \left[ \left(\frac{d_M}{d_A}\right)^{3a} - 1 \right] - \frac{1}{b_0} \left[ 1 - \left(\frac{d_A}{d_M}\right)^{3b_0} \right].$$

**4.3.2. Density  $f_Y$  for the Instantaneous Seeding Model.** Converting the equations for the Instantaneous Seeding model in (3.1) and (3.3) gives:

(1) If  $d_A \leq d_m$ ,

$$f_Y(y) = (C_3 y)^{-1} \left(\frac{d_M}{y}\right)^{3a}, \quad d_m \leq y \leq d_M$$

where

$$C_3 = \frac{1}{3a} \left[ \left(\frac{d_M}{d_m}\right)^{3a} - 1 \right].$$

(2) If  $d_A > d_m$ ,

$$f_Y(y) = (C_4 y)^{-1} \left(\frac{d_M}{y}\right)^{3a}, \quad d_A \leq y \leq d_M$$

where

$$C_4 = \frac{1}{3a} \left[ \left(\frac{d_M}{d_A}\right)^{3a} - 1 \right].$$

**4.3.3. Density  $f_Y$  for the Homogeneous Model.** Converting the equations for the Homogeneous Model in (3.5) and (3.6) gives:

(1) If  $d_A \leq d_m$ ,

$$(4.4) \quad f_Y(y) = (C_5 y)^{-1} \left[ 1 - \left( \frac{y}{d_M} \right)^{3b_0} \right], \quad d_m \leq y \leq d_M$$

where

$$C_5 = \log \frac{d_m}{d_M} - \frac{1}{3b_0} \left[ 1 - \left( \frac{d_m}{d_M} \right)^{3b_0} \right].$$

(2) If  $d_A > d_m$ ,

$$f_Y(y) = \begin{cases} \frac{b_1}{b_0} (C_6 y)^{-1} \left[ 1 - \left( \frac{A}{M} \right)^{3b_0} \right] \left( \frac{y}{A} \right)^{3b_1}, & d_m \leq y < d_A \\ (C_6)^{-1} \left[ 1 - \left( \frac{y}{M} \right)^{3b_0} \right], & d_A \leq y \leq d_M \end{cases}$$

where

$$C_4 = \log \frac{d_M}{d_A} - \frac{1}{3b_0} \left[ 1 - \left( \frac{d_A}{d_M} \right)^{3b_0} \right] \left( \frac{d_m}{d_A} \right)^{3b_1}.$$

**4.3.4. Density  $f_Y$  for the Heavy-Seeding/Long-Latency Model.** Converting the equations for the Heavy-Seeding/Long-Latency model in (3.7) and (3.8), gives:

(1) If  $d_A \leq d_m$

$$f_Y(y) = (C_7 y)^{-1} \left[ \left( \frac{d_M}{y} \right)^{3a} - 1 \right], \quad d_m \leq y \leq d_M,$$

where

$$C_7 = \frac{1}{3a} \left[ \left( \frac{d_M}{d_m} \right)^{3a} - 1 \right] + \log \frac{d_m}{d_M}.$$

(2) If  $d_A > d_m$ ,

$$f_Y(y) = \begin{cases} \frac{b_1}{b_0} (C_8 y)^{-1} \left[ \left( \frac{d_M}{d_A} \right)^{3a} - 1 \right], & d_m \leq y < d_A \\ (C_8 y)^{-1} \left[ \left( \frac{d_M}{y} \right)^{3a} - 1 \right], & d_A \leq y \leq d_M \end{cases},$$

where

$$C_8 = \left( \frac{b_1}{b_0} \log \frac{d_A}{d_m} + \frac{1}{3a} \right) \left[ \left( \frac{d_M}{d_A} \right)^{3a} - 1 \right] + \log \frac{d_A}{d_M}.$$

#### 4.3.5. Density $f_Y$ for the Complete Suppression by Primary Tumor Model.

Converting the expression for the Complete Suppression by Primary Tumor Model in (3.11) gives:

$$(4.5) \quad f_Y(y) = \frac{3b_1(1-v)y^{-1}\left(\frac{y}{d_A}\right)^{3b_1}}{1-\left(\frac{d_m}{d_A}\right)^{3b_1}}, \quad d_m \leq y < d_A,$$

and

$$\Pr(y = d_A) = v$$

**4.3.6. Density  $f_Y$  for the Instantaneous Infinite Shedding and Metastasis Growth Model.** Converting the expressions for the Instantaneous Infinite Shedding and Metastasis Growth Model in (3.15) and (3.16) gives:

(1) If  $d_A < d_m$  then

$$f_Y(y) = 3ad_m^{3a}y^{-1-3a}, \quad d_m \leq y < \infty$$

(2) If  $d_A \geq d_m$  then

$$f_Y(y) = 3y^{-1}\frac{ab_0}{a+b_0}\left(\frac{d_A}{y}\right)^{3a}, \quad d_A < y < \infty$$

and

$$\Pr(y = d_A) = \frac{a}{a+b_0}$$

## 4.4. Expected diameter method

Our first approach is to obtain the conditional distribution of actual diameter,  $y$ , given observed diameter,  $z$ , and use the expected value of  $y$  given  $z$  to transform the data for direct use in our model. For a given metastasis within an organ, we call the cut in which the metastasis' maximum profile diameter is observed the *maximal cut* and denote the distance from the maximal cut to the tumor's center by  $K$ . The maximum value of  $K$  is  $l/2$  (otherwise a different cut would be closer to the center and therefore the maximal cut.) We assume that  $K$  is a uniformly distributed random variable,

i.e.  $K \sim U[0, l/2]$ . As was seen in Figure 1 above, the actual diameter is given by  $y = \sqrt{z^2 + 4K^2}$ , so that the expected value of  $y$  given  $z$  is  $E(y|z) = \frac{2}{l} \int_0^{l/2} \sqrt{z^2 + 4k^2} dk$ .

Making the substitution  $u = 2k$  yields the expression

$$(4.6) \quad \begin{aligned} E(y|z) &= \frac{1}{l} \int_0^l \sqrt{z^2 + u^2} du = \frac{1}{2} \sqrt{z^2 + l^2} + \frac{z^2(\log(l + \sqrt{z^2 + l^2}) - \log(\sqrt{z^2}))}{2l} \\ &= \frac{1}{2} \sqrt{z^2 + l^2} + \frac{z^2}{2l} \log\left(\frac{l}{z} + \sqrt{1 + \left(\frac{l}{z}\right)^2}\right). \end{aligned}$$

Given that a tumor is observed with cross-sectional diameter  $z$ , we can then compute  $E(y|z)$  as a best estimate of the true diameter  $y$  and use this estimate as though it were the true diameter.

**4.4.1. The Distribution,  $f_Z$ , of Circular Profile Diameter,  $Z$ .** Consider  $\Pr(Z \geq z|Y = y)$ , the probability that a tumor of spherical diameter  $y$  will be detected to have a profile diameter of  $z$  or greater for a given  $z \leq d_M$ . First, if  $y < z$ ,  $\Pr(Z \geq z|Y = y) = 0$  because it would be impossible for a metastasis' circular profile to have a diameter greater than its spherical diameter. On the other hand, if  $y > d_M$ , then  $\Pr(Z \geq z|Y = y) = 0$  because it would be impossible for  $y$  to be greater than the maximum diameter,  $d_M$ . Next, if  $z \leq y < \min\{\sqrt{z^2 + l^2}, d_M\}$ , then the circular profile may or may not have a diameter greater than  $z$  depending on whether or not the distance,  $K$ , from the center of the metastasis to the maximal cut is small or large. Specifically, because the circular profile diameter is  $Z = \sqrt{y^2 - 4K^2}$ , we will have  $Z \geq z$  if and only if  $K \leq \frac{1}{2}\sqrt{y^2 - z^2}$ . Because  $K$  has a uniform distribution on the interval  $[0, l/2]$ , the probability that  $K \leq \frac{1}{2}\sqrt{y^2 - z^2}$  is  $\frac{1}{2}\sqrt{y^2 - z^2} \cdot \frac{2}{l}$  and therefore  $\Pr(Z \geq z|Y = y) = \frac{1}{l}\sqrt{y^2 - z^2}$ .

Finally, if  $\sqrt{z^2 + l^2} \leq y < \min\{d_M, \sqrt{z^2 + l^2}\}$ , then no matter where the maximal cut occurs, the profile diameter will be greater than  $z$ .

Thus  $\Pr(Z \geq z|Y = y)$  is given by

$$\Pr(Z \geq z|Y = y) = \begin{cases} 0 & y < z \\ \frac{1}{l} \sqrt{y^2 - z^2} & z \leq y < \min\{\sqrt{z^2 + l^2}, d_M\} \\ 1 & \min\{\sqrt{z^2 + l^2}, d_M\} \leq y \leq d_M \\ 0 & y > d_M. \end{cases}$$

We can compound  $\Pr(Z \geq z|Y = y)$  with the pdf of spherical diameters  $f_Y(y)$  to compute the probability of observing a profile of diameter  $z$  or larger by using

$$(4.7) \quad \begin{aligned} \Pr(Z \geq z) &= \frac{1}{C_d} \int_z^{d_M} \Pr(Z \geq z|Y = y) f_Y(y) dy \\ &= \frac{1}{C_d} \int_{\max\{z, d_m\}}^{\min\{\sqrt{z^2 + l^2}, d_M\}} \frac{1}{l} \sqrt{y^2 - z^2} f_Y(y) dy + \frac{1}{C_d} \int_{\min\{\sqrt{z^2 + l^2}, d_M\}}^{d_M} f_Y(y) dy \end{aligned}$$

where the constant  $C_d$  is found from

$$C_d = \int_{d_m}^{d_M} \Pr(Z \geq z|Y = y) f_Y(y) dy$$

and represents the fraction of metastases that are actually detected. Because the cdf of profile diameters is  $F(z) = \Pr(Z < z) = 1 - \Pr(Z \geq z)$ , we can find the pdf of profile diameters,  $f_Z(z)$ , by computing the derivative with respect to  $z$  of  $-\Pr(Z \geq z)$ .

In doing so, we use the formula

$$\frac{d}{du} \int_{\alpha(u)}^{\beta(u)} f(x, u) dx = f(\beta(u), u) \frac{d\beta}{du} - f(\alpha(u), u) \frac{d\alpha}{du} + \int_{\alpha(u)}^{\beta(u)} \frac{\partial}{\partial u} f(x, u) dx.$$

**4.4.2. Likelihood Estimation of Parameters: True Diameter without Rounding.** In order to use the Douglas data to estimate parameters, we can maximize the likelihood function  $\prod_{i=1}^n f_Z(z_i)$  where  $z_i$  is the  $i$ th observed diameter. We call this the *True Diameter without Rounding* method.

**4.4.3. Likelihood Estimation of Parameters: True Diameter with Rounding.** The previous computation does not account for the fact that the circular profiles

of the tumors were rounded to the nearest 1 mm. One consequence of the rounding is that the smallest observable tumor diameter would be  $d_m = 0.5$  mm. Another consequence of the rounding is that for metastases with diameters near 1 mm (that are numerous in Protocol 10), the amount of rounding with respect to the actual measurement could have been quite large. Because each recorded profile diameter  $z_i$  corresponded to an actual profile diameter between  $z_i - 0.5$  mm and  $z_i + 0.5$  mm, we can compensate for rounding by using

$$\prod_{i=1}^n \Pr(z_i - 0.5 \leq Z \leq z_i + 0.5)$$

which is the same as

$$(4.8) \quad \prod_{i=1}^n [\Pr(Z > z_i - 0.5) - \Pr(Z > z_i + 0.5)]$$

as our likelihood function. We call this the *True Diameter with Rounding* method.

#### 4.5. Computing $F_Z$ and $f_Z$ for the True Diameter Methods

We now give more detailed versions of the cdf  $F_Z$  and pdf  $f_Z$  for the Full model and the submodels that we have discussed.

**4.5.1. Cdf  $F_Z$  and Pdf  $f_Z$ ,  $d_A < d_m$ , for the Full, Homogeneous, and Heavy-Seeding/Long-Latency Models.** For the Full, Homogeneous, and Heavy-Seeding/Long-Latency models, when  $d_A < d_m$ , the distribution of metastasis diameters,  $f_Y(y)$  is a continuous function which we denote by  $f_{YS}(y)$  for “small  $d_A$ .” We use (4.7) to write  $\Pr(Z \geq z)$  and then compute  $f_Z(z)$ . We have two cases to consider:

**Case i)**  $d_A < d_m < d_M < \sqrt{d_m^2 + l^2}$

In this case, we must have  $d_m \leq z \leq d_M \leq \sqrt{d_m^2 + l^2}$ . In view of (4.7), we have

$$\Pr(Z \geq z) = \frac{1}{C_d} \int_z^{d_M} \frac{1}{l} \sqrt{y^2 - z^2} f_{YS}(y) dy$$

and differentiating  $-\Pr(Z \geq z)$  with respect to  $z$  gives

$$f_Z(z) = \frac{1}{C_d} \int_z^{d_M} \frac{z}{l\sqrt{y^2 - z^2}} f_{YS}(y) dy$$

**Case ii)**  $d_A < d_m < \sqrt{d_m^2 + l^2} \leq d_M$

In this case,  $z$  can lie within one of two regions.

Region i)  $d_m \leq z \leq \sqrt{d_M^2 - l^2} < d_M$

$$\Pr(Z \geq z) = \frac{1}{C_d} \int_z^{\sqrt{z^2 + l^2}} \frac{1}{l} \sqrt{y^2 - z^2} f_{YS}(y) dy + \frac{1}{C_d} \int_{\sqrt{z^2 + l^2}}^{d_M} f_{YS}(y) dy$$

$$f_Z(z) = \frac{1}{C_d} \int_z^{\sqrt{z^2 + l^2}} \frac{z}{l\sqrt{y^2 - z^2}} f_{YS}(y) dy$$

Region ii)  $d_m \leq \sqrt{d_M^2 - l^2} < z \leq d_M$

$$\Pr(Z \geq z) = \frac{1}{C_d} \int_z^{d_M} \frac{1}{l} \sqrt{y^2 - z^2} f_{YS}(y) dy$$

$$f_Z(z) = \frac{1}{C_d} \int_z^{d_M} \frac{1}{l} \frac{z}{\sqrt{y^2 - z^2}} f_{YS}(y) dy$$

We can combine the above cases into the single expression

$$f_Z(z) = \frac{1}{C_d} \int_z^{\min\{\sqrt{z^2 + l^2}, d_M\}} \frac{z}{l\sqrt{y^2 - z^2}} f_{YS}(y) dy, d_m \leq z \leq d_M$$

and 0 otherwise.

**4.5.2. Cdf  $F_Z$  and Pdf  $f_Z$ ,  $d_A \geq d_m$ , for the Full, Homogeneous, and Heavy-Seeding/Long-Latency Models.** For the Full, Homogeneous, and Heavy-Seeding/Long-Latency models, when  $d_A \geq d_m$ , the distribution of metastasis diameters,  $f_Y(y)$ , is a piecewise-defined function with a possible discontinuity at  $y = d_A$ . We use  $f_{YL}(y)$  to denote the values of  $f_Y(y)$  for  $y$  on the “left side” of  $d_A$  and  $f_{YR}(y)$  to denote the values of  $f_Y(y)$  for  $y$  on the “right side” of  $d_A$ . With this notation, we can use (4.7) to write  $\Pr(Z \geq z)$ . Based on the proximity of  $d_m$ ,  $d_A$ , and  $d_M$  as measured by the slice width,  $l$ , we have five cases to consider:

$$\text{Case i) } d_m \leq d_A \leq d_M \leq \sqrt{d_m^2 + l^2} \leq \sqrt{d_A^2 + l^2}$$

$$\text{Case ii) } d_m \leq d_A \leq \sqrt{d_m^2 + l^2} \leq d_M \leq \sqrt{d_A^2 + l^2}$$

$$\text{Case iii) } d_m \leq d_A \leq \sqrt{d_m^2 + l^2} \leq \sqrt{d_A^2 + l^2} \leq d_M$$

$$\text{Case iv) } d_m \leq \sqrt{d_m^2 + l^2} \leq d_A \leq d_M \leq \sqrt{d_A^2 + l^2}$$

$$\text{Case v) } d_m \leq \sqrt{d_m^2 + l^2} \leq d_A \leq \sqrt{d_A^2 + l^2} \leq d_M$$

Within each case, we have up to four regions in which  $z$  can lie. We now delineate  $\Pr(Z \geq z)$  for  $z$  within the regions created in each case and find  $f_Z(z)$  by differentiating  $-\Pr(Z \geq z)$  with respect to  $z$ .

$$\text{Case i) } d_m \leq d_A \leq d_M \leq \sqrt{d_m^2 + l^2} \leq \sqrt{d_A^2 + l^2}$$

$$\text{Region i) } d_m \leq z < d_A \leq d_M \leq \sqrt{d_m^2 + l^2} \leq \sqrt{d_A^2 + l^2}$$

$$\Pr(Z \geq z) = \frac{1}{C_d} \left[ \int_z^{d_A} \frac{\sqrt{y^2 - z^2}}{l} f_{YL}(y) dy + \int_{d_A}^{d_M} \frac{\sqrt{y^2 - z^2}}{l} f_{YR}(y) dy \right]$$

$$f_Z(z) = \frac{1}{C_d} \left[ \int_z^{d_A} \frac{z}{l\sqrt{y^2 - z^2}} f_{YL}(y) dy + \int_{d_A}^{d_M} \frac{z}{l\sqrt{y^2 - z^2}} f_{YR}(y) dy \right]$$

$$\text{Region ii) } d_m \leq d_A \leq z \leq d_M \leq \sqrt{d_m^2 + l^2} \leq \sqrt{d_A^2 + l^2}$$

$$\Pr(Z \geq z) = \frac{1}{C_d} \int_z^{d_M} \frac{\sqrt{y^2 - z^2}}{l} f_{YR}(y) dy$$

$$f_Z(z) = \frac{1}{C_d} \int_z^{d_M} \frac{z}{l\sqrt{y^2 - z^2}} f_{YR}(y) dy$$

Combining the results above, we have

$$f_Z(z) = \begin{cases} \frac{1}{C_d} \left[ \int_z^{d_A} \frac{z}{l\sqrt{y^2 - z^2}} f_{YL}(y) dy + \int_{d_A}^{d_M} \frac{z}{l\sqrt{y^2 - z^2}} f_{YR}(y) dy \right] & d_m \leq z < d_A \\ \frac{1}{C_d} \int_z^{d_M} \frac{z}{l\sqrt{y^2 - z^2}} f_{YR}(y) dy & d_A \leq z \leq d_M \end{cases}$$

We choose  $C_d$  so that  $\Pr(Z \geq d_m) = 1$ , so

$$C_d = \int_{d_m}^{d_A} \frac{\sqrt{y^2 - d_m^2}}{l} f_{YL}(y) dy + \int_{d_A}^{d_M} \frac{\sqrt{y^2 - d_m^2}}{l} f_{YR}(y) dy.$$

**Case ii)**  $d_m \leq d_A \leq \sqrt{d_m^2 + l^2} \leq d_M \leq \sqrt{d_A^2 + l^2}$

Region i)  $d_m \leq z < \sqrt{d_M^2 - l^2} < d_A \leq \sqrt{d_m^2 + l^2} \leq d_M \leq \sqrt{d_A^2 + l^2}$

$\Pr(Z \geq z) =$

$$\frac{1}{C_d} \left[ \int_z^{d_A} \frac{\sqrt{y^2 - z^2}}{l} f_{YL}(y) dy + \int_{d_A}^{\sqrt{z^2 + l^2}} \frac{\sqrt{y^2 - z^2}}{l} f_{YR}(y) dy + \int_{\sqrt{z^2 + l^2}}^{d_M} f_{YR}(y) dy \right]$$

$$f_Z(z) = \frac{1}{C_d} \left[ \int_z^{d_A} \frac{z}{l\sqrt{y^2 - z^2}} f_{YL}(y) dy + \int_{d_A}^{\sqrt{z^2 + l^2}} \frac{z}{l\sqrt{y^2 - z^2}} f_{YR}(y) dy \right]$$

Region ii)  $d_m \leq \sqrt{d_M^2 - l^2} \leq z < d_A \leq \sqrt{d_m^2 + l^2} \leq d_M \leq \sqrt{d_A^2 + l^2}$

$$\Pr(Z \geq z) = \frac{1}{C_d} \left[ \int_z^{d_A} \frac{\sqrt{y^2 - z^2}}{l} f_{YL}(y) dy + \int_{d_A}^{d_M} \frac{\sqrt{y^2 - z^2}}{l} f_{YR}(y) dy \right]$$

$$f_Z(z) = \frac{1}{C_d} \left[ \int_z^{d_A} \frac{z}{l\sqrt{y^2 - z^2}} f_{YL}(y) dy + \int_{d_A}^{d_M} \frac{z}{l\sqrt{y^2 - z^2}} f_{YR}(y) dy \right]$$

Region iii)  $d_m \leq d_A \leq z \leq d_M \leq \sqrt{d_A^2 + l^2}$

$$\Pr(Z \geq z) = \frac{1}{C_d} \int_z^{d_M} \frac{\sqrt{y^2 - z^2}}{l} f_{YR}(y) dy$$

$$f_Z(z) = \frac{1}{C_d} \int_z^{d_M} \frac{z}{l\sqrt{y^2 - z^2}} f_{YR}(y) dy$$

Combining the results above, we have

$$f_Z(z) = \begin{cases} \frac{1}{C_d} \left[ \int_z^{d_A} \frac{z}{l\sqrt{y^2 - z^2}} f_{YL}(y) dy + \int_{d_A}^{\sqrt{z^2 + l^2}} \frac{z}{l\sqrt{y^2 - z^2}} f_{YR}(y) dy \right] & d_m \leq z < \sqrt{d_M^2 - l^2} \\ \frac{1}{C_d} \left[ \int_z^{d_A} \frac{z}{l\sqrt{y^2 - z^2}} f_{YL}(y) dy + \int_{d_A}^{d_M} \frac{z}{l\sqrt{y^2 - z^2}} f_{YR}(y) dy \right] & \sqrt{d_M^2 - l^2} \leq z < d_A \\ \frac{1}{C_d} \int_z^{d_M} \frac{z}{l\sqrt{y^2 - z^2}} f_{YR}(y) dy & d_A \leq z \leq d_M \end{cases}$$

We choose  $C_d$  so that  $\Pr(Z \geq d_m) = 1$ , and therefore in this case,

$$C_d = \int_{d_m}^{d_A} \frac{\sqrt{y^2 - d_m^2}}{l} f_{YL}(y) dy + \int_{d_A}^{\sqrt{d_m^2 + l^2}} \frac{\sqrt{y^2 - d_m^2}}{l} f_{YR}(y) dy + \int_{\sqrt{d_m^2 + l^2}}^{d_M} f_{YR}(y) dy.$$

**Case iii)**  $d_m \leq d_A \leq \sqrt{d_m^2 + l^2} \leq \sqrt{d_A^2 + l^2} \leq d_M$

Region i)  $d_m \leq z < d_A \leq \sqrt{d_m^2 + l^2} \leq \sqrt{d_A^2 + l^2} \leq d_M$

$$\Pr(Z \geq z) =$$

$$\frac{1}{C_d} \left[ \int_z^{d_A} \frac{1}{l} \sqrt{y^2 - z^2} f_{YL}(y) dy + \int_{d_A}^{\sqrt{z^2 + l^2}} \frac{1}{l} \sqrt{y^2 - z^2} f_{YR}(y) dy + \int_{\sqrt{z^2 + l^2}}^{d_M} f_{YR}(y) dy \right]$$

$$f_Z(z) = \frac{1}{C_d} \left[ \int_z^{d_A} \frac{z}{l\sqrt{y^2 - z^2}} f_{YL}(y) dy + \int_{d_A}^{\sqrt{z^2 + l^2}} \frac{z}{l\sqrt{y^2 - z^2}} f_{YR}(y) dy \right]$$

Region ii)  $d_m \leq d_A \leq z < \sqrt{z^2 + l^2} \leq d_M$

$$\Pr(Z \geq z) = \frac{1}{C_d} \left[ \int_z^{\sqrt{z^2 + l^2}} \frac{1}{l} \sqrt{y^2 - z^2} f_{YR}(y) dy + \int_{\sqrt{z^2 + l^2}}^{d_M} f_{YR}(y) dy \right]$$

$$f_Z(z) = \frac{1}{C_d} \int_z^{\sqrt{z^2 + l^2}} \frac{z}{l\sqrt{y^2 - z^2}} f_{YR}(y) dy$$

Region iii)  $d_m \leq d_A \leq z \leq d_M \leq \sqrt{z^2 + l^2}$

$$\Pr(Z \geq z) = \frac{1}{C_d} \int_z^{d_M} \frac{1}{l} \sqrt{y^2 - z^2} f_{YR}(y) dy$$

$$f_Z(z) = \frac{1}{C_d} \int_z^{d_M} \frac{z}{l\sqrt{y^2 - z^2}} f_{YR}(y) dy$$

Combining the results above we have

$$f_Z(z) = \begin{cases} \frac{1}{C_d} \left[ \int_z^{d_A} \frac{z}{l\sqrt{y^2 - z^2}} f_{YL}(y) dy + \int_{d_A}^{\sqrt{z^2 + l^2}} \frac{z}{l\sqrt{y^2 - z^2}} f_{YR}(y) dy \right] & d_m \leq z < d_A \\ \frac{1}{C_d} \int_z^{\min\{\sqrt{z^2 + l^2}, d_M\}} \frac{z}{l\sqrt{y^2 - z^2}} f_{YR}(y) dy & d_A \leq z \leq d_M \end{cases}$$

We choose  $C_d$  so that  $\Pr(Z \geq d_m) = 1$ , and therefore

$$C_d = \int_{d_m}^{d_A} \frac{1}{l} \sqrt{y^2 - d_m^2} f_{YL}(y) dy + \int_{d_A}^{\sqrt{d_m^2 + l^2}} \frac{1}{l} \sqrt{y^2 - d_m^2} f_{YR}(y) dy + \int_{\sqrt{d_m^2 + l^2}}^{d_M} f_{YR}(y) dy.$$

**Case iv)**  $d_m < \sqrt{d_m^2 + l^2} \leq d_A \leq d_M \leq \sqrt{d_A^2 + l^2}$

Region i)  $d_m \leq z < \sqrt{d_A^2 - l^2} \leq \sqrt{d_M^2 - l^2} \leq d_A \leq d_M \leq \sqrt{d_A^2 + l^2}$

$$\Pr(Z \geq z) = \frac{1}{C_d} \left[ \int_z^{\sqrt{z^2 + l^2}} \frac{\sqrt{y^2 - z^2}}{l} f_{YL}(y) dy + \int_{\sqrt{z^2 + l^2}}^{d_A} f_{YL}(y) dy + \int_{d_A}^{d_M} f_{YR}(y) dy \right]$$

$$f_Z(z) = \frac{1}{C_d} \int_z^{\sqrt{z^2 + l^2}} \frac{z}{l \sqrt{y^2 - z^2}} f_{YL}(y) dy$$

Region ii)  $d_m \leq \sqrt{d_A^2 - l^2} \leq z < \sqrt{d_M^2 - l^2} \leq d_A \leq d_M \leq \sqrt{d_A^2 + l^2}$

$\Pr(Z \geq z) =$

$$\frac{1}{C_d} \left[ \int_z^{d_A} \frac{\sqrt{y^2 - z^2}}{l} f_{YL}(y) dy + \int_{d_A}^{\sqrt{z^2 + l^2}} \frac{\sqrt{y^2 - z^2}}{l} f_{YR}(y) dy + \int_{\sqrt{z^2 + l^2}}^{d_M} f_{YR}(y) dy \right]$$

$$f_Z(z) = \frac{1}{C_d} \left[ \int_z^{d_A} \frac{z}{l \sqrt{y^2 - z^2}} f_{YL}(y) dy + \int_{d_A}^{\sqrt{z^2 + l^2}} \frac{z}{l \sqrt{y^2 - z^2}} f_{YR}(y) dy \right]$$

Region iii)  $d_m \leq \sqrt{d_A^2 - l^2} \leq \sqrt{d_M^2 - l^2} \leq z < d_A \leq d_M \leq \sqrt{d_A^2 + l^2}$

$$\Pr(Z \geq z) = \frac{1}{C_d} \left[ \int_z^{d_A} \frac{\sqrt{y^2 - z^2}}{l} f_{YL}(y) dy + \int_{d_A}^{d_M} \frac{\sqrt{y^2 - z^2}}{l} f_{YR}(y) dy \right]$$

$$f_Z(z) = \frac{1}{C_d} \left[ \int_z^{d_A} \frac{z}{l \sqrt{y^2 - z^2}} f_{YL}(y) dy + \int_{d_A}^{d_M} \frac{z}{l \sqrt{y^2 - z^2}} f_{YR}(y) dy \right]$$

Region iv)  $d_m \leq \sqrt{d_A^2 - l^2} \leq \sqrt{d_M^2 - l^2} \leq d_A \leq z < d_M \leq \sqrt{d_A^2 + l^2}$

$$\Pr(Z \geq z) = \frac{1}{C_d} \int_z^{d_M} \frac{\sqrt{y^2 - z^2}}{l} f_{YR}(y) dy$$

$$f_Z(z) = \frac{1}{C_d} \int_z^{d_M} \frac{z}{l\sqrt{y^2 - z^2}} f_{YR}(y) dy$$

Combining the results above, we have

$$f_Z(z) = \begin{cases} \frac{1}{C_d} \int_z^{\sqrt{z^2+l^2}} \frac{z}{l\sqrt{y^2-z^2}} f_{YL}(y) dy & d_m \leq z < \sqrt{d_A^2 - l^2} \\ \frac{1}{C_d} \left[ \int_z^{d_A} \frac{z}{l\sqrt{y^2-z^2}} f_{YL}(y) dy + \int_{d_A}^{\sqrt{z^2+l^2}} \frac{z}{l\sqrt{y^2-z^2}} f_{YR}(y) dy \right] & \sqrt{d_A^2 - l^2} \leq z < \sqrt{d_M^2 - l^2} \\ \frac{1}{C_d} \left[ \int_z^{d_A} \frac{z}{l\sqrt{y^2-z^2}} f_{YL}(y) dy + \int_{d_A}^{d_M} \frac{z}{l\sqrt{y^2-z^2}} f_{YR}(y) dy \right] & \sqrt{d_M^2 - l^2} \leq z < d_A \\ \frac{1}{C_d} \int_z^{d_M} \frac{z}{l\sqrt{y^2-z^2}} f_{YR}(y) dy & d_A \leq z \leq d_M \end{cases}$$

We choose  $C_d$  so that  $\Pr(Z \geq d_m) = 1$ , so

$$C_d = \int_{d_m}^{\sqrt{d_m^2+l^2}} \frac{\sqrt{y^2 - d_m^2}}{l} f_{YL}(y) dy + \int_{\sqrt{d_m^2+l^2}}^{d_A} f_{YL}(y) dy + \int_{d_A}^{d_M} f_{YR}(y) dy$$

**Case v)**  $d_m \leq \sqrt{d_m^2 + l^2} \leq d_A \leq \sqrt{d_A^2 + l^2} \leq d_M$

Region i)  $d_m \leq z < \sqrt{d_A^2 - l^2} \leq d_A \leq \sqrt{d_M^2 - l^2} \leq d_M$

$$\Pr(Z \geq z) = \frac{1}{C_d} \left[ \int_z^{\sqrt{z^2+l^2}} \frac{\sqrt{y^2 - z^2}}{l} f_{YL}(y) dy + \int_{\sqrt{z^2+l^2}}^{d_A} f_{YL}(y) dy + \int_{d_A}^{d_M} f_{YR}(y) dy \right]$$

$$f_Z(z) = \frac{1}{C_d} \left[ \int_z^{\sqrt{z^2+l^2}} \frac{z}{l\sqrt{y^2 - z^2}} f_{YL}(y) dy \right]$$

Region ii)  $d_m \leq \sqrt{d_A^2 - l^2} \leq z < d_A \leq \sqrt{d_M^2 - l^2} \leq d_M$

$\Pr(Z \geq z) =$

$$\frac{1}{C_d} \left[ \int_z^{d_A} \frac{\sqrt{y^2 - z^2}}{l} f_{YL}(y) dy + \int_{d_A}^{\sqrt{z^2+l^2}} \frac{\sqrt{y^2 - z^2}}{l} f_{YR}(y) dy + \int_{\sqrt{z^2+l^2}}^{d_M} f_{YR}(y) dy \right]$$

$$f_Z(z) = \frac{1}{C_d} \left[ \int_z^{d_A} \frac{z}{l\sqrt{y^2 - z^2}} f_{YL}(y) dy + \int_{d_A}^{\sqrt{z^2+l^2}} \frac{z}{l\sqrt{y^2 - z^2}} f_{YR}(y) dy \right]$$

Region iii)  $d_m \leq \sqrt{d_A^2 - l^2} \leq d_A \leq z < \sqrt{d_M^2 - l^2} \leq d_M$

$$\Pr(Z \geq z) = \frac{1}{C_d} \left[ \int_z^{\sqrt{z^2+l^2}} \frac{\sqrt{y^2-z^2}}{l} f_{YR}(y) dy + \int_{\sqrt{z^2+l^2}}^{d_M} f_{YR}(y) dy \right]$$

$$f_Z(z) = \frac{1}{C_d} \int_z^{\sqrt{z^2+l^2}} \frac{z}{l\sqrt{y^2-z^2}} f_{YR}(y) dy$$

Region iv)  $d_m \leq \sqrt{d_A^2 - l^2} \leq d_A \leq \sqrt{d_M^2 - l^2} \leq z < d_M$

$$\Pr(Z \geq z) = \frac{1}{C_d} \left[ \int_z^{d_M} \frac{\sqrt{y^2-z^2}}{l} f_{YR}(y) dy \right]$$

$$f_Z(z) = \frac{1}{C_d} \left[ \int_z^{d_M} \frac{z}{l\sqrt{y^2-z^2}} f_{YR}(y) dy \right]$$

Combining the results above, we have

$$f_Z(z) = \begin{cases} \frac{1}{C_d} \left[ \int_z^{\sqrt{z^2+l^2}} \frac{z}{l\sqrt{y^2-z^2}} f_{YL}(y) dy \right] & d_m \leq z < \sqrt{d_A^2 - l^2} \\ \frac{1}{C_d} \left[ \int_z^{d_A} \frac{z}{l\sqrt{y^2-z^2}} f_{YL}(y) dy + \int_{d_A}^{\sqrt{z^2+l^2}} \frac{z}{l\sqrt{y^2-z^2}} f_{YR}(y) dy \right] & \sqrt{d_A^2 - l^2} \leq z < d_A \\ \frac{1}{C_d} \int_z^{\sqrt{z^2+l^2}} \frac{z}{l\sqrt{y^2-z^2}} f_{YR}(y) dy & d_A \leq z < \sqrt{d_M^2 - l^2} \\ \frac{1}{C_d} \left[ \int_z^{d_M} \frac{z}{l\sqrt{y^2-z^2}} f_{YR}(y) dy \right] & \sqrt{d_M^2 - l^2} \leq z \leq d_M \end{cases}$$

We choose  $C_d$  so that  $\Pr(Z \geq d_m) = 1$ , so

$$C_d = \int_{d_m}^{\sqrt{d_m^2+l^2}} \frac{\sqrt{y^2-d_m^2}}{l} f_{YL}(y) dy + \int_{\sqrt{d_m^2+l^2}}^{d_A} f_{YL}(y) dy + \int_{d_A}^{d_M} f_{YR}(y) dy.$$

We note that in all cases reported by Douglas, the section width,  $l$ , is sufficiently small that case i) does not occur.

**4.5.3. Cdf  $F_Z$  and Pdf  $f_Z$  for the Complete Suppression by Primary Tumor Model.** In the case of the CSPT model, we have

$$\Pr(Z \geq z) = \frac{1}{C_d} \int_z^{\min\{\sqrt{z^2+l^2}, d_A\}} \frac{\sqrt{y^2-z^2}}{l} f_Y(y) dy + \frac{1}{C_d} \int_{\min\{\sqrt{z^2+l^2}, d_A\}}^{d_A} f_Y(y) dy + \frac{1}{C_d} \Pr(Y = d_A)$$

if  $d_m \leq z < \sqrt{d_A^2 - l^2}$  and

$$\Pr(Z \geq z) = \frac{1}{C_d} \int_z^{\min\{\sqrt{z^2+l^2}, d_A\}} \frac{\sqrt{y^2-z^2}}{l} f_Y(y) dy + \frac{1}{C_d} \int_{\min\{\sqrt{z^2+l^2}, d_A\}}^{d_A} f_Y(y) dy + \frac{1}{C_d} \frac{\sqrt{d_A^2-z^2}}{l} \Pr(Y = d_A)$$

if  $\sqrt{d_A^2 - l^2} \leq z \leq d_A$ .

Based on the proximity of  $d_m$  and  $d_A$ , as measured by the slice width,  $l$ , we have two cases to consider:

Case i)  $d_m \leq d_A \leq \sqrt{d_m^2 + l^2}$

Case ii)  $d_m \leq \sqrt{d_m^2 + l^2} \leq d_A$

Within each case, we have up to two regions in which  $z$  can lie. We now delineate  $\Pr(Z \geq z)$  for  $z$  within the regions created in each case and find  $f_Z(z)$  by differentiating  $-\Pr(Z \geq z)$  with respect to  $z$ .

**Case i)**  $d_m \leq d_A \leq \sqrt{d_m^2 + l^2}$

Here we must have  $d_m \leq z \leq d_A \leq \sqrt{d_m^2 + l^2}$ , so

$$\Pr(Z \geq z) = \frac{1}{C_d} \left[ \int_z^{d_A} \frac{\sqrt{y^2 - z^2}}{l} f_Y(y) dy + \frac{\sqrt{d_A^2 - z^2}}{l} \Pr(Y = d_A) \right]$$

and

$$f_Z(z) = \frac{1}{C_d} \left[ \int_z^{d_A} \frac{z}{l\sqrt{y^2 - z^2}} f_Y(y) dy + \frac{z}{l\sqrt{d_A^2 - z^2}} \Pr(Y = d_A) \right]$$

We choose  $C_d$  so that  $\Pr(Z \geq d_m) = 1$ , so

$$C_d = \int_{d_m}^{d_A} \frac{\sqrt{y^2 - d_m^2}}{l} f_Y(y) dy + \frac{\sqrt{d_A^2 - d_m^2}}{l} \Pr(Y = d_A)$$

**Case ii)**  $d_m \leq \sqrt{d_m^2 + l^2} \leq d_A$

Region i)  $d_m \leq z < \sqrt{d_A^2 - l^2} \leq d_A$

$$\Pr(Z \geq z) = \frac{1}{C_d} \left[ \int_z^{\sqrt{z^2+l^2}} \frac{\sqrt{y^2 - z^2}}{l} f_Y(y) dy + \int_{\sqrt{z^2+l^2}}^{d_A} f_Y(y) dy + \Pr(Y = d_A) \right]$$

$$f_Z(z) = \frac{1}{C_d} \int_z^{\sqrt{z^2+l^2}} \frac{z}{l\sqrt{y^2-z^2}} f_Y(y) dy$$

Region ii)  $d_m \leq \sqrt{d_A^2 - l^2} \leq z \leq d_A$

$$\Pr(Z \geq z) = \frac{1}{C_d} \left[ \int_z^{d_A} \frac{\sqrt{y^2 - z^2}}{l} f_Y(y) dy + \frac{\sqrt{d_A^2 - z^2}}{l} \Pr(Y = d_A) \right]$$

$$f_Z(z) = \frac{1}{C_d} \left[ \int_z^{d_A} \frac{z}{l\sqrt{y^2 - z^2}} f_Y(y) dy + \frac{z}{l\sqrt{d_A^2 - z^2}} \Pr(Y = d_A) \right]$$

Combining the results for  $f_Z(z)$  in this case, we have

$$(4.9) \quad f_Z(z) = \begin{cases} \frac{1}{C_d} \int_z^{\sqrt{z^2+l^2}} \frac{z}{l\sqrt{y^2-z^2}} f_Y(y) dy & d_m \leq z < \sqrt{d_A^2 - l^2} \\ \frac{1}{C_d} \left[ \int_z^{d_A} \frac{z}{l\sqrt{y^2-z^2}} f_Y(y) dy + \frac{z}{l\sqrt{d_A^2-z^2}} \Pr(Y = d_A) \right] & \sqrt{d_A^2 - l^2} \leq z \leq d_A \end{cases}$$

We choose  $C_d$  so that  $\Pr(Z \geq d_m) = 1$  and therefore

$$C_d = \int_{d_m}^{\sqrt{d_m^2+l^2}} \frac{\sqrt{y^2 - d_m^2}}{l} f_Y(y) dy + \int_{\sqrt{d_m^2+l^2}}^{d_A} f_Y(y) dy + \Pr(Y = d_A).$$

In all cases reported by Douglas, the section width,  $l$ , is sufficiently small that case i) does not occur, so we can focus on case ii) only.

**4.5.4. Cdf  $F_Z$  and Pdf  $f_Z$  for the Instantaneous Infinite Shedding and Metastasis Growth Model.** For the Instantaneous Infinite Shedding and Metastasis Growth model, we have

$$\Pr(Z \geq z) = \frac{1}{C_d} \left[ \Pr(Y = d_A) + \int_{d_A}^{\infty} f_Y(y) dy \right]$$

if  $d_m \leq z < \sqrt{d_A^2 - l^2}$ ,

$$\Pr(Z \geq z) =$$

$$\frac{1}{C_d} \left[ \frac{1}{l} \sqrt{d_A^2 - z^2} \Pr(Y = d_A) + \int_{d_A}^{\sqrt{z^2+l^2}} \frac{\sqrt{y^2 - z^2}}{l} f_Y(y) dy + \int_{\sqrt{z^2+l^2}}^{\infty} f_Y(y) dy \right]$$

if  $\sqrt{d_A^2 - l^2} \leq z < d_A$ , and

$$\Pr(Z \geq z) = \frac{1}{C_d} \left[ \int_z^{\sqrt{z^2+l^2}} \frac{\sqrt{y^2 - z^2}}{l} f_Y(y) dy + \int_{\sqrt{z^2+l^2}}^{\infty} f_Y(y) dy \right]$$

if  $z \geq d_A$ .

Based on the proximity of  $d_m$ , and  $d_A$  as measured by the slice width,  $l$ , we have two cases to consider:

$$\text{Case i) } d_m \leq d_A \leq \sqrt{d_m^2 + l^2}$$

$$\text{Case ii) } d_m \leq \sqrt{d_m^2 + l^2} < d_A$$

Within each case, we have up to three regions in which  $z$  can lie. We now delineate  $\Pr(Z \geq z)$  for  $z$  within the regions created in each case and find  $f_Z(z)$  by differentiating  $-\Pr(Z \geq z)$  with respect to  $z$ .

$$\text{Case i) } d_m \leq d_A \leq \sqrt{d_m^2 + l^2}$$

$$\text{Region i) } d_m \leq z < d_A \leq \sqrt{d_m^2 + l^2}$$

$$\Pr(Z \geq z) = \frac{1}{C_d} \left[ \frac{1}{l} \sqrt{d_A^2 - z^2} \Pr(Y = d_A) + \int_{d_A}^{\sqrt{z^2+l^2}} \frac{\sqrt{y^2 - z^2}}{l} f_Y(y) dy + \int_{\sqrt{z^2+l^2}}^{\infty} f_Y(y) dy \right]$$

$$f_Z(z) = \frac{1}{C_d} \left[ \frac{z}{l \sqrt{d_A^2 - z^2}} \Pr(Y = d_A) + \int_{d_A}^{\sqrt{z^2+l^2}} \frac{z}{l \sqrt{y^2 - z^2}} f_Y(y) dy \right]$$

$$\text{Region ii) } d_m \leq d_A \leq z$$

$$\Pr(Z \geq z) = \frac{1}{C_d} \left[ \int_z^{\sqrt{z^2+l^2}} \frac{\sqrt{y^2 - z^2}}{l} f_Y(y) dy + \int_{\sqrt{z^2+l^2}}^{\infty} f_Y(y) dy \right]$$

$$f_Z(z) = \frac{1}{C_d} \left[ \int_z^{\sqrt{z^2+l^2}} \frac{z}{l\sqrt{y^2-z^2}} f_Y(y) dy \right]$$

Combining the results for  $f_Z$  above, we have

$$f_Z(z) = \begin{cases} \frac{1}{C_d} \left[ \frac{z}{l\sqrt{d_A^2-z^2}} \Pr(Y = d_A) + \int_{d_A}^{\sqrt{z^2+l^2}} \frac{z}{l\sqrt{y^2-z^2}} f_Y(y) dy \right] & d_m \leq z < d_A \\ \frac{1}{C_d} \left[ \int_z^{\sqrt{z^2+l^2}} \frac{z}{l\sqrt{y^2-z^2}} f_Y(y) dy \right] & z \geq d_A \end{cases}$$

The constant  $C_d$  is chosen so that  $\Pr(Z \geq d_m) = 1$ . Therefore

$$C_d = \frac{1}{l} \sqrt{d_A^2 - d_m^2} \Pr(Y = d_A) + \int_{d_A}^{\sqrt{d_m^2+l^2}} \frac{\sqrt{y^2 - d_m^2}}{l} f_Y(y) dy + \int_{\sqrt{d_m^2+l^2}}^{\infty} f_Y(y) dy$$

**Case ii)**  $d_m \leq \sqrt{d_m^2 + l^2} < d_A$

Region i)  $d_m \leq z < \sqrt{d_A^2 - l^2} < d_A$

$$\Pr(Z \geq z) = \frac{1}{C_d} \left[ \Pr(Y = d_A) + \int_{d_A}^{\infty} f_Y(y) dy \right]$$

$$f_Z(z) = 0$$

Region ii)  $d_m \leq \sqrt{d_A^2 - l^2} \leq z < d_A$

$$\Pr(Z \geq z) = \frac{1}{C_d} \left[ \frac{1}{l} \sqrt{d_A^2 - z^2} \Pr(Y = d_A) + \int_{d_A}^{\sqrt{z^2+l^2}} \frac{\sqrt{y^2 - z^2}}{l} f_Y(y) dy + \int_{\sqrt{z^2+l^2}}^{\infty} f_Y(y) dy \right]$$

$$f_Z(z) = \frac{1}{C_d} \left[ \frac{z}{l\sqrt{d_A^2 - z^2}} \Pr(Y = d_A) + \int_{d_A}^{\sqrt{z^2+l^2}} \frac{z}{l\sqrt{y^2 - z^2}} f_Y(y) dy \right]$$

Region iii)  $z \geq d_A$

$$\Pr(Z \geq z) = \frac{1}{C_d} \left[ \int_z^{\sqrt{z^2+l^2}} \frac{\sqrt{y^2 - z^2}}{l} f_Y(y) dy + \int_{\sqrt{z^2+l^2}}^{\infty} f_Y(y) dy \right]$$

$$f_Z(z) = \frac{1}{C_d} \int_z^{\sqrt{z^2+l^2}} \frac{z}{l\sqrt{y^2 - z^2}} f_Y(y) dy$$

Combining the results for  $f_Z$  in this case, we have

$$f_Z(z) = \begin{cases} 0 & d_m \leq z < \sqrt{d_A^2 - l^2} \\ \frac{1}{C_d} \left[ \frac{z}{l\sqrt{d_A^2 - z^2}} \Pr(Y = d_A) + \int_{d_A}^{\sqrt{z^2 + l^2}} \frac{z}{l\sqrt{y^2 - z^2}} f_Y(y) dy \right] & \sqrt{d_A^2 - l^2} \leq z < d_A \\ \frac{1}{C_d} \int_z^{\sqrt{z^2 + l^2}} \frac{z}{l\sqrt{y^2 - z^2}} f_Y(y) dy & z \geq d_A \end{cases}$$

The constant  $C_d$  is chosen so that  $\Pr(Z \geq d_m) = 1$ . Therefore,

$$C_d = \Pr(Y = d_A) + \int_{d_A}^{\infty} f_Y(y) dy = 1$$

(In this case, no metastases can escape detection.) Comparing both cases, we see that we can combine the expressions for  $f_Z(z)$  to obtain

$$(4.10) \quad f_Z(z) = \begin{cases} \frac{1}{C_d} \left[ \frac{z}{l\sqrt{d_A^2 - z^2}} \Pr(Y = d_A) + \int_{d_A}^{\sqrt{z^2 + l^2}} \frac{z}{l\sqrt{y^2 - z^2}} f_Y(y) dy \right], & \max\{d_m, \sqrt{d_A^2 - l^2}\} \leq z < d_A \\ \frac{1}{C_d} \int_z^{\sqrt{z^2 + l^2}} \frac{z}{l\sqrt{y^2 - z^2}} f_Y(y) dy, & z \geq d_A \end{cases}$$

and 0 otherwise where

$$C_d = \frac{1}{l} \sqrt{d_A^2 - \max\{d_m^2, d_A^2 - l^2\}} \Pr(Y = d_A) + \int_{d_A}^{\max\{\sqrt{d_m^2 + l^2}, d_A\}} \frac{\sqrt{y^2 - d_m^2}}{l} f_Y(y) dy \\ + \int_{\max\{\sqrt{d_m^2 + l^2}, d_A\}}^{\infty} f_Y(y) dy.$$

## Perils of Maximum Likelihood Estimation

### 5.1. Infinite Likelihood Using Volume Data and Expected Diameters

In using the method of maximum likelihood with observed volumes (as for Patient A) to estimate parameters in the full model, we observe that the maximum likelihood approaches infinity in at least two cases.

Case i) We have already remarked in (3.9) that taking the limit as  $\gamma_0 \rightarrow 0^+$  of the Full model when  $e^{\gamma_1 R} > m$  causes the pdf to converge pointwise to

$$p(x) = \begin{cases} \frac{1}{\gamma_1} (C_9 x)^{-1} \left(\frac{x}{e^{\gamma_1 R}}\right)^{1/(\rho\gamma_1)} \left(e^{\sigma Q} - e^{-\frac{Q}{\rho}}\right), & m \leq x < e^{\gamma_1 R} \\ \infty & x = e^{\gamma_1 R} \end{cases}$$

where

$$C_9 = \rho \left(e^{\sigma Q} - e^{-\frac{Q}{\rho}}\right) \left(1 - \left(\frac{m}{e^{\gamma_1 R}}\right)^{1/(\gamma_1 \rho)}\right) + \frac{1}{\sigma} (e^{\sigma Q} - 1) + \rho \left(e^{-\frac{Q}{\rho}} - 1\right).$$

Thus the Full model likelihood,  $\prod_{i=1}^n p(x_i)$ , will become infinite if we set  $e^{\gamma_1 R} = x_n$  (or  $\gamma_1 = \frac{\log(x_n)}{R}$ ) and then take the limit as  $\gamma_0 \rightarrow 0^+$ .

Case ii) We have also noted in (3.14) that setting  $\rho = b_0/\gamma_0$  and  $\sigma = a_0\gamma_0$  and taking the limit as  $\gamma_0 \rightarrow \infty$  will cause the Full model to converge pointwise to

$$p(x) = \begin{cases} \infty & x = e^{\gamma_1 R} \\ \frac{ab_0 \left(\frac{e^{\gamma_1 R}}{x}\right)^a}{a+b_0} & x > e^{\gamma_1 R} \end{cases}$$

So, in this case, the Full model likelihood,  $\prod_{i=1}^n p(x_i)$ , will become infinite if we set  $\gamma_1 = \frac{\log(x_1)}{R}$ ,  $\rho = b_0/\gamma_0$ , and  $\sigma = a_0\gamma_0$  and then take the limit as  $\gamma_0 \rightarrow \infty$ . In reality, the lower limits of the cell cycle duration place an upper bound on  $\gamma_0$ , but it is still

instructive to consider this case for what it can tell us about situations when  $\gamma_0$  is large.

In both of these cases, the likelihood is going to infinity because the distribution is changing from a continuous distribution to a mixed continuous/discrete distribution.

In Case i) (CSPT) the mixed distribution is

$$(5.1) \quad p(x) = \frac{1}{\gamma_1} (C_9 x)^{-1} \left( \frac{x}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)} \left( e^{\sigma Q} - e^{-\frac{Q}{\rho}} \right), \quad m \leq x < e^{\gamma_1 R}$$

$$\text{and } \Pr(x = e^{\gamma_1 R}) = (C_9)^{-1} \left[ \frac{1}{\sigma} (e^{\sigma Q} - 1) + \rho \left( e^{-\frac{Q}{\rho}} - 1 \right) \right]$$

$$\text{where } C_9 = \rho \left( e^{\sigma Q} - e^{-\frac{Q}{\rho}} \right) \left[ 1 - \left( \frac{m}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)} \right] + \frac{1}{\sigma} (e^{\sigma Q} - 1) + \rho \left( e^{-\frac{Q}{\rho}} - 1 \right).$$

See (3.11).

In Case ii) (IISMG) the mixed distribution is

$$(5.2) \quad p(x) = (x)^{-1} \frac{ab_0 \left( \frac{e^{\gamma_1 R}}{x} \right)^a}{a+b_0}, \quad e^{\gamma_1 R} < x < \infty$$

$$\text{and } \Pr(x = e^{\gamma_1 R}) = \frac{a}{a+b_0}.$$

See (3.2.2).

Because in these cases the Full model likelihood is approaching infinity, we change to the respective limiting submodels in order to estimate parameters.

**5.1.1. Maximum likelihood estimation for the CSPT model.** When maximizing the CSPT likelihood given observed metastasis volumes  $x_1, x_2, \dots, x_n$ , we see that if  $x_n > e^{\gamma_1 R}$ , the likelihood will be 0. On the other hand, if  $x_n < e^{\gamma_1 R}$ , the likelihood is

$$L = \prod_{i=1}^n \frac{(1-v) x_i^{-1} \left( \frac{x_i}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)}}{\gamma_1 \rho \left[ 1 - \left( \frac{m}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)} \right]}$$

which simplifies to

$$L = \prod_{i=1}^n \frac{(1-v) x_i^{-1} x_i^{1/(\gamma_1 \rho)} e^{-R/\rho}}{\gamma_1 \rho [1 - m^{1/(\gamma_1 \rho)} e^{-R/\rho}]}$$

and the log-likelihood becomes

$$\log L = n \log(1 - v) - \sum_{i=1}^n \log x_i + \frac{1}{\gamma_1 \rho} \sum_{i=1}^n \log x_i - n \frac{R}{\rho} - n \log(\gamma_1 \rho) - n \log \left[ 1 - m^{1/(\gamma_1 \rho)} e^{-R/\rho} \right].$$

The partial of  $\log L$  with respect to  $\gamma_1$  is

$$\frac{\partial}{\partial \gamma_1} \log L = \frac{-1}{\gamma_1^2 \rho} \sum_{i=1}^n \log x_i - \frac{n}{\gamma_1} - n \frac{\frac{1}{\rho \gamma_1} m^{1/(\gamma_1 \rho)} e^{-R/\rho} \log m}{1 - m^{1/(\gamma_1 \rho)} e^{-R/\rho}}$$

which we can see is negative. Thus  $\gamma_1$  will decrease toward the smallest possible value, which occurs as  $e^{\gamma_1 R}$  approaches  $x_n$  from above.

Although it will not be directly comparable to the previous likelihood, we can write the limiting likelihood as

$$L = v^k (1 - v)^{n-k} \prod_{i=1}^{n-k} \frac{x_i^{-1} \left( \frac{x_i}{x_n} \right)^{1/(\gamma_1 \rho)}}{\gamma_1 \rho \left[ 1 - \left( \frac{m}{x_n} \right)^{1/(\gamma_1 \rho)} \right]}$$

where  $k$  is the number of metastases of size  $x_n$  and

$$(5.3) \quad \gamma_1 = \log(x_n)/R.$$

The log-likelihood is then

$$\begin{aligned} \log L = & k \log v + (n - k) \log(1 - v) - \sum_{i=1}^{n-k} \log x_i + \sum_{i=1}^{n-k} \frac{1}{\gamma_1 \rho} \log \frac{x_i}{x_n} \\ & - (n - k) \log(\gamma_1 \rho) - (n - k) \log \left[ 1 - \left( \frac{m}{x_n} \right)^{1/(\gamma_1 \rho)} \right] \end{aligned}$$

Taking the partial with respect to  $v$  gives

$$\frac{\partial}{\partial v} \log L = \frac{k}{v} - \frac{n - k}{1 - v}$$

Setting this partial to zero and solving gives

$$(5.4) \quad v = \frac{k}{n},$$

which gives a maximum of  $L$  with respect to  $v$  because

$$\frac{\partial^2}{\partial v^2} \log L = -\frac{k}{v^2} - \frac{n-k}{(1-v)^2} < 0.$$

Because we have already fixed  $\gamma_1$ , we can find the optimal value for  $\rho$  by finding the optimal value for  $b_1 = \frac{1}{\gamma_1 \rho}$  where

$$\begin{aligned} \log L = & k \log v + (n-k) \log(1-v) - \sum_{i=1}^{n-k} x_i^{-1} + \sum_{i=1}^{n-k} b_1 \log \frac{x_i}{x_n} \\ & + (n-k) \log b_1 - (n-k) \log \left[ 1 - \left( \frac{m}{x_n} \right)^{b_1} \right]. \end{aligned}$$

The partial of the log-likelihood with respect to  $b_1$  is then

$$(5.5) \quad \frac{\partial}{\partial b_1} \log L = \sum_{i=1}^{n-k} \log \frac{x_i}{x_n} + \frac{n-k}{b_1} + (n-k) \frac{\left( \frac{m}{x_n} \right)^{b_1} \log \frac{m}{x_n}}{1 - \left( \frac{m}{x_n} \right)^{b_1}}.$$

Setting this partial to zero, we see that we must solve

$$(5.6) \quad g(b_1) = \frac{\sum_{i=1}^{n-k} \log \frac{x_n}{x_i}}{n-k}$$

where

$$g(b_1) = \frac{1}{b_1} - \frac{\log \frac{x_n}{m}}{\left( \frac{x_n}{m} \right)^{b_1} - 1}$$

We note that

$$\lim_{b_1 \rightarrow -\infty} g(b_1) = \log \frac{x_n}{m} = \log x_n - \log m$$

and that

$$\lim_{b_1 \rightarrow \infty} g(b_1) = 0.$$

Although  $g$  is undefined at 0, if we write  $g$  as

$$g(b_1) = \frac{\left(\frac{x_n}{m}\right)^{b_1} - 1 - b_1 \log \frac{x_n}{m}}{b_1 \left[\left(\frac{x_n}{m}\right)^{b_1} - 1\right]} = \frac{\exp\left(b_1 \log \frac{x_n}{m}\right) - 1 - b_1 \log \frac{x_n}{m}}{b_1 \left[\exp\left(b_1 \log \frac{x_n}{m}\right) - 1\right]}$$

and then use the Maclaurin series for  $e^x$ , we have

$$\begin{aligned} g(b_1) &= \frac{\sum_{i=2}^{\infty} \frac{1}{i!} \left(b_1 \log \frac{x_n}{m}\right)^i}{b_1 \sum_{i=1}^{\infty} \frac{1}{i!} \left(b_1 \log \frac{x_n}{m}\right)^i} \\ &= \frac{\log\left(\frac{x_n}{m}\right) \sum_{i=0}^{\infty} \frac{1}{(i+2)!} \left(b_1 \log \frac{x_n}{m}\right)^i}{\sum_{i=0}^{\infty} \frac{1}{(i+1)!} \left(b_1 \log \frac{x_n}{m}\right)^i} \end{aligned}$$

From this it is apparent that  $g$  has a removable discontinuity at  $b_1 = 0$  where

$$\lim_{b_1 \rightarrow 0} g(b_1) = \frac{1}{2} \log \frac{x_n}{m} = \frac{1}{2} (\log x_n - \log m).$$

Because  $m$  is the smallest detectable metastasis size and  $x_n$  is the size of the largest observed metastasis, we have

$$\lim_{b_1 \rightarrow \infty} g(b_1) = 0 < \frac{\sum_{i=1}^{n-k} \log \frac{x_n}{x_i}}{n-k} < \frac{\sum_{i=1}^{n-k} \log \frac{x_n}{m}}{n-k} = \log \frac{x_n}{m} = \lim_{b_1 \rightarrow -\infty} g(b_1)$$

and therefore (5.6) has at least one solution. In Appendix C, we show that  $g$  is a decreasing function so that the solution to (5.6) is unique. The solution,  $b_1$ , will only be positive if

$$\log \frac{x_n}{m} > 2 \frac{\sum_{i=1}^{n-k} \log \left(\frac{x_n}{x_i}\right)}{n-k},$$

or, equivalently,

$$\frac{1}{n-k} \sum_{i=1}^{n-k} \log x_i > \frac{\log m + \log x_n}{2}$$

5.1.1.1. **Complete Suppression by Primary Using Profile Data.** When we convert profile diameters to their expected diameters and use these to estimate parameters, the maximum likelihood equations for  $\gamma_1$  and  $v$  become

$$(5.7) \quad \gamma_1 = \frac{3}{R} \log \frac{y_n}{c_d}$$

where  $c_d$  is the diameter of a single cell, and

$$(5.8) \quad v = \frac{k}{n}.$$

To find  $b_1$ , we must solve

$$(5.9) \quad \tilde{g}(b_1) = \frac{\sum_{i=1}^{n-k} \log \frac{y_n}{y_i}}{n-k}$$

where

$$\tilde{g}(b_1) = \frac{1}{3b_1} - \frac{\log \frac{y_n}{d_m}}{\left(\frac{y_n}{d_m}\right)^{3b_1} - 1}.$$

Then  $\gamma_1$  and  $b_1$  can be used to recover  $\rho$  and  $\gamma_1$  and, finally,  $\rho$  and  $v$  can be used with (3.13) to determine the relationship between  $\sigma$  and  $Q$ .

**5.1.2. Maximum likelihood estimation for the IISMG model.** In (5.2), if we set  $v = \Pr(x = A) = \frac{a}{a+b_0}$ , then  $1 - v = \frac{b_0}{a+b_0}$  and  $p(x)$  for the IISMG model becomes

$$p(x) = (1 - v) ax^{-1} \left(\frac{A}{x}\right)^a, \quad A < x < \infty$$

In order to maximize the likelihood given observed metastasis volumes  $x_1, x_2, \dots, x_n$ , we see that if  $x_1 < A$ , the likelihood will be zero. On the other hand, if  $x_1 > A$ , the likelihood will be

$$L = (1 - v)^n a^n \prod_{i=1}^n (x_i)^{-1} \left(\frac{A}{x_i}\right)^a$$

and the log-likelihood will be

$$\log L = n \log(1 - v) + n \log a + na \log A - n \log x_i - a \sum_{i=1}^n \log x_i$$

Because  $L$  is an increasing function of  $A$ , the maximum value of  $L$  occurs as  $A$  approaches  $x_1$  from below. Although it will not be directly comparable to the previous likelihood, we can write the limiting likelihood as

$$L = v^k (1 - v)^{n-k} a^{n-k} \prod_{i=k+1}^n x_i^{-1} \left( \frac{x_1}{x_i} \right)^a$$

where  $k$  is the number of metastases of size  $x_1$ . The log-likelihood is then

$$\log L = k \log v + (n - k) \log(1 - v) + (n - k) \log a + (n - k) a \log x_1 - (a + 1) \sum_{i=k+1}^n \log x_i.$$

The partial of  $\log L$  with respect to  $v$  is  $\frac{\partial}{\partial v} \log L = \frac{k}{v} - \frac{(n-k)}{1-v}$ . If we set this partial equal to zero, we find that

$$(5.10) \quad v = \frac{k}{n}$$

which is clearly a point of global max of  $L$  for  $v \in [0, 1]$ .

The partial of  $\log L$  with respect to  $a$  is

$$\frac{\partial}{\partial a} \log L = \frac{n - k}{a} + (n - k) \log x_1 - \sum_{i=k+1}^n \log x_i.$$

If we set this partial to zero, we find that

$$(5.11) \quad \frac{1}{a} = \bar{x}$$

where

$$\bar{x} = \frac{\sum_{i=k+1}^n \log \frac{x_i}{x_1}}{n - k},$$

Which is again a global maximum of  $L$  for  $a > 0$ . Knowing  $a$ , we can solve

$$v = \frac{a}{a + b_0}$$

to obtain

$$(5.12) \quad b_0 = \frac{1-v}{v}a = \frac{n-k}{k}a = \frac{n-k}{k\bar{x}} = \frac{(n-k)^2}{k \sum_{i=k+1}^n \log \frac{x_i}{x}}$$

**5.1.2.1. Instantaneous Infinite Shedding and Metastasis Growth Using Profile Data.** When we convert profile diameters to their expected diameters, and use these to estimate parameters, the IISMG likelihood is

$$L = v^k (1-v)^{n-k} a^{n-k} \prod_{i=k+1}^n y_i^{-1} \left( \frac{y_1}{y_i} \right)^{3a}$$

where  $k$  is the number of metastases of size  $y_1$ . The log-likelihood is then

$$\log L = k \log v + (n-k) \log(1-v) + (n-k) \log a + 3(n-k)a \log y_1 - (3a+1) \sum_{i=k+1}^n \log y_i.$$

We again find that the maximum likelihood estimate for  $v$  is

$$(5.13) \quad v = \frac{k}{n}.$$

The maximum likelihood estimate for  $a$  is

$$(5.14) \quad a = \frac{n-k}{3 \sum_{i=k+1}^n \log \frac{y_i}{y_1}}$$

and the maximum likelihood estimate for  $b_0$  is

$$(5.15) \quad b_0 = \frac{(n-k)^2}{3k \sum_{i=k+1}^n \log \frac{y_i}{y_1}}.$$

## 5.2. CSPT Using the Method of True Diameters

Assuming that  $\sqrt{d_m^2 + l^2} < d_A$ , when we apply the Method of True Diameters (4.9) with the Complete Suppression by Primary Tumor model (3.11), we obtain

$$(5.16) \quad f_Z(z) = \begin{cases} \frac{1}{C_d} \frac{3b_1(1-v)}{d_A^{3b_1} - d_m^{3b_1}} \int_z^{\sqrt{z^2+l^2}} \frac{z}{l\sqrt{y^2-z^2}} y^{3b_1-1} dy, & d_m \leq z < \sqrt{d_A^2 - l^2} \\ \frac{1}{C_d} \left[ \frac{3b_1(1-v)}{d_A^{3b_1} - d_m^{3b_1}} \int_z^{d_A} \frac{z}{l\sqrt{y^2-z^2}} y^{3b_1-1} dy + \frac{zv}{l\sqrt{d_A^2-z^2}} \right], & \sqrt{d_A^2 - l^2} \leq z < d_A \end{cases}$$

and 0 otherwise, where

$$C_d = 1 - \frac{3b_1(1-v)}{d_A^{3b_1} - d_m^{3b_1}} \int_{d_m}^{\sqrt{d_m^2+l^2}} \left( \frac{l - \sqrt{y^2 - d_m^2}}{l} \right) y^{3b_1-1} dy.$$

We note that because of the term  $\frac{zv}{l\sqrt{d_A^2-z^2}}$  in  $f_Z$ , if  $z_n$  is the largest profile diameter, then  $f_Z(z_n)$  will approach infinity as  $d_A \downarrow z_n$ , and so will the likelihood.

Because of this, we expect that if we use the non-rounded likelihood with the Full model, then  $\gamma_0$  will tend to zero so that the Full model will approach the CSPT model and if we then compute the CSPT likelihood,  $d_A$  will decrease toward the largest profile diameter,  $z_n$ , and the CSPT likelihood will tend to infinity.

Although the term  $\frac{zv}{l\sqrt{d_A^2-z^2}}$  approaches infinity as  $d_A \downarrow z$ , it does not become a point mass at  $z$  because

$$\int_z^{d_A} \frac{sv}{l\sqrt{d_A^2-s^2}} ds = \frac{v}{l} \int_z^{d_A} \frac{s}{\sqrt{d_A^2-s^2}} ds = \frac{v}{l} \sqrt{d_A^2 - z^2},$$

which tends to zero as  $d_A \downarrow z$ . Note that if we use the method of True Diameters with rounding, the likelihood will not diverge to infinity because we will be using actual probabilities which are bounded between 0 and 1.

### 5.3. IISMG Using the Method of True Diameters

When we apply the Method of True Diameters (4.10) with the Instantaneous Infinite Shedding and Metastasis Growth model (3.2.2), we have

$$(5.17) \quad f_Z(z) = \begin{cases} \frac{1}{C_d} \left[ \frac{z}{l\sqrt{d_A^2 - z^2}} \frac{a}{a+b_0} + \int_{d_A}^{\sqrt{z^2+l^2}} \frac{z}{l\sqrt{y^2-z^2}} 3y^{-1} \frac{ab_0}{a+b_0} \left(\frac{d_A}{y}\right)^{3a} dy \right], & \max\{d_m, \sqrt{d_A^2 - l^2}\} \leq z < d_A \\ \frac{1}{C_d} \left[ \int_z^{\sqrt{z^2+l^2}} \frac{z}{l\sqrt{y^2-z^2}} 3y^{-1} \frac{ab_0}{a+b_0} \left(\frac{d_A}{y}\right)^{3a} dy \right], & z \geq d_A \end{cases}$$

and 0 otherwise where

$$C_d = \frac{1}{l} \sqrt{d_A^2 - \max\{d_m^2, d_A^2 - l^2\}} \frac{a}{a+b_0} + \int_{d_A}^{\max\{\sqrt{d_m^2+l^2}, d_A\}} \frac{\sqrt{y^2-d_m^2}}{l} 3y^{-1} \frac{ab_0}{a+b_0} \left(\frac{d_A}{y}\right)^{3a} dy \\ + \int_{\max\{\sqrt{d_m^2+l^2}, d_A\}}^{\infty} 3y^{-1} \frac{ab_0}{a+b_0} \left(\frac{d_A}{y}\right)^{3a} dy.$$

We note that because of the term  $\frac{z}{l\sqrt{d_A^2 - z^2}} \frac{a}{a+b_0}$  in  $f_Z$ , if  $z_i$  is a profile diameter, then  $f_Z(z_i)$  will approach infinity as  $d_A \downarrow z_i$ , and, so long as all profiles smaller than  $z_i$  are greater than  $\sqrt{z_i^2 - l^2}$ , the likelihood will also approach infinity. Therefore, if  $z_1$  is the smallest observed profile diameter, and if  $z_2, z_3, \dots, z_i$  are observed profile diameters between  $z_1$  and  $\sqrt{z_1^2 + l^2}$ , the likelihood will approach infinity when  $d_A$  approaches any of  $z_1, z_2, z_3, \dots$ , or  $z_i$  from above. Thus, unless we place a bound on  $\gamma_0$ , if we start with the non-rounded Full model likelihood and allow  $\gamma_0$  to approach infinity as  $\rho$  tends to zero and  $\theta\beta$  tends to infinity, there will be multiple values that  $d_A$  can approach that will create unbounded likelihood. Although the term  $\frac{z}{l\sqrt{d_A^2 - z^2}} \frac{a}{a+b_0}$  approaches infinity as  $d_A \downarrow z$ , it does not become a point mass at  $z$  because

$$\int_z^{d_A} \frac{s}{l\sqrt{d_A^2 - s^2}} \frac{a}{a+b_0} ds = \frac{a}{l(a+b_0)} \int_z^{d_A} \frac{s}{\sqrt{d_A^2 - s^2}} ds = \frac{a}{l(a+b_0)} \sqrt{d_A^2 - z^2}$$

which tends to zero as  $d_A \downarrow z$ . Again, we note that if we use the method of True Diameters with rounding, the likelihood will not diverge to infinity.

## CHAPTER 6

### Data Analysis and Parameter Estimation

In this chapter we present the results of searching for optimal parameters for our models using data from Patient A, Protocol 17, and Protocol 10. Because the likelihoods and  $L^2$  norms differ depending on the type of data, the method used, and the model selected, we adopt the following conventions: (1) We use “ $L$ ” to designate a normalized negative log-likelihood (NNLL) and “ $\Delta$ ” to designate an  $L^2$  distance between empirical and theoretical cdfs, (2) Subscripts of  $V, E, T$ , and  $R$  designate “Volumes”, “Expected Diameters”, “True Diameters”, and “True Diameters with Rounding”, respectively, and (3) Subscripts of  $F, H, S, G, C$ , and  $I$  designate the “Full”, “Homogeneous”, “HSSL”, “IISMG”, “CSPT”, or “Instantaneous Seeding” model, respectively. For example,  $L_{VF}$  indicates the normalized negative log likelihood

$$L_{VF}(A, M, a, b_0, b_1) := -\frac{1}{n} \prod_{k=1}^n \log [p(x_k)]$$

where  $x_k$  is the  $k$ th recorded volume,  $n$  is the total number of metastasis observed in the given location, and  $p$  is the full model pdf given in (2.22). Similarly,  $\Delta_{VF}$  would indicate the  $L_2$  distance between the empirical and theoretical cdfs given by

$$\Delta_{VF}^2 = \sum_{k=1}^{n+1} \int_{x_{k-1}}^{x_k} \left[ P(u) - \frac{1}{n} \sum_{i=1}^{k-1} n_j \right]^2 du$$

where  $x_0 = d_m$ ,  $x_k, 1 \leq k \leq n$  is the  $k$ th recorded volume,  $x_{n+1} = M$ , and  $P(u)$  is the cdf corresponding to the pdf in (2.22).

When using the method of True Diameters with Rounding, we can compute the  $L^2$  distance between the empirical and theoretical cdfs. However, in order to make the  $L^2$  distance computation account for rounding, we can assume that the true cross-sectional

diameter whose reported value after rounding was  $k$  mm is uniformly distributed on the interval  $[k - 0.5, k + 0.5]$ . We can then compute the distance between the model-based theoretical cdf and the corresponding piecewise linear continuous empirical cdf. When we compute an  $L^2$  distance between the theoretical distribution and the piecewise linear continuous empirical cdf, we designate it with an additional subscript “P”. For example,  $\Delta_{RF}$  designates the  $L^2$  distance computed between the full pdf for true diameters and the empirical cdf given by

$$\Delta_{RF}^2 = \sum_{k=1}^{n+1} \int_{z_{k-1}}^{z_k} \left[ F_Z(z) - \frac{1}{n} \sum_{i=1}^{k-1} n_i \right]^2 dz$$

where  $z_0 = d_m$ ;  $z_k$ ,  $1 \leq k \leq n$ , is the  $k$ th observed profile diameter;  $z_{n+1} = d_M$ ; and  $F_Z$  is the cdf corresponding to the pdf obtained by combining (4.3) with (4.7). On the other hand,  $\Delta_{RFP}$  designates the  $L^2$  distance computed between the full pdf for true diameters and the piecewise linear continuous empirical cdf. Specifically,

$$\Delta_{RFP}^2 = \sum_{k=1}^{n+2} \int_{B_{k-1}}^{B_k} [F_Z(z) - G(z)]^2 dz$$

where  $B_0 = 0$ ,  $B_k = k - 0.5$ ,  $1 \leq k \leq n + 1$ ;  $B_{n+2} = d_M$ ;  $G(z)$  is the piecewise linear continuous empirical cdf; and  $F_Z$  is again the cdf corresponding to the pdf obtained by combining (4.3) with (4.7).

## 6.1. Patient A

The MSKCC database contains information from several hundred MSKCC breast cancer patients. For the purposes of fitting our model, the database was searched for candidates for whom (1) treatment included excision of the primary tumor, (2) whole body PET/CT scans were available, (3) a large number of metastases were observed in a single site, (4) primary tumor volume at presentation was available, (5) the volumes of the primary tumor and the secondary tumors were measured at different times and (6) no change of treatment occurred between diagnosis and the time of metastasis surveying. It was found that only one patient, whom we call Patient A, met these criteria. On 4/1/96, at age 74, Patient A was diagnosed with stage III estrogen receptor positive breast cancer. Shortly thereafter, the primary tumor of volume  $10.3 \text{ cm}^3$  was resected and Patient A began an adjuvant hormonal therapy with tamoxifen. On 4/6/04, at age 82, 37 detectable metastases were discovered in Patient A's bone, lung, lymph, and soft tissues. The sizes of these metastases were surveyed by PET/CT imaging. Thus  $U = V = 74$ ,  $W = 82$ , and  $S = 10.3 \text{ cm}^3$ . The prevalent metastatic site was the skeletal system which was found to contain 31 metastases. The sizes of these bone metastases were 26, 31, 31, 31, 33, 34, 38, 47, 49, 51, 52, 54, 54, 55, 65, 67, 73, 78, 78, 81, 84, 87, 98, 101, 114, 139, 142, 172, 196, 213, and 354 voxels. As it would be unlikely for two or more metastases to have exactly the same size, we thought it reasonable to smooth the data by distributing equal sizes uniformly over their respective bins. For example, the three metastases of size 31 voxels must have contained between 30 and 32 voxels and so we spread them uniformly over that interval by assuming that they had sizes 30.5, 31, and 31.5 voxels. The size of a voxel was estimated to be  $0.06485 \text{ cm}^3$  and this was used to convert the voxel counts to the volumes 1.69, 1.98, 2.01, 2.04, 2.14, 2.2, 2.46, 3.05, 3.18, 3.31, 3.37, 3.48, 3.52, 3.57, 4.22, 4.34, 4.73, 5.04, 5.08, 5.25, 5.45, 5.64, 6.36, 6.55, 7.39, 9.01, 9.21, 11.15, 12.71, 13.81, and  $22.96 \text{ cm}^3$ .

For the scanner used, the threshold of measurable volumes was reported as  $5 \times 10^8$  cells. Assuming a volume of  $1 \times 10^{-9} \text{ cm}^3$  for one tumor cell, we obtained a minimum observable volume of  $m = 0.5 \text{ cm}^3$ .

We used the maximum likelihood method to estimate the parameters for the Full model. To accomplish this, we used a Matlab routine based on the non-linear simplex method to minimize the negative log-likelihood. Because of the discontinuity in the pdf at  $A$ , we searched exhaustively with  $A$  restricted first below  $m$ , then between  $m$  and the smallest observed metastasis size, then between successive metastasis sizes, and finally with  $A$  restricted between the largest metastasis size and  $M$ . We then examined the minimum NNLL,  $L_{VF}$ , found in each sub-interval. In every case, we applied the constraints  $a, b_0, b_1, T > 0$  and  $M > x_n$  (where  $x_n$  was the maximum observed metastasis volume) as we searched.

When  $A$  was restricted to lie between the two smallest observed metastasis sizes (1.69 and 1.98  $\text{cm}^3$ ) or between the smallest observable metastasis size and smallest observed metastasis size (0.5 and 1.69  $\text{cm}^3$ ) the likelihood degenerated into the Instantaneous Infinite Shedding and Metastasis Growth (IISMG) Model given in (5.2). This gave  $A = 1.69 \text{ cm}^3$  and using (5.10), (5.11), and (5.12), we found that  $v = \frac{1}{31}$ ,  $a = 0.9705$  and  $b_0 = 29.114$  with  $L_{VG}^* = 2.6443$ .

Comparing the theoretical cdf for the IISMG model to the empirical cdf, we found that the  $L^2$  norm of the difference between the two was  $\Delta_{VG} = 0.5096$ .

When  $A$  was restricted to lie between the two largest metastasis sizes (13.81 and 22.96) the likelihood degenerated into the Complete Suppression by Primary Tumor Model (CSPT) given in (5.1). This gave  $A = 22.96$  and using (5.3), (5.4), and (5.6), we found that  $\gamma_1 = 2.9770$ ,  $v = \frac{1}{31}$ ,  $b_1 = 0.2052$ , and  $\rho = 1.6370$  with  $L_{VC}^* = 2.8376$ . Using equation (5.6), we solved numerically for  $Q$  as a function  $\sigma$  and obtained the graph shown in Figure 1.

Comparing the theoretical cdf for the CSPT model to the empirical cdf, we found that the  $L^2$  norm of the difference between the two was  $\Delta_{VC} = 0.6351$  Plots of the

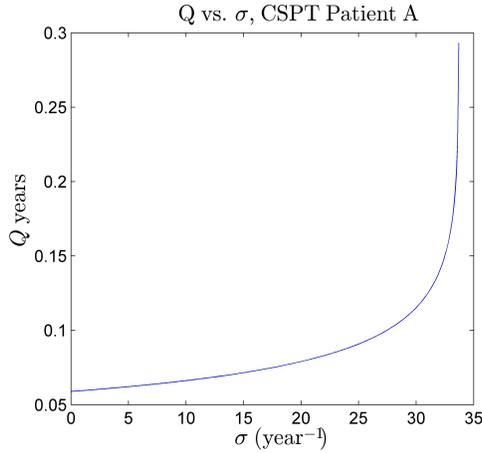


FIGURE 1.  $Q = V - T$  as a function of  $\sigma$  for the CSPT model applied to Patient A.

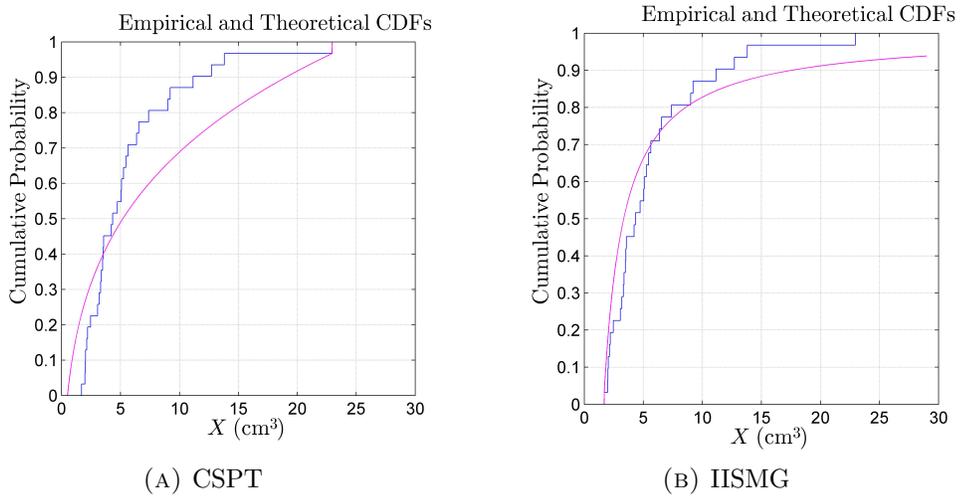


FIGURE 2. Graphs of empirical and theoretical cdfs for the likelihood maximizing parameters for the CSPT and IISMG models applied to Patient A.

empirical and theoretical cdfs for the likelihood maximizing CSPT and IISMG models are shown in Figure 2. That the fits are both visually and numerically poor is not surprising because the model degeneration into the CSPT and IISMG cases is driven more by the transition from a continuous distribution to a mixed discrete/continuous distribution than by an actual increase in likelihood.

After the two cases where infinite likelihood was anticipated because of the transition of the distribution from continuous to mixed discrete-continuous, we found that

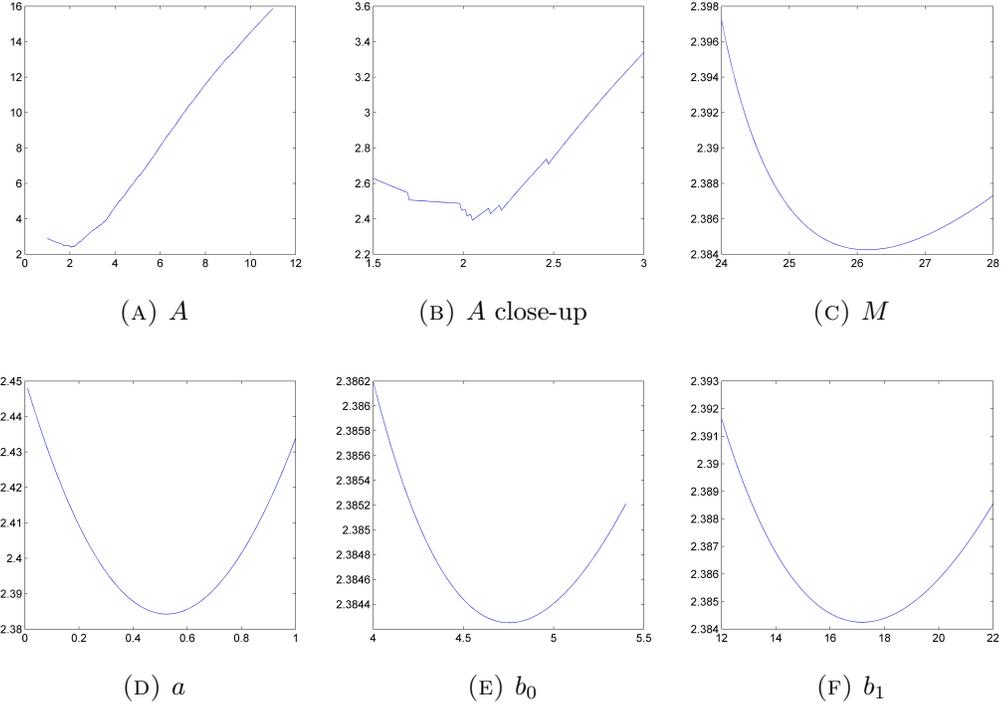


FIGURE 3. Graphs of NLL profile functions for likelihood maximizing parameters with  $A = 2.040$ ,  $M = 26.142$ ,  $a = 0.5232$ ,  $b_0 = 4.7573$ , and  $b_1 = 17.1755$ .

the lowest value of the NLL was  $L_{VF}^* = 2.3844$ , which was obtained when  $A = 2.040$ ,  $M = 26.142$ ,  $a = 0.5232$ ,  $b_0 = 4.7537$ , and  $b_1 = 17.1755$ . Recovering our original parameters, we compute that  $\rho = 0.0218$  years,  $T = 73.7$  years,  $\gamma_0 = 9.6649 \text{ year}^{-1}$ ,  $\gamma_1 = 2.6750 \text{ year}^{-1}$ ,  $\beta = 87.3674 \text{ year}^{-1}$  and  $\theta = 0.0579$ . The plots in Figure 3 show the likelihood as each parameter is varied independently and help us to verify that we have indeed reached a minimum value.

In order to quantify the fit of the model to the data, we computed the  $L^2$  norm of the difference between the empirical and theoretical cdfs and found it to be  $\Delta_{VF} = 0.1885$ , a marked improvement over .5096 and .6351 obtained from the IISMG and CSPT. The graph of the empirical versus theoretical cdfs for the maximum likelihood estimates is shown in Figure 5a and shows a much better fit than when the CSPT and IISMG models were used. By comparison, we compute that the minimum  $L^2$  norm of  $\Delta_{VF}^* = 0.1014$  is achieved when  $A = 3.5455$ ,  $M = 30.7089$ ,  $a = 1.1249$ ,  $b_0 = 4.7537$

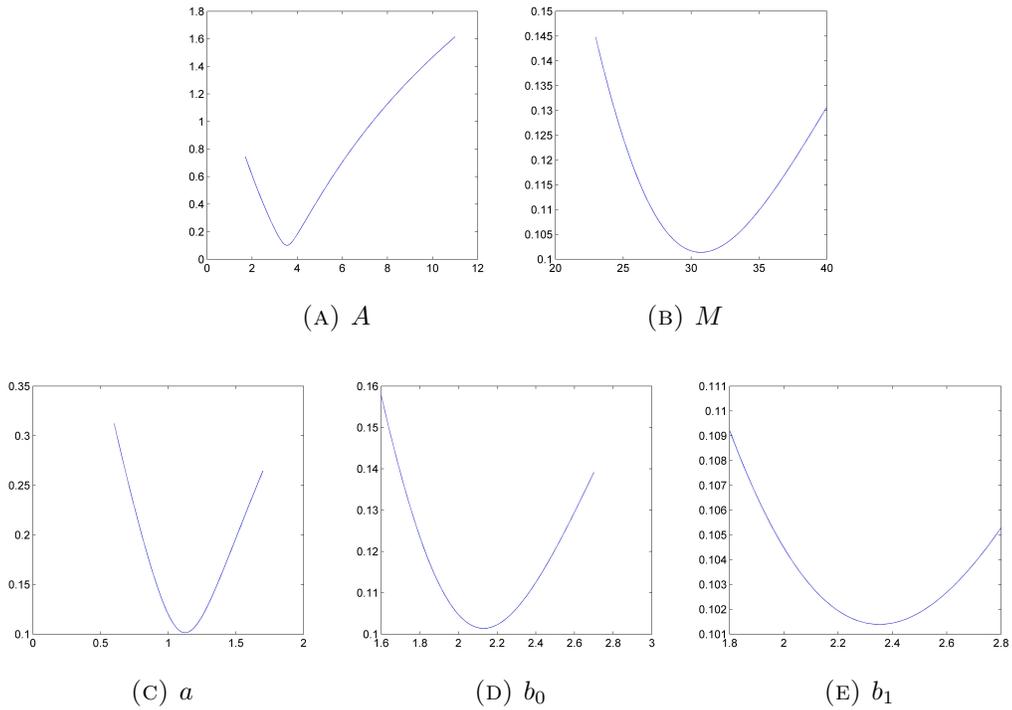


FIGURE 4. Plots of the  $L^2$  norm profile functions for the difference between the empirical and theoretical cdfs for the full model for the  $L^2$  norm minimizing parameter set:  $A = 3.5455$ ,  $M = 30.7089$ ,  $a = 1.1249$ ,  $b_0 = 2.1295$ , and  $b_1 = 2.3515$ .

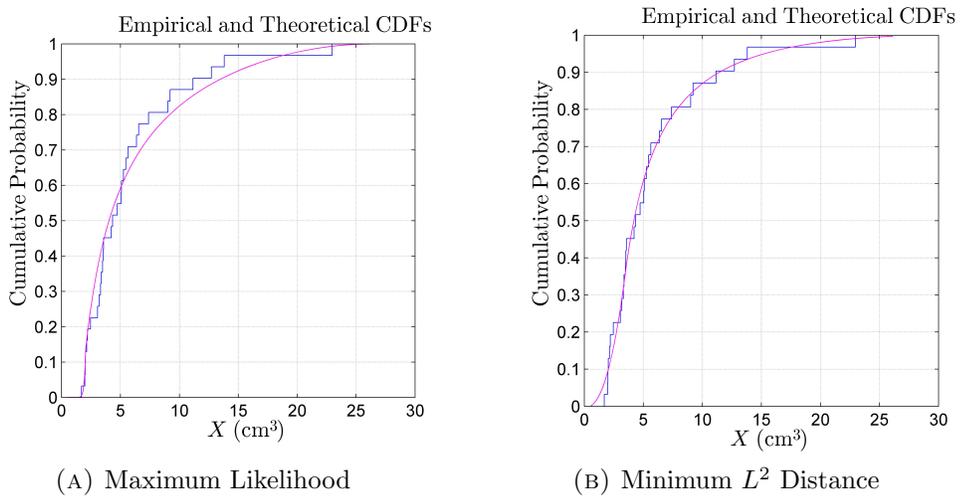


FIGURE 5. Graphs of empirical and theoretical cdfs for (a) the likelihood maximizing and (b)  $L^2$  distance minimizing parameters for the Full model applied to Patient A.

and  $b_1 = 2.3515$ . The plots in Figure 4 show the  $L^2$  norm of the distance between the empirical and theoretical cdfs as each parameter is varied independently and help us to verify that we have indeed reached a minimum value. These minimizing values of the simplifying parameters correspond to the native parameter values  $\gamma_0 = 3.0301$ ,  $\gamma_1 = 2.7439$ ,  $T = 73.2875$ ,  $\rho = 0.1550$ ,  $\beta = 32.2592$ , and  $\theta = 0.1053$ . The normalized negative log-likelihood (NNLL) in this case is naturally somewhat worse at  $L_{VF} = 2.4740$ . The graph of the empirical versus theoretical cdfs for the  $L^2$  norm minimizing estimates is shown in Figure 5b and shows a visible improvement over the fit obtained by maximizing likelihoods.

The IISMG model is not biologically realistic in that it requires  $\gamma_0 \rightarrow \infty$  while the duration of the cell-cycle places an upper bound on the growth rates  $\beta$ ,  $\gamma_0$ , and  $\gamma_1$ . When we imposed this bound, that corresponded to a minimum of 20 hours to complete the cell cycle, we found that the maximum likelihood in the Full model occurred when  $A = 1.69$ ,  $M = 7.0145 \times 10^{17}$ ,  $a = 1.0028$ ,  $b_0 = 9.6947 \times 10^{14}$  and  $b_1 = 1.1101 \times 10^{17}$  with  $L_{VF} = 2.3661$ . These values indicated that the model was attempting to degenerate into the IISMG model, but the bound on cell cycle duration prevented that from happening and the NNLL stayed finite. The plot of empirical vs. theoretical cdf's was nearly identical to the plot obtained with the IISMG model. Through numerical experimentation, we found that the minimum value of  $\Delta_{VF}$  for  $A$  near 1.69 stayed lower than the minimum value of  $\Delta_{VF}$  obtained when  $A = 2.04$  (i.e.  $\Delta_{VF} = 2.3844$ ) until the minimum cell cycle duration was increased to 35.1 hours.

Although the maximum likelihood is better near  $A = 1.69$ , we reject this maximum because it is best explained by the transition of the full-model to a mixed discrete/continuous distribution and does not give a compelling fit to the data when the theoretical cdf is compared to the empirical cdf. Similarly, we reject the maximum likelihood obtained near  $A = 22.96$  because its fit to the data (as measured by the  $L^2$  norm of the distance between the theoretical and empirical cdfs) is even worse.

Our results are summarized in Table 1.

Patient A Parameters							
Model	NNLL	$L^2$ distance	$A$ cm <sup>3</sup>	$M$ cm <sup>3</sup>	$a$	$b_0$	$b_1$
CSPT	$L_{VC}^* = 2.8376$	$\Delta_{VC} = 0.6351$	22.96	22.96	$\infty$	$\infty$	0.2048
IISMG	$L_{VG}^* = 2.6443$	$\Delta_{VG} = 0.5096$	1.69	$\infty$	0.9705	29.114	$\infty$
Full	$L_{VF}^* = 2.3843$	$\Delta_{VF} = 0.1885$	2.040	26.142	0.5232	4.7573	17.1755
Full	$L_{VF} = 2.4740$	$\Delta_{VF}^* = 0.1014$	3.5455	30.7089	1.1249	2.1295	2.3515

Native Parameters						
Model	$\gamma_0$ year <sup>-1</sup>	$\gamma_1$ year <sup>-1</sup>	$T$ years	$\rho$ years	$\beta$	$\theta$
CSPT	0	2.9821	$0.0589 \leq Q \leq 0.3727$	1.6370	$\sigma$ free from 0 to 0.373	
IISMG	$\infty$	2.6496	NA	0	$\sigma = \infty$	
Full	9.6649	2.6749	73.736	0.02177	87.3637	0.05789
Full	3.0301	2.7439	73.2875	0.1550	32.3592	0.1053

TABLE 1. Summary of parameters for Patient A

## 6.2. Protocol 10

The data in Douglas' protocol 10 come from a 65-year-old man who died from an oat cell carcinoma of the lung. The primary tumor, which was *in situ* at autopsy, had a diameter of 105 mm. Thus  $W = V = 65$  and  $S = \frac{\pi(105\text{mm})^3}{6} \approx 606131\text{mm}^3$ . The metastases were measured by slicing the liver into  $l = 7\text{mm}$  sections. The counts of diameters of the observed circular profiles were reported as  $n_1 = 119, n_2 = 96, n_3 = 39, n_4 = 30, n_5 = 31, n_6 = 8, n_7 = 10, n_8 = 18, n_9 = 3, n_{10} = 16, n_{11} = 1, n_{12} = 9, n_{13} = 1, n_{14} = 2, n_{15} = 22, n_{16} = 1, n_{17} = 11, n_{18} = 2, n_{19} = 0, n_{20} = 4, n_{21} = 0, n_{22} = 1, n_{23} = 0, n_{24} = 0, \text{ and } n_{25} = 4$ , where  $n_i$  indicates the number of circular profiles observed to have diameter of  $i$  mm (rounded to the nearest mm.)

**6.2.1. Method of Expected Diameters.** As a first estimate, we used (4.6) to convert the observed profile diameters from Protocol 10 to expected true diameters and obtained the values shown in Table 2.

We then used this synthetic data to compute the NNLL

$$L_{EF}(d_M, a, b) = -\frac{1}{n} \sum_{k=1}^{25} n_k \log f_Y(E(k))$$

Count	119	96	39	30	31	8	10	18	3
Profile Diameter (mm)	1	2	3	4	5	6	7	8	9
Expected Diameter (mm)	3.724	4.202	4.826	5.546	6.333	7.167	8.035	8.927	9.839
Count	16	1	9	1	2	22	1	11	
Profile Diameter (mm)	10	11	12	13	14	15	16	17	
Expected Diameter (mm)	10.765	11.703	12.649	13.603	14.563	15.528	16.497	17.469	
Count	2	0	4	0	1	0	0	4	
Profile Diameter (mm)	18	19	20	21	22	23	24	25	
Expected Diameter (mm)	18.444	19.421	20.401	21.383	22.366	23.350	24.336	25.323	

TABLE 2. Expected diameters (in mm) of metastases for the patient in Protocol 10.

where  $n$  is the total number of observed profiles,  $k$  is the profile diameter (in mm),  $E(k)$  is the expected spherical diameter,  $n_k$  is the number of profiles of diameter  $k$ , and  $f_Y$  is the full model pdf given in (4.2).

The functional  $L_{EF}(d_M, a, b_0)$  was minimized over all  $d_M \geq E(25) = 25.323$  mm,  $a > 0$ , and  $b_0 > 0$ . This led to the estimates  $d_M = 26.808$  mm,  $b_0 = 0.78582$ , and  $a$  essentially zero with NNLL of  $L_{EF} = 2.8832$ . A value of  $a$  near 0 indicated that the Full model was degenerating into the Homogeneous model, so we minimized the functional  $L_{EH}(d_M, b_0)$  (which is the same as  $L_{EF}$  above but with  $f_Y(y)$  given by (4.4).) Fitting this model gave  $d_M = 26.8077$  mm and  $b_0 = 0.7858$  with a slightly higher NNLL value of  $L_{EH}^* = 2.8864$ . To ascertain that what we had found was indeed a minimum, we constructed profile graphs of the functional  $L_{EH}$  where all parameters except for the one of interest were fixed at their optimal values while the selected parameter was varied over the entire range of its admissible values, see Figure 6.

In order to assess the fit of our models to the data, we compared the empirical cdf for the transformed data to the theoretical cdf of tumor diameters obtained by integrating (4.2) for the Full model and (4.4) for the Homogeneous model. We quantified the fit

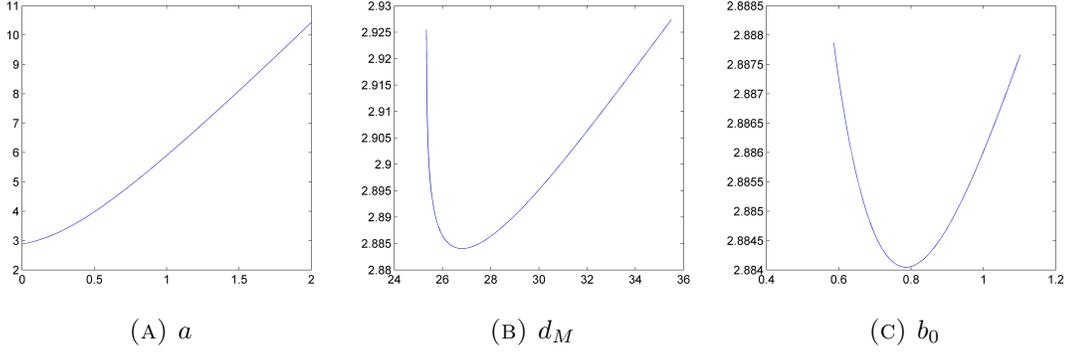


FIGURE 6. Graphs of NNLL profile functions for likelihood maximizing parameters. Plot (A) shows that parameter  $a$  is approaching 0 so that the Full model degenerates into the Homogeneous model. Plots (B) and (C) show  $d_M$  and  $b_0$  when the Homogeneous model is applied to the transformed data.

by computing the  $L^2$  norm of the difference between the empirical and theoretical cdfs. For example, to compare the theoretical cdf for the Full model to the empirical cdf, we computed  $\Delta_{EF}$  where

$$\Delta_{EF}^2 = \sum_{k=1}^{K+1} \int_{E_{k-1}}^{E_k} \left[ H(u) - \frac{1}{n} \sum_{j=1}^{k-1} n_j \right]^2 du.$$

Here  $E_0 := E(d_m)$ ;  $E_k = E(k)$ ,  $1 \leq k \leq K$ ;  $E_{k+1} := E(d_M)$  and  $H$  is the cdf corresponding to the Full model pdf (4.2). The  $L^2$  distance  $\Delta_{EH}$  is similar to  $\Delta_{EF}$  but with the cdf  $H$  corresponding to the Homogeneous model pdf (4.4).

The  $L^2$  norm of the difference between the theoretical and empirical cdfs was  $\Delta_{EF} = 0.7558$  when using the Full model and  $\Delta_{EH} = 0.7558$  for the Homogeneous model as well. The plot of the empirical vs. theoretical cdfs for the Homogeneous model with likelihood maximizing values of the parameters is shown in Figure 7a). By comparison, for the non-homogeneous model, the minimum  $L^2$  norm of  $\Delta_{EF}^* = 0.7074$  was achieved when  $d_M = 25.3231$  mm,  $b_0 = 5.5264$ , and  $a$  essentially zero. For the Homogeneous model, the minimum  $L^2$  norm of 0.7073 was achieved with  $d_M = 25.3230$  mm and  $b_0 = 5.6379$ . The plot of the empirical vs. theoretical cdfs for the Homogeneous model with  $L^2$  distance minimizing values of the parameters is shown in Figure 7b). We

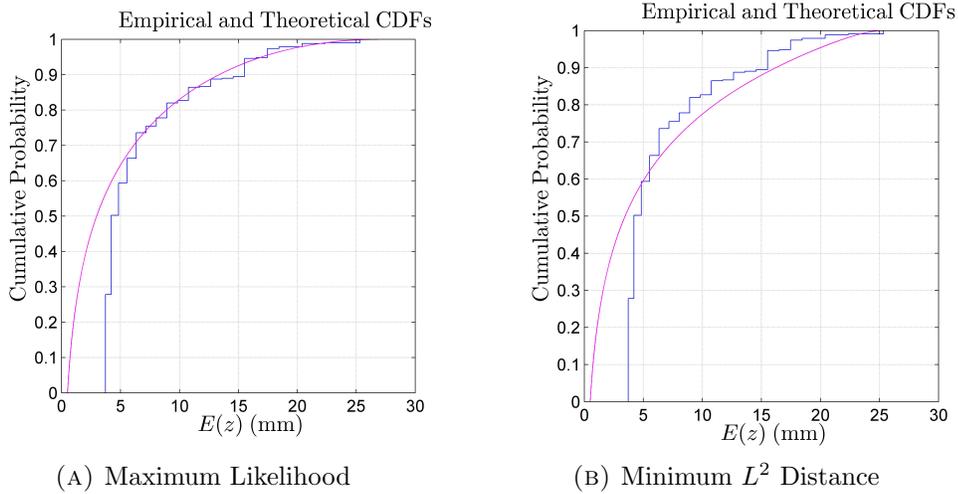


FIGURE 7. Graphs of empirical and theoretical cdfs for transformed diameters of metastases. Plot (A) shows the Homogeneous model with likelihood maximizing values  $d_M = 26.808$  mm and  $b_0 = 0.7858$ . The  $L^2$  norm of the difference between the two cdfs is 0.7558. Plot (B) shows the Homogeneous model with  $L^2$  distance minimizing values  $d_M = 25.3230$  mm and  $b_0 = 5.6379$ . The  $L^2$  norm of the difference between the two cdfs is 0.7073.

notice that even at minimum  $L^2$ , there is a wide divergence between the two cdfs in the lower ranges. One cause is that all 119 profiles of diameter 1 mm were converted to the diameter 3.724 mm, whereas, in reality, observations of profile diameters with diameter 1 mm would have come from tumors with diameters in a range from 0.5 mm to more than 6.5 mm. We also note that tumors of diameter less than the 7 mm section width can escape detection and are therefore under-represented in the empirical cdf.

**6.2.2. Method of True Diameters.** In order to refine our results and better account for the varying sizes of small tumors, we changed to the True Diameter method described in Section 4.4.2. We used the Full model pdf  $f_Y(y)$ , given in (4.2), with the method described in Section 4.5 to create the pdf of observed profile diameters,  $f_Z(z)$ , and then minimized the NNLL  $L_{TF}$  given by

$$L_{TF}(d_M, a, b_0) = -\frac{1}{n} \sum_{k=1}^K n_k \log(f_Z(k)).$$

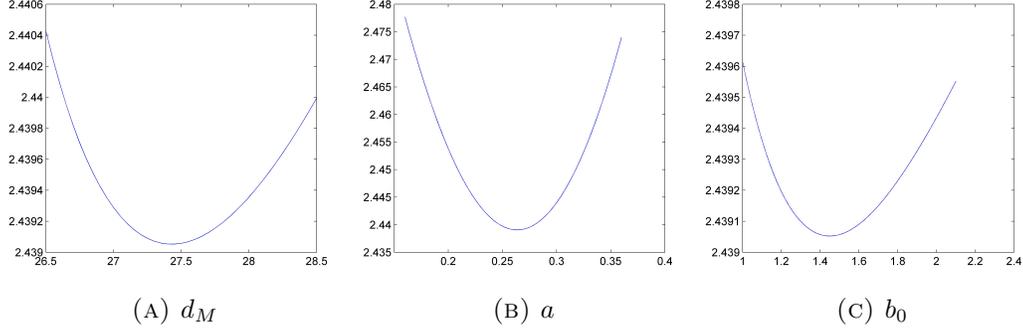


FIGURE 8. Graphs of NNLL profile functions using the method of true diameters without rounding for likelihood maximizing parameters  $d_M$ ,  $a$ , and  $b_0$  when applied to the data for Protocol 10.

We found the minimum NNLL of  $L_{TF}^* = 2.4393$  occurred when  $d_M = 27.4288$  mm,  $a = 0.2636$ , and  $b_0 = 1.4487$ . Graphs of the NNLL profiles as a function of the individual parameters are shown in Figure 8. The profile functions for all three parameters displayed a clear-cut minimum. Of note is a significant improvement in the likelihood, as compared to the Method of Expected Diameters, and the fact that this time the model did not degenerate into its Homogeneous limiting case. The fit between the theoretical cdf and its empirical counterpart in Figure 9a also shows a dramatic improvement compared to the Method of Expected Diameters. This is confirmed by computing the  $L^2$  distance,  $\Delta_{TF}$ , between the theoretical and empirical cdf's given by

$$\Delta_{TF}^2 = \sum_{k=1}^{K+1} \int_{A_{k-1}}^{A_k} \left[ F_Z(u) - \frac{1}{n} \sum_{j=1}^{k-1} n_j \right]^2 du$$

where  $F_Z$  is the cdf corresponding the pdf  $f_Z(z)$ ,  $A_0 := d_m$ ;  $A_k = k, 1 \leq k \leq K$ ; and  $A_{K+1} := d_M$ . The result turned out to be  $\Delta_{TF} = 0.1603$ , a dramatic improvement relative to the value  $\Delta_{EF} = 0.7558$  obtained for the Method of Expected Diameters. Finally, we estimated model parameters by minimizing the functional  $\Delta_{TF}$ , which yielded the following results:  $\Delta_{TF}^* = 0.14398$ ,  $L_{TF} = 2.4426$ ,  $d_M = 26.250$  mm,  $a = 0.2956$ , and  $b = 5.4099$ . Notice a further improvement in  $L^2$  distance compared to the value for the maximum likelihood parameters. The profiles (not shown) of the

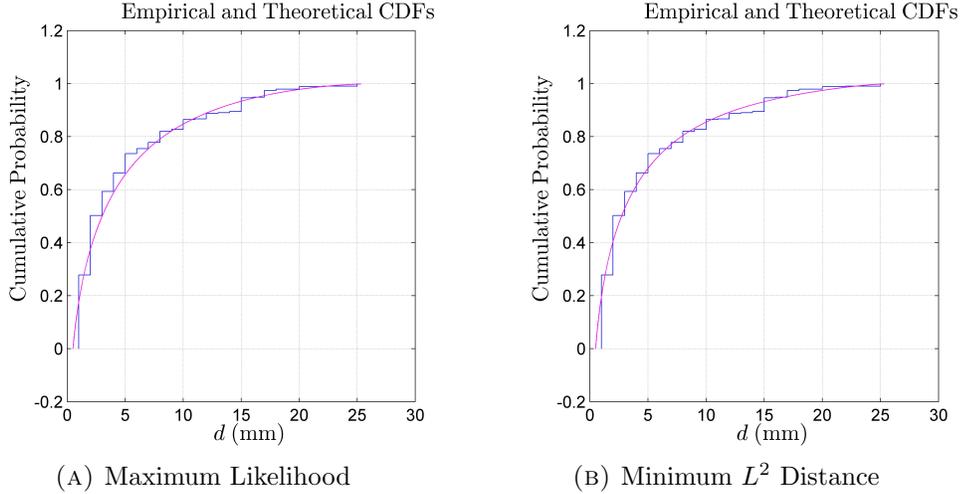


FIGURE 9. Graphs of empirical and theoretical cdfs for the Method of True Diameters. Plot (A) shows the Full model with likelihood maximizing values values  $d_M = 27.4288$  mm,  $a = 0.2636$ , and  $b_0 = 1.4487$ . The  $L^2$  norm of the difference between the two cdfs is 0.1603. Plot (B) shows the Full model with  $L^2$  distance minimizing values  $d_M = 26.520$  mm,  $a = 0.2956$ , and  $b_0 = 5.4099$ . The  $L^2$  norm of the difference between the two cdfs is 0.14398.

functional  $\Delta_{TF}$  displayed a clear-cut minimum. The fit between the two cdf's is given in Figure 9b.

**6.2.3. Protocol 10 True Diameters with Rounding.** As a further refinement to the True Diameter Method, we compensated for the fact that Douglas rounded tumor profile diameters to the nearest 1 mm by using the True Diameter with Rounding method described by (4.8) in section 4.4.3. Specifically, we minimized

$$L_{RF}(d_M, a, b_0) = -\frac{1}{n} \sum_{k=1}^K n_k \log P(k - 0.5 \leq z \leq k + 0.5)$$

where

$$P(k - 0.5 \leq d \leq k + 0.5) = F_Z(k + 0.5) - F_Z(k - 0.5).$$

The minimum NNLL of  $L_{RF}^* = 2.4136$  was obtained when  $d_M = 27.2866$  mm,  $a = 0.2861$ , and  $b_0 = 2.0788$ . Graphs of the NNLL profile functions as the individual parameters are varied are shown in Figure 10 and display a clear-cut minimum.

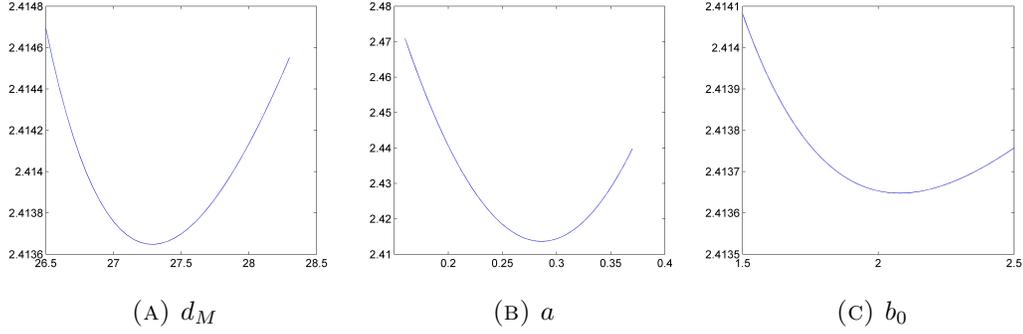


FIGURE 10. Graphs of NLL profile functions of the likelihood maximizing parameters  $d_M$ ,  $a$ , and  $b_0$  when the method of true diameters with rounding is applied to the data for Protocol 10.

As described at the start of this chapter, in order to make our  $L^2$  distance computation account for rounding, we assumed that the true cross-sectional diameter whose reported value after rounding was  $k$  mm was uniformly distributed on the interval  $[k - 0.5, k + 0.5]$ . In this case the  $L^2$  distance,  $\Delta_{RFP}$  between the model-based theoretical cdf  $F_Z$  and the corresponding piecewise linear continuous empirical cdf,  $G$ , is given by the formula

$$\Delta_{RFP}^2 = \sum_{k=1}^{K+2} \int_{B_{k-1}}^{B_k} [F_Z(z) - G(z)]^2 dz,$$

where  $B_0 = 0$ ;  $B_k = k - 0.5, 1 \leq k \leq K + 1$  (here  $K = 25$ ); and  $B_{k+2} = d_M$ .

For the above likelihood maximizing parameters we found  $\Delta_{RFP} = 0.08493$ . A comparison graph of the empirical and theoretical cdfs is shown in Figure 11. For the purpose of comparison to the other methods we have already considered, we computed the  $L^2$  norm of the distance between the empirical cdf and theoretical cdf as well and found that  $\Delta_{RF} = 0.1457$ , which was just slightly higher than the value we obtained by minimizing  $\Delta_{RF}$  in Section 6.2.2

Finally, we estimated model parameters by minimizing  $\Delta_{RFP}$  over  $d_M \geq K + 0.5 = 25.5$  mm,  $a > 0$ , and  $b_0 > 0$ . This led to  $\Delta_{RFP}^* = 0.08264$ ,  $L_{RF} = 2.4146$ ,  $d_M = 26.464$  mm,  $a = 0.2942$  and  $b = 5.2491$ . Notice a further reduction of the  $L^2$  distance between

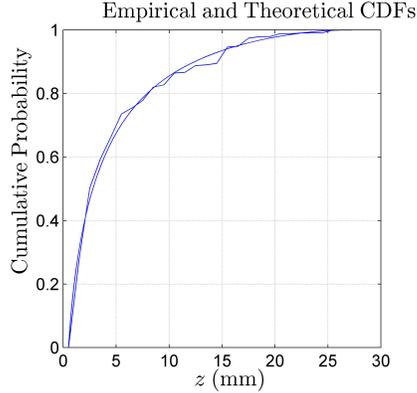


FIGURE 11. Piecewise continuous linear empirical cdf vs. theoretical cdf for the Method of True Diameters with Rounding applied to Protocol 10 with likelihood maximizing values  $d_M = 27.2866$  mm,  $a = 0.2861$ , and  $b_0 = 2.0788$ . The  $L^2$  norm of the difference between the two cdfs is  $\Delta_{RFP} = 00849$ . ( $\Delta_{RF} = 0.1457$ .)

the two cdf's compared to the value obtained for the likelihood-maximizing parameters and that the NNLL deteriorated only slightly. However, these distances are so small that visually the fit between the two cdf's for the  $L^2$  distance minimizing parameters is almost indistinguishable from that in Figure 11. Again, for the purpose of comparison to the other methods we have already considered, we computed the  $L^2$  norm of the distance between the empirical cdf and theoretical cdf as well and found that  $\Delta_{RF} = 0.14402$ , which was nearly identical to the value we obtained by minimizing  $\Delta_{RF}$  in Section 6.2.2 (i.e. 0.14398.)

Table 3 summarizes our results for Protocol 10. Of the maximum likelihood estimators, we see that we obtain the best likelihood with  $L^2$  norm closest to the possible minimum by using the distribution of rounded profiles.

Protocol 10 Parameters					
Method	NLLL	$L^2$ distance	$d_M$ (mm)	$a$	$b$
Expected Diameters	$L_{EH}^* = 2.8864$	$\Delta_{EH} = 0.7558$	26.808	0	0.7858
	$L_{EH} = 2.9930$	$\Delta_{EH}^* = 0.7073$	25.323	0	5.6379
True Diameters	$L_{TF}^* = 2.4393$	$\Delta_{TF} = 0.1603$	27.429	0.2636	1.4487
	$L_{TF} = 2.4426$	$\Delta_{TF}^* = 0.1440$	26.520	0.2956	5.4099
True Diameters w/ Rounding	$L_{RF}^* = 2.4136$	$\Delta_{RFP} = 0.0849$	27.287	0.2861	2.0788
	$L_{RF} = 2.4146$	$\Delta_{RFP}^* = 0.0826$ $\Delta_{RF} = 0.1457$	26.4642	0.2942	5.2491

Native Parameters			
Method	$\theta$	$\gamma_0/\beta$	$\gamma_0\rho$
Expected Diameters	NA	0.8542	1.2726
	NA	0.8542	1.2726
True Diameters	0.2258	0.8567	0.6903
	0.2522	0.8531	0.1848
True Diameters w/Rounding	0.2449	0.8561	0.4811
	0.2509	0.8528	0.1905

TABLE 3. Summary of parameters for Protocol 10

### 6.3. Protocol 17

The data in Douglas' Protocol 17 come from a female who died of metastatic breast cancer at age 81. The size of the primary tumor was not recorded, but it was known to have been removed by a mastectomy five years prior to her death. Thus  $W = 81$  and  $V = 76$ . Metastases were measured by slicing the liver into  $l = 5$  mm sections. The counts of diameters of the observed circular profiles were reported as  $n_1 = 19, n_2 = 18, n_3 = 10, n_4 = 7, n_5 = 17, n_6 = 7, n_7 = 5, n_8 = 4, n_9 = 2, n_{10} = 4, n_{11} = 0, n_{12} = 3, n_{13} = 0, n_{14} = 0, n_{15} = 1, n_{16} = 0, n_{17} = 0, n_{18} = 0, n_{19} = 0, n_{20} = 0, n_{21} = 0, n_{22} = 0, n_{23} = 0, n_{24} = 0, n_{25} = 1$ , where  $n_i$  indicates the number of circular profiles observed to have diameter of  $i$  mm (rounded to the nearest mm.)

**6.3.1. Method of Expected Diameters.** Using (4.6), we convert the data from Protocol 17 to obtain the values shown in Table 4.

We again used the maximum likelihood method to estimate the parameters for the full model by employing a Matlab routine based on the non-linear simplex method to

Count	19	18	10	7	17	7	5	4	2
Profile Diameter (mm)	1	2	3	4	5	6	7	8	9
Expected Diameter (mm)	2.781	3.351	4.071	4.878	5.739	6.636	7.556	8.494	9.444
Count	4	0	3	0	0	1	0	0	
Profile Diameter (mm)	10	11	12	13	14	15	16	17	
Expected Diameter (mm)	10.402	11.368	12.339	13.314	14.292	15.273	16.256	17.242	
Count	0	0	0	0	0	0	0	1	
Profile Diameter (mm)	18	19	20	21	22	23	24	25	
Expected Diameter (mm)	18.229	19.217	20.206	21.196	22.188	23.180	24.172	25.166	

TABLE 4. Expected diameters (in mm) of metastases for the patient in Protocol 17.

minimize the negative log-likelihood. Because of the discontinuity in the pdf at  $d_A$ , we searched exhaustively with  $d_A$  restricted first below  $d_m$ , then between  $d_m$  and the smallest observed metastasis size, then between successive metastasis sizes, and finally with  $d_A$  restricted between the largest metastasis size and  $d_M$ . We then examined the minimum NNLL,  $L_{EF}(d_A, d_M, a, b_0, b_1)$ , found in each sub-interval.

When  $d_A$  was restricted to lie between  $d_m = 0.5$  mm and the smallest expected diameter (2.781 mm) or between the two smallest expected diameters (2.781 and 3.351 mm), the likelihood degenerated into the IISMG Model in (4.10). Using (5.13), (5.14), and (5.15), we found that  $v = 19/98 \approx 0.19388$ ,  $a = 0.47814$ , and  $b_0 = 37.7731$  and then determined that  $\gamma_1 = 3.4400 \text{ year}^{-1}$ .

Comparing the theoretical cdf for the IISMG model to the empirical cdf, we found that the  $L^2$  norm of the difference between the two was  $\Delta_{EG} = 0.3324$ .

When  $d_A$  was restricted to lie between the two largest expected diameters (24.17 and 25.17 mm) the likelihood degenerated into the CSPT Model in (4.10). Using (5.7), (5.8), and (5.6), we found that  $\gamma_1 = 4.7616 \text{ year}^{-1}$ ,  $v = \frac{1}{98} \approx 0.010204$ ,  $b_1 = 0.18194$ ,

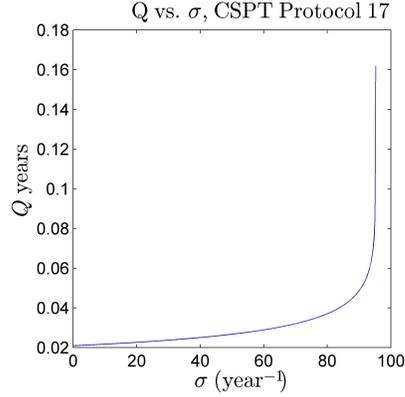


FIGURE 12.  $Q = V - T$  as a function of  $\sigma$  for the CSPT model applied to Expected Diameters from Protocol 17.

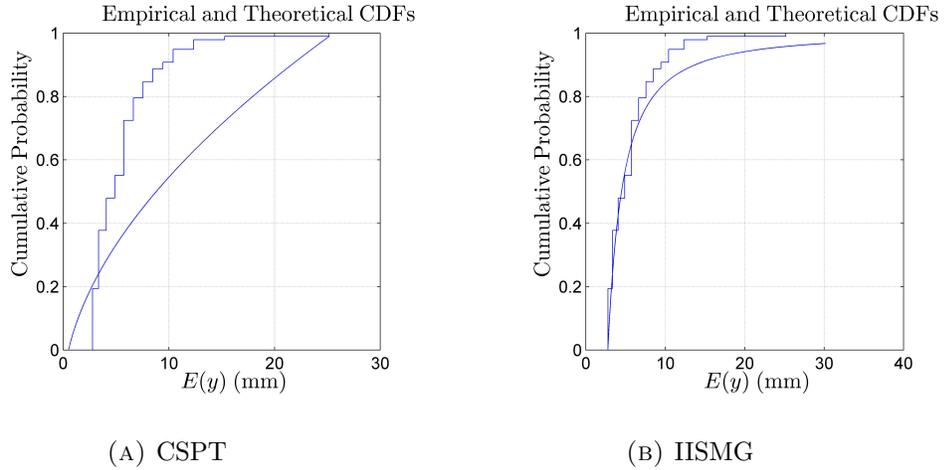


FIGURE 13. Graphs of empirical and theoretical cdfs for the CSPT and IISMG models with likelihood maximizing parameters applied to Protocol 17.

and  $\rho = 1.154284763$  years. Using equation (5.9), we solved numerically for  $Q$  as a function of  $\sigma$  and obtained the graph shown in Figure 12.

Comparing the theoretical cdf for the CSPT model to the empirical cdf, we found that the  $L^2$  norm of the difference between the two was  $\Delta_{EC} = 1.2133$ .

Plots of the empirical and theoretical cdfs for the maximum likelihood fit of the CSPT and IISMG models are shown in Figure 13.

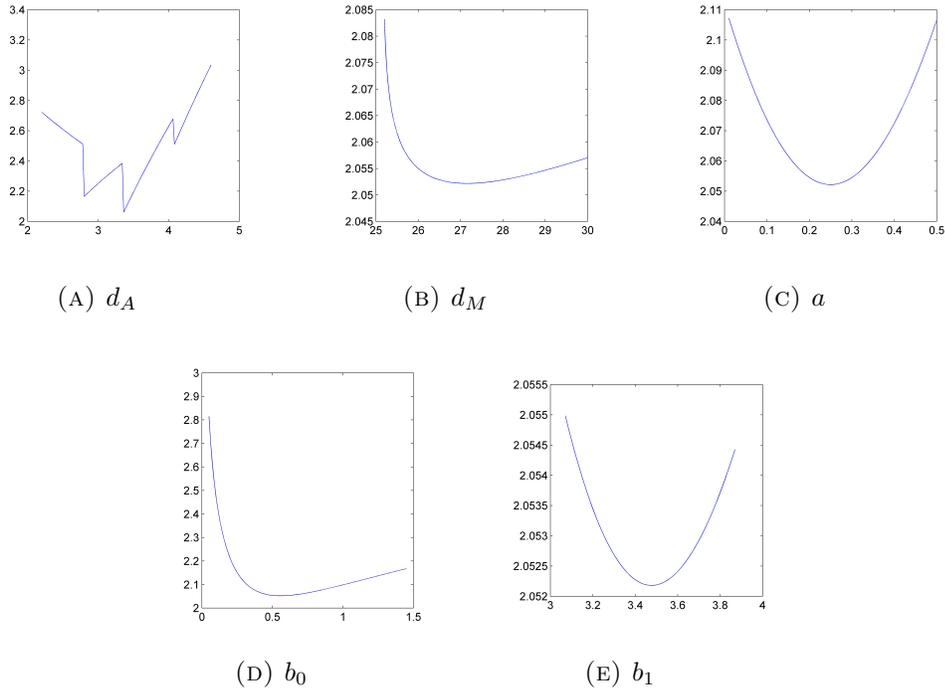


FIGURE 14. Graphs of NNLL profile functions for maximum likelihood parameters  $d_A, d_M, a, b_0,$  and  $b_1$  for the Method of Expected Diameters applied to Protocol 17 .

After the two cases where infinite likelihood was anticipated because of model degeneration, we found that the lowest value of the NNLL was  $L_{EF} = 2.0522$  which was obtained when  $d_A = 3.3515$  mm,  $d_M = 27.1484$  mm,  $a = 0.2497$ ,  $b_0 = 0.5579$ , and  $b_1 = 3.4771$ . Figure 14 shows graphs of the negative log-likelihood profile functions as the individual parameters are varied which indicate that we have achieved a minimum.

When we compute the  $L^2$  norm of the difference between the empirical cdf of expected diameters and theoretical distribution of tumor diameters we obtain  $\Delta_{EF} = 0.2161$ . Figure 15a shows a plot of the two cdfs together.

For comparison, when we fit the model by minimizing the  $L^2$  norm of the difference between the empirical and theoretical cdfs we obtain a minimum  $L^2$  norm of  $\Delta_{EF}^* = 0.15506$  when  $d_A = 3.5589$  mm,  $d_M = 25.16568$  mm,  $a = 0.50169$ ,  $b_0 = 1.26704$ , and  $b_1 = 1.592763$ . Graphs of the  $L^2$  norm profile functions as the individual parameters are varied around their optimal values help to confirm that we have reached a minimum

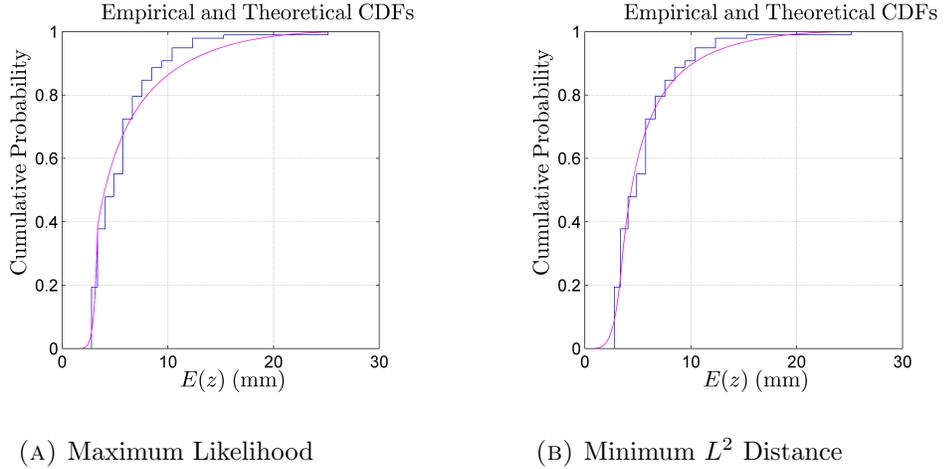


FIGURE 15. Graphs of empirical and theoretical cdfs for the Method of Expected Diameters. Plot (A) shows the Full model with likelihood maximizing values values  $d_A = 3.3515$  mm,  $d_M = 27.1484$  mm,  $a = 0.2497$ ,  $b_0 = 0.5579$ , and  $b_1 = 3.4771$ . The  $L^2$  norm of the difference between the two cdfs is 0.2161. Plot (B) shows the Full model with  $L^2$  distance minimizing values  $d_A = 3.5589$  mm,  $d_M = 25.166$  mm,  $a = 0.50169$ ,  $b_0 = 1.26704$ , and  $b_1 = 1.592763 \times 10^{-11}$ . The  $L^2$  norm of the difference between the two cdfs is 0.1551.

and are shown in Figure 16. Note the monotonicity of the profile function for  $d_M$  indicating that the minimum  $L^2$  distance between the empirical and theoretical cdfs is achieved as  $d_M$  approaches its smallest possible value (the largest expected diameter.) The graph of the empirical and theoretical cdfs together is shown in Figure 15b.

As with Patient A, the IISMG and CSPT models show poor fit. As with Patient 10, we notice that even at minimum  $L^2$ , there is still a divergence between the two cdfs in the lower ranges, although it is better for Protocol 17 because of the added flexibility afforded by  $\gamma_1$ . In this case, all 19 profiles of diameter 1 mm were converted to the diameter 2.781 mm, whereas, in reality, observations of profile diameters with diameter 1 mm would have come from tumors with diameters in a range from 0.5 mm 5.5 mm. We also note that tumors of diameter less than the 5 mm section width can escape detection and are therefore under-represented in the empirical cdf.

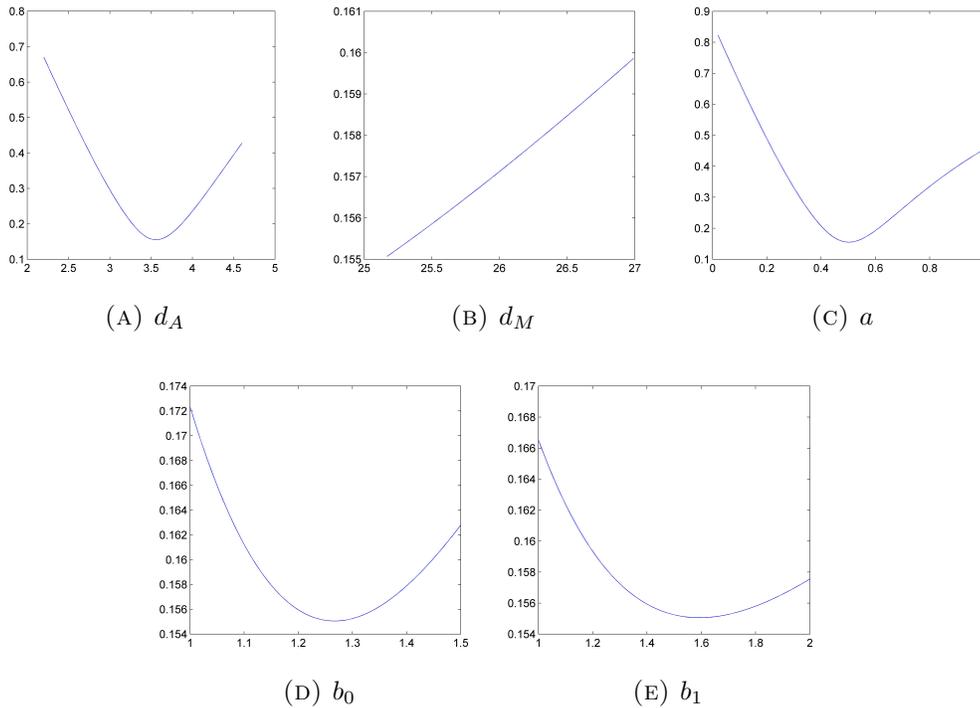


FIGURE 16. Plots of the  $L^2$  norm profile functions for the difference between the empirical and theoretical cdfs for expected diameters for the  $L^2$  norm minimizing parameter set. The graph for  $d_M$  reflects the fact that  $d_M$  cannot be smaller than the largest expected diameter. The  $L^2$  norm is seen to decrease as  $d_M$  approaches the maximum expected diameter obtained from the data.

**6.3.2. Method of True Diameters.** As was noted in (5.16) and (5.17), the method of True Diameters produces infinite likelihood as  $d_A$  approaches the largest observed profile diameter from above and as  $d_A$  approaches each of the observed diameters of 1, 2, 3, 4, and 5 mm. In the case of Patient A, the infinite likelihood was caused by the pdf transitioning from continuous to mixed discrete/continuous and we were able to switch to the mixed CSPT and IISMG models in order to estimate parameters. In the case of the Method of True Diameters with Protocol 17, the spike in the pdf is not carrying a fixed weight and so we can't produce a mixed discrete/continuous model to allow us to fit parameters.

**6.3.3. Method of True Diameters with Rounding.** Although the Method of True Diameters fails, the Method of True Diameters with Rounding, which compensates for the fact that Douglas rounded tumor profile diameters to the nearest 1 mm by using the distribution described in section 4.4.3, is bounded and can be used to find optimal parameter sets. Starting with the Full model and maximizing likelihood led the model to decay into the Heavy-Seeding/Long-latency (HSSL) model described in Section 3.1.3 and Appendix B.4. We found the minimum NNLL of  $L_{RS}^* = 2.3604$  occurred when  $d_A = 4.9464$  mm,  $d_M = 35.0011$  mm,  $a = 0.7854$ , and  $b_1/b_0 = 0.5121$ . Graphs of the NNLL profile functions are displayed in Figure 17 and indicate that we have reached a minimum. A comparison graph of the piecewise linear empirical cdf with the theoretical cdf is shown in Figure 18a. The fit is visually very good and we compute the  $L^2$  distance between the two cdfs to be  $\Delta_{RSP} = 0.0706$ . For the purpose of comparison to the previous cases, we also computed the  $L^2$  norm of the distance between the empirical (piecewise constant) cdf and the theoretical cdf and found that  $\Delta_{RS} = 0.1289$

Finally, we estimated model parameters by minimizing  $\Delta_{RSP}$ . The model again degenerated into the HSSL model and we obtained the minimum value  $\Delta_{RSP} = 0.05651$  when  $d_A = 6.407$  mm,  $d_M = 33.3799$  mm,  $a_0 = 0.8465$  and  $b_1/b_0 = 0.8978$ . Graphs of the  $L^2$  norm profile functions as  $d_A, d_M, a$ , and  $b_1/b_0$  are varied individually are shown in Figure 19 and corroborate that this is a minimum. Graphs of the empirical piecewise linear cdf and the theoretical cdf are shown in Figure 18b and show a slight improvement over the maximum likelihood plot in Figure 18a, particularly in the 2-4 mm range. The NNLL is slightly worse at 2.3675. Again, for the purpose of comparison to the other methods we have already considered, we computed the  $L^2$  norm of the distance between the empirical (piecewise constant) cdf and theoretical cdf and found that  $\Delta_{RS} = 0.1224$ .

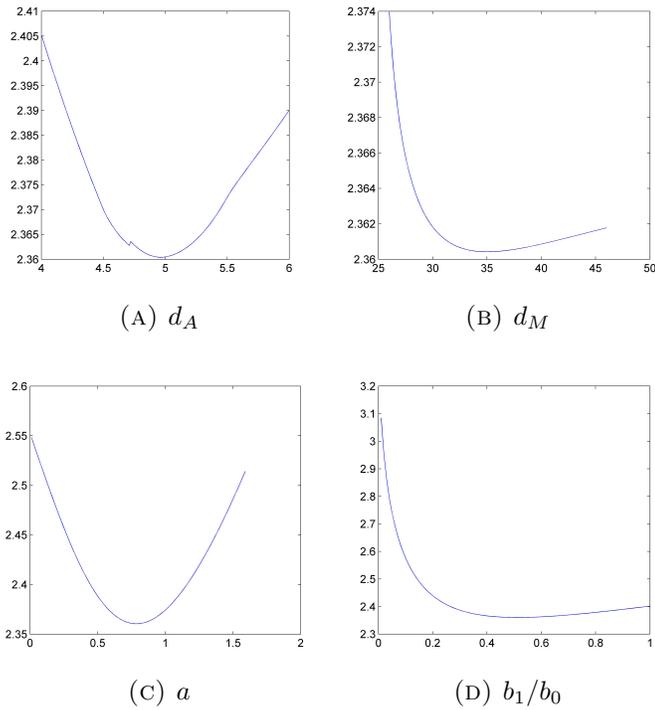


FIGURE 17. Graphs of NNL profile functions for maximum likelihood parameters  $d_A$ ,  $d_M$ ,  $a$ , and  $b_1/b_0$  for the Method of True Diameters with Rounding with the HSSL model applied to Protocol 17 .

Table 5 summarizes our results for Protocol 17. Of the maximum likelihood estimators, we have the lowest value of the  $L^2$  norm when we use the Heavy-Seeding/Long-Latency model with the True Diameter method.

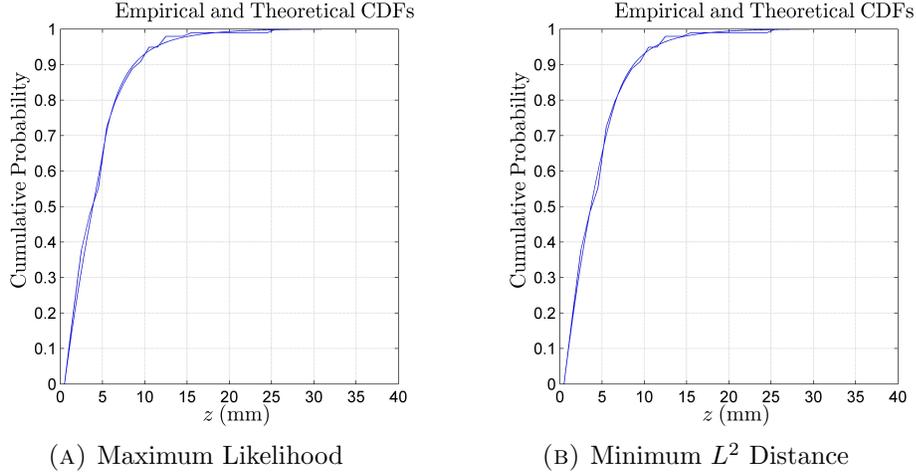


FIGURE 18. Graphs of piecewise continuous empirical cdfs and theoretical cdfs for the Method of True Diameters with Rounding. Plot (A) shows the HSLM model with likelihood maximizing values  $d_A = 4.9464$  mm,  $d_M = 35.0011$  mm,  $a = 0.7854$ , and  $b_1/b_0 = 0.5121$ . The  $L^2$  norm of the difference between the two cdfs is  $\Delta_{RSP} = 0.0706$ . Plot (B) shows the HSLM model with  $L^2$  distance minimizing values  $d_A = 6.4071$  mm,  $d_M = 33.3799$  mm,  $a = 0.8$ , and  $b_1/b_0 = 0.8978$ . The  $L^2$  norm of the difference between the two cdfs is  $\Delta_{RSP} = 0.05651$ .

Protocol 17 Parameters							
Model	NLL	$L^2$ distance	$d_A$ (mm)	$d_M$ (mm)	$a$	$b_0$	$b_1$
CSPT	$L_{EC}^* = 2.9795$	$\Delta_{EC} = 1.2133$	21.1657	21.1657	$\infty$	$\infty$	0.18194
IISMG	$L_{EG}^* = 2.3952$	$\Delta_{EG} = 0.3324$	2.7808	$\infty$	0.47814	37.7731	$\infty$
Exp. D Full	$L_{EF}^* = 2.0522$	$\Delta_{EF} = 0.2161$	3.3515	27.1484	0.2497	0.5579	3.4771
	$L_{EF} = 2.3834$	$\Delta_{EF}^* = 0.1551$	3.5589	25.16568	0.50169	1.26704	1.592763
Round HSLM	$L_{RS}^* = 2.3604$	$\Delta_{RSP} = 0.0706$ $\Delta_{RS} = 0.1289$	4.9464	35.0011	0.7854	$b_1/b_0 = 0.5121$	
	$L_{RS} = 2.3675$	$\Delta_{RSP}^* = 0.05651$ $\Delta_{RS} = 0.1224$	6.407	33.3799	0.8465	$b_1/b_0 = 0.8978$	
Native Parameters							
Model	NLL	$L^2$ distance	$\gamma_0$ year $^{-1}$	$\gamma_1$ year $^{-1}$	$Q$ years	$\rho$ years	$\sigma$ year $^{-1}$
CSPT	$L_{EC}^* = 2.9795$	$\Delta_{EC} = 1.2133$	0	4.7616	$Q = f(\sigma)$	1.1543	
IISMG	$L_{EG}^* = 2.3952$	$\Delta_{EG} = 0.3324$	$\infty$	3.4400	NA	0	$\infty$
Exp. D Full	$L_{EF}^* = 2.0522$	$\Delta_{EF} = 0.2161$	22.1367	3.5520	0.2835	0.08097	5.5266
	$L_{EF} = 2.3834$	$\Delta_{EF}^* = 0.1551$	4.5104	3.5879	1.301	0.1750	2.2628
Round HSLM	$L_{RS}^* = 2.3604$	$\Delta_{RSP} = 0.0706$ $\Delta_{RS} = 0.1289$	1.9384	3.7855	3.0383	$\infty$	1.5223
	$L_{RS} = 2.3675$	$\Delta_{RSP}^* = 0.0565$ $\Delta_{RS} = 0.1224$	3.5380	3.9408	1.3996	$\infty$	3.1764

TABLE 5. Summary of parameters for Protocol 17

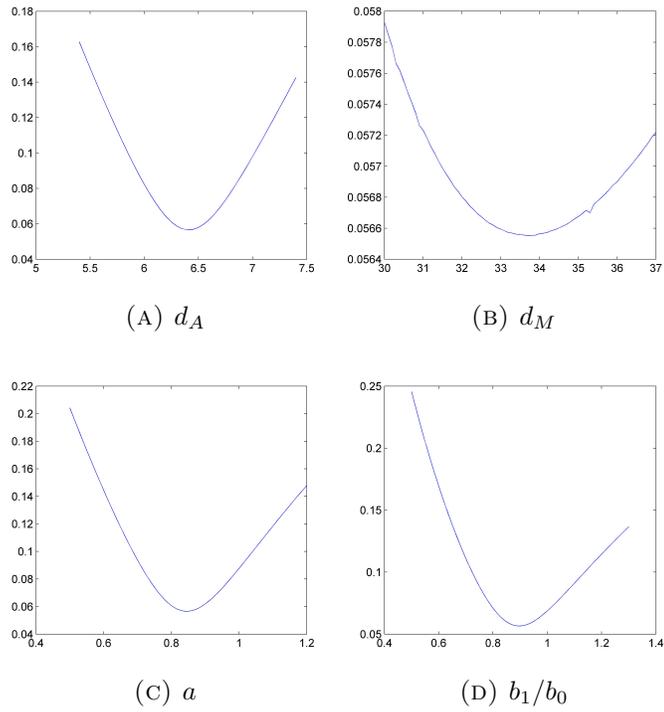


FIGURE 19. Graphs of  $L^2$  distance profile functions for minimum  $L^2$  distance parameters  $d_A, d_M, a$ , and  $b_1/b_0$  for the Method of True Diameters with Rounding with the HSSL model applied to Protocol 17 .

## CHAPTER 7

### Results and Conclusions

#### 7.1. Comparison of the methods of parameter estimation.

In finding parameter estimates for Patient A, we saw that (as expected) the full model degenerated to the mixed discrete/continuous CSPT and IISMG models. Neither provided a compelling fit to the data, and so we resorted to a compromise maximum likelihood that provided finite likelihood but with much better fit between the empirical and theoretical cdfs (though the  $L^2$  norm of 0.1885 obtained at this compromise maximum likelihood location is not nearly as good as the minimum possible  $L^2$  norm, which was 0.1014, see Table 1 in Section 6.1.)

In the discussion that follows of our findings for Patient A, we will report the values of identifiable biological parameters that come from the compromise maximum likelihood estimation and from minimizing the  $L^2$  norm of the distance between the theoretical and empirical cdfs. These values are:

(1) Compromise Likelihood maximization:  $\gamma_0 = 9.6649 \text{ year}^{-1}$ ,  $\gamma_1 = 2.6749 \text{ year}^{-1}$ ,  $T = 73.736 \text{ years}$ ,  $\rho = 0.02177 \text{ years}$ ,  $\beta = 87.3637 \text{ year}^{-1}$ ,  $\theta = 0.05789$  ( $L_{VF}^* = 2.3843$ ,  $\Delta_{VF} = 0.1885$ )

(2)  $L^2$  distance minimization:  $\gamma_0 = 3.0301 \text{ year}^{-1}$ ,  $\gamma_1 = 2.7439 \text{ year}^{-1}$ ,  $T = 73.2875 \text{ years}$ ,  $\rho = 0.1550 \text{ year}$ ,  $\beta = 32.3592 \text{ year}^{-1}$ ,  $\theta = 0.1053$  ( $L_{VF} = 2.4740$ ,  $\Delta_{VF}^* = 0.1014$ ).

In the case of Protocol 10, we found the Method of True Diameters to be vastly superior to the Method of Expected Diameters in terms of both increasing the model-based maximum likelihood of observations and dramatically improving the fit to the observed data, see Table 3 in Section 6.2. Another advantage of the Method of True Diameters over the Method of Expected Diameters is that it did not lead to degeneration of the model of cancer progression into its Homogeneous submodel. Taking into

account the rounding involved in reporting the values of cross-sectional diameters of liver metastases led to further improvement of the likelihood and  $L^2$  distance between the theoretical and empirical cdfs. However, the improvement was not as significant as that due to switching from the Method of Expected Diameters to the Method of True Diameters.

In the case of Protocol 17, we saw that when using the method of expected diameters the likelihood tended to infinity as the model degenerated into either the CSPT or IISMG case. Although the likelihood was infinite, neither model provided a compelling fit when the theoretical and empirical cdfs were compared. (See Table 5 in Section 6.3.) This leads us to discount these points of infinite likelihood as due to the mixed discrete/continuous nature of the CSPT and IISMG models. The third runner up after these two cases (with  $d_A$  held away from  $x_1$  and  $x_n$ ) provided a better fit (with finite likelihood.)

A more difficult problem occurred when we used the method of True Diameters without rounding because the likelihood approached infinity not only for  $d_A$  near the smallest and largest profile diameters,  $z_1$  and  $z_n$ , but also for  $d_A$  near the four other observed profile diameters closest to  $z_1$ . This made parameter estimation using this method essentially meaningless.

When the method of True Diameters with Rounding was used, we obtained a much better fit to the data. The results of parameter estimation with this method were also very consistent. Whether we maximized likelihood or minimized the  $L^2$  distance between the empirical and theoretical cdfs, the NNLLs were 2.36 and 2.37 while the  $L^2$  norms were 0.122 and 0.129, respectively. In these cases we saw that the model parameters approach the HSLL model ( $\rho \rightarrow \infty$ ) with  $\frac{\gamma_1}{\gamma_0}$  greater than 1,  $\gamma_0 = 1.9$  and  $3.6 \text{ year}^{-1}$ ,  $\gamma_1 = 3.75$  and  $3.95 \text{ year}^{-1}$ , disease onset occurring at 1.4 and 3.1 years before resection, and  $\sigma = \theta\beta$  being  $1.5$  and  $3.2 \text{ year}^{-1}$ , respectively.

In the discussion that follows of our findings for Protocols 10 and 17, we will only report the values of identifiable biological parameters found from using the Method of

True Diameters with Rounding. For Protocol 10, these values are:

(1) Likelihood maximization:  $\theta = 0.2449$ ,  $\gamma_0/\beta = 0.8561$  and  $\gamma_0\rho = 0.4811$  ( $L_{RF}^* = 2.4136$ ,  $\Delta_{RFS} = 0.08493$ )

(2)  $L^2$  distance minimization:  $\theta = 0.2509$ ,  $\gamma_0/\beta = 0.8528$  and  $\gamma_\rho = 0.1905$  ( $L_{RF} = 2.4426$ ,  $\Delta_{RF}^* = 0.0826$ ). Of note is a good agreement between the estimates of parameters  $\theta$  and especially  $\gamma_0/\beta$  obtained by the two very different methods of parameter estimation.

For Protocol 17, these values are:

(1) Likelihood maximization:  $\gamma_0 = 1.9384 \text{ year}^{-1}$ ,  $\gamma_1 = 3.7855 \text{ year}^{-1}$ ,  $Q = 3.0383 \text{ years}$  ( $T = 72.9717 \text{ years}$ ),  $\rho = \infty$ ,  $\sigma = \theta\beta = 1.5223$  ( $L_{RS}^* = 2.3604$ ,  $\Delta_{RS} = 0.1289$ ,  $\Delta_{RFS} = 0.0706$ )

(2)  $L^2$  distance minimization:  $\gamma_0 = 3.5380 \text{ year}^{-1}$ ,  $\gamma_1 = 3.9408 \text{ year}^{-1}$ ,  $Q = 1.3996 \text{ years}$  ( $T = 74.6004 \text{ years}$ ),  $\rho = \infty$ ,  $\sigma = \theta\beta = 3.1764$  ( $L_{RF} = 2.3675$ ,  $\Delta_{RSS}^* = 0.05651$ )

## 7.2. Stem-like cancer cells

For both Patient A and Protocol 10, we obtained estimates of  $\theta$ . The estimates were relatively small for Protocol 10 (0.244 and 0.251) and very small for Patient A (0.058 and 0.105.) These small values of parameter  $\theta$  suggest that the rate of metastasis shedding off the primary tumor grows much slower than the size of the primary tumor. Because the said rate is proportional to the number of metastasis-producing cells within the primary tumor, the subpopulation of such cells was also growing much slower than the entire cancer cell population. This points, if only indirectly, to the existence of a small subpopulation of stem-like cancer cells that drives cancer progression and metastasis.

Because the size of the primary tumor in the case of Protocol 17 was not known, we could not estimate  $\theta$  separately, but we did obtain estimates of  $\theta\beta$  that were between 1.5 and 3.2. If the size,  $S$ , of the primary tumor had been known, we would have calculated  $\beta = \frac{\log S}{V-T}$  and then obtained  $\theta$  by computing  $\theta = \frac{\sigma}{\beta} = \frac{\sigma(V-T)}{\log S}$ . From this we

can see that the largest value of  $\theta$  would occur when  $S$  was at the minimum detectable size. Although tumors with diameters as small as 1 mm can occasionally be detected by mammography, the median detectable size with mammography is at a diameter of about 7.5 mm [31]. This would mean that in the case of Protocol 17, that  $\theta$  could be as high as 0.78 to 0.81 (if resected when detected at 1 mm) but, more typically, 0.53 to 0.55 (if resected when detected at 7.5 mm) or lower. Because the publication date of Douglas' paper was 1971 and the patient in Protocol 17 was diagnosed six years prior to death, the primary diagnosis occurred no later than 1965, and probably even earlier. Because mammography was not developed until the late 1950's [10], it is likely that it was not used to diagnose the patient in Protocol 17, in which case the tumor would have been large enough to detect by a physical exam. Recall that Patient A's primary tumor was resected at the much larger volume of 10.3 cm<sup>3</sup>. If Protocol 17's primary tumor had been that large at resection, then  $\theta$  would be between 0.44 and 0.47.

### 7.3. Suppression of metastatic growth by the primary tumor

We can gain some insight into the effect the primary tumor might have on the growth rate of metastases by examining the ratio  $\gamma_0/\beta$ . For Patient A, our two estimates of  $\gamma_0/\beta$  were 0.11 and 0.09, meaning that while the primary tumor was in place, the rate of growth of bone metastases was about one tenth of the rate of growth of the primary tumor in the breast tissue. For Protocol 10, the estimated ratio,  $\gamma_0/\beta$ , of the rate of growth of liver metastases to the rate of growth of the primary tumor was about 0.85. Because we don't know the size,  $S$ , of Protocol 17's primary tumor at the time of resection, we cannot find  $\beta$ . But if, as discussed above, we take 0.47 as a reasonable estimate of  $\theta$ , we can obtain  $\beta$  from  $\sigma/\theta$  and this leads to the two estimates  $\gamma_0/\beta = 0.60$  and 0.52. Some of this difference may be explained by the difference in tissue type (bone versus breast in the case of Patient A, liver versus lung in the case of Protocol 10, and liver versus breast in the case of Protocol 17.) On the other hand, even the largest metastases in each case (Patient A, Protocol 10, and

Protocol 17) were still relatively small (diameters of about 35 mm, 25 mm, and 25 mm respectively.) Hence, at the time of survey, their metastases were most likely still in the exponential phase of their growth in which case one might expect the rate of their growth to be quite high. This points to the possibility of suppression of the growth of metastases by the primary tumor in each case.

#### 7.4. Acceleration of metastatic growth by primary tumor resection

The ratio  $\gamma_1/\gamma_0$  compares the growth rate of metastases after primary tumor resection to the growth rate before.

For Patient A, the growth rate of metastases went down after resection. When we used the maximum likelihood method, the growth rate fell from  $\gamma_0 = 9.6649 \text{ year}^{-1}$  to  $\gamma_1 = 2.6749 \text{ year}^{-1}$ . As we have seen, the maximum likelihood estimates are sensitive to the discontinuity in the pdf that is controlled by  $\gamma_0$ , and we can take the large drop with a degree of skepticism. When we minimized the  $L^2$  distance between the empirical and theoretical cdfs, we found that the growth rate decreased only slightly from  $\gamma_0 = 3.03 \text{ year}^{-1}$  to  $\gamma_1 = 2.74 \text{ year}^{-1}$ . After primary tumor resection, Patient A is known to have taken tamoxifen, which works by suppressing the growth of estrogen receptor positive breast tissue and this could explain the decrease in growth rate.

For Protocol 10, the primary tumor was not resected, so we cannot estimate  $\gamma_1$ , but for Protocol 17, the growth rate of metastases showed an increase from  $\gamma_0$  either 1.94 or 3.54  $\text{year}^{-1}$  to  $\gamma_1$  either 3.8 or 3.9  $\text{year}^{-1}$  respectively. The patient in protocol 17 had breast cancer, but she almost certainly did not receive tamoxifen because tamoxifen was only first synthesized in 1966 [45] and clinical trials did not start until 1971 [24], the year of publication of Douglas' paper.

## 7.5. Duration of metastatic latency

For Patient A , the two methods of parameter estimation provided estimates of  $\rho = 0.022$  and  $0.16$  year, which correspond to average latencies of 1 and 8 weeks, respectively.

For Protocol 17, both methods of parameter estimation selected the HSSL model, where  $\rho = \infty$ . In this model, large numbers of metastases find their way to suitable host sites ( $q$  is large), but very few actually begin to develop. Under this model, the rate of metastasis inception grows exponentially with rate  $\theta\beta$  until the time of resection. Because of the large pool of dormant metastases, after resection, the intensity of metastasis inception remains essentially constant.

For Protocol 10, we could not obtain a direct estimate of  $\rho$ . However, for primary small cell lung cancer (the type of cancer experienced by the patient in Protocol 10) the tumor doubling times has been estimated in the literature to vary within the limits of 25-217 days [23] while a representative average value of 86.3 days was reported in [1]. This places the rate,  $\beta$ , of primary tumor growth within the interval between  $1.17$  and  $10.12 \text{ year}^{-1}$  with a representative value of  $2.93 \text{ year}^{-1}$ . Since the ratio  $\gamma_0/\beta$  was estimated to be about 0.85 we find that the rate,  $\gamma_0$ , of growth of liver metastases varied between  $0.99$  and  $8.60 \text{ year}^{-1}$  and its representative value is  $2.49 \text{ year}^{-1}$ . Using the mean, 0.336, of our two above estimates as an estimate of parameter  $\gamma_0\rho$  we conclude that the expected duration of metastatic latency was somewhere between 2 weeks and 4 months with the representative value of 1.6 months.

Although the duration varied widely among the three patients, we can see that metastatic latency was appreciable for all three.

## 7.6. Timing of cancer onset.

For Patient A, our estimates show that the onset of the primary tumor occurred at a time,  $T$ , of either 73.736 or 73.288 years of age, meaning that onset occurred at either 3 or 9 months prior to the detection of the primary tumor.

For Protocol 10, using the estimates of parameter  $\beta$  derived from the literature described in the previous section, we conclude based on formula (2.20) that the onset of the primary tumor occurred between 2.8 and 24 years prior to the death of the patient with the representative value being 9.6 years.

For Protocol 17, our estimates show that the onset of the primary tumor occurred at a time  $T$  or either 72.97 or 74.60 years of age, meaning that disease onset occurred 3.04 or 1.40 years prior to detection of the primary tumor.

### 7.7. Primary Tumor Growth Rates and Progression Time

Of concern in our parameter estimates are the relatively high values obtained for the growth rate,  $\beta$ , and the relatively short times from disease onset to diagnosis, particularly in the case of Patient A. In this case, we obtained tumor progression times,  $Q = V - T$ , of 0.3 and 0.7 years, that seem too small and our likelihood maximizing primary tumor growth rate,  $\beta$ , is unrealistically high with a doubling time of 2.9 days. This may point to a shortcoming of using an exponential model for primary tumor growth. (There is some evidence that points to other growth models [4], including Gompertz growth [27].)

Note that parameters for Patient A were reported in [20] where a NNLL of  $L_{VF} = 2.43$  was obtained with  $V - T = 32.2$  years,  $\beta = 0.717 \text{ year}^{-1}$ ,  $\gamma_0 = 0.083 \text{ year}^{-1}$ ,  $\gamma_1 = 2.676 \text{ year}^{-1}$ ,  $\theta = 0.025$ ,  $\rho = 79.52$  years and  $\Delta_{VF} = 0.122$ . These parameters clearly represent only a local maximum of the likelihood; however, they do achieve a reasonable compromise between likelihood, goodness of fit, and biological plausibility.

### 7.8. Was the primary tumor detectable at inception of the first metastasis?

For Patient A, the largest metastasis was measured to be of size  $22.96 \text{ cm}^3$  at the time of survey. Using (2.9), we calculate based on the two methods of parameter estimation that the time of inception of the largest metastasis was 5 or 35 days after the onset of the primary tumor. Using the respective estimates of  $\beta$ , we find that

the at the time of inception of the first metastasis, the primary tumor was  $3.25 \cdot 10^{-9}$  or  $2.23 \cdot 10^{-8} \text{ cm}^3$ , which would be roughly 3 or 22 cells, respectively. A tumor of this size would certainly be undetectable. Looking forward to the time of primary tumor detection, the first metastasis would have been either  $1.13 \cdot 10^{-8}$  (from likelihood maximization) or  $6.48 \cdot 10^{-9} \text{ cm}^3$  (from  $L^2$  minimization), which would correspond to 11 or 6 cells, respectively. Thus, at the time of primary tumor diagnosis, the largest metastasis would have certainly been undetectable as well. In summary, for Patient A it would have been impossible to detect the primary tumor before metastatic disease had begun, and by the time the primary tumor was detected, metastatic disease had begun but would have been undetectable. Of course, applying an exponential law (or any other deterministic law) of growth to a small cell population (in this case just a handful of cells) could be misleading. But for this patient, the chance of preventing metastatic disease was so minute that our qualitative conclusion would most likely be the same should a more realistic stochastic model of tumor growth be utilized.

For Protocol 17, we cannot compute the size of the primary tumor at the time of inception of the first metastasis, but we can compute the size of the first metastasis at the time the primary tumor was discovered. Using (2.9) again, we calculate that the time of inception of the largest metastasis (cross-sectional diameter 25 mm at the time of metastasis survey) was 371 or 189 days after the onset of the primary tumor. At the time of primary tumor diagnosis, the first metastasis would have been growing either 2.01 or 0.88 years, respectively, and would have reached a size of  $4.93 \cdot 10^{-8}$  or  $2.26 \cdot 10^{-8} \text{ cm}^3$ , which would be roughly 49 or 23 cells, respectively, and at least 18 smaller metastases would have also begun to grow. So, as with Patient A, at the time of diagnosis of the primary tumor, metastatic disease would have started and yet have been undetectable. We make the same caveat about using predictions of just a few cells from a deterministic model of cellular growth, but we feel confident that our qualitative result is correct.

Answering this question for Protocol 10 requires a little more analysis. First, consider a metastasis with true diameter  $D$ . Then the inception age,  $U$ , for this metastasis is given by

$$(7.1) \quad U = V - 3\gamma_0^{-1} \log \frac{D}{d_c},$$

where  $d_c = 0.009$  mm is the diameter of a single oat cancer cell. Assuming that the primary tumor originated from a single clonogenic cell we find that the size of the primary tumor at age  $U$  was  $\exp[\beta(U - T)]$  which in view of (2.2) and (7.1) equals

$$\left(\frac{d_S}{d_c}\right)^3 \left(\frac{d_c}{D}\right)^{3\beta/\gamma_0}$$

so that the diameter of the primary tumor was

$$(7.2) \quad D_p = d_S \left(\frac{d_c}{D}\right)^{\beta/\gamma_0}$$

Note that, surprisingly, this estimate depends only on the ratio  $\gamma_0/\beta$ , which is an identifiable model parameter. Estimating the diameter of the largest liver metastasis by its expected value computed through formula (4.6) we get  $D = D_{\max} = E(25) = 25.323$  mm in which case formula (7.2) yields  $D_p = 0.0092$  mm. Thus, at the time of the first metastasis inception the primary tumor was extremely small and certainly undetectable. Furthermore, for the smallest metastasis with the expected diameter  $D = D_{\min} = E(1) = 3.724$  mm we find that  $D_p = 0.0876$  mm, i.e. even at the later time of inception of the smallest detected metastasis the primary tumor was still microscopic and undetectable. These estimates of  $D_p$  are so small as to correspond to a primary tumor of just a few cells, in which case a deterministic model of growth would not be very accurate. Still, we can say that even if we had used a more appropriate stochastic model of cellular growth, the qualitative result would be correct, i.e., the primary tumor would have been undetectable at the time of metastasis inception. Thus, removal of the primary tumor by surgery or radiation would not have had

the desired curative effect. In fact, excision of the primary tumor could have even accelerated metastasis growth as discussed in Chapter 1 and as it appears to have done for the patient in Protocol 17.

## 7.9. Summary

One aim was to develop a mathematical and statistical framework for applying our model of cancer progression to data on metastatic volumes or profile diameters. In the case of profile diameter data, we found that the method of true diameters with rounding was superior to the method of expected diameters. Along the way we discovered some pitfalls that arise when applying maximum likelihood methods with a discontinuous pdf.

Another aim was to use our model to answer questions about the natural history of cancer. We saw some evidence for stem-like cancer cells in the small values of  $\theta$  obtained Patient A and Protocol 10. We were not able to obtain  $\theta$  for protocol 17, but we were able to bound  $\theta$  above by 0.81 and perhaps by as low as 0.41.

Our results corroborate observations that the progression of cancer is more “non-linear” than that envisioned by Virchow and Halstead. Rather than a slow spatial spread from primary tumor to lymph nodes and then finally to metastases in distant sites, metastases actually begin to circulate very early in the progression of the disease. In all three of our cases metastasis formation occurred very soon after the onset of the primary tumor. We have also seen evidence that cancer does not behave like a local disease in the sense that removal of the primary tumor can cause a change in the behavior of the disease at metastatic sites. In particular, our results from Protocol 17 indicate that the primary tumor can suppress the growth of metastases and its removal can cause a jump in the rate of metastasis growth.

Finally, we saw that metastatic latency has an appreciable effect with metastases continuing to form weeks, months, or even years after the primary tumor has been removed.

As we analyze these cases, it is important to consider that cancer progression is patient-specific. As cancer researcher Bernard Fisher observed, even within a seemingly narrow class of cancers, like breast cancer, there is still a tremendous heterogeneity of tumor types and courses of metastatic progression [13]. When we consider that the three patients we have analyzed ultimately lost their battle with cancer, it should not surprise us to find in retrospect that their prognosis was dire; that is, that the onset of metastatic disease was early and that resection of the primary (in the case of Protocol 17) lead to an exacerbation of metastatic disease. Neither should we extrapolate that treatment of the primary tumor is hopeless in all cases; many patients enjoy long remissions and new adjuvant therapies may augment the benefits of the excision of the primary.

## Bibliography

1. Takashi Arai, Tetsuo Kuroishi, Yasuki Saito, Yuzo Kurita, Tsuguo Naruke, and Masahiro Kaneko, *Tumor doubling time and prognosis in lung cancer patients: evaluation from chest films and clinical follow-up study*, Japanese journal of clinical oncology **24** (1994), no. 4, 199–204.
2. T.R. Ashworth, *A case of cancer in which cells similar to those in the tumour were seen in blood after death*, Med. J. Australia (1869), no. 14, 146–147.
3. Dominique Bonnet and John E Dick, *Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell*, Nature Medicine **3** (1997), no. 7, 730–737.
4. Antonio Brú, Sonia Albertos, José Luis Subiza, José López García-Asenjo, and Isabel Brú, *The universal dynamics of tumor growth*, Biophysical journal **85** (2003), no. 5, 2948 – 2961.
5. J. C. Burkill, *The Lebesgue integral (Cambridge tracts in mathematics)*, Cambridge University Press, 2004.
6. Aulus Cornelius Celsus, *Of medicine: in eight books*, a new edition ed., Dickinson and Company, Edinburgh, 1814.
7. Michael F. Clarke, John E. Dick, Peter B. Dirks, Connie J. Eaves, Catriona H. M. Jamieson, D. Leanne Jones, Jane Visvader, Irving L. Weissman, and Geoffrey M. Wahl, *Cancer stem cells—perspectives on current status and future directions: AACR workshop on cancer stem cells*, Cancer Research **66** (2006), no. 19, 9339–9344.
8. R. Demicheli, A. Abbattista, R. Miceli, P. Valagussa, and G. Bonadonna, *Time distribution of the recurrence risk for breast cancer patients undergoing mastectomy: further support about the concept of tumor dormancy.*, Breast Cancer Research and Treatment **41** (1996), no. 2, 177–185.
9. J. R. S. Douglas, *Significance of the size distribution of bloodborne metastases*, **27** (1971), no. 2, 390.
10. Robert L. Egan, *Experience with mammography in a tumor institution*, Radiology **75** (1960), no. 6, 894–900.
11. Richard I. Evans, *Making the right choice: treatment options in cancer surgery*, Avery Pub. Group, Garden City Park, N.Y., 1995.
12. Conrad Fischer, *The corpus: The Hippocratic writings (Kaplan classics of medicine)*, Kaplan Publishing, New York, 2008.

13. Bernard Fisher, *Laboratory and clinical research in breast cancer—a personal adventure: The David A. Karnofsky memorial lecture*, *Cancer research* **40** (1980), no. 11, 3863–3874.
14. Judah Folkman, *Tumor angiogenesis factor*, *Cancer Research* **34** (1974), no. 8, 2109–2113.
15. Judah Folkman and Raghu Kalluri, *Tumor angiogenesis*, 6th ed., Holland-Frei Cancer Medicine, BC Decker, Hamilton, Ontario, Canada, 2003.
16. E. Gorelik, S. Segal, and M. Feldman, *On the mechanism of tumor concomitant immunity*, *International Journal of Cancer* **27** (1981), no. 6, 847–847–856.
17. Elieser Gorelik, *Resistance of tumor-bearing mice to a second tumor challenge*, *Cancer Research* **43** (1983), no. 1, 138–145.
18. Early Breast Cancer Trialists Collaborative Group.
19. Leonid Hanin, *Why victory in the war on cancer remains elusive: Biomedical hypotheses and mathematical models*, *Cancers* **3** (2011), no. 1, 340–367.
20. Leonid Hanin and Olga Korosteleva, *Does extirpation of the primary breast tumor give boost to growth of metastases? evidence revealed by mathematical modeling*, *Mathematical biosciences* **223** (2010), no. 2, 133–141.
21. Leonid Hanin, Jason Rose, and Marco Zaider, *A stochastic model for the sizes of detectable metastases*, *Journal of Theoretical Biology* **243** (2006), pp. 407–417 (English).
22. Leonid G Hanin, *Identification problem for stochastic models with application to carcinogenesis, cancer detection and radiation biology*, *Discrete Dynamics in Nature and Society* **7** (2002), no. 3, 177–189.
23. Kassem Harris, Inga Khachaturova, Basem Azab, Theodore Maniatis, Srujitha Murukutla, Michel Chalhoub, Hassan Hatoum, Thomas Kilkenny, Dany Elsayegh, Rabih Maroun, et al., *small cell lung cancer doubling time and its effect on clinical presentation: A concise review*, *Clinical medicine insights. Oncology* **6** (2012), 199.
24. V. Craig Jordan, *Tamoxifen: a most unlikely pioneering medicine*, *Nat Rev Drug Discov* **2** (2003), no. 3, 205–213.
25. Kazuharu Kai, Yoshimi Arima, Toshio Kamiya, and Hideyuki Saya, *Breast cancer stem cells*, *Breast Cancer* **17** (2010), 80–85.
26. Bruno Kisch, *Forgotten leaders in modern medicine. Valentin, Gruby, Remak, Auerbach*, *Transactions of the American Philosophical Society* **44** (1954), no. 2, 139.
27. Anna Kane Laird, *Dynamics of tumour growth*, *British journal of cancer* **18** (1964), no. 3, 490.
28. SH Lang, FM Frame, and AT Collins, *Prostate cancer stem cells*, *The Journal of Pathology* **217** (2009), no. 2, 299–306.

29. Paul H. Lange, Kiumars Hekmat, George Bosl, B. J. Kennedy, and Elwin E. Fraley, *Accelerated growth of testicular cancer after cytoreductive surgery*, *Cancer* **45** (1980), no. 6, 1498–1506.
30. Radu Marches, Richard Scheuermann, and Jonathan Uhr, *Cancer dormancy from mice to man: A review*, *Cell Cycle* **5** (2006), no. 16, 1772–1778.
31. James Michaelson, Sameer Satija, Richard Moore, Griffin Weber, Elkan Halpern, Andrew Garland, Daniel B Kopans, and Kevin Hughes, *Estimates of the sizes at which breast cancers become detectable on mammographic and clinical grounds*, *Journal of Womens Imaging* **5** (2003), 3–10.
32. Siddhartha Mukherjee, *The emperor of all maladies: A biography of cancer*, first ed., Scribner, New York, 2010.
33. George N. Naumov, Ian C. MacDonald, Pascal M. Weinmeister, Nancy Kerkvliet, Kishore V. Nadkarni, Sylvia M. Wilson, Vincent L. Morris, Alan C. Groom, and Ann F. Chambers, *Persistence of solitary mammary carcinoma cells in a secondary site*, *Cancer Research* **62** (2002), no. 7, 2162–2168.
34. Michael S. O’Reilly, Thomas Boehm, Yuen Shing, Naomi Fukai, George Vasios, William S. Lane, Evelyn Flynn, James R. Birkhead, Bjorn R. Olsen, and Judah Folkman, *Endostatin: An endogenous inhibitor of angiogenesis and tumor growth*, *Cell* **88** (1997), no. 2, 277–285.
35. Michael S. O’Reilly, Lars Holmgren, Yuen Shing, Catherine Chen, Rosalind A. Rosenthal, Marsha Moses, William S. Lane, Yihai Cao, E.Helene Sage, and Judah Folkman, *Angiostatin: A novel angiogenesis inhibitor that mediates the suppression of metastases by a lewis lung carcinoma*, *Cell* **79** (1994), no. 2, 315 – 328.
36. World Health Organization, *Who — cancer fact sheet 297*, <http://www.who.int/mediacentre/factsheets/fs297/en/index.html>, id: 1.
37. Amrith Raj Rao, Hanif G. Motiwala, and Omer M.A. Karim, *The discovery of prostate-specific antigen*, *BJU International* **101** (2008), no. 1, 5–10.
38. M. Retsky, R. Demicheli, and W. Hrushesky, *Premenopausal status accelerates relapse in node positive breast cancer: Hypothesis links angiogenesis, screening controversy*, *Breast Cancer Research and Treatment* **65** (2001), 217–224, 10.1023/A:1010626302152.
39. Michael Retsky, Romano Demicheli, William Hrushesky, Michael Baum, and Isaac Gukas, *Surgery triggers outgrowth of latent distant disease in breast cancer: An inconvenient truth?*, *Cancers* **2** (2010), no. 2, 305–337.
40. Michael Retsky, Rick Rogers, Romano Demicheli, William JM Hrushesky, Isaac Gukas, Jayant S Vaidya, Michael Baum, Patrice Forget, Marc DeKock, and Katharina Pachmann, *Nsaid analgesic*

- ketorolac used perioperatively may suppress early breast cancer relapse: particular relevance to triple negative subgroup*, *Breast cancer research and treatment* **134** (2012), no. 2, 881–888.
41. Tannishtha Reya, Sean J. Morrison, Michael F. Clarke, and Irving L. Weissman, *Stem cells, cancer, and cancer stem cells*, *Nature* **414** (2001), no. 6859, 105, M3: Article.
  42. V. K. Rohatgi, *An introduction to probability theory and mathematical statistics*, Wiley, 1976.
  43. H. L. Royden, *Real analysis*, 3rd ed., Macmillan Publishing Company, New York, 1988.
  44. G. Sandblom, E. Varenhorst, J. Rosell, O. Lofman, and P. Carlsson, *Randomised prostate cancer screening trial: 20 year follow-up*, *BMJ* **342** (2011), no. mar31-1, d1539–d1539.
  45. Walter Sneader, *Drug discovery: A history*, Wiley, New York, 2005.
  46. American Cancer Society, *Cancer facts & figures 2011*, <http://www.cancer.org/acs/groups/content/@epidemiologysurveillance/documents/document/acspc-029771.pdf>, 2011.

## APPENDIX A

### Some Useful Results Concerning Poisson Processes

#### A.1. Definition of a Poisson process

A Poisson process of non-negative locally-integrable intensity  $\lambda(t)$  is an integer-valued stochastic process  $X(t), t \geq 0$  such that

(i)  $X(0) = 0$ ,

(ii) for  $t > 0$ , the random variable  $X(t)$  has the Poisson distribution

$$\Pr(X(t) = k) = \frac{(\Lambda(t))^k e^{-\Lambda(t)}}{k!} \text{ for } k = 0, 1, \dots; \text{ where } \Lambda(t) = \int_0^t \lambda(u) du \text{ and}$$

(iii) for any time points  $t_0 = 0 < t_1 < t_2 < \dots < t_n$ , the process increments  $X(t_1) - X(t_0), X(t_2) - X(t_1), \dots, X(t_n) - X(t_{n-1})$  are independent random variables.

It is possible to use ii) and iii) to obtain an alternate version of ii) that applies to the number of arrivals between two times:

(ii') for  $t \geq 0$  and  $s > 0$ , the random variable  $X(t+s) - X(t)$  has the Poisson distribution

$$\Pr(X(t+s) - X(t) = k) = \frac{(\Lambda(t+s) - \Lambda(t))^k}{k!} e^{-(\Lambda(t+s) - \Lambda(t))} \text{ for } k = 0, 1, \dots$$

CLAIM. Properties i), ii), and iii) of a Poisson Process are equivalent to properties i), ii'), and iii).

PROOF. That ii') and i)  $\Rightarrow$  ii) is clear if we set  $t = 0$  in ii') and apply i). The proof that ii) and iii)  $\Rightarrow$  ii') uses strong mathematical induction on  $k$ . We first observe that

$$(A.1) \quad \Pr(X(t+s) = k) = \sum_{j=0}^k \Pr(X(t+s) - X(t) = k-j | X(t) = j) \Pr(X(t) = j)$$

$$\underbrace{\hspace{10em}}_{\text{independence}} = \sum_{j=0}^k \Pr(X(t+s) - X(t) = k-j) \Pr(X(t) = j)$$

Step 1: If  $k = 0$ , then (A.1) becomes

$$\Pr(X(t+s) = 0) = \Pr(X(t+s) - X(t) = 0) \Pr(X(t) = 0)$$

and solving for  $\Pr(X(t+s) - X(t) = 0)$  gives

$$\begin{aligned} \Pr(X(t+s) - X(t) = 0) &= \frac{\Pr(X(t+s) = 0)}{\Pr(X(t) = 0)} = \frac{(\Lambda(t+s))^0 e^{-\Lambda(t+s)} 0!}{0! (\Lambda(t))^0 e^{-\Lambda(t)}} = e^{-(\Lambda(t+s) - \Lambda(t))} \\ &= \frac{(\Lambda(t+s) - \Lambda(t))^0 e^{-(\Lambda(t+s) - \Lambda(t))}}{0!} \end{aligned}$$

so that ii') holds when  $k = 0$ .

Step 2: If  $k \geq 1$ , solving (A.1) for  $\Pr(X(t+s) - X(t) = k)$  gives

$$(A.2) \quad \Pr(X(t+s) - X(t) = k)$$

$$= \frac{\Pr(X(t+s) = k)}{\Pr(X(t) = 0)} - \sum_{j=1}^k \Pr(X(t+s) - X(t) = k-j) \frac{\Pr(X(t) = j)}{\Pr(X(t) = 0)}$$

By ii), we know that

$$(A.3) \quad \frac{\Pr(X(t+s) = k)}{\Pr(X(t) = 0)} = \frac{(\Lambda(t+s))^k e^{-\Lambda(t+s)}}{k! e^{-\Lambda(t)}} = \frac{e^{-(\Lambda(t+s) - \Lambda(t))}}{k!} (\Lambda(t+s))^k.$$

Assuming that for  $0 \leq j < k$ ,

$$\Pr(X(t+s) - X(t) = j) = \frac{(\Lambda(t+s) - \Lambda(t))^j}{j!} e^{-(\Lambda(t+s) - \Lambda(t))}$$

we have

$$\begin{aligned}
\text{(A.4)} \quad \sum_{j=1}^k \Pr(X(t+s) - X(t) = k-j) \frac{\Pr(X(t) = j)}{\Pr(X(t) = 0)} \\
= \sum_{j=1}^k \frac{(\Lambda(t+s) - \Lambda(t))^{k-j}}{(k-j)!} e^{-(\Lambda(t+s) - \Lambda(t))} \frac{(\Lambda(t))^j e^{-\Lambda(t)}}{j! e^{-\Lambda(t)}} \\
= \sum_{j=1}^k \frac{(\Lambda(t+s) - \Lambda(t))^{k-j} (\Lambda(t))^j}{j! (k-j)!} e^{-(\Lambda(t+s) - \Lambda(t))}.
\end{aligned}$$

Using the binomial theorem, we have

$$\begin{aligned}
\frac{e^{-(\Lambda(t+s) - \Lambda(t))}}{k!} (\Lambda(t+s))^k &= \frac{e^{-(\Lambda(t+s) - \Lambda(t))}}{k!} [(\Lambda(t+s) - \Lambda(t)) + \Lambda(t)]^k \\
&= \frac{e^{-(\Lambda(t+s) - \Lambda(t))}}{k!} \sum_{j=0}^k \frac{k!}{j! (k-j)!} (\Lambda(t+s) - \Lambda(t))^{k-j} (\Lambda(t))^j.
\end{aligned}$$

Therefore the expression in (A.4) can be rewritten as

$$\begin{aligned}
\sum_{j=1}^k \Pr(X(t+s) - X(t) = k-j) \frac{\Pr(X(t) = j)}{\Pr(X(t) = 0)} \\
= \frac{e^{-(\Lambda(t+s) - \Lambda(t))}}{k!} (\Lambda(t+s))^k - \frac{e^{-(\Lambda(t+s) - \Lambda(t))}}{k!} (\Lambda(t+s) - \Lambda(t))^k,
\end{aligned}$$

which, in view of (A.2) and (A.3) implies that

$$\Pr(X(t+s) - X(t) = k) = \frac{(\Lambda(t+s) - \Lambda(t))^k}{k!} e^{-(\Lambda(t+s) - \Lambda(t))}.$$

By strong mathematical induction, property ii') follows for all non-negative integers  $k$ . □

## A.2. Joint distribution of the time of occurrence of events in a Poisson process given their number

Our goal is to prove the following theorem.

THEOREM 1. For a Poisson process  $X(t)$  with time-dependent locally integrable non-negative rate  $\lambda(t)$ , the pdf of the occurrence times given the number,  $n$ , of events at time  $t$  is

$$(A.5) \quad f_{W_1 \dots W_n | X(t)=n}(w_1, \dots, w_n) = \frac{n! \prod_{i=1}^n \lambda(w_i)}{\left(\int_0^t \lambda(s) ds\right)^n}$$

for  $t \geq w_n > w_{n-1} > \dots > w_1 > 0$  and zero otherwise

For brevity, we can write

$$f_{W_1 \dots W_n | X(t)=n}(w_1, \dots, w_n) = n! \omega(w_1) \dots \omega(w_n), \text{ where } \omega(w) = \frac{\lambda(w)}{\int_0^t \lambda(s) ds}, \omega \in (0, t].$$

In view of the theorem, it is clear that because there are  $n!$  rearrangements of the ordered recurrence times  $W_1, W_2, \dots, W_n$ , if we consider the occurrence times without regard to order, then the pdf of the unordered occurrence times  $U_1, U_2, \dots, U_n$  is

$$f_{U_1 \dots U_n | X(t)=n}(u_1, \dots, u_n) = \frac{\prod_{i=1}^n \lambda(u_i)}{\left(\int_0^t \lambda(s) ds\right)^n}$$

where  $u_1, u_2, \dots, u_n \in (0, t]$  are distinct numbers. Thus the unordered occurrence times are equidistributed with a random sample from the distribution with pdf

$$(A.6) \quad \omega(u) = \frac{\lambda(u)}{\int_0^t \lambda(s) ds}, u \in (0, t].$$

We begin the proof of Theorem 1 with the following definitions. We say that a closed box,  $P$ , is “positioned at  $x$ ” (or “ $x$  is in  $P$ ’s southwest corner”) if the point  $x = (x_1, x_2, \dots, x_n)$  is such that  $x \in P$  and that for any point  $p = (p_1, p_2, \dots, p_n) \in P$ ,  $x_1 \leq p_1, x_2 \leq p_2, \dots, x_n \leq p_n$ .

For a measurable set  $A \subset \mathbb{R}^n$ , we denote the Lebesgue measure of  $A$  by  $|A|$ .

Recall that a *Vitali covering* of a set  $E$  is a covering  $\mathcal{V}$  such that for each  $x \in E$  and  $\delta > 0$ , there is a set  $V \in \mathcal{V}$  such that  $x \in V$  and the diameter of  $V$  is non-zero and

less than  $\delta$ . The Vitali Covering Theorem [5] says that given a measurable set  $E \subset \mathbb{R}^n$  with finite Lebesgue measure and given a collection,  $\mathcal{V}$ , of closed boxes in  $\mathbb{R}^n$  that is a Vitali covering for  $E$ , there exists a finite or countably infinite disjoint subcollection  $\{U_j\} \subseteq \mathcal{V}$  such that  $|E \setminus (\cup_i U_i)| = 0$ .

In order to prove the theorem, we need the following lemma:

LEMMA 1. Let  $W = (W_1, W_2, \dots, W_n)$  be a random vector supported on an open set  $G \subset \mathbb{R}^n$  (i.e.  $\Pr(W \in G) = 1$ .) If there is a non-negative measurable function  $f$  on  $G$  such that  $\int_G f dx = 1$  and for almost every point  $x \in G$ ,

$$(A.7) \quad \lim_{\text{diam}(P) \rightarrow 0} \frac{\Pr(W \in P)}{|P|} = f(x),$$

where the limit is taken as the diameters of boxes  $P$  positioned at  $x$  shrink to zero, then  $f$  is the pdf of  $W$ .

PROOF. We need to show that for any open box  $Q \subset G$ ,

$$\Pr(W \in Q) = \int_Q f dx.$$

Let  $D$  be the set of all points in  $Q$  at which

$$\lim_{\text{diam}(P) \rightarrow 0} \frac{\Pr(W \in P)}{|P|} = f(x)$$

does not hold and set  $Q' = Q \setminus D$ . Let  $\epsilon > 0$ . For every  $x \in Q'$ , let  $\mathcal{K}_x$  be the set of all closed boxes  $K$  with ‘‘southwest corner’’ at  $x$  for which  $K \subset Q$  and

$$\left| \frac{\Pr(W \in K)}{|K|} - f(x) \right| < \frac{\epsilon}{|Q|}.$$

Now set  $\mathcal{K} = \bigcup_{x \in Q'} \mathcal{K}_x$ .

It is easy to check that our covering  $\mathcal{K}$  is a Vitali covering of  $Q'$ . Thus, by the Vitali Covering Theorem, we can find a disjoint countable subcollection  $\{K_j\} \subset \mathcal{K}$ , such that  $|Q' \setminus (\cup_j K_j)| = 0$ , and, because  $D$  is a set of measure zero,  $|Q \setminus (\cup_j K_j)| = |(Q' \cup D) \setminus (\cup_j K_j)| = 0$ .

Because of our choice of covering,

$$\begin{aligned}
& \left| \Pr \left( W \in \bigcup_j K_j \right) - \int_{\bigcup_j K_j} f(x) dx \right| = \left| \sum_j \Pr(W \in K_j) - \sum_j \int_{K_j} f(x) dx \right| \\
& = \left| \sum_j \int_{K_j} \frac{\Pr(W \in K_j)}{|K_j|} dx - \sum_j \int_{K_j} f(x) dx \right| = \left| \sum_j \int_{K_j} \left[ \frac{\Pr(W \in K_j)}{|K_j|} - f(x) \right] dx \right| \\
& < \sum_j \int_{K_j} \frac{\varepsilon}{|Q|} dx = \frac{\varepsilon}{|Q|} \sum_j |K_j| = \frac{\varepsilon}{|Q|} \left| \bigcup_j K_j \right| = \frac{\varepsilon}{|Q|} |Q| = \varepsilon.
\end{aligned}$$

Thus, due to the arbitrariness of  $\varepsilon$ ,

$$\Pr \left( W \in \bigcup_j K_j \right) = \int_{\bigcup_j K_j} f(x) dx.$$

Therefore, since  $\bigcup_j K_j \subset Q$ , we have  $\int_Q f(x) dx = \int_{\bigcup_j K_j} f(x) dx = \Pr \left( W \in \bigcup_j K_j \right) \leq \Pr(W \in Q)$  for all boxes  $Q \in G$ .

To show the reverse inequality, suppose that  $\Pr(W \in Q_0) > \int_{Q_0} f(x) dx$  for some box  $Q_0 \subset G$ . Because the boundary of  $Q_0$  is a set of measure 0,

$$\int_{Q_0} f(x) dx = \int_{Q_0} f(x) dx < \Pr(W \in Q_0) \leq \Pr(W \in \overline{Q_0}).$$

Because  $G \setminus \overline{Q_0}$  is open, we can find a countable collection of disjoint open boxes  $Q_i, i = 1, 2, 3, \dots$ , such that  $G \setminus \overline{Q_0} = \bigcup_{i=1}^{\infty} Q_i$  and thus  $G = \overline{Q_0} \cup \bigcup_{i=1}^{\infty} Q_i$ . Now we see that

$$\begin{aligned}
1 &= \Pr(W \in G) = \Pr \left( W \in \overline{Q_0} \cup \bigcup_{i=1}^{\infty} Q_i \right) = \Pr(W \in \overline{Q_0}) + \sum_{i=1}^{\infty} \Pr(W \in Q_i) \\
&\geq \Pr(W \in \overline{Q_0}) + \sum_{i=1}^{\infty} \int_{Q_i} f(x) dx > \int_{\overline{Q_0}} f(x) dx + \sum_{i=1}^{\infty} \int_{Q_i} f(x) dx = \int_G f(x) dx = 1
\end{aligned}$$

which is a contradiction. Thus it must be that

$$\Pr(x \in Q) = \int_Q f(x) dx$$

for every open box  $Q \subset G$  and this completes the proof of the lemma.  $\square$

In order to prove the theorem, we apply the foregoing lemma in the case where the random vector  $W = (W_1, W_2, \dots, W_n)$  is the vector of times of occurrence of the first  $n$  events of the Poisson process.

First, we will use induction on  $n$  to show that for the function

$$(A.8) \quad f(v_1, v_2, \dots, v_n) = \frac{n! \prod_{i=1}^n \lambda(v_i)}{\left(\int_0^t \lambda(s) ds\right)^n}, \quad t \geq v_n > \dots > v_2 > v_1 \geq 0$$

and the open set  $G$  defined by

$$G = \{(x_1, x_2, \dots, x_n) | 0 < x_1 < x_2 < x_3 < \dots < x_n < t\},$$

we have

$$\int_G f dw = \int_0^t \int_0^{v_n} \dots \int_0^{v_2} f(v_1, v_2, \dots, v_n) dv_1 \dots dv_{n-1} dv_n = 1.$$

Let  $I_n$  denote the iterated integral,

$$(A.9) \quad I_n(v_n) \equiv \int_0^{v_n} \dots \int_0^{v_3} \int_0^{v_2} \lambda(v_1) \lambda(v_2) \dots \lambda(v_{n-1}) dv_1 dv_2 \dots dv_{n-1}$$

Claim: For every integer  $n \geq 2$ ,

$$I_n(v_n) = \frac{1}{(n-1)!} \left( \int_0^{v_n} \lambda(s) ds \right)^{n-1}$$

Step 1: If  $n = 2$ , then (A.9) reduces to

$$I_2(v_2) = \int_0^{v_2} \lambda(v_1) dv_1,$$

which is equal to

$$\frac{1}{(2-1)!} \left( \int_0^{v_2} \lambda(s) ds \right)^{2-1}.$$

and therefore the claim holds when  $n = 2$ .

Step 2: Assume that the claim is true for  $n = k$ . Then, in the case when  $n = k + 1$ , we have to investigate the integral

$$I_{k+1} = \int_0^{v_{k+1}} \int_0^{v_k} \cdots \int_0^{v_3} \int_0^{v_2} \lambda(v_1) \lambda(v_2) \cdots \lambda(v_{k-1}) \lambda(v_k) dv_1 \cdots dv_{k-1} dv_k$$

which can also be written in the form

$$I_{k+1}(v_{k+1}) = \int_0^{v_{k+1}} \lambda(v_k) I_k(v_k) dv_k$$

By our supposition,

$$I_k(v_k) = \frac{1}{(k-1)!} \left( \int_0^{v_k} \lambda(s) ds \right)^{k-1},$$

and we obtain

$$I_{k+1}(v_{k+1}) = \int_0^{v_{k+1}} \lambda(v_k) \frac{1}{(k-1)!} \left( \int_0^{v_k} \lambda(s) ds \right)^{k-1} dv_k.$$

Letting  $u(v_k) = \int_0^{v_k} \lambda(s) ds$  so that  $du = \lambda(v_k) dv_k$ , the integral becomes

$$I_{k+1}(v_{k+1}) = \int_0^{u(v_{k+1})} \frac{1}{(k-1)!} u^{k-1} du = \frac{u(v_{k+1})^k}{k!} - 0 = \frac{1}{k!} \left( \int_0^{v_{k+1}} \lambda(s) ds \right)^k$$

and thus the claim holds by mathematical induction.

Applying this to the function in (A.8), we have

$$\begin{aligned} \int_G f dw &= \frac{n! \int_0^t \cdots \int_0^{v_3} \int_0^{v_2} \lambda(v_1) \lambda(v_2) \cdots \lambda(v_n) dv_1 dv_2 \cdots dv_n}{\left( \int_0^t \lambda(s) ds \right)^n} \\ &= \frac{n! I_{n+1}(t)}{\left( \int_0^t \lambda(s) ds \right)^n} = \frac{n! \frac{1}{n!} \left( \int_0^t \lambda(s) ds \right)^n}{\left( \int_0^t \lambda(s) ds \right)^n} = 1. \end{aligned}$$

We next verify that the random vector  $(W_1, W_2, \dots, W_n)$  satisfies, conditional on  $X(t) = n$ , condition (A.7) of the above Lemma with function  $f$  given by (A.8). Given a point  $x = (x_1, x_2, \dots, x_n) \in G$ , if  $P$  is a closed box positioned at  $x$  with  $\text{diam} P < \min_{i \neq j} |x_i - x_j|$ , then  $P$  has the form  $P = \{(p_1, p_2, \dots, p_n) | x_1 \leq p_1 \leq x_1 + \Delta x_1, \dots, x_n \leq$

$p_n \leq x_n + \Delta x_n\}$  where for each  $i \in \{1, \dots, n\}$ ,  $\Delta x_i$  is such that  $(x_i, x_i + \Delta x_i) \cap (x_j, x_j + \Delta x_j) = \emptyset$  for  $i \neq j$ .

Now  $\Pr(W \in P, X(t) = n)$  is  $\Pr(x_1 < W_1 \leq x_1 + \Delta x_1, x_2 < W_2 \leq x_2 + \Delta x_2, \dots, x_n < W_n \leq x_n + \Delta x_n, W_{n+1} > t)$  which can be re-written in terms of our Poisson process as

$$\Pr(X(x_1) = 0, X(x_1 + \Delta x_1) - X(x_1) = 1, X(x_2) - X(x_1 + \Delta x_1) = 0, \dots, X(x_n + \Delta x_n) - X(x_n) = 1, X(t) - X(x_n + \Delta x_n) = 0).$$

Applying the property of independence of numbers of events occurring on non-overlapping intervals we obtain

$$\begin{aligned} \Pr(W \in P, X(t) = n) &= \\ &\Pr(X(x_1) = 0) \cdot \Pr(X(x_1 + \Delta x_1) - X(x_1) = 1) \cdot \Pr(X(x_2) - X(x_1 + \Delta x_1) = 0) \\ &\quad \dots \Pr(X(x_n + \Delta x_n) - X(x_n) = 1) \cdot \Pr(X(t) - X(x_n + \Delta x_n) = 0) \\ &= \Pr(X(x_1) = 0) \cdot \Pr(X(t) - X(x_n + \Delta x_n) = 0) \cdot \left[ \prod_{i=2}^n \Pr(X(x_i) - X(x_{i-1} + \Delta x_{i-1}) = 0) \right] \\ &\quad \cdot \left[ \prod_{i=1}^n \Pr(X(x_i + \Delta x_i) - X(x_i) = 1) \right] \\ &= e^{-\int_0^{x_1} \lambda(s) ds} \cdot e^{-\int_{x_n + \Delta x_n}^t \lambda(s) ds} \left[ \prod_{i=2}^n e^{-\int_{x_i + \Delta x_{i-1}}^{x_i} \lambda(s) ds} \right] \cdot \left[ \prod_{i=1}^n \left( \int_{x_i}^{x_i + \Delta x_i} \lambda(s) ds \right) e^{-\int_{x_i}^{x_i + \Delta x_i} \lambda(s) ds} \right] \\ &= e^{-\int_0^t \lambda(s) ds} \cdot \prod_{i=1}^n \int_{x_i}^{x_i + \Delta x_i} \lambda(s) ds. \end{aligned}$$

Thus,

$$\begin{aligned} \frac{\Pr(W \in P | X(t) = n)}{|P|} &= \frac{\Pr(W \in P, X(t) = n)}{|P| \Pr(X(t) = n)} = \frac{e^{-\int_0^t \lambda(s) ds} \cdot \left[ \prod_{i=1}^n \int_{x_i}^{x_i + \Delta x_i} \lambda(s) ds \right]}{|P| \frac{\left( \int_0^t \lambda(s) ds \right)^n}{n!} e^{-\int_0^t \lambda(s) ds}} \\ &= \frac{n! \left[ \prod_{i=1}^n \int_{x_i}^{x_i + \Delta x_i} \lambda(s) ds \right]}{|P| \left( \int_0^t \lambda(s) ds \right)^n} = \frac{n!}{\left( \int_0^t \lambda(s) ds \right)^n} \cdot \left[ \prod_{i=1}^n \frac{\int_{x_i}^{x_i + \Delta x_i} \lambda(s) ds}{\Delta x_i} \right] \end{aligned}$$

By the fundamental theorem of calculus (see, e.g. [43]),

$$\lim_{\Delta x_i \rightarrow 0} \frac{\int_{x_i}^{x_i + \Delta x_i} \lambda(s) ds}{\Delta x_i} = \frac{d}{dx_i} \int_0^{x_i} \lambda(s) ds = \lambda(x_i)$$

for almost all  $x_i$ . Therefore

$$\lim_{\text{diam}(P) \rightarrow 0} \frac{\Pr(W \in P | X(t) = n)}{|P|} = \frac{n!}{\left( \int_0^t \lambda(s) ds \right)^n} \cdot \prod_{i=1}^n \lambda(x_i) = \frac{n! \prod_{i=1}^n (\lambda(x_i))}{\left( \int_0^t \lambda(s) ds \right)^n}$$

Applying Lemma 1 now shows that Theorem 1 holds.

### A.3. Independent Classification of Events

Suppose that each event of a Poisson process with rate  $\lambda(t)$  is classified randomly as being exactly one of  $n$  types, and suppose that the probability of an event being classified as type- $i$ ,  $i = 1, 2, \dots, n$ , may depend on the time at which it occurs. Further, suppose that if an event occurs at time  $s$ , then, independently of all else, it is classified as being a type- $i$  event with probability  $P_i(s)$  where  $\sum_{i=1}^n P_i(s) = 1$  and each  $P_i$  is a measurable function.

**THEOREM 2.** If  $X_i(t)$  represents the number of type- $i$  events that occur by time  $t$ ,  $i = 1, 2, \dots, n$ , then  $X_i(t), i = 1, \dots, n$  are independent Poisson random variables having respective cumulative rates  $\Lambda_i(t) = \int_0^t P_i(s) \lambda(s) ds$ .

PROOF. We compute the joint distribution of the  $X_i(t)$  by conditioning on  $X(t) = m$ , the number of events to have occurred by time  $t$ :

$$\begin{aligned}
\text{(A.10)} \quad & \Pr \{X_1(t) = k_1, X_2(t) = k_2, \dots, X_n(t) = k_n\} \\
&= \sum_{m=0}^{\infty} \Pr \{X_1(t) = k_1, X_2(t) = k_2, \dots, X_n(t) = k_n | X(t) = m\} \Pr \{X(t) = m\} \\
&= \Pr \left\{ X_1(t) = k_1, X_2(t) = k_2, \dots, X_n(t) = k_n | X(t) = \sum_{i=1}^n k_i \right\} \Pr \left\{ X(t) = \sum_{i=1}^n k_i \right\},
\end{aligned}$$

the other conditional probabilities in the sum having been zero when  $m \neq \sum_i^n k_i$ . In order to simplify notation, we will use  $m = \sum_i^n k_i$  for the remainder of the proof.

Consider an arbitrary event that occurred in the interval  $[0, t]$ . If it had occurred at time  $s$ , then the probability that it would be a type- $i$  event is  $P_i(s)$ . By Theorem 1, the occurrence time distribution has pdf  $\frac{\lambda(s)}{\int_0^t \lambda(u) du}$ ,  $s \in (0, t]$ . Thus the probability that it would be a type- $i$  event is

$$p_i = \int_0^t P_i(s) \frac{\lambda(s)}{\int_0^t \lambda(u) du} ds = \frac{\int_0^t P_i(s) \lambda(s) ds}{\int_0^t \lambda(s) ds} = \frac{\Lambda_i(t)}{\Lambda(t)}$$

where  $\Lambda(t) = \int_0^t \lambda(s) ds$ .

Hence,  $\Pr \{X_1(t) = k_1, X_2(t) = k_2, \dots, X_n(t) = k_n | X(t) = m\}$  has the multinomial distribution

$$\Pr \{X_1(t) = k_1, X_2(t) = k_2, \dots, X_n(t) = k_n | X(t) = m\} = m! \prod_{i=1}^n \frac{p_i^{k_i}}{k_i!}$$

Then by (A.10),

$$\begin{aligned}
\Pr \{X_1(t) = k_1, X_2(t) = k_2, \dots, X_n(t) = k_n\} &= m! \prod_{i=1}^n \frac{p_i^{k_i}}{k_i!} \Pr \{X(t) = m\} \\
&= m! \left( \prod_{i=1}^n \frac{p_i^{k_i}}{k_i!} \right) \frac{\Lambda(t)^m}{m!} e^{-\Lambda(t)} = \left( \prod_{i=1}^n \frac{(p_i \Lambda(t))^{k_i}}{k_i!} \right) e^{-\Lambda(t)} = \left( \prod_{i=1}^n \frac{(p_i \Lambda(t))^{k_i}}{k_i!} \right) e^{-\int_0^t \left( \sum_{i=1}^n P_i(s) \right) \lambda(s) ds} \\
&= \left( \prod_{i=1}^n \frac{(p_i \Lambda(t))^{k_i}}{k_i!} \right) e^{-\int_0^t \left( \sum_{i=1}^n P_i(s) \right) \lambda(s) ds} \\
&= \prod_{i=1}^n \frac{\Lambda_i(t)^{k_i} e^{-\int_0^t P_i(s) \lambda(s) ds}}{k_i!} = \prod_{i=1}^n \frac{\Lambda_i(t)^{k_i} e^{-\Lambda_i(t)}}{k_i!}
\end{aligned}$$

from which we observe that each of the  $X_i(t)$  is independent and has a Poisson distribution

$$\Pr \{X_i(t) = k_i\} = \frac{(\Lambda_i(t))^{k_i} e^{-\Lambda_i(t)}}{k_i!}$$

with cumulative rate

$$\Lambda_i(t) = \int_0^t P_i(s) \lambda(s) ds$$

as desired. □

#### A.4. Intensity of a filtered Poisson process

**THEOREM 3.** If events are initiated according to a Poisson process with intensity  $\lambda(t)$  but are removed (or filtered out) independently of each other with fixed probability  $1 - q$ , then the resulting stochastic process is Poisson with intensity  $q\lambda(t)$ .

**PROOF.** Let  $X(t)$  be the number of events initiated by time  $t > 0$  in the original Poisson process and let  $X_s(t)$  be the corresponding number of events that have survived the filtering process. We now verify that the conditions given in the definition of the Poisson process hold for  $X_s(t)$  (see section A.1)

i) Because  $X$  is a Poisson process,  $X(0) = 0$ . Because  $0 \leq X_s(t) \leq X(t)$ , we must have  $X_s(0) = 0$  as well.

ii) Applying Theorem 2, we designate an event as type-1 if it is not filtered out and as type-2 if it is filtered out. In this context,  $X_s(t) = X_1(t)$  and  $P_1(s) = q$  and  $P_2(s) = 1 - q$  and we see that

$$\Pr \{X_s(t) = k\} = \frac{\left(\int_0^t P_1(s) \lambda(s) ds\right)^k}{k!} e^{-\int_0^t P_1(s) \lambda(s) ds}$$

so that  $X_s(t)$  has a Poisson distribution with mean

$$\Lambda_s(t) = \int_0^t P_1(s) \lambda(s) ds = \int_0^t q \lambda(s) ds = q \int_0^t \lambda(s) ds$$

iii) For independence, let  $I_i = [t_{i-1}, t_i], i = 1, 2, \dots, n$  denote disjoint time intervals and call an event type- $i, i = 1, \dots, n$  if it occurs in  $I_i$  and is not filtered out, or of type- $(n+1)$  otherwise. Denote the respective number of arrivals of events of type- $i$  up to time  $t$  by  $X_i(t), i = 1, 2, \dots, n, n+1$ . Then the probability that an event is classified as type- $i$  is

$$P_i(s) = \begin{cases} q, & s \in I_i \\ 0 & \text{otherwise} \end{cases}, i = 1, 2, \dots, n,$$

and the probability an event is classified as type- $(n+1)$  is

$$P_{n+1}(s) = \begin{cases} 1 - q, & s \in \cup_{i=1}^n I_i \\ 1 & \text{otherwise} \end{cases}.$$

By Theorem 2, the counts  $X_1(t), X_2(t), \dots, X_n(t)$ , are independent, so

$$\Pr \{X_1(t) = k_1, X_2(t) = k_2, \dots, X_n(t) = k_n\} = \prod_{i=1}^n \Pr \{X_i(t) = k_i\}.$$

In the context of our filtered Poisson process, for some time  $t$  beyond intervals  $I_i, i = 1, \dots, n$ , the event that  $X_i(t) = k_i$  is the event that there are  $k_i$  arrivals in interval  $I_i$  and thus the number of arrivals in disjoint intervals is independent. Thus  $X_s(t)$  is a Poisson process with rate

$$\lambda_s(t) = \frac{d}{dt} \Lambda_s(t) = \frac{d}{dt} q \int_0^t \lambda(s) ds = q \lambda(t).$$

□

### A.5. Intensity of a delayed Poisson process

**THEOREM 4.** If events are initiated according to a Poisson process with intensity  $\nu(t)$  with individual events delayed by random times that are independent between themselves as well as of the Poisson process and identically distributed with pdf  $f$ , then the resulting stochastic process for completed events is Poisson with intensity  $\lambda(t) = \int_0^t \nu(s) f(t-s) ds$ .

**PROOF.** Let  $X(t)$  denote the number of events initiated by time  $t > 0$  in the original Poisson process and let  $X_p(t)$  be the corresponding number of events that have completed their delay by time  $t$ . We now verify that  $X_p(t)$  satisfies the conditions of a Poisson process.

i) Because  $X$  is a Poisson process,  $X(0) = 0$ . Because  $0 \leq X_p(t) \leq X(t)$ , we must have  $X_p(0) = 0$  as well.

ii) At time  $t$ , we designate an event that has occurred by time  $t$  as type-1 if it has completed its delay by time  $t$  and as type-2 if it has not completed its delay. Note that if  $X_1(t)$  is the number of type-1 events that have occurred by time  $t$ , then  $X_1(t) = X_p(t)$ . The probability that an event that occurs at time  $s$  will complete its delay by time  $t > s$  is

$$P_1(s) = \Pr \{ \text{delay} \leq t - s \} = F(t - s),$$

where  $F$  is the cdf associated with pdf  $f$ . Applying Theorem 2,  $X_p(t)$  (which is also  $X_1(t)$ ) has a Poisson distribution with mean

$$\Lambda_p(t) = \int_0^t P_1(s) \nu(s) ds = \int_0^t F(t-s) \nu(s) ds.$$

iii) For independence, let  $I_i = [t_{i-1}, t_i], i = 1, 2, \dots, n$  denote disjoint time intervals. We classify an event as type- $i, i = 1, \dots, n$  if it completes its delay in  $I_i$ , and as type- $(n+1)$  otherwise. Denote the respective number of arrivals of events of type- $i$  up to

time  $t$  by  $X_i(t)$ ,  $i = 1, 2, \dots, n, n + 1$ . Denote the delay by the random variable  $Y$ . If an event occurs at time  $s$ , then the probability that it will be classified as type- $i$ ,  $i = 1, \dots, n$  is the probability that  $s + Y \in I_i$  which is the probability that  $t_{i-1} < s + Y < t_i$  or  $t_{i-1} - s < Y < t_i - s$  which is  $P_i(s) = F(t_i - s) - F(t_{i-1} - s)$ .

Because the time intervals are disjoint, we see that  $P_{n+1}(s) + \sum_{i=1}^n P_i(s) = 1$  and thus by Theorem 2, the counts  $X_1(t), X_2(t), \dots, X_n(t)$ , are independent, and

$$\Pr \{X_1(t) = k_1, X_2(t) = k_2, \dots, X_n(t) = k_n\} = \prod_{i=1}^n \Pr \{X_i(t) = k_i\}.$$

In the context of our delayed Poisson process, for some time  $t$ , the event that  $X_i(t) = k_i$  is the event that there are  $k_i$  completions in interval  $I_i$  by time  $t$  and thus the numbers of completions in disjoint intervals are independent.

Thus  $X_p(t)$  is a Poisson process with rate

$$\begin{aligned} \lambda_p(t) &= \frac{d}{dt} \Lambda_p(t) = \frac{d}{dt} \int_0^t F(t-s) \nu(s) ds = F(t-t) \nu(t) + \int_0^t f(t-s) \nu(s) ds \\ &= \int_0^t f(t-s) \nu(s) ds \end{aligned}$$

as desired. □

## APPENDIX B

### Catalog of Models

Writing our models in terms of the “Native Parameters”  $\gamma_0, \gamma_1, \sigma = \theta\beta, \rho, Q = V - T$ , and  $R = W - V$ , preserves the original meaning of these quantities and is helpful in understanding limiting cases and submodels whereas writing them in terms of the “Simplifying Parameters”  $a_0 = \frac{\theta\beta}{\gamma_0}, b_0 = \frac{1}{\rho\gamma_0}, b_1 = \frac{1}{\rho\gamma_1}, A = e^{\gamma_1 R}$  and  $M = e^{\gamma_0 Q + \gamma_1 R}$  helps clarify the functional form of the models. For ease of reference, we present a catalog of models in both forms.

#### B.1. Full Model

**B.1.1.**  $\gamma_1 \leq \log(m)/R$ .

B.1.1.1. *Native Parameters.*

$$p(x) = (C_1 x)^{-1} \left[ \left( \frac{e^{(\gamma_0 Q + \gamma_1 R)}}{x} \right)^{\frac{\sigma}{\gamma_0}} - \left( \frac{x}{e^{(\gamma_0 Q + \gamma_1 R)}} \right)^{\frac{1}{\gamma_0 \rho}} \right], m \leq x \leq e^{\gamma_0 Q + \gamma_1 R}$$

$$C_1 = \frac{\gamma_0}{\sigma} \left[ \left( \frac{e^{\gamma_0 Q + \gamma_1 R}}{m} \right)^{\sigma/\gamma_0} - 1 \right] + \gamma_0 \rho \left[ \left( \frac{m}{e^{\gamma_0 Q + \gamma_1 R}} \right)^{1/\gamma_0 \rho} - 1 \right].$$

B.1.1.2. *Simplifying Parameters.*

$$p(x) = (C_1 x)^{-1} \left[ \left( \frac{M}{x} \right)^a - \left( \frac{x}{M} \right)^{b_0} \right], m \leq x \leq M$$

$$C_1 = \frac{1}{a} \left[ \left( \frac{M}{m} \right)^a - 1 \right] + \frac{1}{b_0} \left[ \left( \frac{m}{M} \right)^{b_0} - 1 \right].$$

B.1.1.3. *Cdf.*

$$P(x) = (C_1)^{-1} \left[ \begin{aligned} & \frac{\gamma_0}{\sigma} \left( \frac{e^{(\gamma_0 Q + \gamma_1 R)}}{m} \right)^{\frac{\sigma}{\gamma_0}} - \frac{\gamma_0}{\sigma} \left( \frac{e^{(\gamma_0 Q + \gamma_1 R)}}{x} \right)^{\frac{\sigma}{\gamma_0}} \\ & + \rho \gamma_0 \left( \frac{m}{e^{(\gamma_0 Q + \gamma_1 R)}} \right)^{\frac{1}{\gamma_0 \rho}} - \rho \gamma_0 \left( \frac{x}{e^{(\gamma_0 Q + \gamma_1 R)}} \right)^{\frac{1}{\gamma_0 \rho}} \end{aligned} \right], m \leq x \leq e^{\gamma_0 Q + \gamma_1 R}$$

**B.1.2.**  $\gamma_1 > \log(m)/R$ .

B.1.2.1. *Native Parameters.*

$$p(x) = \begin{cases} (C_2x)^{-1} \frac{\gamma_0}{\gamma_1} \left(\frac{x}{e^{\gamma_1 R}}\right)^{\frac{1}{\gamma_1 \rho}} \left(e^{\sigma Q} - e^{-\frac{Q}{\rho}}\right), & m \leq x \leq e^{\gamma_1 R}, \\ (C_2x)^{-1} \left[ \left(\frac{e^{\gamma_0 Q + \gamma_1 R}}{x}\right)^{\frac{\sigma}{\gamma_0}} - \left(\frac{x}{e^{\gamma_0 Q + \gamma_1 R}}\right)^{\frac{1}{\gamma_0 \rho}} \right], & e^{\gamma_1 R} < x \leq e^{\gamma_0 Q + \gamma_1 R} \end{cases}$$

$$C_2 = \gamma_0 \rho \left(e^{\sigma Q} - e^{-\frac{Q}{\rho}}\right) \left[1 - \left(\frac{m}{e^{\gamma_1 R}}\right)^{\frac{1}{\gamma_1 \rho}}\right] + \frac{\gamma_0}{\sigma} (e^{\sigma Q} - 1) + \gamma_0 \rho \left(e^{-\frac{Q}{\rho}} - 1\right)$$

B.1.2.2. *Simplifying Parameters.*

$$p(x) = \begin{cases} (C_2x)^{-1} \frac{b_1}{b_0} \left(\frac{x}{A}\right)^{b_1} \left[\left(\frac{M}{A}\right)^a - \left(\frac{A}{M}\right)^{b_0}\right], & m \leq x \leq A, \\ (C_2x)^{-1} \left[\left(\frac{M}{x}\right)^a - \left(\frac{x}{M}\right)^{b_0}\right], & A < x \leq M \end{cases}$$

$$C_2 = \frac{1}{b_0} \left[\left(\frac{M}{A}\right)^a - \left(\frac{A}{M}\right)^{b_0}\right] \left[1 - \left(\frac{m}{A}\right)^{b_1}\right] + \frac{1}{a} \left[\left(\frac{M}{A}\right)^a - 1\right] + \frac{1}{b_0} \left[\left(\frac{A}{M}\right)^{b_0} - 1\right]$$

B.1.2.3. *Cdf Native Parameters.*

$$P(x) = \begin{cases} (C_2)^{-1} \gamma_0 \rho \left[\left(\frac{x}{e^{\gamma_1 R}}\right)^{\frac{1}{\gamma_1 \rho}} - \left(\frac{m}{e^{\gamma_1 R}}\right)^{\frac{1}{\gamma_1 \rho}}\right] \left(e^{\sigma Q} - e^{-\frac{Q}{\rho}}\right), & m \leq x \leq e^{\gamma_1 R}, \\ (C_2)^{-1} \gamma_0 \left\{ \begin{array}{l} \frac{1}{\sigma} e^{\sigma Q} + \rho e^{-\frac{Q}{\rho}} - \frac{1}{\sigma} \left(\frac{e^{\gamma_0 Q + \gamma_1 R}}{x}\right)^{\frac{\sigma}{\gamma_0}} \\ - \rho \left(\frac{x}{e^{\gamma_0 Q + \gamma_1 R}}\right)^{\frac{1}{\gamma_0 \rho}} \\ + \rho \left[1 - \left(\frac{m}{e^{\gamma_1 R}}\right)^{\frac{1}{\gamma_1 \rho}}\right] \left(e^{\sigma Q} - e^{-\frac{Q}{\rho}}\right) \end{array} \right\}, & e^{\gamma_1 R} < x \leq e^{\gamma_0 Q + \gamma_1 R} \end{cases}$$

B.1.2.4. *Cdf Simplifying Parameters.*

$$P(x) = \begin{cases} (C_2)^{-1} \frac{1}{b_0} \left[\left(\frac{x}{A}\right)^{b_1} - \left(\frac{m}{A}\right)^{b_1}\right] \left[\left(\frac{M}{A}\right)^a - \left(\frac{A}{M}\right)^{b_0}\right], & m \leq x \leq A \\ (C_2)^{-1} \left\{ \begin{array}{l} \frac{1}{a} \left(\frac{M}{A}\right)^a + \frac{1}{b_0} \left(\frac{A}{M}\right)^{b_0} - \frac{1}{a} \left(\frac{M}{x}\right)^a - \frac{1}{b_0} \left(\frac{x}{M}\right)^{b_0} \\ + \frac{1}{b_0} \left[1 - \left(\frac{m}{A}\right)^{b_1}\right] \left[\left(\frac{M}{A}\right)^a - \left(\frac{A}{M}\right)^{b_0}\right] \end{array} \right\}, & A \leq x \leq M \end{cases}$$

## B.2. Instantaneous Seeding

**B.2.1.**  $\gamma_1 \leq \log(m)/R$ .

B.2.1.1. *Native Parameters.*

$$p(x) = C_3 x^{-\frac{\sigma}{\gamma_0}-1}, m \leq x \leq e^{\gamma_0 Q + \gamma_1 R}$$

$$C_3 = \frac{\sigma/\gamma_0}{m^{-\frac{\sigma}{\gamma_0}} - (e^{\gamma_0 Q + \gamma_1 R})^{-\frac{\sigma}{\gamma_0}}}.$$

B.2.1.2. *Simplifying Parameters.*

$$p(x) = C_3 x^{-a-1}, m \leq x \leq M$$

$$C_3 = \frac{a}{m^{-a} - M^{-a}}.$$

B.2.1.3. *Cdf.*

$$P(x) = C_3 \frac{\gamma_0}{\sigma} \left( m^{-\frac{\sigma}{\gamma_0}} - x^{-\frac{\sigma}{\gamma_0}} \right), m \leq x \leq e^{\gamma_0 Q + \gamma_1 R}$$

**B.2.2.**  $\gamma_1 > \log(m)/R$ .

B.2.2.1. *Native Parameters.*

$$p(x) = C_4 x^{-\frac{\sigma}{\gamma_0}-1}, e^{\gamma_1 R} \leq x \leq e^{\gamma_0 Q + \gamma_1 R}$$

$$C_4 = \frac{\sigma/\gamma_0}{(e^{\gamma_1 R})^{-\frac{\sigma}{\gamma_0}} - (e^{\gamma_0 Q + \gamma_1 R})^{-\frac{\sigma}{\gamma_0}}}.$$

B.2.2.2. *Simplifying Parameters.*

$$p(x) = C_4 x^{-a-1}, A \leq x \leq M$$

$$C_4 = \frac{a}{A^{-a} - M^{-a}}.$$

B.2.2.3. *Cdf.*

$$P(x) = C_4 \frac{\gamma_0}{\sigma} \left( e^{-\frac{\gamma_1 \sigma}{\gamma_0} R} - x^{-\frac{\sigma}{\gamma_0}} \right), e^{\gamma_1 R} \leq x \leq e^{\gamma_0 Q + \gamma_1 R}$$

### B.3. Homogeneous

**B.3.1.**  $\gamma_1 \leq \log(m)/R$ .

B.3.1.1. *Native Parameters.*

$$p(x) = (C_5 x)^{-1} \left[ 1 - \left( \frac{x}{e^{\gamma_0 Q + \gamma_1 R}} \right)^{\frac{1}{\gamma_0 \rho}} \right], m \leq x \leq e^{\gamma_0 Q + \gamma_1 R}$$

$$C_5 = \gamma_0 Q + \gamma_1 R - \log m + \gamma_0 \rho \left[ \left( \frac{m}{e^{\gamma_0 Q + \gamma_1 R}} \right)^{\frac{1}{\gamma_0 \rho}} - 1 \right]$$

B.3.1.2. *Simplifying Parameters.*

$$p(x) = (C_5 x)^{-1} \left[ 1 - \left( \frac{x}{M} \right)^{b_0} \right], m \leq x \leq M$$

$$C_5 = \log \frac{M}{m} + \frac{1}{b_0} \left[ \left( \frac{m}{M} \right)^{b_0} - 1 \right]$$

B.3.1.3. *Cdf.*

$$P(x) = (C_5)^{-1} \left[ \log \frac{x}{m} + \gamma_0 \rho \left( \frac{m}{e^{\gamma_0 Q + \gamma_1 R}} \right)^{\frac{1}{\gamma_0 \rho}} - \gamma_0 \rho \left( \frac{x}{e^{\gamma_0 Q + \gamma_1 R}} \right)^{\frac{1}{\gamma_0 \rho}} \right], m \leq x \leq e^{\gamma_0 Q + \gamma_1 R}$$

**B.3.2.**  $\gamma_1 > \log(m)/R$ .

B.3.2.1. *Native Parameters.*

$$p(x) = \begin{cases} (C_6 x)^{-1} \frac{\gamma_0}{\gamma_1} \left( 1 - e^{-\frac{Q}{\rho}} \right) \left( \frac{x}{e^{\gamma_1 R}} \right)^{\frac{1}{\gamma_1 \rho}} & m \leq x < e^{\gamma_1 R} \\ (C_6 x)^{-1} \left[ 1 - \left( \frac{x}{e^{\gamma_0 Q + \gamma_1 R}} \right)^{\frac{1}{\gamma_0 \rho}} \right] & e^{\gamma_1 R} \leq x \leq e^{\gamma_0 Q + \gamma_1 R} \end{cases}$$

$$C_6 = \gamma_0 Q - \gamma_0 \rho \left( 1 - e^{-\frac{Q}{\rho}} \right) \left( \frac{m}{e^{\gamma_1 R}} \right)^{\frac{1}{\gamma_1 \rho}}$$

B.3.2.2. *Simplifying Parameters.*

$$p(x) = \begin{cases} (C_6 x)^{-1} \frac{b_1}{b_0} \left[ 1 - \left( \frac{A}{M} \right)^{b_0} \right] \left( \frac{x}{A} \right)^{b_1} & m \leq x < A \\ (C_6 x)^{-1} \left[ 1 - \left( \frac{x}{M} \right)^{b_0} \right] & A \leq x \leq M \end{cases}$$

$$C_6 = \log \frac{M}{A} - \frac{1}{b_0} \left[ 1 - \left( \frac{A}{M} \right)^{b_0} \right] \left( \frac{m}{A} \right)^{b_1}$$

B.3.2.3. *Cdf.*

$$P(x) = \begin{cases} (C_6)^{-1} \gamma_0 \rho \left( 1 - e^{-\frac{Q}{\rho}} \right) e^{-\frac{R}{\rho}} \left( x^{\frac{1}{\gamma_1 \rho}} - m^{\frac{1}{\gamma_1 \rho}} \right), & m \leq x < e^{\gamma_1 R} \\ (C_6)^{-1} \left[ \begin{array}{l} \log \frac{x}{m} + \gamma_0 \rho \left( e^{\frac{\gamma_1 R}{\gamma_0 \rho}} - x^{\frac{1}{\gamma_0 \rho}} \right) \left( e^{-\gamma_0 Q - \gamma_1 R} \right)^{\frac{1}{\gamma_0 \rho}} \\ + \gamma_0 \rho \left( 1 - e^{-\frac{Q}{\rho}} \right) e^{-\frac{R}{\rho}} \left( e^{\frac{R}{\rho}} - m^{\frac{1}{\gamma_1 \rho}} \right) \end{array} \right], & e^{\gamma_1 R} \leq x \leq e^{\gamma_0 Q + \gamma_1 R} \end{cases}$$

## B.4. Heavy-Seeding/Long-Latency (HSLL)

**B.4.1.**  $\gamma_1 \leq \log(m)/R$ .

B.4.1.1. *Native Parameters.*

$$p(x) = (C_7 x)^{-1} \left[ \left( \frac{e^{\gamma_0 Q + \gamma_1 R}}{x} \right)^{\frac{\sigma}{\gamma_0}} - 1 \right], m \leq x \leq e^{\gamma_0 Q + \gamma_1 R}$$

$$C_7 = \frac{\gamma_0}{\sigma} \left[ \left( \frac{e^{\gamma_0 Q + \gamma_1 R}}{m} \right)^{\frac{\sigma}{\gamma_0}} - 1 \right] + \log m - \gamma_0 Q - \gamma_1 R$$

B.4.1.2. *Simplifying Parameters.*

$$p(x) = (C_7 x)^{-1} \left[ \left( \frac{M}{x} \right)^a - 1 \right], m \leq x \leq M$$

$$C_7 = \frac{1}{a} \left( \frac{M}{m} - 1 \right)^a + \log \frac{m}{M}$$

B.4.1.3. *Cdf.*

$$P(x) = (C_7)^{-1} \left[ \frac{\gamma_0}{\sigma} \left( \frac{e^{\gamma_0 Q + \gamma_1 R}}{m} \right)^{\frac{\sigma}{\gamma_0}} - \frac{\gamma_0}{\sigma} \left( \frac{e^{\gamma_0 Q + \gamma_1 R}}{x} \right)^{\frac{\sigma}{\gamma_0}} - \log \frac{x}{m} \right], m \leq x \leq e^{\gamma_0 Q + \gamma_1 R}$$

**B.4.2.**  $\gamma_1 > \log(m)/R$ .

B.4.2.1. *Native Parameters.*

$$p(x) = \begin{cases} \frac{\gamma_0}{\gamma_1} (C_8 x)^{-1} (e^{\sigma Q} - 1), & m \leq x < e^{\gamma_1 R} \\ (C_8 x)^{-1} \left[ \left( \frac{e^{\gamma_0 Q + \gamma_1 R}}{x} \right)^{\frac{\sigma}{\gamma_0}} - 1 \right], & e^{\gamma_1 R} \leq x \leq e^{\gamma_0 Q + \gamma_1 R} \end{cases}$$

$$C_8 = -\gamma_0 Q + (e^{\sigma Q} - 1) \left( \frac{\gamma_0}{\sigma} + \gamma_0 R - \frac{\gamma_0}{\gamma_1} \log m \right)$$

B.4.2.2. *Simplifying Parameters.*

$$p(x) = \begin{cases} \frac{b_1}{b_0} (C_8 x)^{-1} \left[ \left( \frac{M}{A} \right)^a - 1 \right], & m \leq x < A \\ (C_8 x)^{-1} \left[ \left( \frac{M}{x} \right)^a - 1 \right], & A \leq x \leq M \end{cases}$$

$$C_8 = -\log \frac{M}{A} + \left[ \left( \frac{M}{A} \right)^a - 1 \right] \left( \frac{1}{a} + \frac{b_1}{b_0} \log \frac{A}{m} \right)$$

B.4.2.3. *Cdf.*

$$P(x) = \begin{cases} \frac{\gamma_0}{\gamma_1} (C_8)^{-1} (e^{\sigma Q} - 1) \log \frac{x}{m}, & m \leq x < e^{\gamma_1 R} \\ (C_8)^{-1} \begin{bmatrix} \frac{\gamma_0}{\sigma} e^{\sigma Q} + \gamma_1 R - \frac{\gamma_0}{\sigma} \left( \frac{e^{\gamma_0 Q + \gamma_1 R}}{x} \right)^{\frac{\sigma}{\gamma_0}} \\ -\log x + \frac{\gamma_0}{\gamma_1} (e^{\sigma Q} - 1) \log \frac{e^{\gamma_1 R}}{m} \end{bmatrix}, & e^{\gamma_1 R} \leq x \leq e^{\gamma_0 Q + \gamma_1 R} \end{cases}$$

## B.5. Complete Suppression by Primary Tumor (CSPT)

In this case, we must have  $\gamma_1 > \log(m)/R$  or no metastases will be detectable.

### B.5.1. $\gamma_1 > \log(m)/R$ .

#### B.5.1.1. *Native Parameters.*

$$p(x) = \frac{(1-v)x^{-1} \left(\frac{x}{e^{\gamma_1 R}}\right)^{1/(\gamma_1 \rho)}}{\gamma_1 \rho \left[1 - \left(\frac{m}{e^{\gamma_1 R}}\right)^{1/(\gamma_1 \rho)}\right]}, \text{ if } m \leq x < e^{\gamma_1 R},$$

$$\Pr(x = e^{\gamma_1 R}) = v$$

where

$$v = \frac{\left[\frac{1}{\sigma}(e^{\sigma Q} - 1) + \rho \left(e^{-\frac{Q}{\rho}} - 1\right)\right]}{\rho \left(e^{\sigma Q} - e^{-\frac{Q}{\rho}}\right) \left[1 - \left(\frac{m}{e^{\gamma_1 R}}\right)^{1/(\gamma_1 \rho)}\right] + \frac{1}{\sigma}(e^{\sigma Q} - 1) + \rho \left(e^{-\frac{Q}{\rho}} - 1\right)}.$$

Given  $\gamma_1$ ,  $\rho$ , and  $v$ , (3.12) can be solved for  $Q$  in terms of  $\sigma$  and has a unique positive solution in terms of  $\sigma$  if and only if

$$0 \leq \sigma < \frac{1-v}{\rho v \left(1 - \left(\frac{m}{e^{\gamma_1 R}}\right)^{1/(\gamma_1 \rho)}\right)}.$$

When it exists, that solution satisfies

$$\frac{e^{\sigma Q} - 1}{e^{-\frac{Q}{\rho}} - 1} = \frac{\sigma \rho \left[1 - v \left(\frac{m}{e^{\gamma_1 R}}\right)^{1/(\gamma_1 \rho)}\right]}{\sigma \rho v \left[1 - \left(\frac{m}{e^{\gamma_1 R}}\right)^{1/(\gamma_1 \rho)}\right] + v - 1}.$$

#### B.5.1.2. *Simplifying Parameters.*

$$p(x) = \frac{b_1(1-v)x^{-1} \left(\frac{x}{A}\right)^{b_1}}{1 - \left(\frac{m}{A}\right)^{b_1}}, \text{ } m \leq x < A$$

$$\Pr(x = A) = v$$

where  $v$  is the same as in the native parameters case.

B.5.1.3. *Cdf.*

$$P(x) = \begin{cases} (C_9)^{-1} \rho \left[ \left( \frac{x}{e^{\gamma_1 R}} \right)^{\frac{1}{\gamma_1 \rho}} - \left( \frac{m}{e^{\gamma_1 R}} \right)^{\frac{1}{\gamma_1 \rho}} \right] \left( e^{\sigma Q} - e^{-\frac{Q}{\rho}} \right), & m \leq x < e^{\gamma_1 R} \\ 1, & x = e^{\gamma_1 R} \end{cases}$$

## B.6. Instantaneous Infinite Shedding and Metastasis Growth (IISMG)

In this model, we set  $\sigma = a\gamma_0$  and  $\rho = \frac{1}{b_0\gamma_0}$  and take the limit as  $\gamma_0 \rightarrow \infty$ .

**B.6.1.**  $\gamma_1 \leq \log(m)/R$ .

$$p(x) = am^a x^{-1-a}, m \leq x < \infty$$

B.6.1.1. *Cdf.*

$$P(x) = 1 - m^a x^{-a}, m \leq x < \infty$$

**B.6.2.**  $\gamma_1 > \log(m)/R$ .

B.6.2.1. *Native Parameters.*

$$p(x) = (x)^{-1} \frac{ab_0 \left(\frac{e^{\gamma_1 R}}{x}\right)^a}{a + b_0}, e^{\gamma_1 R} < x < \infty$$

and

$$\Pr(x = e^{\gamma_1 R}) = \frac{a}{a + b_0}.$$

B.6.2.2. *Simplifying Parameters.*

$$p(x) = (x)^{-1} a \left(1 - \frac{a}{a + b_0}\right) \left(\frac{A}{x}\right)^a, A < x < \infty$$

and

$$\Pr(x = A) = \frac{a}{a + b_0}.$$

B.6.2.3. *Cdf.*

$$P(x) = 1 - \left(\frac{b_0}{a + b_0}\right) \left(\frac{e^{\gamma_1 R}}{x}\right)^a, e^{\gamma_1 R} \leq x < \infty$$

## APPENDIX C

### Some Computations for the CSPT and IISMG Models

#### C.1. Proof that $Q$ can be determined from $\sigma$ in the CSPT model

CLAIM. Given  $\gamma_1, \rho, \sigma > 0$  and  $0 < v < 1$ , we can solve the equation

$$v = \frac{\frac{1}{\sigma} (e^{\sigma Q} - 1) + \rho \left( e^{-\frac{Q}{\rho}} - 1 \right)}{\rho \left( e^{\sigma Q} - e^{-\frac{Q}{\rho}} \right) \left[ 1 - \left( \frac{m}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)} \right] + \frac{1}{\sigma} (e^{\sigma Q} - 1) + \rho \left( e^{-\frac{Q}{\rho}} - 1 \right)}$$

to obtain a positive value for  $Q$  and this solution is unique if

$$0 < \sigma < \frac{1 - v}{\rho v \left[ 1 - \left( \frac{m}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)} \right]}.$$

If

$$\sigma \geq \frac{1 - v}{\rho v \left[ 1 - \left( \frac{m}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)} \right]},$$

there is no solution for  $Q$ .

PROOF. We rewrite  $e^{\sigma Q} - e^{-\frac{Q}{\rho}}$  in the expression for  $v$  above as  $e^{\sigma Q} - 1 + 1 - e^{-\frac{Q}{\rho}}$  and then separate terms involving  $e^{\sigma Q} - 1$  and  $e^{-\frac{Q}{\rho}} - 1$  in order to obtain

$$\frac{e^{\sigma Q} - 1}{e^{-\frac{Q}{\rho}} - 1} = \frac{\sigma \rho \left[ 1 - v \left( \frac{m}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)} \right]}{\sigma \rho v \left[ 1 - \left( \frac{m}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)} \right] + v - 1}.$$

The fraction on the left is negative so long as  $Q \neq 0$ . The numerator of the fraction on the right is positive. Therefore, we will only have a solution for  $Q$  if

$$\sigma \rho v \left[ 1 - \left( \frac{m}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)} \right] + v - 1 < 0$$

and this is equivalent to

$$\sigma < \frac{1 - v}{\rho v \left[ 1 - \left( \frac{m}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)} \right]}.$$

If we separate the exponentials in our equation for  $Q$ , we have

$$(C.1) \quad e^{\sigma Q} - 1 = \frac{\sigma \rho \left[ 1 - v \left( \frac{m}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)} \right]}{\sigma \rho v \left[ 1 - \left( \frac{m}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)} \right] + v - 1} \left( e^{-\frac{Q}{\rho}} - 1 \right).$$

If we use  $k$  to denote the positive quantity

$$\frac{\sigma \rho \left[ 1 - v \left( \frac{m}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)} \right]}{1 - v - \sigma \rho v \left[ 1 - \left( \frac{m}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)} \right]},$$

then solving (C.1) is equivalent to solving  $h(Q) = 0$  where

$$h(Q) = e^{\sigma Q} - 1 - k \left( 1 - e^{-\frac{Q}{\rho}} \right).$$

We note that  $h(0) = 0$ ,  $h'(Q) = \sigma e^{\sigma Q} - \frac{k}{\rho} e^{-\frac{Q}{\rho}}$ , and  $h''(Q) = \sigma^2 e^{\sigma Q} + \frac{k}{\rho^2} e^{-\frac{Q}{\rho}} > 0$ .

It is also true that  $k > \sigma \rho$  because

$$\begin{aligned} 1 - v \left( \frac{m}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)} &> 1 - v \left( \frac{m}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)} - v \left[ 1 - \left( \frac{m}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)} \right] - \sigma \rho v \left[ 1 - \left( \frac{m}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)} \right] \\ &= 1 - v - \sigma \rho v \left[ 1 - \left( \frac{m}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)} \right] \end{aligned}$$

and therefore

$$\frac{1 - v \left( \frac{m}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)}}{1 - v - \sigma \rho v \left[ 1 - \left( \frac{m}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)} \right]} > 1$$

or

$$k = \frac{\sigma \rho \left[ 1 - v \left( \frac{m}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)} \right]}{1 - v - \sigma \rho v \left[ 1 - \left( \frac{m}{e^{\gamma_1 R}} \right)^{1/(\gamma_1 \rho)} \right]} > \sigma \rho.$$

Because  $\sigma \rho < k$ ,

$$h'(0) = \sigma - \frac{k}{\rho} < 0,$$

and therefore  $h$  is initially decreasing. Because  $h(0) = 0$ , there must exist  $Q_n > 0$  such that  $h(Q_n) < 0$ . But

$$\lim_{Q \rightarrow \infty} h(Q) = \infty,$$

and by the intermediate value theorem, there must exist  $Q > 0$  such that  $h(Q) = 0$ . Suppose that this  $Q$  is not unique. Then we can find  $Q_2 > Q_1 > 0$  such that  $h(Q_2) = h(Q_1) = 0$ . By Rolle's theorem, there would exist  $c_1 \in (0, Q_1)$  and  $c_2 \in (Q_1, Q_2)$  such that  $h'(c_1) = h'(c_2) = 0$ . But  $h'' > 0$  so  $h'(c_1) < h'(c_2)$ , a contradiction. Therefore,  $Q$  must be unique.  $\square$

## C.2. Proof that $g(b_1)$ is decreasing

CLAIM. The function  $g(b_1)$  given by

$$g(b_1) = \frac{1}{b_1} - \frac{\log\left(\frac{x_n}{m}\right)}{\left(\frac{x_n}{m}\right)^{b_1} - 1}$$

is decreasing

PROOF. In order to simplify the expression for  $g(b_1)$ , we set  $a = \frac{x_n}{m}$  and  $a^{b_1} = e^t, t \in \mathbb{R}$ . Then  $g(b_1) = \frac{\log a}{t} - \frac{\log a}{e^t - 1} = \log a \left( \frac{1}{t} - \frac{1}{e^t - 1} \right)$ .

We have to show that  $h(t) = \frac{1}{t} - \frac{1}{e^t - 1}$  is decreasing, i.e. that  $h'(t) = -\frac{1}{t^2} - \frac{e^t}{(e^t - 1)^2} \leq 0$  and that could only happen if  $t^2 e^t \leq (e^t - 1)^2$ . Because taking the square root is an order-preserving operation, the inequality  $t^2 e^t \leq (e^t - 1)^2$  is equivalent to  $|t|e^{t/2} \leq |e^t - 1|$  which can be written as  $te^t \leq e^t - 1$  for  $t \geq 0$  and  $-te^t < 1 - e^t$  for  $t < 0$ . These two inequalities could also be written as  $t/2 \leq \sinh t$  for  $t \geq 0$  and  $\sinh t < t/2$  for  $t < 0$ .

We set  $\alpha(t) = t/2$  and  $\beta(t) = \sinh(t)$ , and note that  $\alpha(0) = 0 = \beta(0)$  and  $\alpha'(t) = 1/2 < 1 \leq \cosh t = \beta'(t)$ . Then when  $t > 0$ ,  $\alpha(t) = \int_0^t \alpha'(\tau) d\tau \leq \int_0^t \beta'(\tau) d\tau = \beta(t)$ , so  $t/2 \leq \sinh t$ . On the other hand, when  $t < 0$ ,  $-\alpha(t) = \int_t^0 \alpha'(\tau) d\tau \leq \int_t^0 \beta'(\tau) d\tau = -\beta(t)$ , so  $-t/2 \leq -\sinh t$  so  $\sinh t \leq t/2$ . It then follows that  $g(b_1)$  is decreasing  $\square$