Spring 2018

The Mathematics of Cancer: Fitting the Gompertz Equation to Tumor Growth

Dyjuan Tatro
Bard College, dt1038@bard.edu

The Mathematics of Cancer: Fitting the Gompertz Equation to Tumor Growth

A Senior Project submitted to
The Division of Science, Mathematics, and Computing
of
Bard College

by
Dyjuan Tatro

Annandale-on-Hudson, New York
May, 2018

Abstract

Mathematical models are finding increased use in biology, and partuculary in the field of cancer research. In relation to cancer, systems of differential equations have been proven to model tumor growth for many types of cancer while taking into account one or many features of tumor growth. One feature of tumor growth that models must take into account is that tumors do not grow exponentially. One model that embodies this feature is the Gomperts Model of Cell Growth. By fitting this model to longterm breast cancer study data, this project ascertains gompertzian parameters that can be used to predicts tumor growth as a function of time.

Contents

List of Figures

Dedication

I dedicate this project to the prisoners in New York State who could accomplish similar work if only provided an educational opportunity comparable to the one Bard College has granted me.

Acknowledgments

I must thank, first and foremost, my adviser Dr. Micheal Tibbetts for his guidance and constant encouragement throughout this process. His believed that I would finish despite my own doubt and uncertainty. Further, thank you to Professors Keri-Ann Norton and Stefan Mendez-Diez for their work advising on this project. Thank you to Max Kenner, the Bard Prison Initiative, and Bard College for having created the opportunity for me to be a Bard student. Thank you to my family for having supported me throughout my imprisonment, and having granted me the peace of mind to focus on my studies under adverse circumstances. Thank you to Lynn Novick and Sarah Botstein for their encoragement and support as well as for the BPI documentary they have under production. Your film shall change the world.

1
Introduction

At the intersection of mathematics and biology resounds a call to use math in this century to do what it did for physics in the twentieth-century [1]. That is, to apply math to biology in such a manner as to increase the predictive power of the biological field [1]. Let us recall that Physicists mathematically predicted the Higgs Bosun before having materially discovered it. Such predictive power can be harnessed toward biology to aid in areas of cancer research, bioinformatics, and drug testing, to name a few. In the arena of cancer research, mathematical models have become more and more common as a means to model cancer cell proliferation, drug scheduling, and tumor development. Researchers have erected many different models for cancer cell proliferation, ranging form those that model with one degree of freedom, say tumor size, to those that model cancer at several degrees of freedom, say growth, necrosis, and angiogenesis [10]. One of the problems that has arisen with modeling tumor growth is that tumors do not grow exponentially. In other words, tumor cell proliferation cannot be reduced to the biology of cell divison, mitosis.

This problem means that, even at one degree of freedom, tumor growth has to be modeled with mathematically nuanced constructs.

Models constructed with differential eaquations have been proven to accurately predict tumor growth curves for many types of tumor [10]. One differential equation model tha manages such predictions with a high degree of fidelity is the Gompertz Model of cell growth. The Gompertz Model's key feature is that it accounts for exponientail decay. In relation to tumor growth, this feature means that the Gompertz Model captures how tumour growth rates decrease as a the mass of the tumour increases. This project fits the Gompertz Model to a dataset prudeced by Bloom et. al. that tracked untreated breast cancer in 250 over about twenty years. By fitting the model to this datatset, this project can ascertain gompertzian pararmeters to predict breast tumor growth rates as a fuction of time.

2
An Introduction to Cancer

2.1 Cancer

Cancer constitutes a highly dynamic set of diseases grounded in genomic mutations that cause malignant cellular growth [2]. In a classic review article titled "The Hallmarks of Cancer," Hanahan and Weinberg compile the traits that make a disease a cancer [2]. They identify six physiological characteristics of cell biology that converge to cause malignant growth: self-sufficiency in growth signals, insensitivity to growth-inhibitory signals, evasion of apoptosis, limitless replicative potential, sustained angiogenesis, and tissue invasion coupled with metastasis [2]. As a tumor develops, it must acquire each of these physiologic changes in order to breach cellular anticancer mechanisms [2]. It is thus helpful to view cancer as the outcome of a multistep process wherein the disease emerges from an interplay of factors resulting from progressive genomic alterations [2]. Hanahan and Weinberg's traits, undergird the interplay of factors that cause cancer, and the vast majority, if not all, human cancers share those six traits [2].

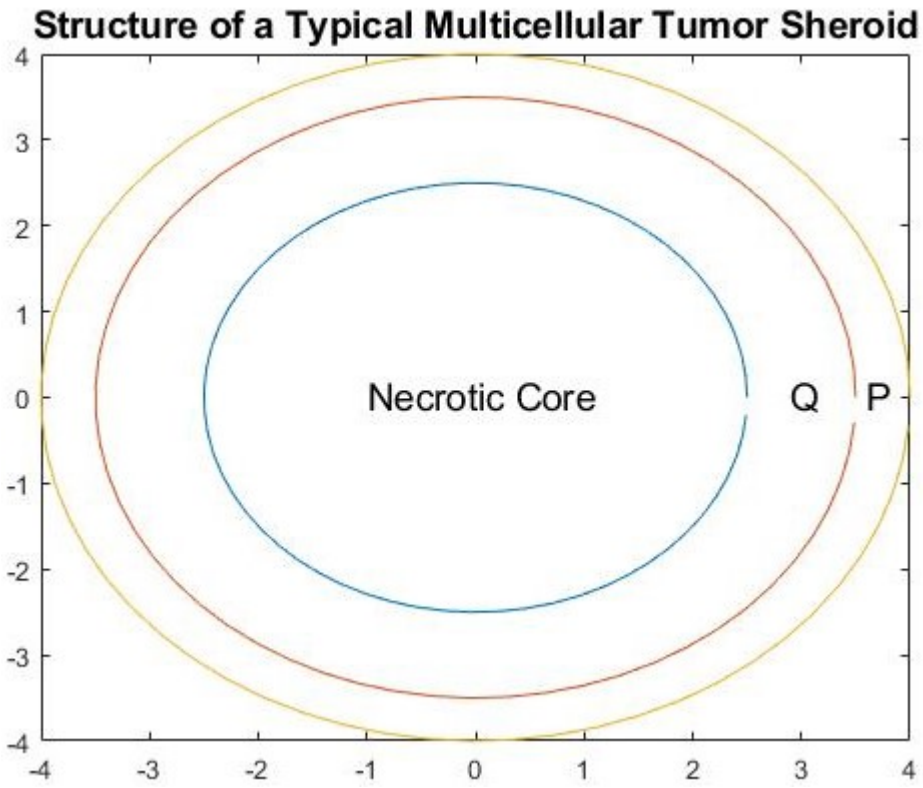

Figure 2.1.1. Structure of a Tumor Spheroid

2.2 Tumor Biology

Cancer manifests in the form of a tumor spheroid, an abnormal mass of uncontrollably dividing cells. A tumor spheroid generally develops three cell regions, as seen in Figure 1.1.1, that differ in accordance with nutrient supply. The cells in the proliferative zone, P, undergo constant cell division as they have access to the nutrients necessary for proliferation [2]. The quiescent layer of cells, Q, lose their ability to proliferate due to an insufficient nutrient supply, but receive enough sustenance to remain alive. The cells that constitute the necrotic core, N, however, die due to prolonged nutrient deficiency. As a tumor spheroid grows, it becomes harder for the cells toward its centre and core to receive nutrients.

3
Mathematical Models

3.1 Overview

Over the past several decades, the factors and forces that drive tumor growth—tumor growth kinetics—have come under biological study, and, hence, have garnered extensive experimental study [10]. Researchers have found that both human and animal relative tumor growth rates diminish with time [10]. Or, to restate this finding, the time it takes for human and animal tumors to double in size increases with time [10]. Such characteristics, the differential relation of tumor growth rates to time, make tumorigenesis favorable to expression in the form of differential equations.

A differential equation constitutes a mathematical formalism relating a function to its derivatives. It is a relational concept in the form of an equation. The function usually represents a quantifiable physical entity while its derivatives describe that quantities which that entity changes with time, its rate of change [18]. When related with an equation, of tumors many mathematical constructs have been developed to model tumorigenesis. At the most basic, such constructs model cancer-cell proliferation. The more complicated constructs model cancer-cell proliferation in

conjunction with with factors such as, necrosis, angiogenesis, and tumor response to drug treatment.

Mathematical modeling provides for the expression of which effects one wants to consider to the exclusion of those one may not understand or find of interest [18]. This degree of flexibility has great value to the study of cancer in two ways, 1) models can provide a vehicle to test theoretical hypotheses of tumor growth by assessing a theories descriptive power against experimental data, and 2) models can estimate the pattern on tumor development for use as a prognostic clinical tool or to assess the clinical efficacy of drug therapies under development [10].

3.2 Bloom Data

One of the datasets Norton applies the Gompertz model to is that of Bloom et. al. from their article, "Natural History of Untreated Breast Cancer (1805-1933) Comparison of Untreated and Treated Cases According to Histological Grade Malignancy". At Middlesex Hospital, London, Bloom et. al. tracked 250 untreated cases of breast cancer in woman. The study occurred during the period 1803-1820 from the patients onset of symptoms through death [22]. Figure 3.2.1 constitutes a reproduction of the Bloom study data. As seen in Figure 3.2.1, all the patients in the study had died with twenty years [22].

Graphing the data in Figure 3.2.1. produces the curve seen in Figure 3.2.2.

3.3 The Gompertz Model of Cell Growth

As discussed in Section 1.2, a tumor will grow more slowly the bigger it becomes as less and less of its cell mass will be concentrated within the proliferative zone relative to the number of cells at the tumor's centre and core. Mathematically speaking, this decrease in the growth rate embodies means that tumor's exhibit non-exponential

Deaths Occurring Each Year, Calculated from Onset of Symptoms for 250 Untreated Cases of Breast Cancer				
Year of Death	No. of Cases	Cumulative Total	Survivals	
1	34	34	216	86%
2	52	86	164	66%
3	55	141	109	43.6%
4	38	179	71	28%
5	25	204	46	18.4%
6	13	217	33	13%
7	10	227	23	9%
8	6	233	17	7%
9	5	238	12	5%
10	3	241	9	3.6%
11	4	245	5	2%
13	1	246	4	1.6%
15	2	248	2	0.8%
16	1	249	1	0.4%
19	1	250	0	0.0%
Survival of Treated Breast Cancer, Middlesex Hospital, 1805-1933 (250 Cases).				

Figure 3.2.1. Bloom Data (Reproduced from [22]).

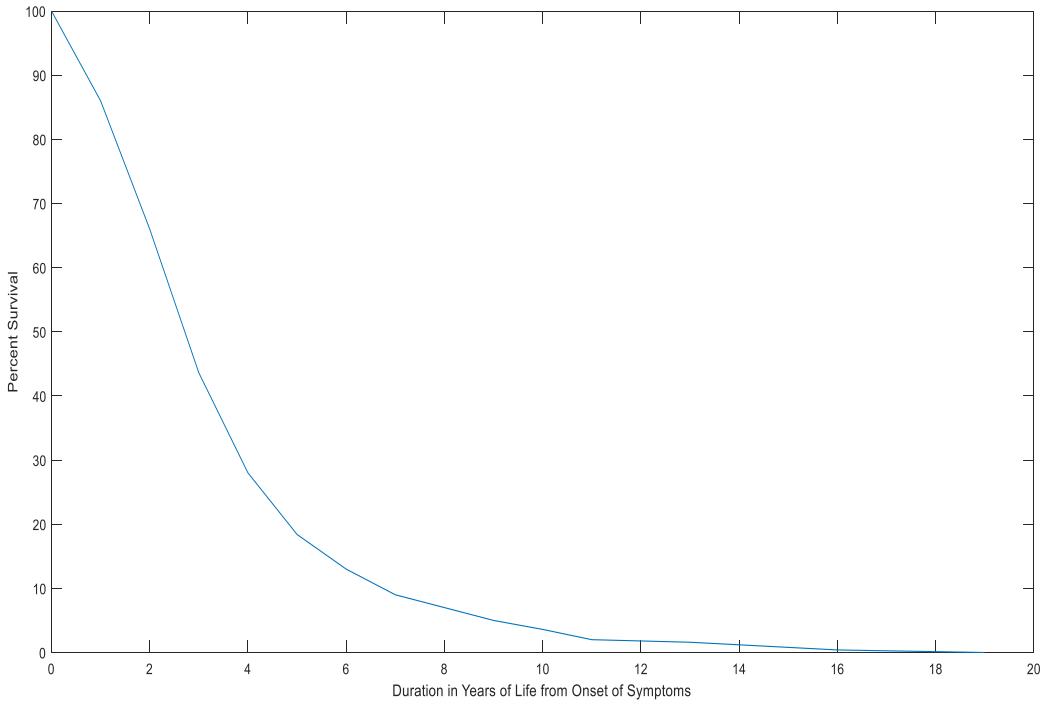

Figure 3.2.2. Graphical Representation of Bloom Data.

growth patterns, which the Gompertz model takes into account. The Gompertz model's essential characteristic is its ability to exhibit exponential decay of relative tumor growth rates (SB). In other words, the Gompertz model embodies the fact that tumor cell growth rates decrease as a function of time.

The Gompertz equation had originally been constructed for the purpose of actuarial analysis, but later came into use a a growth curve. In a 1984 paper,"A Stochastic Numerical Model of Breast Cancer that Simulates Clinical Data", Speer et. al. propose that every individual tumor initially grows with identical Gompertzian parameters, and then becomes kinetically heterogenous through a random time-dependent process [4, 20]. In his 1988 paper, "A Gompertzian Model of Human Breast Cancer Growth", Larry Norton points out that neither theory nor data support Speer et. .al.'s assumption of uniform nascent tumor growth, and, further, that the complex growth curves produced by Speer et. al.'s model fail to fit individual cancer growth patterns. Considering the clinical and theoretical importance of human breast cancer, and in response to Speer et. al., Norton fits a parsimonious, unadorned Gompertzian Model to three historical datasets by taking kinetic heterogeneity as an intrinsic property of neoplasia.

3.4 Norton's Fit

Norton lays out the gompertzian equation as:

$$N(t) = N(0)e^{k[1-e^{(-bt)}]} \tag{3.4.1}$$

where $k = \log_e \left[\frac{N(\infty)}{N(0)}\right]$.

$N(t)$ is a function of $N(0)$, t, and b as well as a limiting size N(∞). As the Gompertz model has been adapted to the kinetics of tumor growth, Norton analyses his data with $N(t)$ as the size of an individual patient's tumor at time t, which he

measures from the onset of symptoms [4]. He defines tumor size at the onset of symptoms as $N(0)$ and holds $N(\infty)$ constant [4]. By rearrangement of of equation 3.4.1:

$$t_i = \left(\frac{-1}{b_i} \log_e 1 - \frac{1}{k} \log -e \left(\frac{N_L}{N(0)} \right) \right) \tag{3.4.2}$$

where $P_t(t_i)$ equals a portion of the 250 cancers from the Bloom data with Gompertzian parameters $b \leq b_i$.

Using these methods, Norton Norton was ablw to define the probability distribution of b ! [21, ?Ln]. He then randomly chose a value of b_i from the normal-log distribution and calculated t_i for $N(t_i) = N_L$. Using Values of t_i, he estimated $P_L(t)$, which he graphically compared with $P_L(t)$ from the Bloom Data [4].

3.5 Norton's Parameters

Norton finds that the Bloom data reflects the Gompertzian equation with $N(0) = 4.8 \times 10^9$ cells, $N(\infty) = 3.1 \times 10^{12}$ cells, and lethal tumor cell count of $N_L = 10^{12}$ cells '[4].

4
Project Fit

4.1 Poject Model

Equation 3.2.1 is an analytic solution to the following system of differential equations [20]:

$$\begin{cases} \frac{dN}{dt} = -rNln\left(\frac{N}{K}\right); N(0) = N_0 \\ N(t=0) = 1mm^3 \end{cases} \tag{4.1.1}$$

which we can solve exactly as

$$N(t) = K\left(\frac{N_0}{k}\right)^{e^{-rt}} \tag{4.1.2}$$

from which we can see that the volume asymptotically converges to a carrying given by $K = N_0 e^{\frac{r_0}{r}}$. r specfies growth in proportion to cell population size. The cell population will only grow faster as it becomes larger if it is below the carrying capacity.

4.2 $lsqcurvefit$

In Matlab, I fit the equation to the Bloom data using $lsqcurvefit$, which solves nonlinear equations in the least-squares sense. More specifically, $lsqcurvefit$ takes

the given $xdata$, coupled with observed $ydata$, and finds coefficients best-fitted to the equation

$$\frac{min}{x}\frac{1}{2}\left\|F(x,xdata)-ydata\right\|_2^2=\frac{1}{2}\sum_i(F(x,xdata_i)-ydata_i)^2 \tag{4.2.1}$$

where the $xdata$ and $ydata$ constitute vectors and $F(x,data)$ constitutes a vector valued function [18]. The $Matlab$ code for using this equation with the Bloom data is in Appendix A.2.

Fitting the Gompertz model to the Bloom data, using lsqcurve fit via the $Matlab$ code in Apendix A.1 , returns the following fit and parameters:

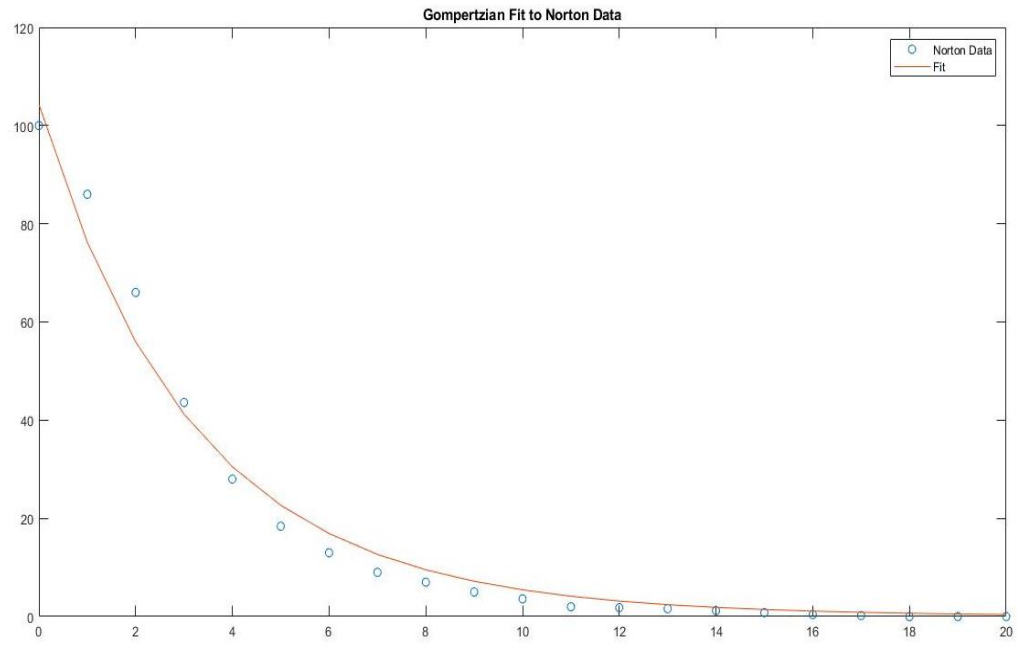

Figure 4.2.1. A Gompertzian Fit to the Bloom Data.

$r = 8.032444124239194e-09$

$K = 1.037246983488218e+02$ billion cells

$N_0 = 1.34$ million cells.

5
Conclusion

5.1 Conclusion

This project demonstrates that the Gompertz Model of Cell Growth accurately models tumor growth in accordance with Norton findings that an unadorned gompertzian model is sufficient to predict tumor growth. Taking the Bloom data and fitting it with the gompertzian model returned parameters r, K and,N_0, consistent with Norton et. al.'s findings.

Appendix A
Matlab Code

A.1 Gompertz Mode

$functionV = Gompertz(p, t)$

$V = p(1). * (p(2)/p(1)).^{e}xp(-p(3) * t);$

$withparametersp(1) = K, p(2) = initialpopulation, andp(3) = r$

End

A.2 Code to Fit Gompertz Model to Study Data

$years = [012345678910111315 1619];$

$percent = [100866643.62818.4139753.621.6.8.40];$

$p0 = [.11000];// \ options = optimset('MaxFunEvals', 100000,' Maxiter', 5000);$

$[p, error] = lsqcurvefit(@gompertz, p0, years, percent, [], [], options)$

$modelpercent = gompertz(p, years);$

$plot(years, percent,' o', years, modelpercent)$

Bibliography

[1] Bellomo, N. et. al. (2007). On The Foundations of Cancer Modelling: Selected Topics, Speculations, and Perspectives. *Mathematical Models and Methods in Applied Sciences*. 18:4, 593-646.

[2] Weinberg, R.A. (2014). *The Biology of Cancer*, Second Edition. New York: Garland Science.

[3] Altrock, P.M. et. al. (2015). The Mathematics of Cancer: integrating quantitive models. *Nature*. 15, 730-745.

[4] Norton, L. (1988). A Gompertzian Model of Human Breast Cancer Growth. *Cancer Research*, 48, 7067-7071.

[5] Zhang, P. and Brusic, V. (2014). Mathematical Modeling for Novel Cancer Drug Discovery and Development. *Expert Opinion on Drug Discovery*, 9:10, 1133-1150.

[6] Folkman, J. (2002). Role of Angiogenesis in Tumor Growth and Metastasis. *Seminars in Biology*, 29:6, 15-18.

[7] Baylin, S.B. and Ohm, J.E. (2006). Epigenetic Gene Silencing in Cancer–a mechanism for early oncogenic pathway addiction. *Nature Reviews*, 6, 107-116.

[8] Vogelstein, B. and Kinzler, K.W. (2004). Cancer Genes and the Pathways they Control. *Nature Medicine*, 10:8, 789-799.

[9] Friedl, P. and Wolf, K. (2003). Tumour-Cell Invasion and Migration: Diversity and Escape Mechanisms. *Nature Reviews*, 3, 362-374.

[10] Benzerky, S. et al. (2014). Classical Mathematical Models for the Description and Prediction of Experimental Tumor Growth. *PLOS Computational Biology*, 10:8, e1003800.

[11] Engelhart, M. et al. (2010). Optimal Control for Selected Cancer Chemotherapy ODE Models: A View on the Potential of Optimal Schedules and Choice of Objective Function. *Mathematical Biosciences*, 229, 123-134.

[12] Masood, S.A. et. al. (2006). Specific Killing of Multiple Myeloma Cells by (-)-epigallocatechin-3-gallate Extracted from Green Tea: Biologic Activity and Therapeutic Implications. *Blood*, 108:8, 2804-2810.

[13] Hanahan, D. and Weinberg, R.A. (2000). The Hallmarks of Cancer. *Cell*, 100, 57-70.

[14] Anderson, A.R.A .et. al. (2000). Mathematical Modelling of Tumour Invasion and Metastasis. *Journal of Theoretical Medicine*, 2, 129-54.

[15] Byrne, H.M. and Chaplain, M.A.J. (1996). Modelling the Role of Cell-Cell Adhesion in the Growth and Development of Carcinomas. *Mathematical Computational Biology*, 24:12, 1-17.

[16] Magni, P. et. al. (2006). A Mathematical Model to Study the Effects of Drugs Administration on Tumour Growth Dynamics. *Mathematical Biosciences* 200, 127-157.

[17] Kirscher, D. and Panetta, J.C. (1998). Modeling Immunotherapy of the Tumor-Immune. *Journal of Mathematical Biology*, 37, 235-252.

[18] Davis, P. (1999). *Differential Equations: Modeling With Matlab*. New Jersey: Prentice Hall.

[19] Pelengaris, S. and Khan, M. (2013). Introduction. In S. Pelengaris and M. Khan (Eds.), *The Molculear Biology of Cancer: A Bridge from Bench to Bedside*, Second Edition. West Sussex: Wiley-Blackwell.

[20] Speer, J.F. et. al. (1984). A Stochastic Numerical Model of Breast Cancer that Simulates Clinical Data. *Cancer Research*, 44, 4124-4130.

[21] Norton, L. et. al. (1976). Predicting the Course of Gompertzian Growth. *Nature*, 264, 542-545.

[22] Bloom, H.J.G. (1962). Natural History of Untreated Breast Cancer (1805-1933) Comparison of Untreated and Treated Cases According to Histological Grade of Malignancy. *British Medical Journal*, 2, 5299, 213-222.