Estimating tumor growth rates in vivo

Anne Talkington and Rick Durrett
Dept. of Math, Duke University, Durham, NC

November 19, 2014

Abstract

In this paper we develop methods for inferring tumor growth rates from the observation of tumor volumes at two time points. We fit power law, exponential, Gompertz, and Spratt's generalized logistic model to five data sets. Though the data sets are small and there are biases due to the way the samples were ascertained, several interesting conclusions come from our analyses.

1 Introduction

Finding formulas to predict the growth of tumors has been of interest since the early days of cancer research. Many models have been proposed, but there is still no consensus about the growth patterns that solid tumors exhibit [7]. This is an important problem because an accurate model of tumor growth is needed for evaluating screening strategies [18], optimizing radiation treatment protocols [27, 2], and making decisions about patient treatment [5, 6].

Recently, Sarapata and de Pillis [29] have examined the effectiveness of a half-dozen different models in fitting the growth rates of in vitro tumor growth in ten different types of cancer. While the survey in [29] is impressive for its scope, the behavior of cells grown in a laboratory setting where they always have an ample supply of nutrients is not the same as that of tumors in a human body.

One cannot have a very long time series of observations of tumor size in human patients because, in most cases, soon after the tumor is detected it will be treated, and that will change the dynamics. However, we have found five studies where tumor sizes of different types of cancers were measured two times before treatment and the measurements were given in the paper, [11], [13], [28], [21], and [22]. We describe the data in more detail in Section 4. Another data set gives the time until death of 250 untreated cases observed from 1805 to 1933, see [1]. That data is not useful for us because there is no information on tumor sizes.

In the next section, we review the models that we will consider. Each model has a growth rate r. Given the volumes V_1 and V_2 at two time points t_1 and t_2, there is a unique value of r that makes the tumor grow from volume V_1 to V_2 in time $t_2 - t_1$. We use the average of the growth rates that we compute in this way as an estimate for the growth rate. Chingola and Foroni [3] used this approach to fit the Gompertz model to data on the growth of multicellular tumor spheroids. Here, we extend their method to other commonly used growth models.

A new feature of our analysis is that in order to find the best model we plot the estimated values of r versus the initial tumor volume V_1 and look at trends in the sizes of the rates. To explain our method, we begin by noting that all of our models have the form

$$\frac{dV}{dt} = rV(t)f(V(t))$$

We call f the *correction factor* because it gives the deviation from exponential growth. If the true tumor growth law has $f_0 < f$ then the estimated growth rates will tend to decrease as the tumor volume increases. For example, this will occur if we fit the exponential, $f \equiv 1$ but the true tumor growth law has $f_0(v) \downarrow 0$. Conversely, if the true growth law has $f_0 > f$ then the estimated growth rates will tend to increase as the tumor volume increases. This will occur if growth follows a power law, which corresponds to $f_0(v) = v^{\alpha-1}$, and we fit a power law with a value of α that is too small.

2 Tumor growth laws

In writing this section we have relied heavily on the surveys in [7] and [26]. This material can also be found in Chapter 4 of the excellent recent book by Wodarz and Komarova [40].

1. Exponential growth is the most commonly used tumor growth model. Cells divide at a constant rate independent of tumor size, so the tumor volume V satisfies

$$\frac{dV}{dt} = rV \tag{1}$$

The solution is $V(t) = V_0 e^{rt}$, where V_0 is the size at time 0. This model was first applied to cancer in 1956 by Collins et al [4]. Their work introduced the tumor doubling time, $DT = (\ln 2)/r$, to quantify the rate of growth. The exponential growth law has been used to model leukemia [30]. Friberg and Mattson [6] found exponential growth in a study of more than 300 untreated lung cancers.

Exponential growth describes the ideal scenario in which cells divide without constraint, and continue to double indefinitely. This should be a good model of early tumor growth. However, limitations of the availability of nutrients, oxygen, and space imply that exponential growth is not appropriate for the long term growth of solid tumors, so we must consider alternative formulations.

2. The power-law differential equation generalizes the exponential:

$$\frac{dV}{dt} = rV(t)^{\alpha} \tag{2}$$

When $\alpha = 1$ this reduces to the exponential. The solution when $\alpha < 1$ is

$$V(t) = (V_0^{1-\alpha} + (1-\alpha)rt)^{1/(1-\alpha)} \tag{3}$$

If we assume that growth only occurs at the surface of a three dimensional solid tumor then $\alpha = 2/3$. This value of α was suggested in 1932 by Mayenord [16]. This choice is supported by the observation of linear growth of the diameter of 27 glioma patients, see [15] .

3. The power-law with linear death has the form

$$\frac{dV}{dt} = rV(t)^\alpha - r\frac{V(t)}{K^{1-\alpha}} = rV^\alpha\left(1 - \left(\frac{V}{K}\right)^{1-\alpha}\right) \tag{4}$$

When $\alpha = 2/3$ this is van Bertalanffy model [37]. When $\alpha = 3/4$ this is the universal curve of West, Brown, and Enquist [39], who used it to fit the growth of 13 different organisms. Guiot et al [9] used this model to fit the growth of tumor spheroids in vitro and patient data. Castorina et al. [2] have investigated the implications of this growth law for radiotherapy. We mention this model here for completeness. We will not fit it to our data.

4. The Gompertz model was put forward by Benjamin Gompertz in 1825 as a means to explain human mortality curves [8] and hence determine the value of life insurances. A hundred years later, it was proposed as a model for biologic growth by the geneticist Sewall Wright. On page 494 of [41], he observes that "the average growth power, as measured by the percentage rate of increase, tends to fall at a more or less uniform percentage rate." In other words, the growth rate of an organism or organ tends to decrease exponentially. This model became popular in the cancer literature after Anna Laird [14] used it to successfully fit the growth of 19 tumor cell lines. Larry Norton [23, 24, 25] has for many years championed the use of the Gompertz in modeling breast cancer growth.

One way of thinking about this model, which is close to Wright's description, is to write

$$\frac{dV}{dt} = \alpha(t)V(t) \quad \text{where} \quad \frac{d\alpha}{dt} = -r\alpha(t).$$

This leads to a solution

$$V(t) = V_0 \exp\left(\frac{\alpha_0}{r}(1 - e^{-rt})\right) \tag{5}$$

where α_0 is the initial growth rate. To bring out the analogy with the logistic, we will take a second approach. If we start with the differential equation

$$\frac{dV}{dt} = rV(t)\log(K/V(t)) \tag{6}$$

where $K = V_\infty = \lim_{t\to\infty} V(t)$, then the solution is

$$V(t) = V_0 \exp(A(1 - \exp(-rt)) \tag{7}$$

with $A = \log(V_\infty/V_0)$.

For our method to work, the rate r must be the only parameter in the model, so we will fix the value of the carrying capacity. In [3] the authors take $K = 10^{12}$. They use $V_0 = 10^{-6}$ mm^3, i.e., one the volume of one cell, so $V_\infty = 10^6$ mm^3 or 10^3 cm^3. Independent of the units used,

$$A = \log(10^{12}) = 27.631. \tag{8}$$

Norton [23] took the lethal tumor volume to be $N_L = 10^{12}$ cells, but used a carrying capacity of 3.1×10^{12} cells so the tumor size would actually reach N_L. To fit the Gompertz model to the Bloom data set [1] on mortality from untreated breast cancers, he took the number of cells at detection to be $N(0) = 4.8 \times 10^9$ and assumed a lognormally distributed growth rate

with mean $\ln(r) = -2.9$ and standard deviation 0.71. With these choices his survival curve fit the Bloom data almost perfectly. See Figure 1 in [23].

5. The generalized logistic interpolates between the logistic and the Gompertz:

$$\frac{dV}{dt} = rV(t)\left(1 - (V(t)/K)^{\beta}\right). \tag{9}$$

If we let $\beta \to \infty$ we get the exponential. If we take $\beta = 1$ we get the logistic, while if we replace r by r/β and let $\beta \to 0$ we get the Gompertz, see page 1928 in [27]. The solution is

$$V(t) = K[1 + Q\exp(-\beta rt)]^{-1/\beta} \tag{10}$$

where $Q = [(K/V_0)^{\beta} - 1]$. When $\beta = 1$ this reduces to the familiar formula for the solution of the logistic.

$$V(t) = \frac{KV_0e^{rt}}{K + V_0(e^{rt} - 1)}$$

Spratt et al. [32] took $K = 2^{40} \approx 10^{12}$ and found that the best fit of this model to cancer data came from $\beta = 1/4$, see their Table 1. If we set $\beta = 1/4$ in (10) we get

$$V(t) = \frac{V_\infty}{\left[1 + ((V_\infty/V_0)^{1/4} - 1)\, e^{-0.25rt}\right]^4},$$

which is the formula on page 5 of [38], except that their r is random and has a lognormal distribution with mean 1.07 and variance 1.37. Spratt et al. [32, 33] give a similar formula

$$V(t) = (1.1 \times 10^6)[1 + 1023e^{-0.25rt}]^{-4}.$$

To explain the constant in front note that they give 10^{-6} mm^3 as the volume of one cell, and use a maximum tumor size of $V_\infty/V_0 = 2^{40} = 1.0995 \times 10^{12}$ cells. Based on data on 335 women with two mammograms and another 113 with an average of 3.4 mammograms, they found that this model fit better than the Gompertz and the exponential, and that the rate r had roughly a lognormal distribution. See page 2016 in [33].

Figure 1 compares the correction factors $f(V)$ for the different models by plotting them against $\log(V)$. The fact that the correction factors, which we think of as a modification of the exponential rate are often > 1 (except for the Spratt model) highlights the fact that various quantities we have called r have different interpretations. Figure 2 gives a visual comparison of our growth models by plotting their solutions with $V(0) = 10^{-9}$ cm^3 (one cell), and r chosen so that $V(10) = 10$ cm^3.

3 Estimating r from two time points

Our data will give the tumor volume at two time points, t_1 and t_2. In each case this allows us to solve for the value of r. We will estimate the growth rate by averaging the values of r computed for all of the tumors in the data.

3.1 Exponential growth

Since $V(t) = V_0 e^{rt}$, we have

$$\hat{r}_E = \frac{\log(V(t_2)) - \log(V(t_1))}{t_2 - t_1}. \tag{11}$$

Note that since we look at the logarithm of the ratio, the rate is independent of the units in which the volume is measured.

3.2 Power law

The solution in (3) has $V(t)^{1-\alpha} = V_0^{1-\alpha} + (1-\alpha)rt$, so we have

$$\hat{r}_\alpha = \frac{V(t_2)^{1-\alpha} - V(t_1)^{1-\alpha}}{(1-\alpha)(t_2 - t_1)} \tag{12}$$

The estimate can be rewritten as

$$\hat{r}_\alpha = \frac{\exp[(1-\alpha)\log(V(t_2))] - \exp[(1-\alpha)\log(V(t_1))]}{(1-\alpha)(t_2 - t_1)}.$$

Using the fact that $e^x \approx 1 + x$ when x is small, we see that as $\alpha \to 1$

$$\hat{r}_\alpha \to \frac{\log(V(t_2)) - \log(V(t_1))}{t_2 - t_1},$$

the rate estimate of the exponential.

When we use this estimate on a data set we will get very different values of $\hat{r}_\alpha$ for different α's. The reason for this is that $\hat{r}_\alpha$ has units of (volume)$^{1-\alpha}$/time. All of the other estimates described in this section are independent of the units volume is measured in. However, as we will see in Table 1, the values of those rate estimates can vary considerably.

3.3 Gompertz

To estimate r, Chingola and Foroni [3] start with the solution in (7), and take logs of both sides

$$\frac{1}{A}\log(V(t)/V_0) = 1 - e^{-rt}$$

Rearranging gives

$$t_i = -\frac{1}{r}\log\left(1 - \frac{1}{A}\log(V(t_i)/V_0)\right)$$

Using the fact that $A = \log(V_\infty/V_0)$, we can rewrite

$$\begin{aligned} 1 - \frac{1}{A}\log(V(t_i)/V_0) &= \frac{A - \log(V(t_i)/V_0)}{A} \\ &= \frac{\log(V_\infty/V_0) - \log(V(t_i)/V_0)}{A} = \frac{1}{A}\log(V_\infty/V(t_i)). \end{aligned}$$

If we let $\Omega_i = \log(V_\infty/V(t_i))$ then we have

$$t_i = -\frac{1}{r}\log(\Omega_i/A).$$

Taking the equation for $i = 2$ and subtracting the one for $i = 1$ we have we get estimator

$$\hat{r}_G = \frac{\log(\Omega_1) - \log(\Omega_2)}{t_2 - t_1}. \tag{13}$$

Since Ω_1 involves a ratio of two volumes, it is independent of the units in which the volumes are measured.

To connect with the calculation in the appendix of [3] note that, as in (5), they write the Gompertz as

$$V(t) = V_0 \exp\left(\frac{\alpha_0}{\beta}[1 - \exp(-\beta t)]\right)$$

so $r = \beta$, and $A = \alpha_0/\beta$ is what they call K. As $t \to 0$, $1 - \exp(-\beta t) \sim \beta t$. So α_0 is the exponential growth rate when t is small.

3.4 Generalized Logistic

Changing K to V_∞, the start time to t_1 and rearranging (10) we have

$$((V_\infty/V(t_1))^\beta - 1)e^{-\beta r(t_2-t_1)} = (V_\infty/V(t_2))^\beta - 1.$$

If we let $\Gamma_i = (V_\infty/V(t_i))^\beta - 1$ then we can write the above as $e^{-\beta r(t_2-t_1)} = \Gamma_2/\Gamma_1$. Taking logs and rearranging, we have

$$\hat{r}_{GL} = \frac{\log(\Gamma_1) - \log(\Gamma_2)}{\beta(t_2 - t_1)} \tag{14}$$

which is similar to the Gompertz estimator in (13). Writing

$$\Gamma_i = e^{\beta \log(V_\infty/V(t_i))} - 1 \approx \beta\Omega_i$$

when β is small, we see that letting $\beta \to 0$ gives the Gompertz estimate.

To summarize and compare our estimates we note that

$$\begin{aligned}
\hat{r}_E &= \tfrac{\log(V(t_2)) - \log(V(t_1))}{t_2 - t_1} \\
\hat{r}_G &= \tfrac{\log(\Omega_1) - \log(\Omega_2)}{t_2 - t_1} \qquad \Omega_i = \log(V_\infty/V(t_i)) \\
\hat{r}_{GL} &= \tfrac{\log(\Gamma_1) - \log(\Gamma_2)}{\beta(t_2 - t_1)} \qquad \Gamma_i = (V_\infty/V(t_i))^\beta - 1 \\
\hat{r}_\alpha &= \tfrac{V(t_2)^{1-\alpha} - V(t_1)^{1-\alpha}}{(1-\alpha)(t_2 - t_1)}
\end{aligned}$$

4 Data sets

Heuser et al. [10] discovered 109 breast cancer tumors in 108 women in a screening population of 10,120 women receiving over 30,000 mammograms over three years. Forty-five of the cancers were diagnosed on the initial screening. However, of the remaining 64, there were 32 women who had an earlier mammogram on which the tumor could be seen in retrospect. Nine of these breast cancers did not grow in size between the two measurements, leaving us 23 data points. For each tumor they reported the size of the major axis b and the minor axis a measured in mm, e.g., 22×17.

Three methods were used in [10] to convert a and b to a volume. The fourth given below is from Nakajiima et al. [21]. Let $q = a/2$ and $r = b/2$ be the minor and major radii. Let $s_1 = (2/3)q + (1/2)r$ and $s_2 = (qr)^{1/2}$ be the geometric mean.

Sphere	$V = (4\pi/3)s_1^3$
Cylinder	$V = \pi qr^2$
Spheroid-1	$V = (4\pi/3)qr^2$
Spheroid-2	$V = (4\pi/3)s_2^3$

The cylinder and the first spheroid volumes differ by a constant, so the rate estimates will be the same. We do not expect drastic differences between the other three methods, so we will work with the first spheroid formula.

Of 79 acoustic neurinomas seen by Laasonen and Troupp [13], no operation was performed on 21 of these patients or it was delayed for at least six months, so a second CT scan was available for these patients. The reasons for not operating were as follows: 7 patients had bilateral tumors so there was a delay on the operation for the other one, 9 patients wanted more time to decide in favor or against an operation, 4 were too old or too ill with some other disease, and 1 has a 0.38 cm^3 tumor not diagnosed at first in another hospital. Volume measurements were done with a program built into the scanner. They report the initial and final volume in cm^3.

At the Nordstate Hospital between 1978 and 2000, a total of 1954 patients seen had meningiomas and 1700 were operated on. Between 1990 and 2001, a total of 80 asymptomatic patients were diagnosed by computed tomography or MRI. Among them were 7 patients with associated neurofibromatosis Type 2, 4 patients had multiple meningiomas, 22 patients underwent surgery immediately after diagnosis, and 6 had surgery later due to significant tumor growth. Nakamura et al. [22] examined the natural history of the remaining 41 "incidental" meningiomas, which occurred at a wide variety of different locations in the brain. Again the initial and final volumes were reported in cm^3.

Nakajima et al [21] studied 34 hepatocellular carcinomas (HCCs) in patients who initially refused therapy, giving data, as [10] did, on the major and minor axis. The tumors varied in their clinical stage: 18 were stage I, 14 stage II and 3 stage III, and histology: 19 were well-differentiated, 9 moderately differentiated, and 6 poorly differentiated. See the paper for more on the clinical classification.

Saito et al [28] studied the tumor volume doubling times of 21 HCCs. Patients were only selected if their tumors were less than 3 cm in diameter at the start of observation, and two abdominal ultrasounds were available. This occurred because 3 patients refused treatment,

6 had clinical complications that prevented surgery, and in 12 cases the initial diagnosis was uncertain. They report only the initial and final diameter in mm. They talk about the major and minor axes when they discuss the doubling time, but do not give the measurements, so we have calculated volumes by assuming the tumors are spherical.

5 Rate Estimates

The next table gives the rate estimates for our five data sets first for the various power laws and the exponential (which corresponds to power 1), then continues with the generalized logistic (the Logistic $\beta = 1$ and Spratt's model $\beta = 1/4$), and the Gompertz (which is the limit $\beta \to 0$). In the last case, we give both the rate estimate $\hat{r}_G$ and the initial growth rate $\alpha_0 = Ar$ defined in (8). Note that even though Saito and Nakamura both study HCC, all of the rate estimates differ roughly by a factor of 2.

	Nakamura	Laasonen	Heuser	Saito	Nakajima
$\hat{r}_{0.5}$	0.2121	0.7547	1.1488	5.6704	11.324
$\hat{r}_{2/3}$	0.1518	0.6931	1.0705	3.8570	7.4361
$\hat{r}_{0.8}$	0.1192	0.6669	1.0399	2.9064	5.4554
$\hat{r}_{0.9}$	0.1026	0.6593	1.0365	2.3867	4.3904
$\hat{r}_E$	0.0856	0.6618	1.1515	1.9896	3.5837
$\hat{r}_L$	0.0867	0.6275	1.1533	2.0117	3.6329
$\hat{r}_S$	0.1203	0.7731	1.4061	2.8359	5.2360
$\hat{r}_G$	0.0074	0.0936	0.1671	0.4084	0.7713
α_0	0.2045	2.5863	4.6171	11.284	21.312

Table 1: Rate estimates for our five data sets.

The polynomial rate estimates decrease as the power increases. The rates for the logistic are always close to those of the exponential. This should not be surprising. We fit the logistic, as we do the Gompertz and Spratt models, using a carrying capacity of $K = 1000\ cm^3$, while all the final tumor volumes in all the datasets are $< 102.16\ cm^3$, and in the first three data sets all are smaller than $10\ cm^3$. Because of this, the correction factor $1 - V(t_1)/K$ is close to 1 for $t_1 \le t \le t_2$.

The correction factor $1 - (V(t)/K)^{1/4}$ in the Spratt model has a significant effect when $V(t) = 1\ cm^3$, which corresponds to 10^9 cells, so its rate estimates are larger than the exponential. The correction factor $\log(K/V)$ in the Gompertz has a much stronger effect, but compared to the other estimates the Gompertz rate $\hat{r}_G$ is much smaller. This can be traced to the fact that to get from the generalized logistic to the Gompertz we must replace r by r/β and let $\beta \to 0$. If we look instead at the initial growth rate α_0, it is much larger than the exponential rate estimate.

6 Comparing the rate estimates

While all of our differential equations have a constant denoted by r, these rates do not all have the same meaning, e.g., as mentioned in Section 3 the power laws they have different

units. To try to reconcile the various estimates, we recall that all of our models have

$$\frac{d}{dt}\log V(t) = \frac{V'(t)}{V(t)} = rf(V(t)) \tag{15}$$

where $f(V) = V^{\alpha-1}$ for power laws, $f(V) = \log(K/V)$ for the Gompertz, and $f(V) = 1-(V/K)^{\beta}$ for the generalized logistic.

If we multiply the rate estimate r by the average value of $f(V_1)$ in the data, we get an estimate for the average exponential growth rate of the tumor at the initial time. If we use the average value of $f(V_2)$ in the data set, we get an estimate for the average exponential growth rate of the tumor at the final time. Integrating the differential equation

$$\hat{r}_E = \frac{\log V(t_2) - \log V(t_1)}{t_2 - t_1} = r \cdot \frac{1}{t_2 - t_1}\int_{t_1}^{t_2} f(V(s))\, ds$$

From this we see that if we were to multiply our rate estimate by the average value of $f(V(s)$ over the interval $[t_1, t_2]$ then we would get exactly the exponential rate. This suggests that the best correction is to multiply by the average value of $(f(V_1)+f(V_2))/2$, which is a simple approximation of the integral.

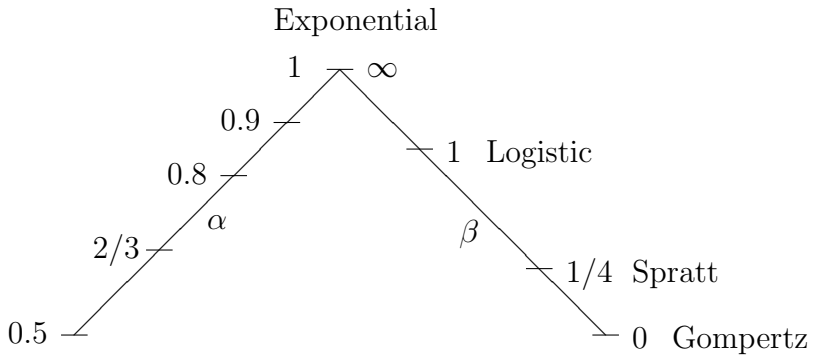
As the reader will see in Table 2, when we do this transformation to the estimates from the Gompertz and generalized logistic, then the result is close to the exponential rate. However, when we apply the same procedure to the power law estimates the estimates get worse. The reason for this unfortunate outcome is that in several data sets there are tumors with a volume of $< 1\, cm^3$, and $x^{\alpha-1}$ is strongly convex there, so the average value $(f(V_1)+f(V_2))/2$ is much different from the integral. To avoid this problem, we convert the tumor volumes to mm^3 before we transform them. (Note that the rate estimate in mm^3 is $10^{3(1-\alpha)}$ times the one in cm^3, so we don't have to do any work to compute the new rate.) The rescaling to mm^3 has made the range of values larger by a factor of 10^3 but the second derivative is now smaller by a factor of $10^{3(\alpha-3)} < 10^{-6}$ so $v^{\alpha-1}$ is now very flat over the range of observed values. Because of this, the average value of $V^{\alpha-1}$ over $[t_1, t_2]$ is close to $(\bar{V})^{\alpha-1}$, where $\bar{V}$ is the average of V over $[t_1, t_2]$.

The next table gives the result of multiplying our earlier estimates by the average value of $(f(V_1)+f(V_2))/2$ to convert them into estimates of the average exponential growth rate of the tumors in the sample. The translated Gompertz estimates are now similar to the others, with the striking exception of the Nakamura data set, while the translated Spratt estimates are very close to the exponentials. Note that the Nakajima and Saito power law estimates have been reduced by the translation while the Heuser, Laasonen, and Nakamura estimates, which are based on smaller tumors, have become larger.

	Nakamura	Laasonen	Heuser	Saito	Nakajima
$\tilde{r}_{0.5}$	0.1034	0.8592	1.3442	2.6656	4.8215
$\tilde{r}_{2/3}$	0.0833	0.7238	1.0969	2.2387	4.0708
$\tilde{r}_{0.8}$	0.0845	0.6704	0.9948	2.0585	3.7296
$\tilde{r}_{0.9}$	0.0856	0.6593	0.9607	1.9953	3.6074
$\tilde{r}_E$	0.0896	0.6617	1.0514	1.9896	3.5957
$\tilde{r}_S$	0.0841	0.6700	1.1533	2.0408	3.6888
$\tilde{r}_G$	0.0158	0.6588	1.1847	2.1666	3.9164

Table 2: Rate estimates translated to be the average exponential growth rate of the tumor during the time between observations.

To help understand the trends in the fits, we find the next picture useful.



The exponential is at the top because for a fixed value of r it grows the fastest. As the power law α decreases from 1, or the generalized logistic β decreases from infinity, growth slows down for a fixed r so the rate estimates increase (in most cases).

7 Model Comparisons

As explained in the introduction, if the rate of growth $rV_t f(V_t)$ in the model being fitted is slower than the actual tumor dynamics, we would expect $\hat{r}$ to increase when plotted against the initial volume V_1, or if we use linear regression to fit a straight line, then the slope would be positive. Similarly, if the rate of growth $rV_t f(V_t)$ in the model being fitted is faster than the actual tumor dynamics then we would expect $\hat{r}$ to decrease when plotted against the initial volume V_1, or if we use linear regression to fit a straight line, then the slope would be negative. To test to see if the linear relationship is not flat we use a t-test to see if the rate estimate $\hat{r}$ and the initial tumor volume V_1 have a significant correlation. In the tables below *'s indicate a p-value of < 0.1 and ** one with $p < 0.05$. When there is a significant correlation, and the slope of the regression line is positive, we will say informally that the positive slope is significant. In the tables that follow, we also give R^2, the fraction of the variance explained by the line to provide another perspective on the fit.

7.1 Heuser et al. [11]

The time between observations for more than half of the women in the data set was close to one year. The maximum time between observations was 2.02 years with an average of 0.89. We discarded one tumor that was 7 cm in diameter (and not detected!) on the initial mammogram because its initial volume of 179.6 cm^3 was 25 times larger than the second largest tumor, making it impossible to fit a line to $\hat{r}$ versus V_1. In the results presented below, we also eliminated one tumor (the square in Figure 3) with an initial volume of 6.28 and final volume of 28.73, which was an outlier compared the average values were 0.89 cm^3 and 1.05 cm^3. However, including this tumor does not change the results very much.

In the next table, we see that the slopes of $\hat{r}$ versus V_1 are positive for all power law rates, as well as for the Spratt and Gompertz growth laws, but that the trends are significant only for the first three power laws. The fact that the slope is the smallest for the exponential suggests that it may give the best fit to the data. Here, and in other comparisons against the volume, we must keep in mind that tumors with larger growth rates may be larger when detected, creating a positive correlation that our analysis would attribute to the growth rate in the model being slower than that in the tumor.

model	signif.	R^2	slope vs. V_1	r_{max}	ss vs. V_1
0.5	0.001**	0.4398	0.5945	4.391	0.478
2/3	0.011**	0.2947	0.3751	3.431	0.386
0.8	0.092*	0.1425	0.2186	2.820	0.274
0.9	0.396	0.0382	0.1022	2.435	0.148
Exp	0.915	0.0006	−0.0127	2.397	−0.019
Spratt	0.669	0.0098	0.0627	2.829	0.078
Gompertz	0.206	0.0827	0.0229	0.387	0.209

As we explained in the last section, one problem with comparing the slopes is that the estimated rates vary considerably in magnitude. To compensate for this, we will scale the rates by the maximum rate observed and the maximum initial volume to have graphs where the x and y values range over $[0, 1]$. Note that now the scaled slopes (ss) versus V_1 more clearly show a minimum at the exponential.

Norton [23] analyzed the Heuser data set using the Gompertz model with the log-normal r that he got from fitting the Bloom data set, i.e., $\ln(r)$ is normal with mean -2.9 and standard deviation. He plotted the $\log_{10}$ of the 23 rate Gompertz estimates for the Heuser data along the lognormal density function, see his Figure 2. Norton judges the fit to be good because "the circles are expected to belong to the lower 23% of the data and indeed do fall in the lower 15%." The figure of 23% comes from the fact mentioned in Section 4 that of the 109 tumors (in 108 women), 45 were detected on the initial screening, and 32 of the remaining 64 tumors were not detectable on the previous mammogram. Discarding the 9 tumors that did not grow, this means that only 23 out of the 100 grew slowly enough to be visible on two successive mammograms.

Our analysis does not match with his findings. The $\ln(r_G)$ for our 21 data points range from -3.56582 to 1.22169. Only two are below his mean -2.9 and 13 are more than 3 of his standard deviations above the mean. Furthermore the positive slope when the Gompertz

rate is plotted against the initial volume suggests that the true growth rate is larger than the Gompertz.

7.2 Saito el al. [28]

In this data set of the largest initial tumor volume was 14.14 cm^3 (which corresponds to a diameter of 3 cm), while the largest final volume was 57.91 cm^3. Average values of initial and final volumes were 5.19 cm^3 and 15.47 cm^3 respectively. The maximum time between observations was 2.09 years, with an average value of 0.65 years.

When we consider the slopes versus V_1, the four power laws have significant positive slopes. The minimum slope occurs for the exponential model. Figure 4 shows the rate estimates for the exponential and Spratt model and the straight line fit.

model	r_{max}	signif	ss vs. V_1	R^2
0.5	19.80	0.001**	0.537	0.448
2/3	11.63	0.004**	0.466	0.366
0.8	7.62	0.021**	0.380	0.261
0.9	5.55	0.032**	0.293	0.158
Exp	4.05	0.327	0.170	0.053
Spratt	6.67	0.880	0.291	0.166
Gompertz	1.08	0.379	0.382	0.263

Table 3: Rate estimates for the data set from Saito el al. [28]

7.3 Nakajima et al. [21]

In this data set, four of the tumors, which were stage I but moderately or poorly differentiated showed much larger growth rates than the others, so we removed them from our analysis. Their doubling times were from 17–31 days, see the diamonds in Figure 5. Among the remaining tumors the maximum time between observations was 0.69 years, with an average of 0.35 years. The largest initial volume was 31.52 cm^3 with an average of 6.80 cm^3. The largest final volume was 102.16 cm^3 with an average value of 19.28 cm^3. When we consider the slopes versus V_1, the power laws 0.5, 2/3, and 0.8 have significant positive slopes as does the Gompertz. The minimum slope, here as in the Saito data set, occurs for the exponential. The slopes are not very small (0.170 and 0.107 respectively) but there are not significantly different from zero.

Model	r_{max}	signif.	ss vs. V_1	R^2
0.5	40.23	0.0005**	0.616	0.354
2/3	22.03	0.0057**	0.522	0.243
0.8	13.65	0.0463**	0.409	0.134
0.9	9.62	0.196	0.286	0.589
Exp	7.96	0.6061	0.107	0.0096
Spratt	11.84	0.1503	0.312	0.073
Gompertz	2.01	0.0324**	0.428	0.153

Table 4: Rate estimates for the data set of Nakajima et al. [21].

7.4 Laasonen and Troupp [13]

We removed the two points from this data set because the second volume was smaller than the first. The maximum time between observations 2.66 years with an average of 1.24 years. Two of the tumors had initial volumes of 7.05 cm^3 and 7.55 cm^3, while all of the others were smaller than 3 cm^3 with an average size of 0.878 cm^3. These two outliers had a large effect on the slope of the fitted linear relationship between $\hat{r}$ and V_1(see Figure 6) so we removed them in order to compute the regression line. Among the remaining tumors the largest final volume was 4.52 cm^3, with an average size of 1.71 cm^3.

Model	r_{max}	signif.	ss vs. V_1	R^2
0.5	1.648	0.577	0.155	0.0198
2/3	1.356	0.997	0.00128	1×10^{-10}
0.8	1.315	0.565	−0.1622	0.0211
1	1.258	0.091*	−0.454	0.0776
Spratt	1.550	0.1647	−0.375	0.117
Gompertz	0.188	0.3794	−0.244	0.0486

Table 5: Rate estimates for the Laasonen and Troupp [13] data.

Here, there is some evidence that the exponential fit has negative slope. The 1/2 power law fit has positive slope, the 2/3's power law has a slope close to 0, while the other fits have negative slopes, suggesting that the 2/3's power gives the best fit. Figure 7 shows how the rate estimates compare when the powers are 0.5, 2/3, and 0.8, and the three regression lines.

7.5 Nakamura et al. [22]

Five tumors that showed very large growth over a short amount of time were removed from this data set. The 0.5 power law rate estimates for these five tumors are indicated with circles in Figure 8. If these points are included then all regression lines will have a negative slope. The largest time between observations was 8.75 years with an average of 3.82 years. The largest initial volume 29.31 cm^3, with an average of 9.90 cm^3. The largest final volume, which occurred in the same patient. was 35.71 cm^3, with an average of 12.67 cm^3.

Model	r_{max}	signif.	slope vs. V_1	R^2
0.5	0.516	0.377	0.1582	0.023
2/3	0.395	0.635	−0.074	0.0067
0.8	0.320	0.116	−0.246	0.0712
Exp	0.388	0.009**	−0.263	0.1806
Spratt	0.352	0.041**	−0.316	0.1164
Gompertz	0.021	0.151	−0.233	0.0595

Table 6: Rate estimates for the Nakamura et al. [22] data set.

When plotted against V_1 the exponential line and Spratt lines have significant negative slopes. Again the 1/2 power law fit has positive slope, the 2/3's power law has the slope close to 0, while the other fits have negative slopes, suggesting that the 2/3's power gives the best fit.

We are not the only ones to have investigated trends in the exponential rate estimates $\hat{r}_E$ as a function of the volume. Mehrara and Forsell-Aronson [17] did this for seven data sets including those of Nakamura, Nakajima, and Saito considered here. They plotted $\hat{r}_E$ as a function of $\log(V_1)$ to see if the observed rate heterogeneity was due to the fact that the apropriate model was the Gompertz. Of the three data sets that we both have considered, only Nakamura plot has a significant correlation between $\hat{r}_E$ and $\log(V_1)$, with a p value of 0.0005 and an $R^2 = 0.2424$.

7.6 Variability of the rate estimates

To give an idea of the variability in the rate estimates we have plotted the log of the rate estimates minus the average value of the logs in Figure 9. Several researchers have suggested that the rates should be lognormal. Our rate estimates do not show the symmetry characteristic of the normal, but have a skew consistent with the idea that the way the data was collected creates a bias against tumor with large growth rate. Unfortunately, the number of observations is not large enough to test this hypothesis.

8 Conclusions

Here we have developed methods to estimate tumor growth rates for observations of the volume at two different times, and use trends in the values of the rate estimates versus initial tumor volume to decide which model is the best fit. The breast cancer and hepatocellular carcinoma data sets show signs of exponential growth, while our analysis of the data for the neurological tumors (acoustic neurinomas and incidental meningiomas) suggests a 2/3 power law. Due to the small size of the data sets and the large heterogeneity in tumor characteristics, these conclusions are far from definitive. In particular, given the small sizes of tumors in some data sets, it would be hard to detect the presence of growth limitations.

The rates that appear in the models have much different interpretations, so one of our contributions has been to make the rate estimates comparable by transforming them into estimates of the average exponential growth rate of tumors in the sample. The exponential

rate estimates and the rescaled 2/3's power law rates are given in the next table:

Nakamura	0.083
Lasonen	0.724
Heuser	1.052
Saito	1.990
Nakamura	3.584

Our rate estimates are consistent with the observation that acoustic neurinomas grow very slowly. It is interesting to note that while Saito and Nakamura each studied hepatocellular carcinoma, the rate estimates differ by a factor of 2. It is somewhat puzzling that the smaller rate estimate comes from the study where the initial tumor diameters were all $< 3\ cm$, since one expects tumor growth to decelerate as size increases. However, this may simply reflect the fact that two measurements were only available for patients who did not undergo treatment skews the estimates in favor of smaller growth rates.

A similar remark applies to the breast cancer data which only shows tumors that were detected on two successive mammograms. In addition to these biases our results are also effected by small sample size. Our estimated doubling time of $365/1.052 = 343$ days is 37% larger than the 250 days found by Spratt et al. [33] based on multiple mammograms in 448 patients, and almost twice the doubling time of 173.6 days found by Kuroshi et al. [12] based on two mammograms in 122 patients.

Despite the biases introduced by the way the data was ascertained, we believe that the results reported here show that the method is useful for quantifying tumor growth rates and comparing different models. The analysis of larger studies should allow our methods to give important new insights into the long studied question of the growth patterns of solid tumors *in vivo*.

Acknowledgements

This work was begun during an REU in the summer of 2013 associated with an NSF Research Training Grant at Duke University in mathematical biology. Both authors were partially supported by DMS 1305997 from the probability program at NSF. They would also like to thank Natalia Komarova and Marc Ryser who read an earlier draft of this paper and made a number of helpful suggestions.

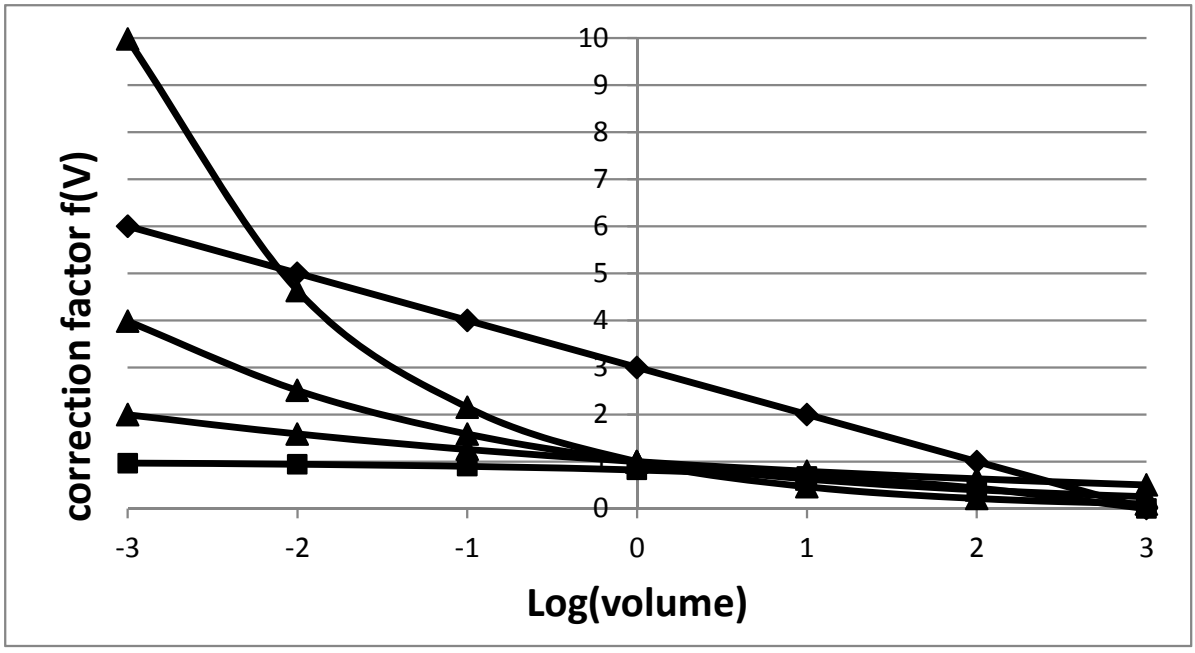


Figure 1: Correction factors $f(V)$ for the models plotted against $\log(V)$. The Gompertz is the straight line of diamonds, Spratt are the squares, while power laws 2/3, 0.8 and 0.9 are in decreasing order the curves marked with triangles.

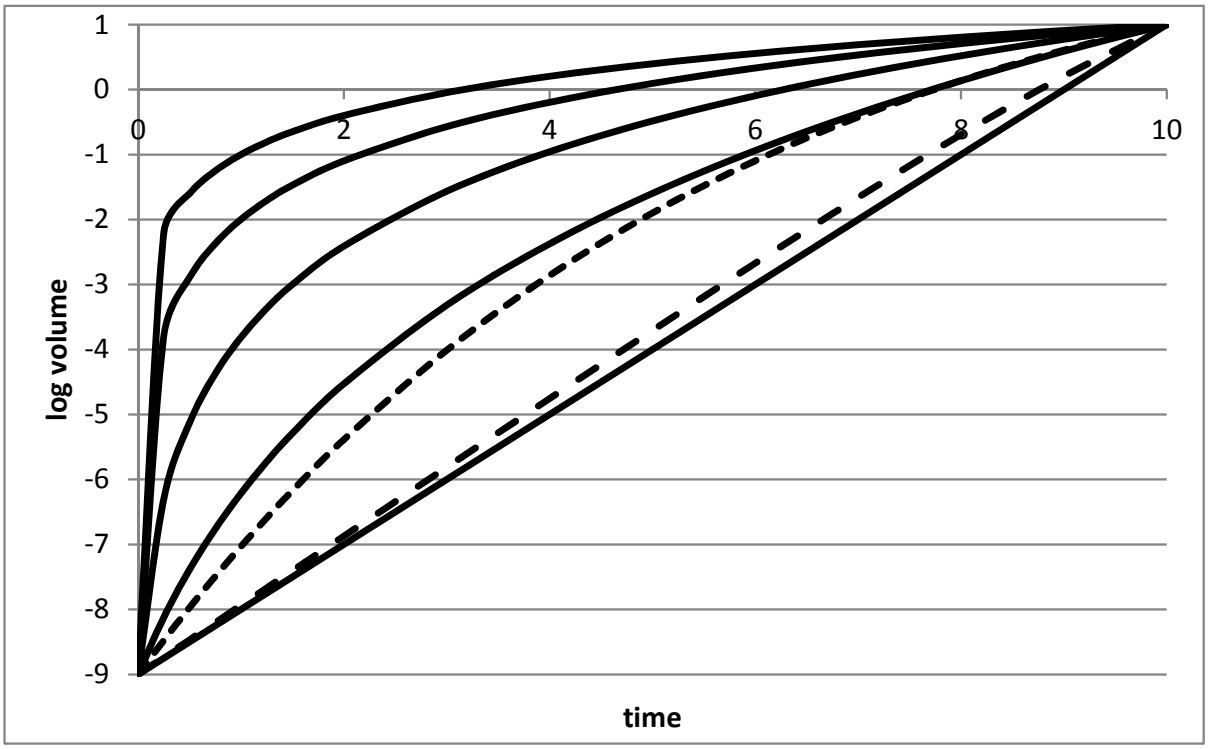


Figure 2: Comparison of $\log_{10} V(t)$ for our models. All start with one cell, $V(0) = 10^{-9}\ cm^3$. Rates are chosen so that $V(10) = 10\ cm^3$. The solid lines in decreasing order are power laws $\alpha = 0.5,\ 2/3,\ 0.8,\ 0.9$. The exponential is the straight line $-9 + t$. The dashed lines are Gompertz (larger) and Spratt (close to exponential).

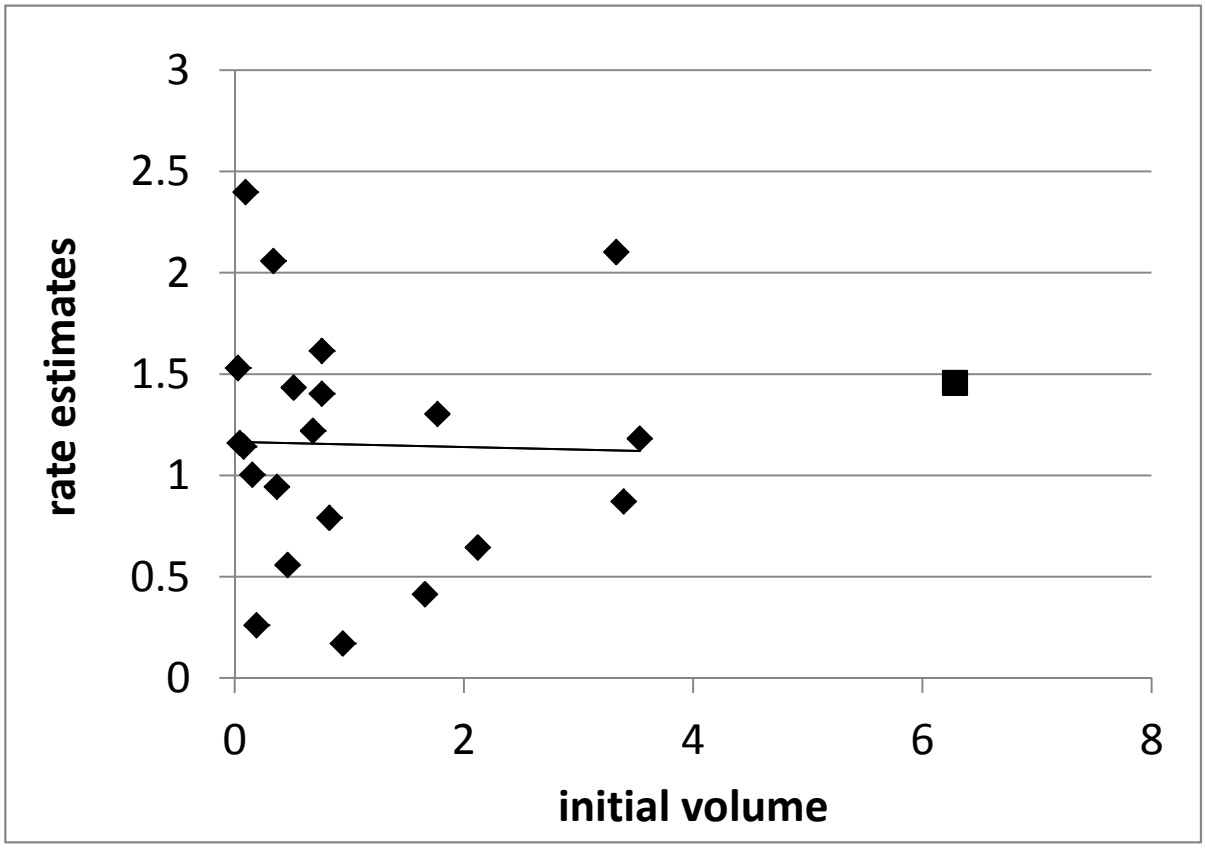


Figure 3: Heuser data set: Exponential growth rate estimates plotted versus initial volume. The tumor indicated by the square was not used to compute the regression line.

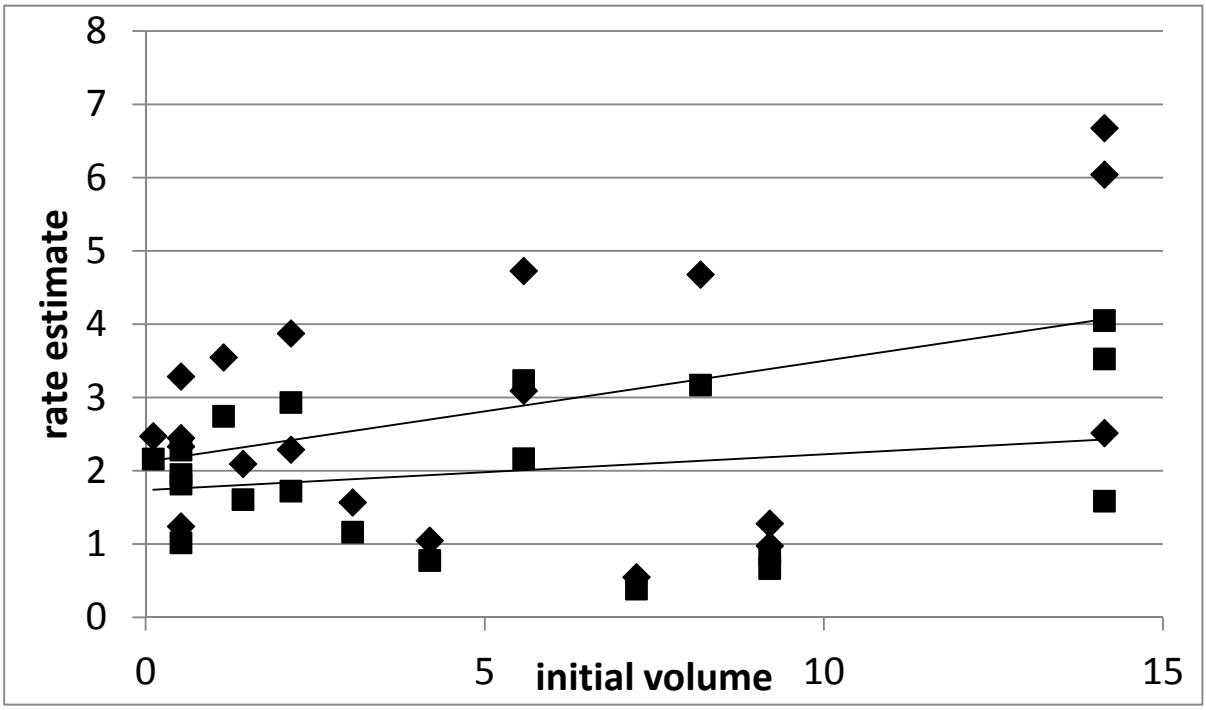


Figure 4: Exponential (squares) and Spratt (diamonds) rate estimates for the Saito data set, and the corresponding regression lines that have slopes 0.0488 and 0.1337. As to be expected the Spratt rate estimate are larger, and the discrepancy increase with the initial size of the tumor.

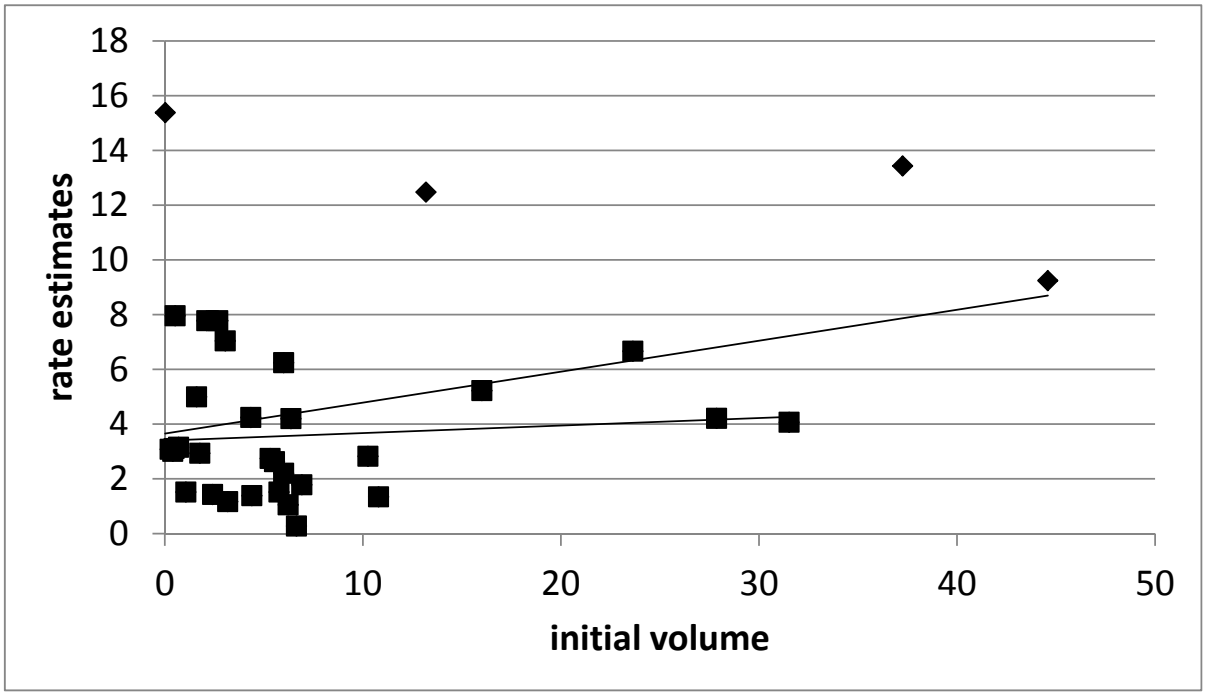


Figure 5: Nakajima data set. The diamonds are the four points we discarded. Also shown are the regression line fits for the entire data (slope 0.1131) and reduced data sets (slope 0.0274). The three diamonds with initial volume > 10 cm^3 are responsible for the large increase in slope.

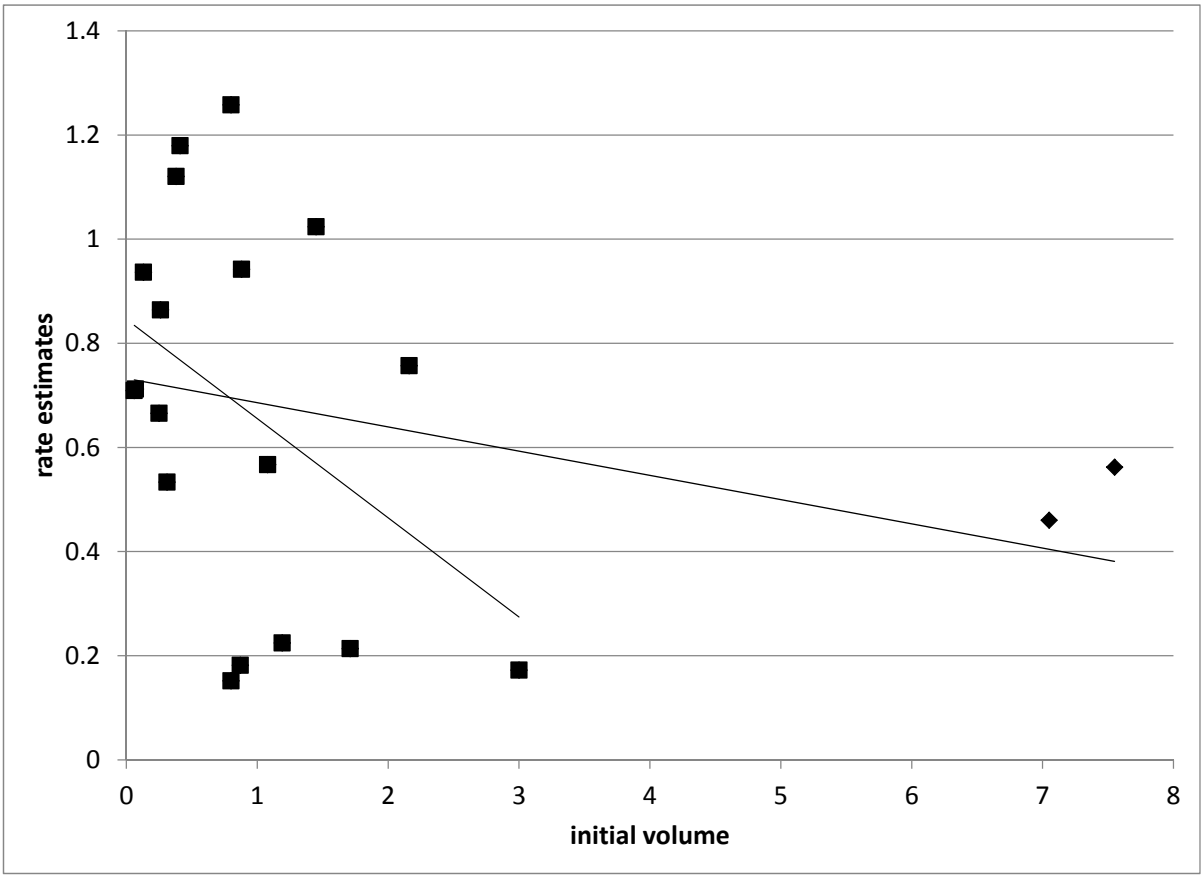


Figure 6: Exponential rate estimates for the Laasonen and Troupp [13] data, showing the large influence of the two points with initial volume $> 7\ cm^3$.

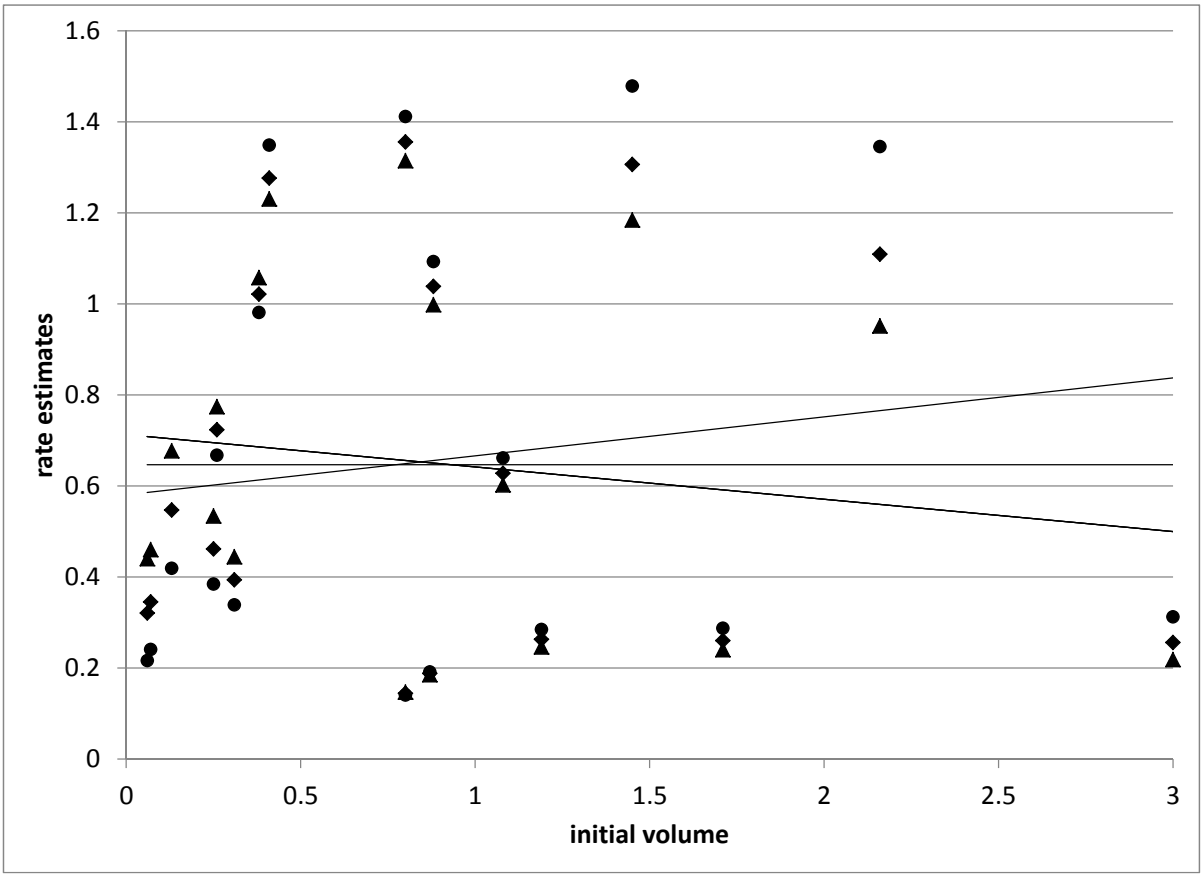


Figure 7: Rate estimates for the power laws 0.5 (circle), 2/3 (diamond), and 0.8 (triangle) for the Laasonen and Troup data plotted versus initial tumor size V_1, as well as the three least squares lines, which have slopes 0.0855, -6×10^{-6}, and -0.0711. Note that in most cases (but not all) the rate estimate decreases as the power increases.

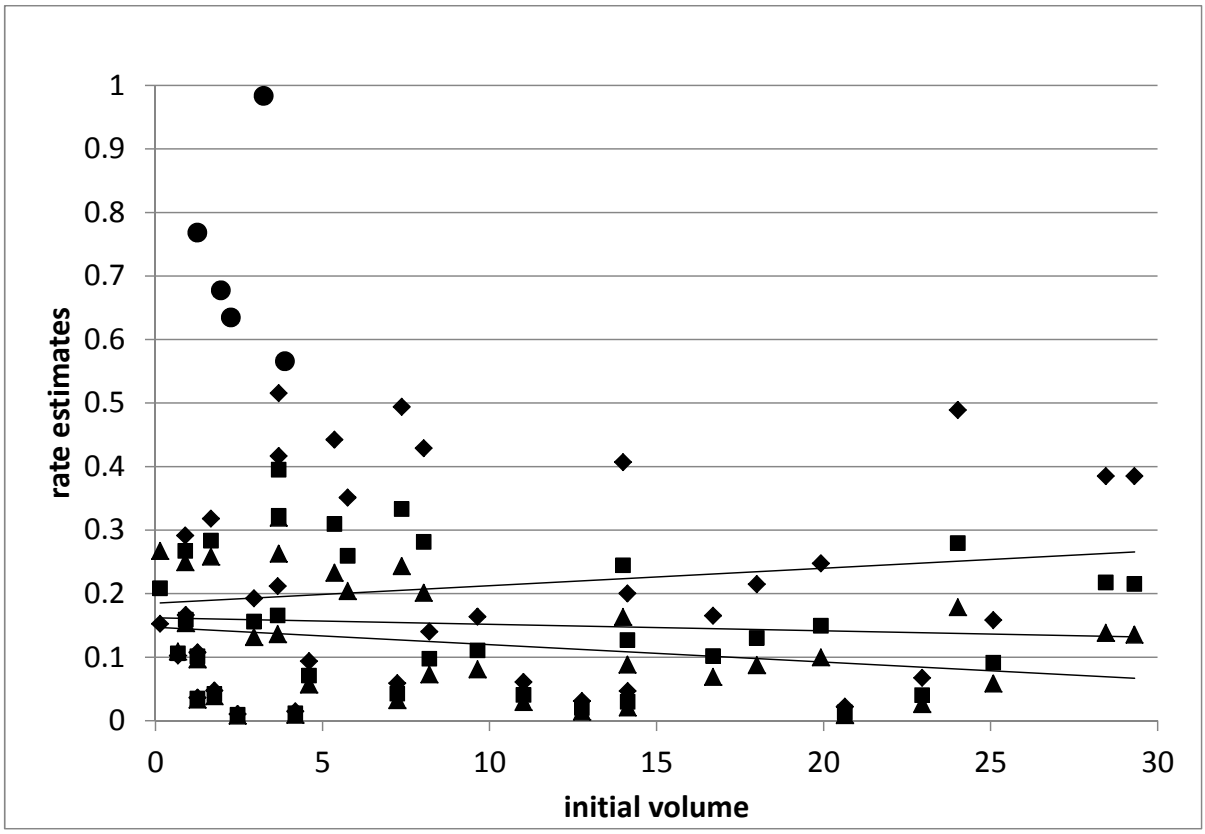


Figure 8: Rate estimates for the power laws 0.5 (diamond), 2/3 (square), and 0.8 (triangle) for the Nakamura data plotted versus initial tumor size V_1, as well as the three regression lines, which which have slopes 0.0028, -0.0010, and -0.0027. The circles are the 0.5 rate estimates for the five tumors we have excluded.

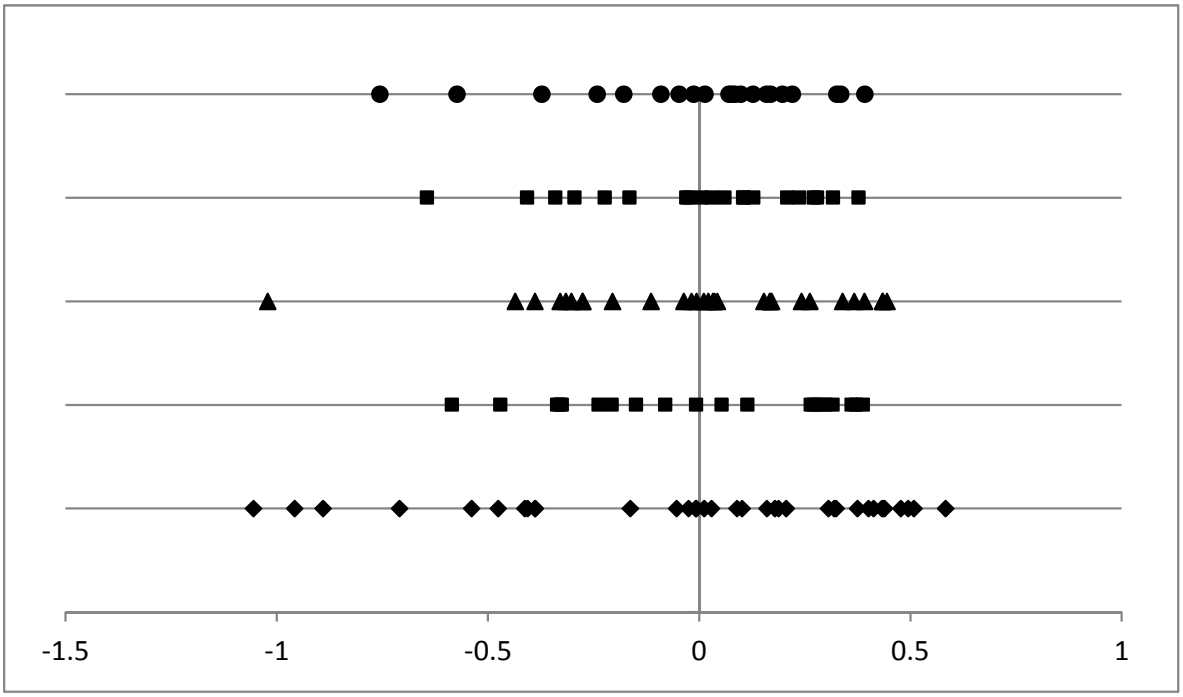


Figure 9: Variability in the logarithm of the rate estimates centered by subtracting the mean of the logarithms. From top to bottom we have the exponential rate estimates for Heuser, Saito, and Nakajima, followed by the 2/3 power law estimates for Laasonen and Tropp, and Nakamura.

References

[1] Bloom, H.J.G., Richardson, W.W., and Harries, E.J. (1962) Natural history of untreated breast cancer (1805–1933) *British Medical Journal.* 2, 213–221

[2] Castorina, P., Diesboeck, T.S, Gabriele, P., and Guiot, C. (2007) Growth laws in cancer: Implications for radiotherapy. *Radiation Research.* 168, 349-356

[3] Chingola, R., and Foroni, R.I. (2005) Estimating the growth kinetics of experimental tumors from as few as two determinations of tumor size: Implications for clicnical oncology. *IEEE Transactions on Biomedical Engineering.* 52, 808–815

[4] Collins, V.P. et al. (1956) Observations on growth rates of human tumors. *Am. J. Roentgeol. Radium Ther. Nucl. Med.* 76, 988–1000

[5] Comen, E., Morris, P.G., and Norton, L. (2012) Translating mathematical modeling of tumor growth patterns into novel therapeutic approaches to breast cancer. *J. Mammay Gland Biol. Noeplasia.* 17, 241–249

[6] Friberg, S., and Mattson, S. (1997) On the growth rates of human malignant tumors: Implications for decision making. *Journal of Surgical Oncology.* 65, 284–297

[7] Gerlee, P. (2103) The model muddle: in search of tumor growth laws. *Cancer Research.* 73, 2407-2411

[8] Gompertz, B. (1825) On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of contingencies. *Phil. Trans. Roy. Soc.* 115, 513-583

[9] Guiot, C., Degioris, P.G., Delsanto, P.P., Gabriele, P., and Deisboeck, T.S. (2003) Does tumor growth follow a "universal law"? *J. Theor. Biol.* 225, 147–151

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

[11] Hueser, L., Spratt, J.S., and Polk, H.C. Jr. (1979) Growth rates of primary breast cancers. *Cancer.* 43, 1888-1894

[12] Kuroshi, T., et al. (1990) Tumor growth rate rate and preognosis of breast cancer mainly detected by mass screening. *Japanese J. Cancer Research.* 81, 454–462

[13] Laasonen, E.M., and Troupp, H. (1986) Volume growth of acoustic neurinomas. *Neuro-radiology.* 28, 203–207

[14] Laird, A.K. (1963) Dynamics of tumor growth. *British J. Cancer.* 18, 490–502

[15] Mandonnet, E., et al. (2003) Continuous growth of mean tumor diameter in a subset of grade II gliomas. *Annals of Neurology.* 53, 524–528

[16] Mayenord, W.V. (1932) On a law of growth of Jensen's rat sarcoma. *American Journal of Cancer.* 16, 687–693

[17] Mehara, E., and Forsell-Aronsson, E. (2014) Analysis of inter-patient variability in tumor growth rates. *Theoretical Biology and Medical Modeling.* 11, paper 21

[18] Michaelson, J.S., Halpern, E., and Kopans, D.B. (1999) Breast cancer: Computer simulation method for estimating optimal intervals for screening. *Radiology.* 212, 551-560

[19] Millet, I., Bouic-Pages, E., Hoa, D., Azira, D., and Taourel, P. (2011) Growth of breast cancer recurrences assessed by consecutive MRI. *BMC Cancer.* 11, paper 155

[20] Mueller-Klieser, W. (1997) Three diemnsional cell cultures: from molecular mechanisms to clinical applications. *American Journal of Physiology.* 274, C1109–C1123

[21] Nakajima, T., et al. (2002) Simple tumor profile chart based on cell kinetic parameters and histologic grade is useful for estimating the natual growth rate of hepatocellular carcinoma. *Human Pathology.* 33, 92–99

[22] Nakamura, M., Roseer, F., Michel, J., Jacobs, C., and Samii, M. (2003) The natural history of incidental meningiomas. *Neurosurgery.* 53, 62–71

[23] Norton, L. (1988) A Gompertzian model of human breast cancer growth. *Cancer Research.* 48, 7067–7071

[24] Norton, L. (2005) Conceptual and practical implications of breast tissue geometry: Toward a more effective, less toxic therapy. *The Oncologist.* 10, 370–381

[25] Norton, L, Simon, R., Brereton, H.D., and Bogden, A.E. (1976) Predicting the course of Gompertzian growth. *Nature.* 264, 542–545

[26] Rodriguez-Brenes, I.A., Komarova, N.J., and Wodarz, D. (2013) Tumor growth dyanmics: insights nto evolutionary processes. *Trends in Ecology and Evolution.* 28, 597–604

[27] Sachs, R.K., Hlatky, L.R., and Hanfeldt, P. (2001) Simple ODE models of tumor growth and anti-angiogenic or radiation treatment. *Mathematical and Computer Modeling.* 33, 1297–1305

[28] Saito, Y., et al. (1998) Multiple regression analysis for assessing the growth of small hepatocellular carcinomas. *J. Gastroenterology.* 33, 229–235

[29] Sarapata, E.A., and de Piliis, L.G. (2014) A comparison and catalog of intrinsic tumor growth models. *Bulletin of Mathematical Biology*, to appear

[30] Shackeney, S.E. (1970) A computer model for tumor growth with chemotherapy, and its application to *l*1210 leukemia treated with cytosine arabinose (nsc-63878). *Cancer Chemotherapy Reports.* 54, 399–429

[31] Speer, J.F., Petrovsky, V.E., Rtesky, M.W., and Wardwell, R.H. (1984) A stochastic numerical model of breast cancer growth that simulates clinical data. *Cancer Research.* 44, 4124–4130

[32] Spratt, J.A., von Fournier, D., Spratt, J.S., and Weber, E.E. (1992) Decelerating growth and human breast cancer. *Cancer.* 71, 2013-2019

[33] Spratt, J.A., von Fournier, D., Spratt, J.S., and Weber, E.E. (1992) Mammographic assessment of human breast cancer growth and duration. *Cancer.* 71, 2020–2026

[34] Sutherland, R.M. (1988) Cell and einvironmental interactions in tumor microregions: the multicell tumor spheroid model. *Science.* 240, 177–184

[35] Tabar, L., et al. (1992) Update of the Swedish two-county program of mammographic screening for breast cancer. *Radiol. Clin. North Am.* 30, 187–210

[36] Tabar, L., et al. (1995) Efficacy of breast cancer screening by age. New results from the Swedish two-county trial. *Cancer.* 75, 2507–2517

[37] von Bertalanffy, L. (1949) Problems of organic growth. *Nature.* 163, 156–158

[38] Weedon-Fekjaer, Lindquist, B.H., Vatten, L.J., Aalen, O.O., and Tretli, S. (2008) Breast cancer growth estimated through mammography screening data. *Breast Cancer Research.* 10:R41

[39] West, G.B., Brown, J.H., and Enquist, B.J. (2001) A general model for ontogenic growth. *Nature.* 413, 628-631

[40] Wodarz, D., and Komarova, N. (2014) *Dynamics of Cancer: Mathematical Foundations of Oncology.* World Scientific

[41] Wright, S. (1926) Book Review. *J Amer. Stat. Assoc.* 21, 493–497