Analyzing the Growth Dynamics of Bladder Cancer Using Ordinary Differential Equation Modeling

Joel Greenfield
advised by Daniel Bergman, Ph.D.

August 5, 2022

1 Introduction

Understanding the effects of exposing cancer cells to changes to its microenvironment, including the presence of medical drugs and other cell populations, is an important tool for investigating malignancy in cancers and the efficacy of anticancer therapeutics. The ability to quantify various aspects of the growth of cancer and of the interactions between cancer and its microenvironment aids in our understanding of cancer growth dynamics. In this paper, we implement the tools of ordinary differential equation (ODE) modeling and parameter extraction to interpret data from experiments that study the effect that changes in the microenvironment have on cancer growth.

ODE modeling is a useful tool for analyzing the growth dynamics of various cell populations present in cancer growth experiments. These models consist of one or more ODEs, each of which describes the growth dynamics of a single population. The ODEs contain a number of unspecified parameters which, along with the initial value of each cell population, can be varied to fit the solution of the model equations to experimental data. The model ODEs are constructed by setting the rate of change of each cell population to a sum of terms each describing a feature of the growth dynamics of that population, such as the rates of growth, apoptosis, and killing by other cells.

In most cases, we can draw a direct connection between the value of a parameter and a biologically significant quantity, such as the growth rate of a cell population. Determining values of model parameters that yield a satisfactory agreement between the model solution and experimental data allows us to quantify important aspects of the underlying cell growth dynamics that are not immediately measurable from the data.

In this paper we develop a method to determine the parameter values of an arbitrary ODE model given experimental data which yield the best agreement between the model solution and the data. We then apply this method to two ODE models aimed at describing the growth dynamics of MB49 mouse bladder

cancer, a cell line used extensively to study bladder cancer in *in vivo* and *in vitro* experiments, under various experimental conditions.

The first model consists of a single ODE describing the exponential growth of healthy cancer cells. This model is applied to data from experiments that track the growth of cancer cells expressing different levels of an immunogenic antigen while being exposed to different dosages of an anticancer drug. We use this model alongside experimental data to quantify the growth rate of various cell lines, allowing us to investigate the effects of drug dosage and antigenicity on cancer cell growth.

The second model consists of a system of ODEs describing the interactions between cancer cells and T cells, incorporating phenomena such as the immune response of T cells and T-cell exhaustion. The binding of antigens to T-cell receptors triggers an immune response that incites T cells to kill cancer cells. T-cell exhaustion refers to the incapacitation of T cells in response to chronic exposure to antigens.

We run tests confirming that this model is structurally identifiable. This allows us to asses the extent to which, given a sufficiently good initial guess, we can theoretically extract the value of each model parameter from data. However, to gain information about the *practical* identifiability of the model, we investigate the model using the tools of Bayesian inference. By studying this model in a Bayesian framework, we gain information about the practical identifiability of several model parameters (i.e., to what extent we can use experimental data to determine the values of the parameters) and about correlations between certain parameters.

2 Methods

2.1 A Method of Extracting the Best-Fit Parameter Values from Experimental Data

Past research suggests that exponential growth is a reasonable model of *in vitro* cancer cell population growth before carrying capacity effects become significant. We construct a simple ODE model describing the exponential growth of a cancer cell population consisting of a single ODE, namely

$$L' = rL \tag{1}$$

where L denotes the size of the cancer cell population in confluence. *Confluence* refers to the percentage of a culture dish that is covered by cells and is understood to be directly proportional to the cell population. The solution to Eq. (1) depends on the values of two parameters: the growth rate r and the initial confluence value L_0.

Consider the synthetic data set plotted in blue in each panel of Figure 1. Given this data, we are interested in determining the values of r and L_0 that yield the best agreement between the solution to Eq. (1) and the data. Figure 1 showcases varying levels of agreement between the solution to Eq. (1) (red)

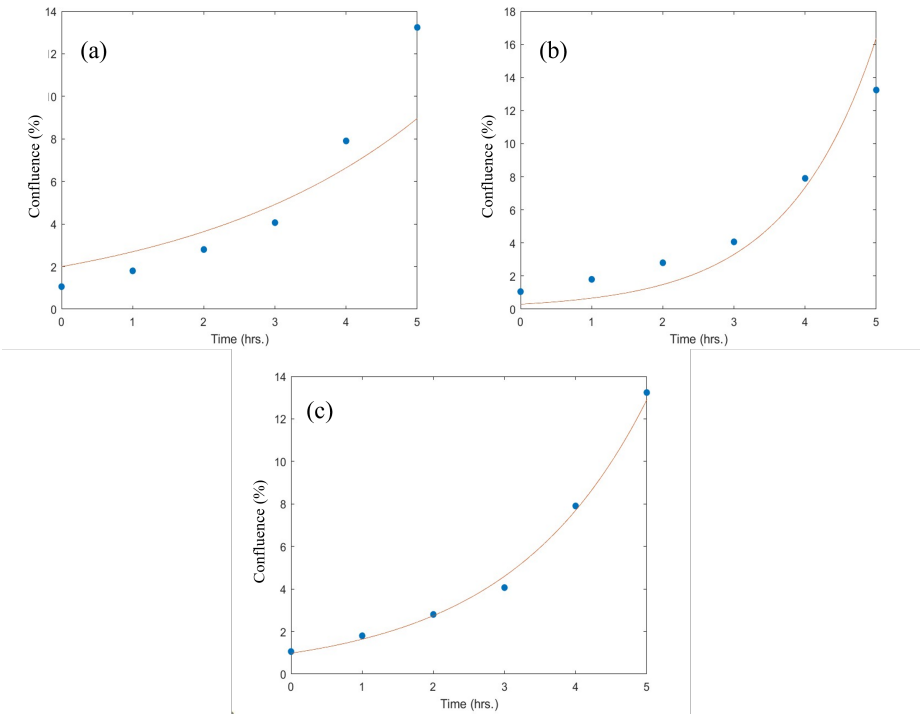


Figure 1: Three realizations of the solution to Eq. (1) (red) overlaid with data (blue). *(a)*: The chosen value of the growth rate r is too small, and the chosen initial value L_0 is too large. *(b)*: r is too large, and L_0 is too small. *(c)*: The values of r and L_0 yield a satisfactory level of agreement between the model solution and the data.

and the data (blue). In Figure 1(a), the chosen value of the growth rate r is too small, and the chosen initial value L_0 is too large. In Figure 1(b), r is too large, and L_0 is too small. In Figure 1(c), the values of r and L_0 yield a satisfactory level of agreement between the model solution and the data.

To quantify how well the solution to an ODE model agrees with data, we use a variation on the method of least squares. We construct objective functions which, given a set of parameter values, compute the sum of the square of the Euclidean norm of the error between the model solution and the data for each state variable over all data points. The first objective function computes the sum of the square of the *absolute* error between the model solution and the data over all data points (see Eq. (2)), whereas the second objective function computes the sum of the square of the *relative* error (relative to the data value) over all data points (see Eq. (3)). Note that the division in Eq. (3) is performed element-wise. The functional forms of both versions of the objective function

are listed in Eq. (2)-(3).

$$O_1(p_1, p_2, \ldots, p_n) = \sum_{i=1}^{m} \|\mathbf{f}(t_i; p_1, p_2, \ldots, p_n) - \mathbf{y_i}\|^2 \tag{2}$$

$$O_2(p_1, p_2, \ldots, p_n) = \sum_{i=1}^{m} \left\| \frac{\mathbf{f}(t_i; p_1, p_2, \ldots, p_n) - \mathbf{y_i}}{\mathbf{y_i}} \right\|^2 \tag{3}$$

where $(t_1, \mathbf{y_1}), (t_2, \mathbf{y_2}), \ldots, (t_m, \mathbf{y_m})$ is the data and $\mathbf{f}(t; p_1, p_2, \ldots, p_n)$ is the solution to the model with parameters $p_1, p_2, \ldots, p_n$ at time t.

The differences in the functional forms of Eq. (2) and Eq. (3) affect certain features of the model solutions generated with the parameter values that minimize their respective objective function. We find that the parameter values generated by minimizing Eq. (2) yield a model solution that favors fitting data points at larger confluences at the cost of a worse fit for data points at smaller confluences (see Figure 2(a)). On the other hand, the values generated by minimizing Eq. (3) yield a solution that favors fitting data points at smaller confluences over those with larger confluences (see Figure 2(b)).

We claim that both realizations of the model solution in Figure 2 are satisfactory, agreeing sufficiently well with the data. In general, we expect that if an ODE model captures the underlying growth dynamics observed in experimental data sufficiently well, then choosing the parameter values $p_1, p_2, \ldots, p_n$ that minimize the value of *either* Eq. (2) or Eq. (3) will yield a satisfactory level of agreement between the solution to an ODE model and given data. Since both minimizing Eq. (2) and minimizing Eq. (3) generate parameter values that yield a good fit between the model solution and the data, in this analysis we define the *best-fit* parameter values to be the single set of parameter values computed by averaging the values of each parameter determined using Eq. (2) and Eq. (3). In other words, we weight the outputs of absolute (see Eq. (2)) and relative (see Eq. (3)) error equally.

For example, see Table 1, which displays the best-fit parameter values of r and L_0 generated by both minimizing Eq. (2) and minimizing Eq. (3). The solutions to Eq. (1) using these sets of parameter values are graphed in Figure 2. As defined above, the best-fit parameter values are

$$L_{0,\text{ best-fit}} = \frac{0.92\% + 1.04\%}{2} = 0.98\% \tag{4}$$

and

$$r_{\text{best-fit}} = \frac{0.53 \text{ hr.}^{-1} + 0.50 \text{ hr.}^{-1}}{2} = 0.52 \text{ hr.}^{-1}, \tag{5}$$

which yield the solution to Eq. (1) graphed in Figure 1(c) (and reprinted in Figure 2(c) for comparison).

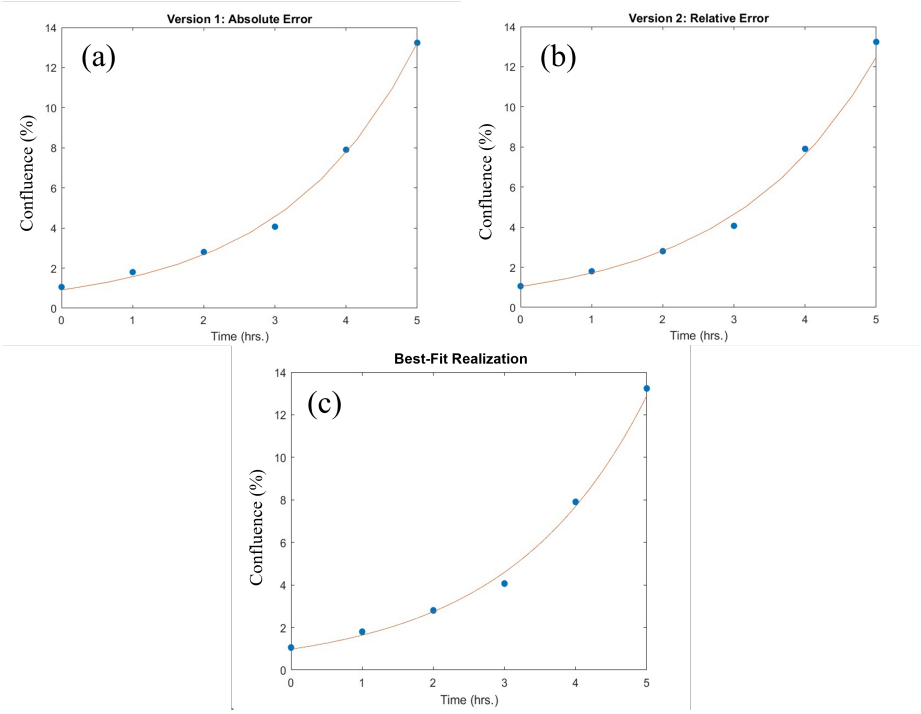


Figure 2: Three realizations of the solution to Eq. (1) using the best-fit parameter values determined by minimizing the objective functions Eq. (2)-(3) overlaid with data. *(a)*: Minimizing Eq. (2) favors fitting data points at larger confluences at the cost of a worse fit for data points at smaller confluences. *(b)*: Minimizing Eq. (3) favors fitting data points at smaller confluences over those with larger confluences. *(c)*: The best-fit realization of the solution to Eq. (1) using the average of the best-fit parameter values determined by minimizing Eq. (2) and Eq. (3).

Objective Function	Parameter	Best-Fit Parameter Value
Absolute error (Eq. (2))	L_0	0.92 %
	r	0.53 hr.$^{-1}$
Relative error (Eq. (3))	L_0	1.04 %
	r	0.50 hr.$^{-1}$

Table 1: The best-fit parameter values of r and L_0 generated by both minimizing Eq. (2) and minimizing Eq. (3).

2.2 Assessing the Efficacy of an Anticancer Drug Targeting a Mutated FGFR3 Pathway in Bladder Cancer

This analysis aims to explore the on- and off-target effects of a fibroblast growth factor receptor 3 (FGFR3) inhibitor aimed at slowing the growth of MB49 bladder cancer cells with a mutation in the FGFR3 gene that increases proliferation and decreases apoptosis. Data is collected from an experiment in which the confluences of three cell lines are recorded periodically. The three cell lines are

1. **MB49**, the parental bladder cancer cell line,
2. **G370C**, MB49 cells with a mutation in the FGFR3 gene, and
3. **2G**, MB49 cells that have undergone the laboratory process of inserting the mutation in the FGFR3 without receiving the mutation.

The 2G cell line serves as an appropriate control group for the G370C cell line since it allows us to control for any effects that the process of inserting the mutation has on the growth dynamics of the G370C cell line.

In this experiment, the FGFR3 inhibitor is administered in four dosages, control (none), low, medium, and high, to each of the three cell lines. The samples are allowed to grow freely in cell culture dishes. The confluences of the samples of the MB49 and 2G cell lines receiving each of the four drug dosages are recorded at 0, 3, 48, and 72 hours. The confluences of the samples of the G370C cell line receiving each drug dosage are recorded at 0, 45, and 69 hours. Furthermore, this experiment is performed in replicate to account for the inherent noise present in the processes underlying population growth. Specifically, the confluences of six samples of each control-dosage cell culture and three samples of each low-, medium-, and high-dosage cell culture are recorded at each time point. The procedure used to measure confluence values in this experiment interferes significantly with the growth of the sample; therefore, in this experiment each sample yields a single data point.

Since the samples grow unimpeded, we expect their growth to be approximately exponential, following the growth dynamics outlined in Eq. (1). We plot the data and confirm that exponential growth does a sufficiently good job of capturing the trend of the data for each combination of cell line and drug dosage. We then apply the parameter extraction techniques outlined in Section 2.1 to yield a best-fit growth rate r and initial value L_0 for each combination of cell line and drug dosage. Comparing the best-fit growth rates across cell lines and dosages provides information about how the FGFR3 inhibitor affects the growth dynamics of bladder cancer cells both with and without the mutation in the FGFR3 gene. The intended, targeted effect of the drug is to decrease the growth rate of cell populations with the mutation while leaving non-mutated cell populations unaffected. Other changes, such as a decrease in the growth rates of both mutated and non-mutated cell populations, are undesirable, off-target effects.

2.3 Analyzing the Effects of SIY Antigenicity on Bladder Cancer Growth Dynamics

This analysis aims to characterize the effects of the SIY antigen on the growth dynamics of MB49 bladder cancer cells in the absence of T cells. The effects of SIY antigenicity on cancer cell growth dynamics will be important to consider in future experiments in which MB49 cells will be grown in the presence of T cells. Data is collected from an experiment in which the confluences of five cells lines are recorded periodically. The five cell lines are

1. **MB49**, the parental cell bladder cancer cell line,
2. **L14**, MB49 cells expressing low levels of the SIY antigen,
3. **L19**, MB49 cells expressing medium levels of the SIY antigen,
4. **H1**, MB49 cells expressing high levels of the SIY antigen, and
5. **ZS2**, MB49 cells that have undergone the laboratory process of inducing SIY antigen expression without adding any antigen.

The ZS2 cell line serves as an appropriate control group for the three antigenic cell lines (L14, L19, and H1) since it allows us to control for any effects that the process of inducing SIY antigen expression has on the growth dynamics of the antigenic cell lines.

In this experiment, six samples of each cell line are allowed to grow freely in cell culture dishes. The confluences of the samples are measured every four hours over the course of 28 hours. Unlike the procedure used in the FGFR3 inhibitor experiment (see Section 2.2), the procedure used to measure confluence values in this experiment does not significantly interfere with the growth of the sample; therefore, each sample is measured at each time point during the experiment.

Since the samples grow unimpeded, we similarly expect their growth to be approximately exponential, following the growth dynamics outlined in Eq. (1). We plot these data on a logarithmic scale and use a measure of linear correlation to confirm that exponential growth does a sufficiently good job of capturing the trend of the data for each sample of each cell line. We then apply the parameter extraction techniques outlined in Section 2.1 to yield a best-fit growth rate r and initial value L_0 for each sample of each cell line. We expect that SIY antigenicity has no noticeable effect on the growth dynamics of bladder cancer cells, since in the absence of T cells we expect that the SIY antigen will not inhibit cancer cell population growth. Comparing the best-fit growth rates across cell lines provides information about the effect of SIY antigenicity on bladder cancer growth, allowing us to evaluate our expectation that bladder cancer growth is unaffected by SIY antigenicity.

2.4 Constructing an ODE Model to Describe the Interactions Between Bladder Cancer Cells and T Cells

We construct a model consisting of four ODEs describing the interactions between healthy cancer cells (L), apoptotic cancer cells (A), healthy T cells (T), and apoptotic T cells (U). The model incorporates the natural reproduction and death of cancer and T cells, the killing of cancer cells by T cells, and the exhaustion of T cells in response to chronic antigen exposure. Eq. (6) provides a description of the biological processes corresponding to each term of the model ODEs.

$$
\begin{aligned}
L' = {} & \boxed{\text{Growth rate of healthy cancer cells}} \\
& - \boxed{\text{Rate at which healthy cancer cells become apoptotic}} \\
& - \boxed{\text{Rate at which cancer cells are killed by T cells}} \\
A' = {} & \boxed{\text{Rate at which healthy cancer cells become apoptotic}} \\
& + \boxed{\text{Rate at which cancer cells are killed by T cells}} \\
& - \boxed{\text{Rate at which apoptotic cancer cells are cleared}} \\
T' = {} & \boxed{\text{Growth rate of healthy T cells}} \\
& - \boxed{\text{Rate at which healthy T cells become apoptotic}} \\
& - \boxed{\text{Rate at which healthy T cells die due to exhaustion}} \\
U' = {} & \boxed{\text{Rate at which healthy T cells become apoptotic}} \\
& + \boxed{\text{Rate at which healthy T cells die due to exhaustion}} \\
& - \boxed{\text{Rate at which apoptotic T cells are cleared}}
\end{aligned}
\tag{6}
$$

The model ODEs are listed in Eq. (7).

$$
\begin{aligned}
L' &= (r-a)L - \theta \cdot k_i(L,T) \\
A' &= aL + \theta \cdot k_i(L,T) - dA \\
T' &= (\alpha - \beta)T - \phi \cdot k_i(L,T) \\
U' &= \beta T + \phi \cdot k_i(L,T) - dU.
\end{aligned}
\tag{7}
$$

We suggest four functional forms of $k_i(L,T)$, which describes the mechanism by which cancer cells are killed by T cells and by which T cells die due to

exhaustion:

$$k_1(L,T) = LT \tag{8}$$

$$k_2(L,T) = \frac{(T/L)^\lambda}{s + (T/L)^\lambda} \tag{9}$$

$$k_3(L,T) = \frac{T^\lambda}{s + T^\lambda} \tag{10}$$

$$k_4(L,T) = LT^\lambda. \tag{11}$$

Eq. (8) implements the law of mass action, whereas Eq. (9) and Eq. (10) are increasing Hill functions of T/L and T, respectively. Eq. (11), a variation on the law of mass action, did not yield usable results in this analysis.

When applying the best-fit parameter extraction method described in Section 2.1, we fix the net growth rates of cancer cells and T cells to typical values which are well-known from previous experiments:

$$r_{net} = r - a = \text{constant} \tag{12}$$

$$\alpha_{net} = \alpha - \beta = \text{constant}. \tag{13}$$

A description of the biological significance of each model parameter along with an assumed value of each model parameter is given in Table 2.

We analyze the model parameters r, α, θ, ϕ, and d, along with a noise estimate parameter σ, in a Bayesian framework. A similar method has been used in past research [1]. We perform Bayesian inference with uninformative prior distributions and noisy synthetic data by sampling sets of parameter values from prior distributions and computing posterior distributions according to Bayes' theorem. This helps us understand to what extent the model parameters are practically identifiable. Using the sets of parameter values sampled in this process, we generate pair plots to investigate correlations between pairs of parameters.

To assess the feasibility of performing this parameter extraction in practice, we investigate the practical identifiability of several model parameters of interest. We perform Bayesian inference on the cancer-T-cell model using three versions of $k_i(L,T)$ (see Eq. (8)-(10)) with noisy synthetic data to look for trends in the posterior distributions of these parameters and to look for correlations between pairs of parameters. The fourth version of $k_i(L,T)$ (see Eq. (11)) did not yield usable results. The synthetic data are generated by recording the solution to the cancer-T-cell model equations at 0, 5, 10, and 15 hours for a set of reasonable input parameters (see Table 3). Uniformly distributed noise is added to these data over a range of $\pm 15\%$ of the data value; this is meant to simulate the measurement error present in experimental data.

The input parameter values used in this analysis are the same as the assumed values listed in Table 2. These values, along with the prior distributions for the parameters of interest, are listed are listed in Table 3. Note that each prior distribution is a uniform distribution, except for the the prior distribution for the noise estimate parameter σ, which is an inverse gamma distribution:

Parameter	Description	Assumed Value	Units
L_0	Initial confluence of healthy cancer cells	10	%
A_0	Initial confluence of apoptotic cancer cells	0	%
T_0	Initial confluence of healthy T cells	5	%
U_0	Initial confluence of apoptotic T cells	0	%
r	Growth rate of cancer cells	0.2	hr.^{-1}
a	Death rate of cancer cells	0.1	hr.^{-1}
θ	Rate at which cancer cells are killed by T cells	0.01	$(\% \cdot \text{ hr.})^{-1}$ or $\% \cdot \text{hr.}^{-1}$
d	Rate at which apoptotic cells are cleared	0.25	hr.^{-1}
α	Growth rate of T cells	0.06	hr.^{-1}
β	Death rate of T cells	0.05	hr.^{-1}
ϕ	Rate at which T cells die due to exhaustion	0.01	$(\% \cdot \text{ hr.})^{-1}$ or $\% \cdot \text{hr.}^{-1}$
λ	Exponent in cancer killing/T-cell exhaustion term	1	-
s	Steepness coefficient in cancer killing/T-cell exhaustion term	1	-

Table 2: A description of the biological significance of each model parameter of the cancer-T-cell ODE model (see Eq. (7)-(10)), along with assumed values of each model parameter. Note that the units of θ and ϕ are $(\% \cdot \text{ hr.})^{-1}$ when using Eq. (8) and $\% \cdot \text{hr.}^{-1}$ when using Eq. (9) or Eq. (10).

Parameter	Input Value	Prior Distribution
L_0	10 %	-
A_0	0 %	-
T_0	5 %	-
U_0	0 %	-
r	0.2 hr.$^{-1}$	Uniform$[0, 0.5]$
a	0.1 hr.$^{-1}$	-
θ	0.01 (% $\cdot$ hr.)$^{-1}$ or % $\cdot$ hr.$^{-1}$	Uniform$[0, 0.1]$
d	0.25 hr.$^{-1}$	Uniform$[0, 0.5]$
α	0.06 hr.$^{-1}$	Uniform$[0, 0.5]$
β	0.05 hr.$^{-1}$	-
ϕ	0.01 (% $\cdot$ hr.)$^{-1}$ or % $\cdot$ hr.$^{-1}$	Uniform$[0, 0.1]$
λ	1	-
s	1	-
σ	-	InverseGamma($\alpha = 2$, $\beta = 3$)

Table 3: Input values and prior distributions for the parameters of the cancer-T-cell model (see Eq. (7)-(10)), along with the noise estimate parameter σ. Bayesian inference was not performed on the initial conditions, a, β, λ, or s; therefore these parameters do not require prior distributions.

$\sigma \sim$ InverseGamma($\alpha = 2$, $\beta = 3$). Eight chains of 4,000 samples each are obtained from the prior distributions using a No U-Turn Sampler (NUTS) with a target acceptance ratio of 0.65.

Parameter values such as the intrinsic growth rates of the cancer cells (r) and T cells (α) are well known from previous experiments, so discovering strong correlations between these and other less well-known parameters (in this case, θ, ϕ, and d) will allow for the identification of the values of these less well-known parameters based on their correlation with well-known parameter values.

3 Results

3.1 Parameter Extraction from an ODE Model to Assess the Efficacy of an Anticancer Drug

We apply the best-fit parameter extraction method outlined in Section 2.1 to the data from the FGFR3 inhibitor experiment (see Section 2.2). Fitting the parameters of the exponential growth ODE model (see Eq. (3)) to the experimental data yields a best-fit growth rate and initial confluence for each combination of cell line and FGFR3 inhibitor dosage. Comparing the best-fit growth rates across cell lines and drug dosages allows us to investigate the on- and off-target effects of an FGFR3 inhibitor aimed at slowing the growth of MB49 bladder cancer cells with a mutation in the FGFR3 gene.

Figures 3-5 show the experimental data (blue) overlaid with the solution to

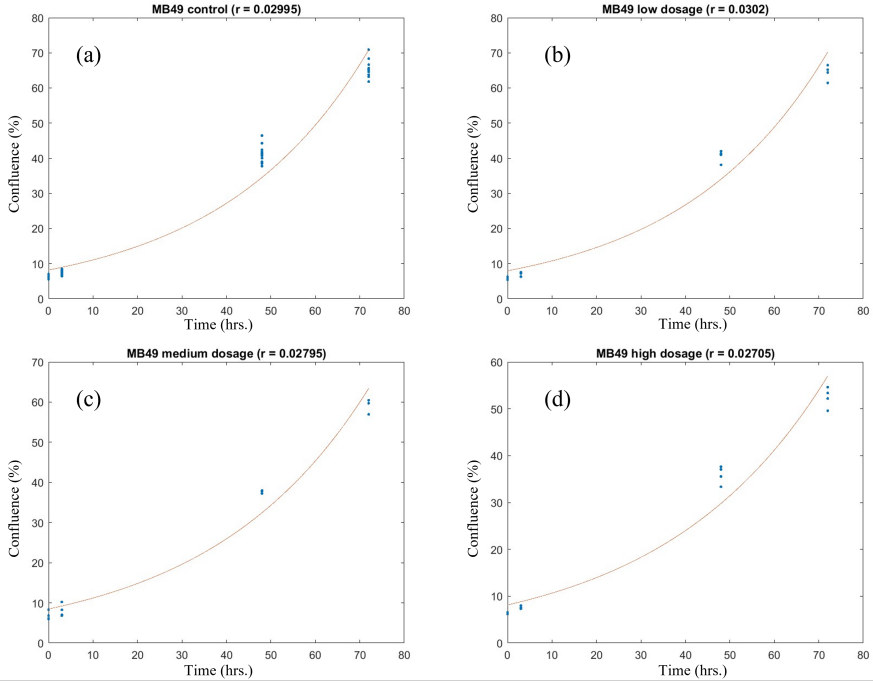


Figure 3: Experimental data for the MB49 cell line (blue) overlaid with the solution to Eq. (1) using the best-fit parameter values (red) for control *(a)*, low *(b)*, medium *(c)*, and high *(d)* dosages of an FGFR3 inhibitor.

Eq. (1) using the best-fit parameter values (red) for the MB49, G370C, and 2G cell lines, respectively, treated with four different dosages of the drug. The best-fit growth rate in hr.$^{-1}$ is displayed in the title of each panel of Figures 3-5. As noted in Section 2.2, the procedure used to measure confluence values in these trials interferes significantly with the growth of the cell culture. Therefore, each data point in Figures 3-5 is taken from a different sample.

For each cell line, we normalize the best-fit growth rates to that of the control-dosage trial for that cell line. This allows us to better observe the effects of the drug on the cancer cell growth dynamics across cell lines and dosages by controlling for the differences in the cell lines' intrinsic growth rates (see Table 4 and Figure 6).

In assessing the efficacy of the FGFR3 inhibitor, we are especially interested in comparing the G370C and 2G cell lines, as the 2G cell line is intended to serve as a control for the effects of inserting the mutation in the FGFR3 gene of the G370C cells (see Section 2.2). Noticeable on-target effects (i.e., a decrease in the growth rate of the G370C cell line while leaving the growth rate of the 2G cell line unaffected) emerge only in the high-dosage trial. Even for the high-dosage trial, off-target effects (i.e., a decrease in the growth rates of both the

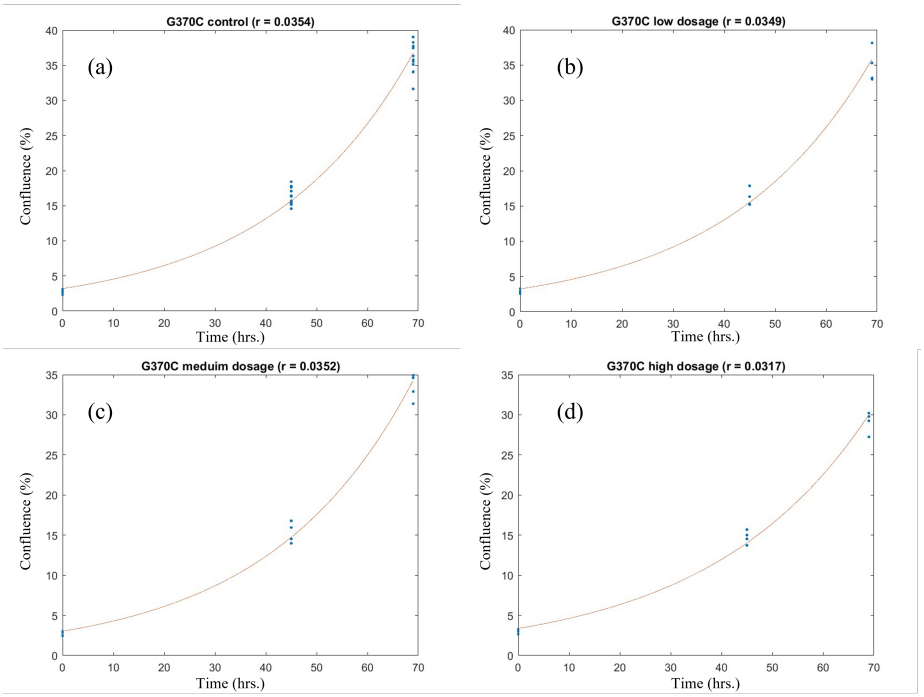


Figure 4: Experimental data for the G370C cell line overlaid with the solution to Eq. (1) using the best-fit parameter values for control *(a)*, low *(b)*, medium *(c)*, and high *(d)* dosages of an FGFR3 inhibitor.

	Control Dosage	Low Dosage	Medium Dosage	High Dosage
2G	1	0.985	0.996	0.918
G370C	1	0.986	0.994	0.895
MB49	1	1.008	0.933	0.903

Table 4: The best-fit growth rates across cell lines and FGFR3 dosages, normalized to the control-dosage trial for each cell line.

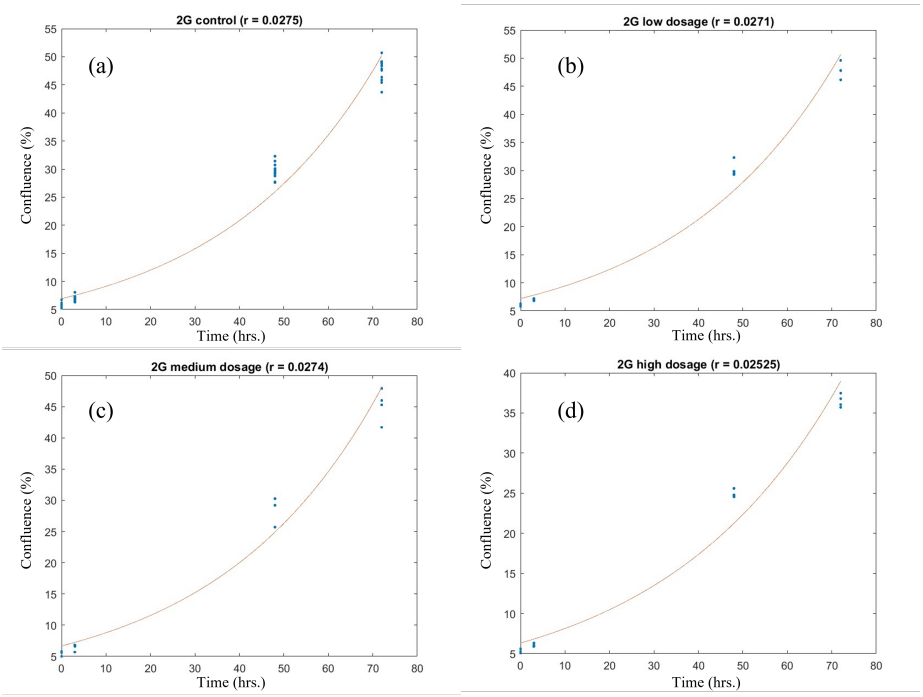


Figure 5: Experimental data for the 2G cell line overlaid with the solution to Eq. (1) using the best-fit parameter values for control *(a)*, low *(b)*, medium *(c)*, and high *(d)* dosages of an FGFR3 inhibitor.

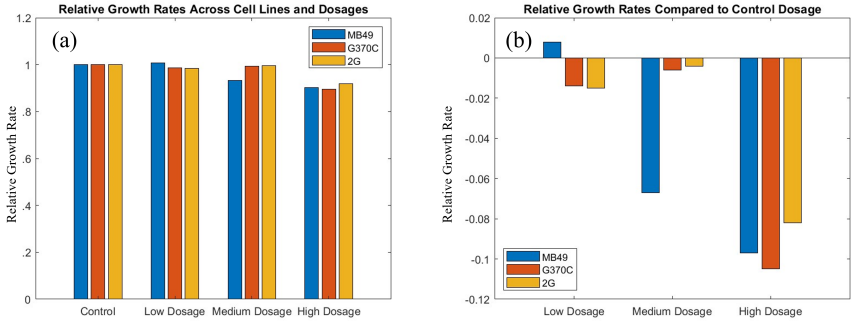


Figure 6: *(a)*: The best-fit growth rates across cell lines and FGFR3 inhibitor dosages, normalized to the control-dosage trial for each cell line. *(b)*: The differences between the best-fit growth rates and the control-dosage growth rates for the low-, medium-, and high-dosage trials, normalized to the control-dosage trial for each cell line.

	MB49	ZS2	L14	L19	H1
Sample 1	0.998	0.994	0.992	0.999	0.999
Sample 2	0.999	0.993	0.997	1.000	0.999
Sample 3	0.999	0.992	0.994	0.999	0.999
Sample 4	1.000	0.997	0.995	0.999	0.999
Sample 5	1.000	0.997	0.995	0.999	0.999
Sample 6	0.998	0.996	0.995	0.999	0.999

Table 5: Pearson correlation coefficient (PCC) values for the data from the SIY antigenicity experiment plotted on a logarithmic scale (see Figure 7). A PCC value close to 1 indicates a strong, positive correlation between the independent and dependent variables.

G370C and 2G cell lines) dominate over on-target effects. The relatively small effect the FGFR3 inhibitor has on slowing the growth of bladder cancer cells with the mutation to the FGFR3 gene even at high dosages suggests that the mutation may promote cancer growth in a way other than by increasing the intrinsic growth rate of the cancer cells. One possibility is that the mutation affects the *in vivo* microenvironment in a way that increases the growth rate of bladder cancer cells and decreases the the rate at which they become apoptotic.

3.2 Parameter Extraction from an ODE Model to Analyze the Effect of SIY Antigenicity on Bladder Cancer Growth Dynamics

We expect that the exponential growth model outlined in Eq. (1) is a good description of the data from the SIY antigenicity experiment (see Section 2.3). To confirm this expectation, we plot the data on a logarithmic scale and compute the Pearson correlation coefficient (PCC) for each cell culture data set (see Table 5 and Figure 7). The PCC is a measure of linear correlation; a PCC value of 1 indicates that there is a perfect positive, linear relationship between the independent and dependent variable, whereas a PCC value of 0 indicates that there is no correlation between the independent and dependent variable. If the logarithmic-scale plots are approximately linear with a positive slope (and therefore yield PCC values close to 1), we can be confident that the data approximately follow exponential growth. Since the PCC values for each sample of each cell line are greater than 0.99, we can be confident that the exponential growth ODE model Eq. (1) will yield a close fit to the data.

We again apply the best-fit parameter extraction method outlined in Section 2.1 to the experimental data. Fitting the parameters of Eq. (1) to the experimental data yields a best-fit growth rate and initial confluence for each sample of each cell line. Figure 8 shows the experimental data (blue) overlaid with the solution to Eq. (1) using the best fit parameter values for one sample of MB49 cells (red). This procedure was repeated for the other 29 samples used in the experiment.

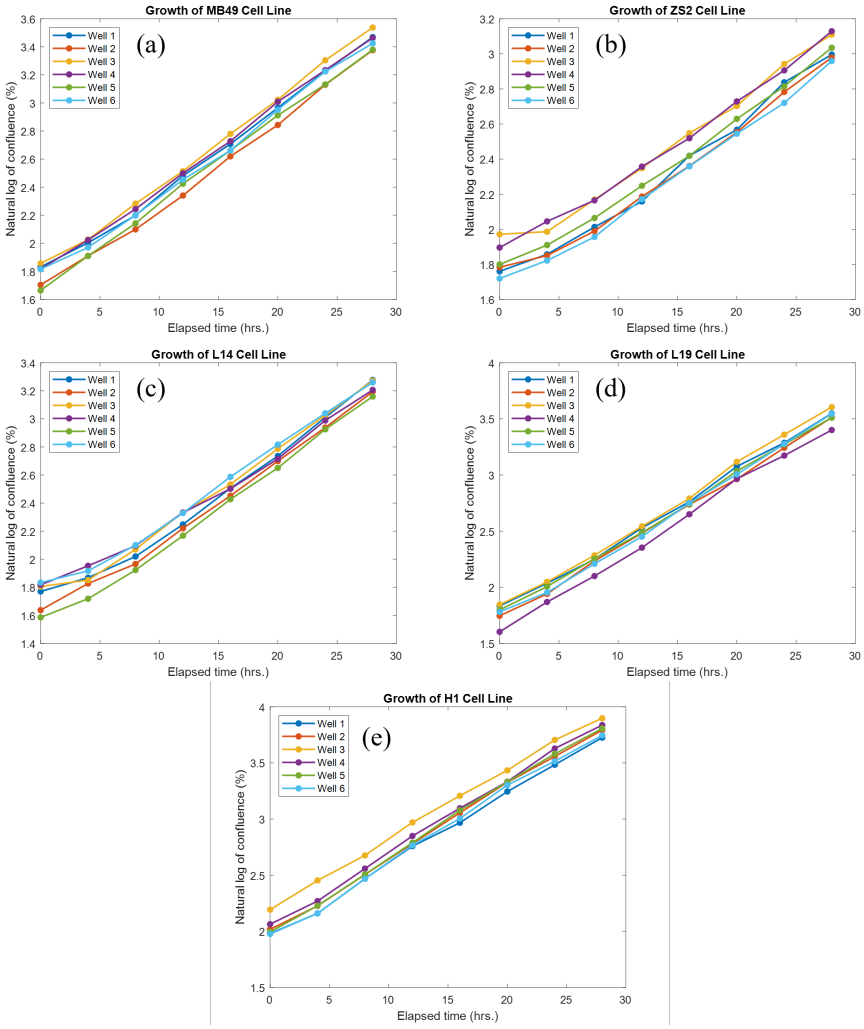


Figure 7: Experimental data plotted on a logarithmic scale for the MB49 *(a)*, ZS2 *(b)*, L14 *(c)*, L19 *(d)*, and H1 *(e)* cell lines. If the growth dynamics of the samples approximately follow exponential growth, then we expect to see nearly linear graphs when the data are plotted on a logarithmic scale.

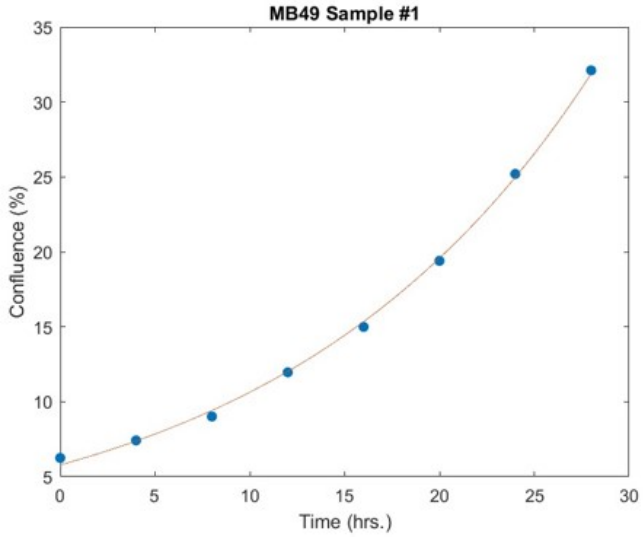


Figure 8: Experimental data (blue) overlaid with the solution to Eq. (1) using the best fit parameter values for one sample of MB49 cells (red). This procedure was repeated for the other 29 samples used in the experiment (not shown).

As noted in Section 2.3, this experiment yields true time-series data (i.e., the confluence of each sample is measured multiple times without being destroyed). Therefore, we take the mean of the six best-fit growth rates for each cell line to yield a single growth rate associated with each cell line. Table 6 shows that the standard deviation of the growth rates among samples of the same cell line is small compared to the mean growth rate for each cell line; therefore, it is reasonable to characterize each cell line with a single intrinsic growth rate. Comparing the best-fit growth rates across cell lines allows us to investigate the effect of SIY antigenicity on the growth dynamics of cancer cells.

The means and standard deviations of the best-fit growth rates, along with

	Mean Growth Rate (days^{-1})	Standard Deviation of Growth Rates (days^{-1})	Doubling Time (hrs.)
MB49	1.46	0.02	11.4
ZS2	1.10	0.02	15.1
L14	1.36	0.06	12.3
L19	1.54	0.03	10.8
H1	1.53	0.04	10.9

Table 6: The means and standard deviations of the best-fit growth rates, along with their associated doubling times, for each cell line.

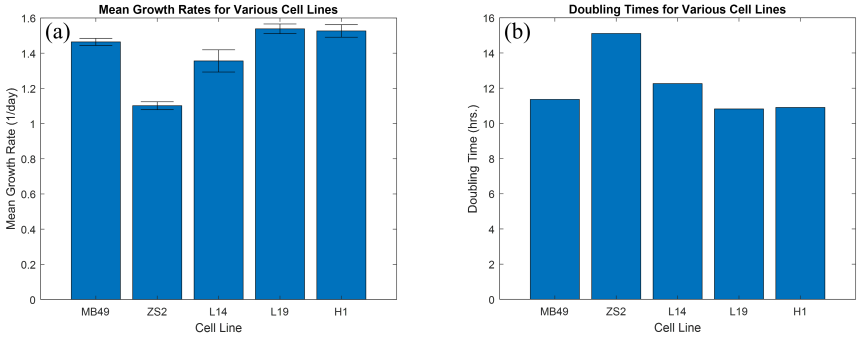


Figure 9: *(a)*: The mean best-fit growth rate for each cell line. The error bars span ± 1 standard deviation. *(b)*: The doubling time associated with the mean growth rate for each cell line.

their associated doubling times, are shown for each cell line in Table 6 and plotted in Figure 9. In analyzing the effect of antigenicity on cancer cell population growth dynamics, we are especially interested in comparing the ZS2 cell line with the antigenic cell lines (L14, L19, and H1), as the ZS2 cell line is intended to serve as a control for the effects of inducing SIY antigen expression in the antigenic cell lines (see Section 2.3). Contrary to our expectation, the growth dynamics of MB49 cells appear to depend strongly on the amount of SIY antigen expressed by cancer cells. Specifically, the L19 (medium-antigen) and H1 (high-antigen) cell lines exhibit even larger growth rates than the L14 (low-antigen) cell line, which exhibits a larger growth rate than the ZS2 (control) cell line. The growth rate of the MB49 parental cell line is comparable to those of the L19 and H1 cell lines.

3.3 Investigating a Cancer-T-Cell ODE Model in a Bayesian Framework

We use the Generating Series for Testing Structural Identifiability (GenSSI) 2.0 software package to perform structural identifiability analysis on the cancer-T-cell model outlined in Eq. (7) for each of the four versions of $k_i(L,T)$ (see Eq. (8)-(11)). Using Eq. (8) (the law of mass action), we find that the model is globally structurally identifiable. Using Eq. (9)-(11), the model is locally structurally identifiable. For any choice of $k_i(L,T)$, the model is structurally identifiable, which informs us that we can theoretically extract the value of each model parameter from data given a sufficiently good initial guess.

To assess the feasibility of performing this parameter extraction in practice, we investigate the practical identifiability of several model parameters of interest. We are most interested in simultaneously analyzing the parameters r and α, whose values are well known from previous experiments, with the less well-known parameters θ, d, and ϕ. We perform the Bayesian inference procedure

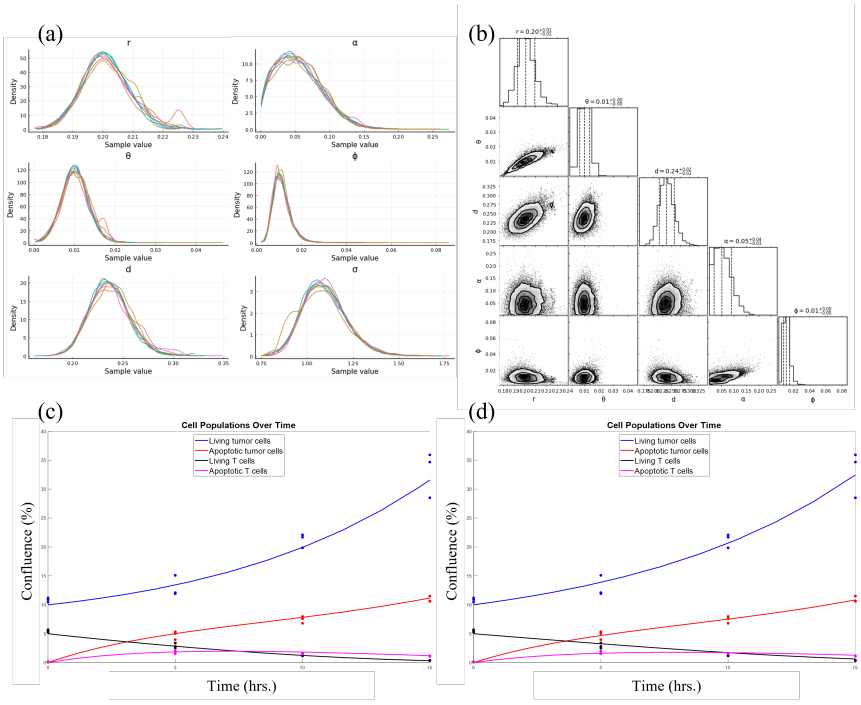


Figure 10: Bayesian inference results using Eq. (8) (the law of mass action). *(a)*: Posterior distributions for five model parameters and the noise estimate parameter σ. *(b)*: Pair plots for five model parameters. *(c)-(d)*: Realizations of the solution to (7) using (8) using two of the 8000 sampled sets of parameter values used in this analysis, overlaid with synthetic data.

on these five parameters along with the noise estimate parameter σ for each of three versions of $k_i(L,T)$ (see Eq. (8)-(10)).

The results of this analysis using Eq. (8) (the law of mass action) are displayed in Figure 10. Each posterior distribution in Figure 10(a) appears to be approximately normal, indicating that the model parameters are practically identifiable to a large extent. We note the strong, positive correlation between r and θ and the weaker positive correlation between α and ϕ shown in the pair plots (see Figure 10(b)). Figure 10(c)-(d) show two realizations of the solution to the cancer-T-cell model outlined in Eq. (7) using Eq. (8) using two of the 8000 sampled sets of parameter values used in this analysis overlaid with synthetic data. We note the high level of agreement between the best-fit model solutions and the data across all four cell populations.

The results using Eq. (9) (an increasing Hill function of T/L) are displayed in Figure 11. The posterior distributions of r, α, d, and σ in Figure 11(a) appear to be approximately normal, indicating that these parameters are practically identifiable to a large extent. There is a strong, positive correlation between r

Parameter	Input Value	Posterior Distribution Mean Using		
		Eq. (8)	Eq. (9)	Eq. (10)
r	0.2	0.2014	0.2057	0.2380
θ	0.01	0.0102	0.0290	0.0562
d	0.25	0.2376	0.2976	0.3653
α	0.06	0.0554	0.0768	0.0741
ϕ	0.01	0.011	0.0325	0.0152

Table 7: The means of the posterior distributions for each model parameter analyzed for the three versions of $k_i(L,T)$ used, along with the input parameter values.

and θ and between α and ϕ (see Figure 11(b)). Since r appears to be practically identifiable and r and θ are highly correlated, we can be confident in our ability to identify the value of θ given the value of r. A similar relationship exists between α and ϕ. These results indicate that these parameters are practically identifiable to a large extent.

The results using Eq. (10) (an increasing Hill function of T) are displayed in Figure 12. We note a strong, positive correlation between α and ϕ. There is also a strong, positive, three-way correlation between r, θ, and d; however, the lack of sharp peaks in the posterior distributions of these parameters greatly limits their practical identifiability. One way of overcoming the limitations imposed by the unidentifiability of r, θ, and d is by using the fact that r, the intrinsic growth rate of cancer cells, is well-known from previous experiments. Assuming the value of r allows us to use the correlation between r and θ and between r and d to closely identify each parameter.

The means of the posterior distributions describing each model parameter analyzed for the three versions of $k_i(L,T)$ used are shown alongside the input parameter values in Table 7.

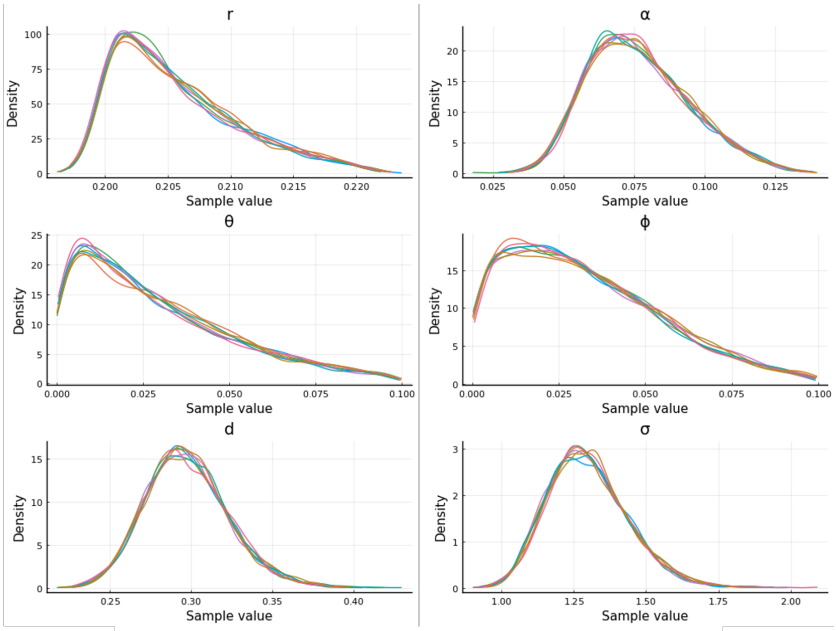
4 Discussion

This paper showcases some basic methods used to apply ODE models to experimental data in order to obtain a more detailed description of the growth dynamics governing MB49 bladder cancer. Extracting the best-fit values of model parameters from experimental data reveals information about the underlying growth dynamics of a system that is not immediately measurable from the data alone.

We first define a method to generate the best-fit realization of an ODE model to experimental data by choosing model parameter values that minimize the outputs two objective functions, one considering absolute error and the other considering relative error. This yields two reasonable sets of parameter values; the average of the the two values of each parameter is defined to be the best-fit value of that parameter.

We apply this method using a simple model of exponential growth to assess

(a)



(b)

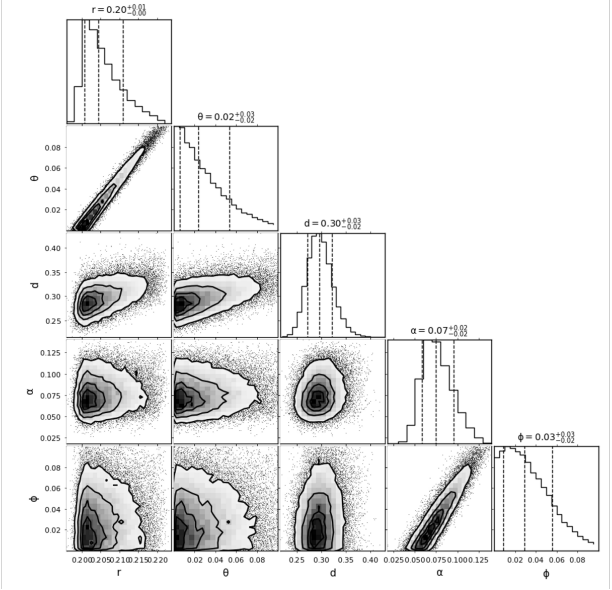
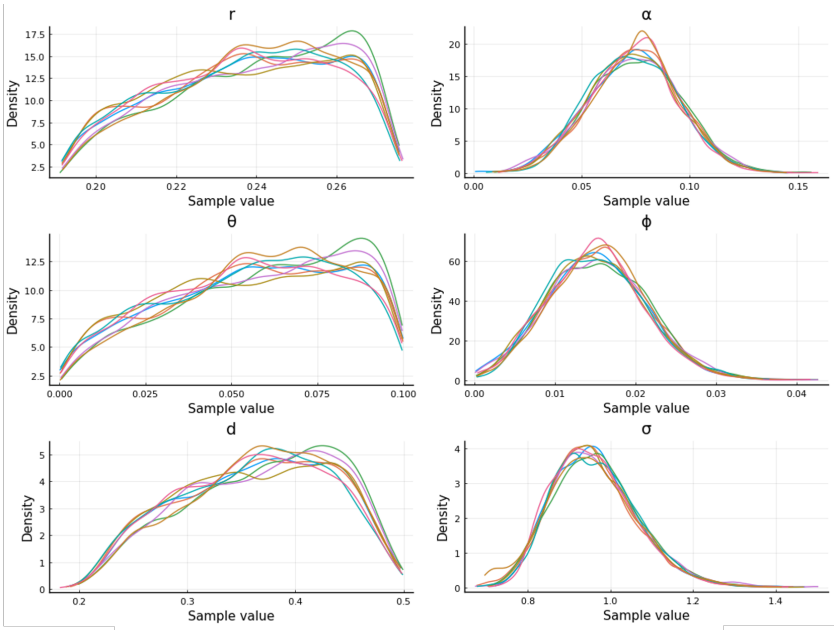


Figure 11: Bayesian inference results using Eq. (9) (an increasing Hill function of T/L). *(a)*: Posterior distributions for five model parameters and the noise estimate parameter σ. *(b)*: Pair plots for five model parameters.

(a)



(b)

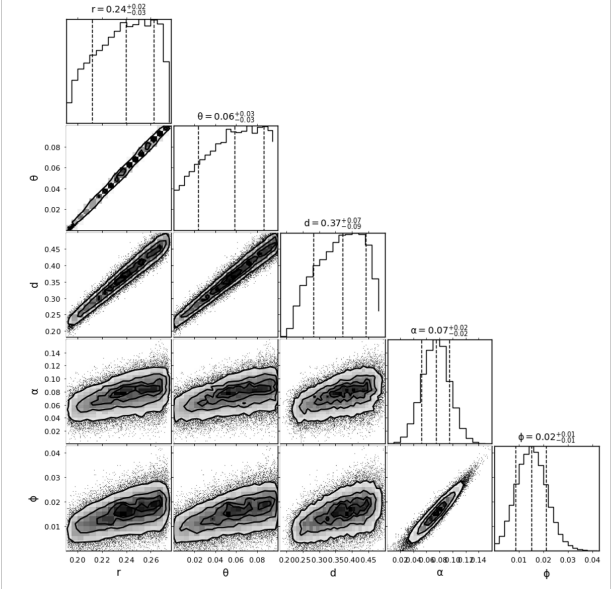


Figure 12: Bayesian inference results using Eq. (10) (an increasing Hill function of T). *(a)*: Posterior distributions for five model parameters and the noise estimate parameter σ. *(b)*: Pair plots for five model parameters.

the efficacy of an FGFR3 inhibitor intended to slow the growth of MB49 bladder cancer cells with a mutation in the FGFR3 gene. We find that the majority of the drug's effects are off-target, slowing the growth of cancer cell populations both with and without the mutation. However, at high dosages, on-target effects begin to emerge, and mutated populations grow at slower rates than the non-mutated population. We conclude that the mutation to the FGFR3 gene likely promotes cancer cell population growth in a way other than by increasing the intrinsic growth rate of cancer cells, possibly by affecting the *in vivo* microenvironment in a way that is not captured by the *in vitro* experiment considered here.

Next, we analyze the effect of the amount of SIY antigen expression on the growth dynamics of MB49 cells in the absence of T cells. Using the same exponential growth model as before, we determine a best-fit growth rate for several cell lines corresponding to different levels of SIY antigenicity. We find that higher levels of antigen expression generally correspond to larger growth rates. These results contradict our expectation that SIY antigenicity has no effect on the growth dynamics of populations of bladder cancer cells.

Finally, we investigate the behavior of an ODE model describing cancer-T-cell interactions. We use a software package to show that the model is structurally identifiable for any choice of $k_i(L, T)$. This informs us that, given a sufficiently good initial guess, we can theoretically extract the value of each model parameter from data. To assess the feasibility of performing this parameter extraction in practice, we investigate the model in a Bayesian framework. The tools of Bayesian inference allow us to gain information about the practical identifiability of several model parameters of interest and investigate correlations between model parameters with well-known values and those whose values are not well known. We aim to apply this cancer-T-cell model to future experimental data that study the interactions between bladder cancer and T-cell populations over time. Fitting the parameters of the cancer-T-cell model to these data will allow us to gain a clearer picture of the underlying dynamics that govern the interactions between bladder cancer and T-cell populations.

This approach to applying the mathematical tools of ODE modeling to data obtained through biological experimentation should prove to be a valuable tool for collaboration between laboratory researchers and mathematicians, assisting in our understanding of malignancy in cancers and eventually leading to improved therapeutics.

References

[1] Jesse Kreger, Evanthia T Roussos Torres, and Adam L MacLean. Myeloid-derived suppressor cell dynamics control outcomes in the metastatic niche. *bioRxiv*, 2022.