Modeling Asymmetric Cell Division of Pathogenic Yeast within Host Macrophages for Comparison to Fluorescence Confocal Microscopy Data

MTSU MATHEMATICAL SCIENCES DEPARTMENT

THESIS FORMAT

A Thesis

Presented to the Faculty of the Department of Mathematical Sciences

Middle Tennessee State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Mathematical Sciences

---

by

Anna Davis

April 4, 2024

Thesis Committee:

Member 1, Chair, Dr. Rachel Leander

Member 2, Dr. James Hart

Member 3, Dr. David Nelson

Member 4, Dr. Yixiang Wu

DEDICATION

I would like to dedicate this thesis to my family, especially my husband Thomas Davis, my parents, and my sister, and to Bethany Maynard, Cassidy Fleenor, and Grace Beckner.

ACKNOWLEDGMENTS

I would like to thank the following individuals for their contribution to this work. First, thank you to Dr. David Nelson for providing much of the biological context in this thesis. Thank you also for welcoming me into your lab space and giving me the opportunity to learn so many new things. Thank you to Derek Wiggins for taking his time to teach me how to culture cells and allowing me to be present for multiple infections and imagings. Also, thank you for the counting rules presented in this thesis.

Thank you to all of my peers in the mathematics department, especially Morgan Chappell, Breanne Jones, Joy Lewis, Joy Rhodes, Sarvani Pemmaraju, Nada Srour, Moustafa Shamdeen, Ivan Lozano and Wil Jones. A special thank you to Morgan Chappell for working on this modeling project with me and creating many of the figures found in this paper. Many thanks to my thesis committee for taking the time to read, listen to, and improve this thesis. Finally, I would like to acknowledge my advisor, Dr. Rachel Leander, for teaching me so much throughout this process. None of this would be possible without her.

ABSTRACT

Cryptococcus neoformans is a fungal pathogen found most commonly in soil and bird droppings. This pathogen causes pulmonary infections and can establish long-lasting intracellular infections in macrophages. While most healthy individuals are not impacted by exposure to these fungi, those who are immunocompromised may experience pulmonary infections. If *Cryptococcus neoformans* disseminates from the lung, it can infect the central nervous system and cause fungal meningitis.

Our ultimate objective is to study the replication of this fungal pathogen in two macrophage cell lines, J774 macrophages and fetal liver-derived alveolar-like macrophages (FLAMs) by using florescence confocal microscopy. Intracellular growth of *Cryptococcus neoformans* is usually studied in J774 macrophages. We are interested in comparing the growth of *Cryptococcus neoformans* in FLAMs to that in J774 cells because the lung is the most common site of initial *Cryptococcus neoformans* infection, and FLAMs are more similar to alveolar macrophages from the lung than J774 cells.

In this work, we develop a stage-structured model for the asymmetric cell division of *Cryptococcus neoformans* in macrophages which is compatible with empirical data. Quantifying changes in growth and division under perturbations and assigning statistical significance requires a model of the expected growth and division pattern. A small population size combined with asymmetric division makes the go-to model of continuous exponential growth inappropriate for quantifying changes in growth rate and assigning statistical significance. Our stage-structured model is useful because it more accurately captures the dynamics of a small and asymmetrically dividing cellular population. We couple this model with an error model for the process of detecting and counting fungal cells in order to counteract some of the difficulties of collecting the empirical data, especially the challenge of reliably detecting fungal cells that are very small.

CONTENTS

List of Figures

CHAPTER 1

INTRODUCTION

Cryptococcus neoformans

Cryptococcus neoformans is a pathogenic yeast that is prevalent across the world but most commonly found in soil, bird droppings, and trees. Most healthy individuals are not impacted by exposure to these fungi, but those who are immunocompromised may experience lung infections and meningitis [6] . We are studying the replication of this fungal pathogen in fetal liver-derived alveolar-like macrophages (FLAMs) and J774 macrophages by using fluorescence confocal microscopy and modeling. [1][14]

Cryptococcus neoformans are unique among fungal pathogens in that they are encapsulated in a thick polysaccharide (PS) shell [5]. Like *Saccharomyces cerevisiae* (budding yeast), *Cryptococcus neoformans* reproduces by asymmetric budding. This means the adult cell, which we will refer to as the mother cell, grows a small bud that increases in size before eventually splitting from the mother cell. During budding, the cell wall deforms to generate a specialized region, known as a scar, from which the daughter cell will bud [5]. Unlike *Saccharomyces cerevisiae*, which form a new scar for each bud, *Cryptococcus neoformans* always buds from the same scar [5]. Budding is different than the more familiar process of symmetrical cell division, in which a mother cell splits in half to create two identical daughter cells. The process of asymmetrical budding is shown in Figures 1, 2, and 3, which can be viewed at the end of this chapter. Asymmetric cell division makes it more difficult to accurately count the number of cells in the population because small buds are not reliably imaged in the microscopy process, and it is difficult to discern if a bud has separated from its mother.

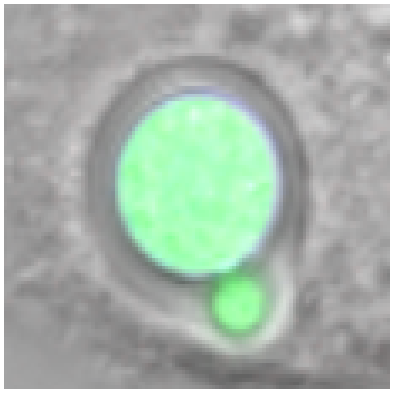

Figure 1: *Intracellular Cryptococcus neoformans Budding in FLAM Macrophage*

Cell Lines

We studied the replication of *Cryptococcus neoformans* in two macrophage cell lines, J774 macrophage-like cells and fetal liver-derived alveolar-like macrophages (FLAMs) by using florescent confocal microscopy. The J774 macrophages are derived from sarcoma found in mice [1]. These macrophages are commonly used as an in vitro macrophage model in the *C. neoformans* research field [2, 15]. The J774 macrophages featured in the microscopy data are considered to be male J774 cells, meaning they are derived from male mice. As the name suggests, FLAMs are similar to alveolar macrophages from the lung [14]. Similar to the J774 macrophages, the FLAMs featured in the microscopy images are male FLAM cells, being derived from male mice.

Confocal Microscopy Data

Images were collected using a Zeiss AxioObserver LSM700 confocal laser scanning microscope. The microscope was equipped with a Plan-Apochromat 63x magnification/1.40 numerical aperture oil immersion DIC M27 objective lens, stage top CO_2

incubator, and XL enclosure. Images were collected every thirty minutes for a maximum of 24 hours.

Images from each field were collected at multiple vertical distanced above the base of the dish. The images for a single field are referred to as z-slices or collectively as the z-stack. There were 5 to 7 z-slices per field of view depending on the experiment. To accommodate differences in cell thickness, J774 images were collected as 5 z-slices at 2 μm intervals across an 8 μm range, and FLAM images were collected as 7 z-slices at 2 μm intervals across a 12 μm range.

Fields were included in the study if at least one macrophage in the field contained a small number (one or two) of intracellular *Cryptococcus neoformans* at the initial time of imaging. Macrophages that had one or two fungal cells at the initial time were then followed and their fungal burden was recorded for up to 10 hours. A small multiplicity of infection (0.25:1 - 0.5:1) was used to ensure sufficient numbers of macrophages had only one or two *C. neoformans* at the onset of imaging. The objective was to obtain infected macrophages that were spatially separated with a low fungal burden.

The software that was used to image the pathogens during image acquisition was Zeiss Zen Black. In order to analyze the confocal microscopy data and estimate the *C. neoformans* radius we utilized a software called Zeiss Zen Lite. These measurements informed our model parameterization.

Now, we will discuss some of the mathematical components involved with the replication of *Cryptococcus neoformans*.

Stage Structured Population

Asymmetric cell division adds complexity to the system because it results in a stage-structured population; it will take longer for a bud to produce its own bud than it will for the original mother to produce a new bud. Therefore, in order to model

population growth, we have to keep track of each cell's status as an immature bud or actively budding mother cell.

Asymmetric cell division coupled with the small population size makes the standard exponential model for population growth inappropriate. The continuous exponential growth model is as follows:

$$N(t) = N_0 e^{rt}$$

.

In this model, r is the growth rate, N_0 is the initial population size, and $N(t)$ is the population size at time t. It is also worth noting $\frac{dN}{dr} = rN$. The doubling time for this model $\frac{log2}{r}$ and the average time to divide, $\frac{1}{r}$, are uniform throughout the population.

Age- and stage-structured population modeling is a well-developed field within mathematical biology with over forty years of results. Age structure occurs when the probability of an important life event, like budding or death, depends on the age of the individual. In the most extreme case, the event occurs at a specific age. See [7] for a classical review of age-structured population modeling, including references for early works on age-structured modeling of *Saccharomyces cerevisiae*. Stage structure occurs when individuals pass through distinct life stages which are characterized by stage-specific events or event timing. Models with stage- and age-structure are often used to study cellular populations due to the existence of distinct cell cycle phases [7, 12, 11].

In recent years, the timing and variability of cell stage events has been a topic of intense interest. Previously, variability in the times between division and budding (G1) and budding and division in *Saccharomyces cerevisiae* was studied using single-cell imaging of strains expressing Myo1 tagged with green fluorescent protein [13]. This tagging results in the formation of a green ring at the location of a new bud which disappears at cytokinesis, in other words, when the bud separates from the mother cell. Results showed that G1 is more variable in daughter cells than mother

cells due to noisy size control, and that variability in G1 time in both mothers and daughters is dominated by molecular noise.

Although images were taken frequently (every 3 minutes) in [13] z-stacks were not used. Hence, errors in budding ring counts due to small rings, rings near the top of the mother cell, or crowding are present in this work (See section S1.C of [13]). These sources of error impact our data as well. We will discuss how modeling can be used to quantify, compensate, and limit these types of error in the following sections.

Our methods also rely on estimates of the rate at which small buds increase in size. When quantifying the growth of individual cells through time, the authors of [13] found that bi-linear and exponential models are equally accurate, implying a linear model might be appropriate for describing the growth of small buds. In this work, we also found that a linear model provides a good fit to the growth of the radius of a small bud.

CHAPTER 2

BIOLOGICAL METHODS

2.1 Biological Methods Related to Infection

Macrophages were plated at a density of $4.0x10^5$ cells in 35 mm glass-bottom dishes containing $2mL$ of growth medium approximately 24 hours prior to infection. For infection, macrophages were co-incubated with H99S-GFP *Cryptococcus neoformans* that had been opsonized with 18B7 anti-GXM antibodies and complement to enhance phagocytosis. The multiplicity of infection was 0.25-1 to 0.5:1, meaning there was between 0.25-0.5 yeast to every macrophage.

A low multiplicity of infection improved the accuracy of the counting process. For example, low multiplicity of infection limited the number of macrophages which were infected so there was more likely to be spatial separation between infected cells. If we use too many *Cryptococcus neoformans* and have many infected cells in the culture, this could make it hard to discriminate between adjacent cells, especially when time intervals are large.

After two hours of co-incubation, residual extracellular yeast were removed by washing. Fresh growth media was added containing two fluorescent stains, calcofluor white and propidium iodide. The concentration of both the calcoflour white and the propidium iodide was 7.5μg/mL. Calcofluor white, which is a blue fluorescent dye, labels the cell wall of extracellular yeast. It is not membrane permeable, which is why it only labels extracellular yeast and not intracellular yeast. Propidium iodide is a red fluorescent DNA intercalating dye which is also not membrane permeable. Both live extracellular in intracellular yeast which were living expressed EGFP, which labeled live *C. neoformans* with green fluorescence. The green color seen in Figures 1, 2 and 3 comes from EGFP proteins. The yeast are modified to express this EGFP gene. When yeast transitioned from exhibiting green fluorescence to no green fluorescence, it most

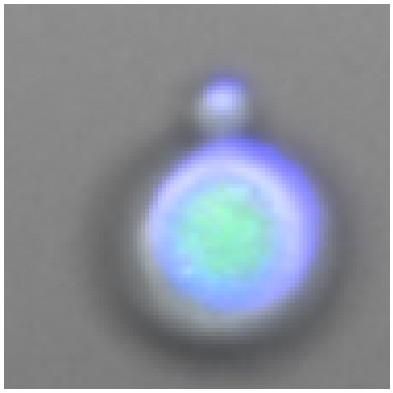

Figure 2: *Cryptococcus neoformans Extracellular Budding*

likely meant that that yeast was dead, having been killed by the host macrophage.

2.2 Biological Methods Related to Counting

Counts of yeast per macrophage were recorded manually. Because it is difficult to determine if a bud is detached, all visible cells, including buds, were counted as individual yeast cells. (Figure 3 illustrates this potential difficulty of assessing if a bud is detached.) In addition, the following criteria were used to determine if individual macrophages should be included in the study.

In order to be included in the study:

- Macrophages must be completely inside the field of view, and the full thickness of the macrophage must be captured within the z-stack at the start of counting. If a macrophage violates this rule at a later time, counting stops, but the count data collected up until the frame at which the rule is violated is included in the subsequent data analysis.

- All *C. neoformans* within the macrophage must be completely in frame at the start of counting.

- Macrophages must have just 1 or 2 intracellular (green) *C. neoformans* at the start of counting. If a *C. neoformans* cell is partially blue, indicating uncertainty as to its location, the cell is monitored through time to determine whether it is intracellular or extracellular.

After a macrophage was selected for inclusion in the study, its intracellular burden was recorded every thirty minutes until the count could no longer be confirmed. There are several reasons why a count might not be confirmed.

- A macrophage or yeast within the macrophage could experience cell death via lysis. (In our images, lytic cells appear red.)
- The yeast within the macrophage might be too numerous and closely packed to to be clearly distinguished as individual cells.
- A yeast could appear to be exocytosed by its host macrophage. This is called non-lytic expulsion.
- An intracellular yeast could appear to be obscured by another yeast or macrophage.
- A macrophage or intracellular yeast could become out of focus or move outside the field of view.
- A macrophage could appear to donate one of its intracellular yeast cells to another macrophage.

When a count could no longer be confirmed, counting stopped.The data recorded up until that time was included in the subsequent analysis. Due to variability in the fate of individual macrophages, the duration of time over which the intracellular fungal burden was recorded varied between macrophages.

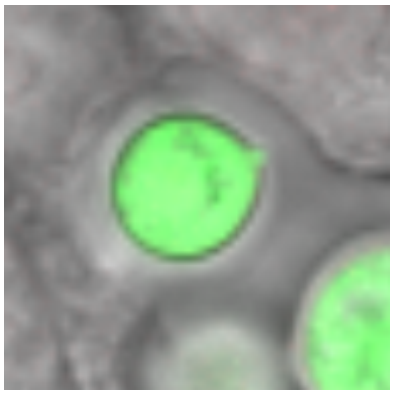

Figure 3: *Small Intracellular Cryptococcus neoformans Budding in FLAM Macrophage*

CHAPTER 3

ANALYTIC MODEL

3.1 Model Parameters

In this section, we describe the model parameters and their means of estimation. For convenience, the parameters and their symbols are listed below.

Model Parameters

- R_L: radius of the smallest observable disc
- z: depth of a z-stack
- λ: growth rate of a small cell
- $M(0,a),\ B(0,a)$: initial distributions of cell ages
- T_b: inter-budding time
- T_m: maturation time

The initial distribution of cell ages, inter-budding time, and maturation time are fit to the cell count data by minimizing the negative likelihood of the data given the parameters. The remaining parameters, R_L, and z, and λ are directly estimated from the microscope settings or images as described below.

The parameter R_L is the radius of the smallest visible disc. Hence, R_L is limited by the resolution of the microscope, which is the smallest distance over which distinct objects can be discriminated [10]. The lateral resolution (R_l), or resolution in the x, y-plane parallel to the specimen, is generally better than the axial resolution (R_a), or resolution along the z-axis perpendicular to the specimen. In this work we suppose that $R_L = R_l$. In order to confirm that this parameterization is reasonable we found

some very small buds and measured their radius. These measurements were near the lateral resolution of the microscope, so $R_L = R_l$ seems a reasonable approximation of the minimal radius of an observable bud.

The theoretical lateral resolution can be calculated using the following formula [10]:

$$R_l := \frac{.61\mu}{\text{NA}}, \tag{1}$$

where μ is the wavelength of emission light and NA is the numerical aperture of the objective lens [10]. In our case $\mu = 518$ nm and NA $= 1.4$, so the lateral resolution is 0.2257 μm.

The theoretical axial resolution can be calculated with the following formula:

$$R_a := 2\frac{n\mu}{\text{NA}^2}, \tag{2}$$

where n is the diffraction index of the specimen. In our case, n is the diffraction index of the cell culture media, which is approximated as that of water, 1.33 [9, 8]. This gives an axial resolution of 0.703 μm. Since this resolution is much smaller than the distance between z-slices (2 μm), very little light is mislocated to the wrong z-slice. For this reason, the impact of a limited axial resolution is not included in our model.

3.1.1 Calculating the Growth Rate of Small Cells

The parameter λ is the growth rate of a small bud. To compute λ we use Zenlite software to estimate the radii of small cells through time and fit a linear model to the resulting curve. In detail, the radius of a visible bud cross section, as determined by EGFP, which is found in the cytoplasm and the signal used for assessing intracellular fungal burden, is first estimated by fitting a circle to the perimeter of cross section in Zenlite. The perimeter is fit by hand, and the radius of the fit circle is provided by Zenlite. When possible, the radius of the cross section of EGFP at two different z-slices is used to infer the actual radius of the EGFP of the bud. This can be done as follows.

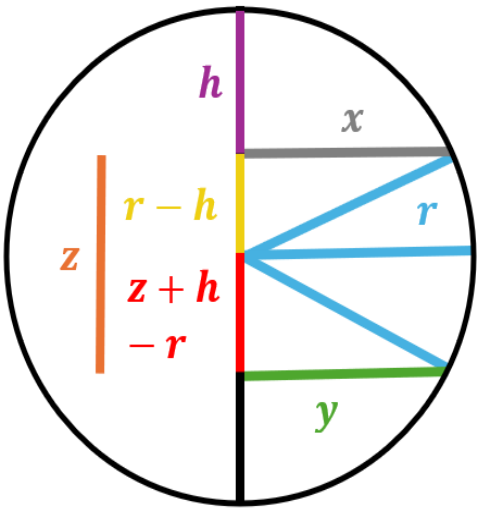

Figure 4: *Variables for Calculating Radius Given Two Slices*

Let x represent the radius of the top slice imaged by the microscope, and y represent the radius of the bottom slice imaged by the microscope. Since we are interested in assessing the growth rate of small buds, we only consider cells with at most two visible z-slices. This forces the top and bottom slice to be on opposite sides of the equator of the bud, as seen in Figure 4. Let h denote the distance from the top of the cell to the first slice, and r denote the radius of the bud. Note that the width of the z stack, z, is the distance between x and y.

Since x, y, and z are known from the image and microscope settings, our task is to determine r in terms of these variables. We first solve for the radius of the bud, r, as a function of h and x:

$$x^2 + (r-h)^2 = r^2 \longleftrightarrow \frac{x^2+h^2}{2h} = r \tag{3}$$

. Next, we solve for h in terms of x, y, and z:

$$(z+h-r)^2 + y^2 = r^2 \longleftrightarrow hz^2 + h(z^2+y^2-x^2) - x^2z = 0$$

.

$$h = \frac{-(z^2 + y^2 - x^2) + \sqrt{(z^2 + y^2 - x^2)^2 + 4x^2z^2}}{2z} \tag{4}$$

. Substituting (4) into (3) gives an expression for r in terms of x, y, and z.

After examining plots of the radii of several small buds through time, we selected a linear model of bud growth. The growth rate of several small buds was then fit to a linear model by minimizing the sum of the squared error between the model and the data in MATLAB. Figure 5 illustrates an example of how we fit the radius of a bud to a linear growth rate. Figure 6 displays the fit growth rate statistics for J774 cells and FLAMs.

At the time of writing, the growth rate had been estimated in only a small number of J774 macrophages, so we fit the model using a value 0.017 μm min^{-1}, which reflects the growth rate in FLAMS. In future work, we plan to measure λ in many more J774 macrophages, and use the results to define a range over which to fit λ to the cell growth data. Our decision to fit λ is based on the observations that measurements of lambda are subject to error and that the results of model fitting are sensitive to λ.

3.2 Stage- and age-structured population growth model

In this section we develop a continuous-time, stage- and age-structured model for population growth. The model describes the evolution of the density of buds, $B(t, a)$, and mothers, $M(t, a)$ through time where a represents the age of the cell and t represents time. Cells age continuously through time. When buds reach an age of T_m, they produce their own observable bud and become mother cells of age zero. When mothers reach an age of T_b, they produce a new observable bud, and their age resets to zero.

From these assumptions, we derive partial differential equations for the evolution of B and M.

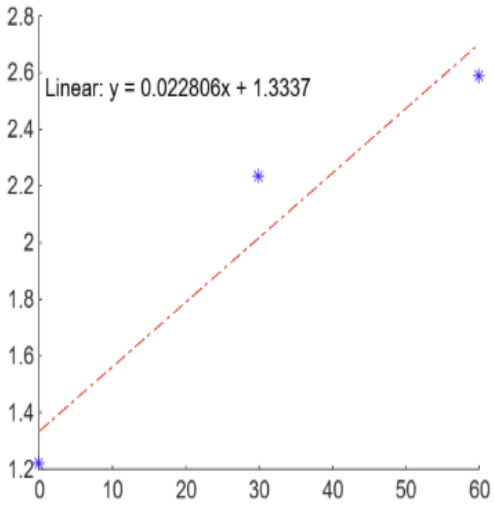

Figure 5: *Linear Fit*

Column	n	Mean	Variance	Std. dev.	Std. err.	Median	Range	Min	Max
J774s	3	0.0069194333	0.00001644567	0.0040553261	0.0023413436	0.0086508	0.0075359	0.0022858	0.0098217
FLAMs	18	0.029045333	0.0025205814	0.050205392	0.011833524	0.014537	0.217338	0.010722	0.22806

Figure 6: *Statistical Analysis on Slope Results*

Since cells age at the same rate time passes:

$$B(t+\delta, a) = B(t, \delta - a) \rightarrow B(t+\delta, a) - B(t, a) = B(t, \delta - a) - B(t, a)$$

.

Hence,

$$\lim_{\delta \to 0} \frac{B(t+\delta, a) - B(t, a)}{\delta} = \lim_{\delta \to 0} \frac{B(t, a-\delta) - B(t, a)}{\delta}.$$

In summary,

$$\frac{\partial B}{\partial t} = \frac{-\partial B}{\partial a} \tag{5}$$

.

Similarly,

$$\frac{\partial M}{\partial t} = \frac{-\partial M}{\partial a} \tag{6}$$

.

The boundary condition for B is

$$B(t, 0) = B(t, T_m) + M(t, T_b), \quad t > 0 \tag{7}$$

.

The boundary condition for M is

$$M(t, 0) = B(t, T_m) + M(t, T_b), \quad t > 0 \tag{8}$$

.

These boundary conditions can be interpreted biologically as follows. The density of new buds, $B(t, 0)$, is equal to the density of new mothers, $M(t, 0)$, and both are given by the sum of the mothers that just budded, $M(t, Tb)$, and the buds that just matured, $B(t, Tm)$.

The initial condition for B is

$$B(0, a) = B_0(a), \quad 0 < a < T_m, \tag{9}$$

where B_0 is given.

The initial condition for M is

$$M(0, a) = M_0(a), \quad 0 < a < T_m, \tag{10}$$

where M_0 is given.

This system of first order partial differential equation can be solved using the method of characteristics [4].

For $0 < t < a < T_m$, and $t < T_b$

$$B(t, a) = B_0(a - t). \tag{11}$$

For $0 < t < a < T_b$,

$$M(t, a) = M_0(a - t). \tag{12}$$

Defining $B_2(t) := B(t, T_m) + M(t, T_b)$ for $0 < t < T_b$ and $M_2(t) := B(t, T_m) + M(t, T_b)$ for $0 < t < T_b$, we have

For $0 < a < t < T_m$, and $t < T_b$

$$B(t, a) = B_2(t - a). \tag{13}$$

For $0 < a < t < T_b$,

$$M(t, a) = M_2(t - a). \tag{14}$$

The previous equations solve the system for B and M when $0 \leq t \leq T_b$ and $0 \leq a \leq T_m$, $0 \leq a \leq T_b$, respectively.

The process can be repeated to solve the system for all time, substituting $B(T_b, a)$ and $M(T_b, a)$ for $B_0(a)$ and $M_0(a)$, respectively.

3.2.1 Modifying the Empirical Population

Since our empirical cell population is finite, the density of buds and mothers is not continuous or differentiable. Instead, it is a sum of point-mass or Dirac delta distributions. Although this density does not satisfy any partial differential equation, since

cells age continuously as time passes, the solution is still constant along characteristics. This means the solution formula provided above via the characteristic equation is still valid. However, the boundary condition must be updated to avoid double-counting cells as they reach the age of maturity or budding counting. Letting $B(t,a)$ and $M(t,a)$ give the number of buds and mothers of age a at time t, respectively, the new boundary condition is as follows: For $0 < \delta < \min\{t, T_m, T_b\}$

$$\begin{aligned} B(t,0) &= B(t-\delta, Tm-\delta) + M(t-\delta, T_b-\delta) && (15)\\ M(t,0) &= B(t-\delta, T_m-\delta) + M(t-\delta, T_b-\delta) && (16)\\ M(t,T_b) &= 0 && (17)\\ B(t,T_m) &= 0, && (18) \end{aligned}$$

Note that the final condition avoids double counting point masses; when cells reach the age of budding, their age instantly resets to zero.

CHAPTER 4

NUMERICAL METHODS

4.1 Solving the Numerical Model

The cell population growth model is solved numerically in MATLAB. This requires discretization of the variables. We let time and age be measured in minutes and use a discrete time and age step size of $\delta = 0.1$ minutes for each variable.

In addition to the model variables described above, the numerical model makes use of the age at which a bud is 100 % observable, denoted by age_{ob}. Assuming a linear growth model, age_{ob} can be computed from the radius above which a bud is observable 100% of the time, R_{ob}. See section 5.1 for the computation of R_{ob} and section 3.2-3.3 for details on the linear growth model.

The MATLAB code assumes $T_m > T_b$. This makes sense in the biological context, since it can be determined from the empirical observations that it takes a bud longer to have its own bud than it takes for the mother to produce a new bud.

The model parameters, T_b, T_m and a_0 are then initialized as follows.

Initial guesses for T_b and T_m are selected. The initial guess for a_0, the initial distribution of cell ages, is computed from an auxiliary variable m_0, which is a vector encoding how far each cell counted at the initial time penetrates the age range defined by T_m and age_{ob}: For each cell counted at time zero, there is associated a single parameter $m_0(i)$, such that $0 \leq m_0(i) < 1$. These parameters are converted to ages as follows.

$$a_0 = m_0 T_m + \text{age}_{\text{ob}} \tag{19}$$

As written, the code assumes each initially visible cell is a bud, or a mature cell with a partially observable (though not visible) bud.

- If $0 < a_0(i) < T_m$, cell i is an immature bud of age $a_0(i)$.
- If $T_m \leq a_0(i) < Tm + \text{age}_{\text{ob}}$, cell i is a mother cell with a partially observable bud. In this case, both the mother cell and the bud have age $a_0(i) - Tm < \text{age}_{\text{ob}}$.

Although this method does not exactly account for mothers without buds, it effectively accounts for these cells since $T_b < T_m$. That is, as far as the model is concerned, a bud of age $T_m - x$ is equivalent to a mother of age $T_b - x$.

We next discretize all three initial parameters, T_m, T_b, and m_0 by rounding them to the nearest $\delta = 0.1$.

Assuming we want to solve the system up to time T_f, we split the system into time chunks of width T_b. The number of such chunks is T_f/T_b rounded up to the nearest integer.

Next, we determine how many time steps we need to traverse one time chunk. This is the steps per chunk: $\frac{T_b}{\delta}$ rounded up to the nearest integer. Therefore, the total number of time steps (tstepts) is the number of time chunks (Tsteps) multiplied by the steps per chunk (spc).

Next we create two vectors to capture the number of cells in each age class at the initial time. Define M_0 as a $1 \times (\text{spc} + 1)$ vector giving the number of mother cells in each age category at time zero, and define B_0 as a $1 \times (1 + Tm/\delta)$ vector giving the number of buds in each age category at time zero.

Next, Figures 7 and 8 demonstrate how we cycle through the time chunks to solve the population growth model. First, we will reference Figure 7. The diagram on the left of Figure 7 shows how we solve for the distribution of buds, and the diagram on the right of Figure 7 shows how we solve for the distribution of mothers.

The axis labeled B_0 corresponds to the distribution of buds at the initial time. This distribution is effectively shifted forward in time and rightward in age, to solve for $B(t, a)$ for $0 < t < a$ and $t < T_b$, using the fact that B is constant on the characteristic curves $a - t = c$. That is, $B(t, a) = B_0(a - t)$ in region R_2. The solution on one such characteristic curve is shown in purple in Figures 7 and 8. Note

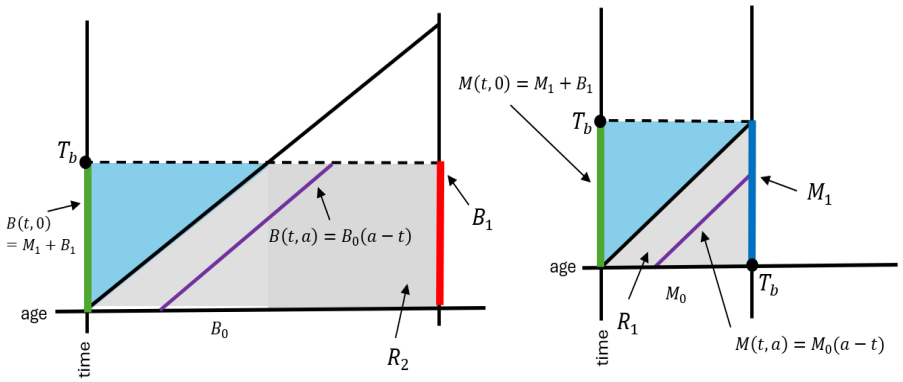

Figure 7: *Solving from Initial Data with the Characteristic Equation*

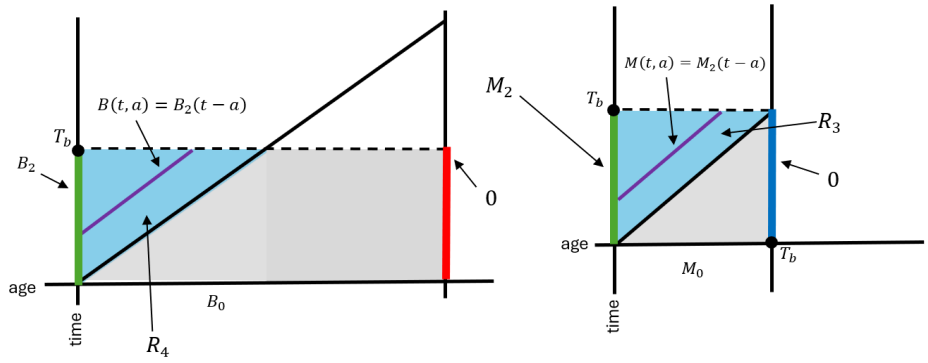

Figure 8: *Solving from Initial Data with the Characteristic Equation, Part Two*

that new cells enter only through the zero age class, or left-hand boundary in the diagram.

Similarly we solve for $M(t,a)$ for $0 < t < a$ and $t < T_b$, using the initial data M_0. That is $M(t,a) = M_0(a-t)$ in region R_1.

Through this process, we have now solved the system in regions $R1$ and $R2$ in Figure 7. The numbers of mature buds, $B(t,T_m)$, for $0 < t < T_b$ (represented by B_1 in Figure 7) and the numbers of budding mothers, $M(t,T_b)$, for $0 < t \leq T_b$ (represented by M_1 in Figure 7) are then used to compute the numbers of new buds and mothers with new buds for $0 < t \leq T_b : B(t,0) = M(t,T_b) + B(t,T_m) = M(t,0)$.

Now, looking at Figure 8, the right hand boundaries, which represent new buds and mothers with new buds and are set to zero to avoid double-counting cells. That is, new buds and mothers with new buds were assigned to the left boundaries in the previous step, so the right boundary is set to zero.

Next, the diagram on the left of Figure 8 shows how the distribution of new buds (denoted by B_2 on the vertical axis) is effectively shifted forward in time and rightward in age, to solve for $B(t,a)$ in R_4, where $0 < a < t < T_b$, using the fact that B is constant on the characteristic curves $t - a = c$. That is, $B(t,a) = B_2(t-a)$ in region R_4.

Similarly, the diagram on the right of Figure 8 shows how the distribution of new mothers (denoted by M_2 of the vertical axis) is effectively shifted upward in time and rightward in age, to solve for $M(t,a)$ in R_3, where $0 < a < t < T_b$, using the fact that M is constant on the characteristic curves $t - a = c$. That is, $M(t,a) = M_2(t-a)$ in region R_3.

Finally, the top boundaries of the diagrams represent the numbers of cells in all age classes at time T_b. The process described above is then repeated with $M(T_b,a)$ and $B(T_b,a)$ in place of B_0 and M_0.

CHAPTER 5

PROBABILITY

5.1 Probability a Cell is Observed Given its Radius

The counting process is subject to error in several ways. First, there are buds present at each time which cannot be seen with 100% certainty. Specifically, it can happen that the cross section of the bud that is captured by the microscope is too small to be seen even though the bud is large enough to generate an observable cross section. This would happen if the bud is sliced at too high or too low a point, resulting in a cross section with a radius that falls below the lateral resolution of the microscope. Second, observable cross sections might be missed due to human counting error. As seen by Figure 3, buds can be hard to see visually. Finally, it could happen that visual artifacts might be misclassified as cross sections of buds or mother cells. For example, these visual artifacts could be caused by *Cryptococcus neoformans* that are not imaged in the z-stack of interest.

As a result of these errors, even a perfect model might fail to match the cell count data. As such, we model the error in the counting process so that we can estimate the likelihood of the data, given the parameters.

First, we solve for the probability a cell is observed given its radius, R. We suppose there is a minimal visible radius, which is defined as R_L. A cross section with radius less than R_L cannot be seen, while one with radius greater than R_L is visible. Using R_L and the width of the z stack we compute the probability a bud is observable from the microscope image, given it has radius R and assuming a spherical bud shape.

First, we must consider the conditions under which a bud is visible. Let h be the distance between the top of the bud and the first slice of the bud, and denote the radius of the first slice by d. See Figure 9.

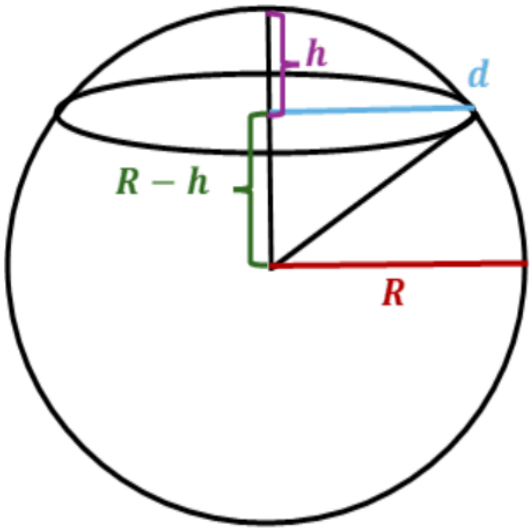

Figure 9: *Probability Variables*

Using the Pythagorean theorem, we can find $d^2 = 2Rh - h^2$. So,

$$d^2 > R_L^2 \longleftrightarrow 2Rh - h^2 > R_L^2 \longleftrightarrow 0 > h^2 - 2Rh + R_L^2. \tag{20}$$

The roots of this equation are then $h = \frac{1}{2}(2R \pm \sqrt{4R^2 - 4R_L^2}) = R \pm \sqrt{R^2 - R_L^2}$

So, for a slice with height h to be visible, $R - \sqrt{R^2 - R_L^2} < h < R$ must hold. As such, the width of the visible domain on the top of the sphere is $\sqrt{R^2 - R_L^2}$ and by symmetry the total width of the visible domain is $2\sqrt{R^2 - R_L^2}$. Thus, if $z > 2\sqrt{R^2 - R_L^2}$, the probability of observing the sphere is $\frac{1}{z}\sqrt{R^2 - R_L^2}$.

If $z \leq 2\sqrt{R^2 - R_L^2}$, there must be a slice in the visible domain, so the cell is 100% observable.

In summary, the probability of observing a cell given its radius is $R < \mathrm{R_{ob}}$ is

$$P_{ob}(R) = \frac{2\sqrt{R^2 - R_L^2}}{z} \tag{21}$$

where R_{ob} is determined by $z = 2\sqrt{R^2 - R_L^2}$, that is,

$$R_{ob} = \sqrt{\left(\frac{z}{2}\right)^2 + R_L^2}. \tag{22}$$

is the radius at which a cell is 100% observable.

We also include a model for counting error. This model assumes each cell with a visible cross section has a small chance of being missed. For simplicity, we suppose the probability of missing a visible cross-section (p_m) is not dependent on the radius of the cross section. Hence, the probability of counting k cells given $n \geq k$ cells are visible is:

$$P_c(k|n) = \binom{n}{k} p_m^{n-k}(1-p_m)^k, \tag{23}$$

and the probability of counting k cells given $n < k$ cells are visible is zero. That is $P_c(k|n) = 0$ for $k > n$. Hence, the probability of counting j cells is

$$f_c(j) := \sum_{k=0}^{N} P_c(j|k) f_v(k), \tag{24}$$

where $f_v(x)$ denotes the probability x cells were visible from the microscope image given the population state, and N denotes the current population size. The computation of f_v is described in the next section.

We also suppose there is a small chance of counting a cell that is not present, for example, due to visual artifacts. Specifically, we suppose the probability of counting n extra cells is r^n. Since the use of fluorescent stains greatly decreases the likelihood of seeing an artifact, r is selected so the probability of counting at least one artifact: $\frac{r}{1-r}$ is very small. Specifically, we set $\frac{r}{1-r} = 10^{-10}$. Finally, since r is very small, we approximate the probability of introducing more than two artifacts as zero. Note that the chance of introducing an artificial cell into the count is independent of the number of cells present in this model, so the probability of recording j cells given the population state is

$$f_r(j) = f_c(j)\frac{1-2r}{1-r} + f_c(j-1)r + f_c(j-2)r^2 \tag{25}$$

This probability mass function for the number of cells recorded (f_r) is then used to compute the likelihood of the cell count data given the state of the model population.

5.2 Computing the probability a cell is visible

In this section we describe how to compute the probability mass function for the number of visible cells, given the state of the model population.

Let the state of the discretized population at the i^{th} observation time be described by the vectors $B(i,j)$ and $M(i,j)$, so that $B(i,j)$ and $M(i,j)$ denote the numbers of buds and mothers, respectively of age $j\delta$ after $t(i) = 30 \times i$ minutes. Let

$$N(i) = \sum_j (M(i,j) + B(i,j))$$

be the size of the population at time $t(i)$.

First, we need to find the number of cells that are 100% observable at each observation time given the parameters. These cells are guaranteed to be visible. Recall that age_{ob} is the age of the smallest cell that is always observable. Hence, $ind_{ob} = \frac{\text{age}_{\text{ob}}}{\delta}$ (rounded up to the nearest integer) is the age index of the smallest observable bud.

Given that all mother cells are large enough to be observable 100% of the time, the number of large (100% observable cells) at each time step is the number of mother cells plus the number of buds that are 100% observable, that is, those with age index greater than or equal to ind_{ob}. Denote this number of cells as NLC(i)

Next we compute the probability that a small cell, that is, one with an age index less than ind_{ob} is visible: Assuming small cells grow linearly with growth rate λ, the radius of a cell with age index $j < \text{ind}_{\text{ob}}$ is

$$R_j = R_L + j\lambda\delta,$$

and the probability the cell is visible is $P_v(R_j)$, as in equation (21). Hence the number of buds that are visible from the j^{th} age class follows a binomial distribution, and the probability of observing x buds from this age class is

$$f_j(x) = \binom{B(i,j)}{x} (P_v(R_j))^x (1 - P_v(R_j))^{B(i,j)-x} \tag{26}$$

The probability mass function for the number of small cells visible is the convolution of the probability mass functions for each small age class:

$$\hat{f}_v(x) = (f_1 * \ldots * f_{ind_{ob}-1})(x), \tag{27}$$

where $*$ denotes the convolution,

$$f_1 * f_2(x) := \sum_{i=0}^{x} f_1(i) f_2(x-i). \tag{28}$$

Finally, the probability mass function for the number of cells that are visible at time $t(i)$ from the entire population is

$$f_v(x) = \hat{f}_v(x - NLC(i)), \tag{29}$$

since $NLC(i)$ cells are 100% observable.

CHAPTER 6

FITTING THE MODEL AND INITIAL RESULTS

6.1 Fitting the Parameters

To fit the model population growth model to the data, we use the probabilistic model described in the previous section to minimize the negative likelihood of the data.

This process is as follows. First, the cell count data is entered as a vector, N, which gives the number of cells counted at each observation time. Next the minimal observable radius, R_L, width between z-slices, z, and linear cell growth rate, λ are estimated from the microscope settings and images. R_L and z are then used to compute the radius at which a cell is visible 100% of the time, R_{ob}. Next, R_{ob} and λ are used to find the age at which a bud is visible 100% of the time, a_{ob}, where an age of zero corresponds to a radius of R_L: Recall $R_{ob} = \frac{1}{2}((2R_L)^2 + (z)^2)^{\frac{1}{2}})$. Assuming linear growth, we compute the age of observation according to $R_{ob} = R_L + \lambda a_{ob}$. Solving this equation for a_{ob} yields $a_{ob} = \frac{(R_{ob}-R_L)}{\lambda}$. As parameterize, $R_{ob} = 1.02$ μm, and $a_{ob} = 47.3$ min.

Because the likelihood function is continuous but not differentiable, we use fminsearch (MATLAB) with several initial guesses to find the optimal parameters. The results shown in Figures 11 and 12 were generated from a mesh of initial guesses for T_b, T_m and m_0, where all possible initial combinations with $m_0 \in \{0.6, 0.8\}$, $T_b \in \{0.8(99.3), 99.3, 1.2(99.3)\}$, and $T_m \in \{2.2(99.3), 2.4(99.3)\}$ where considered. Since fminsearch does not allow for parameter constraints, the initial parameter guesses are transformed as $l_1 = log(m0./(1 - m0))$, $l_2 = log(T_b - a_{ob})$ and $l_3 = log(Tm - Tb)$. These transformations are reversed on completion of the optimization routine to find the corresponding values of the parameters while meeting the constraints, $0 \leq m_0 \leq 1$, $0 \leq T_b, T_m$, $T_b \leq T_m$. One run of the optimization algorithm is shown in Figure 10.

Fitting the model to intracellular replication that occurs in a single J774 macrophage,

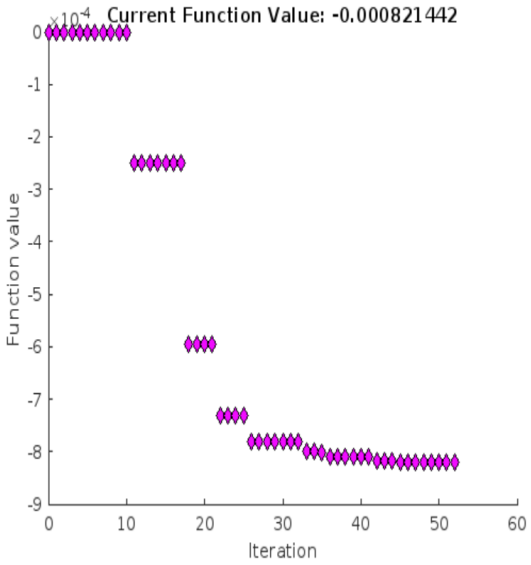

Figure 10: *Minimizing Negative Likelihood*

we generated the following results. First, after running the code with an initial guess for T_m, T_b, and m_0, the following values were obtained: $T_m = 183$ min, $T_b = 78$ min, $m_0 = 0.823$.

6.2 Initial Results

Entering the parameters above into the population growth model, we generate Figures 11 and 12, which illustrate the results of the model compared to the empirical data it was fit to. The graph in Figure 11 shows the actual count and number of cells predicted by the model at each observation time. Figure 12 shows the count along with the predicted number of cells in the population at each point in time.

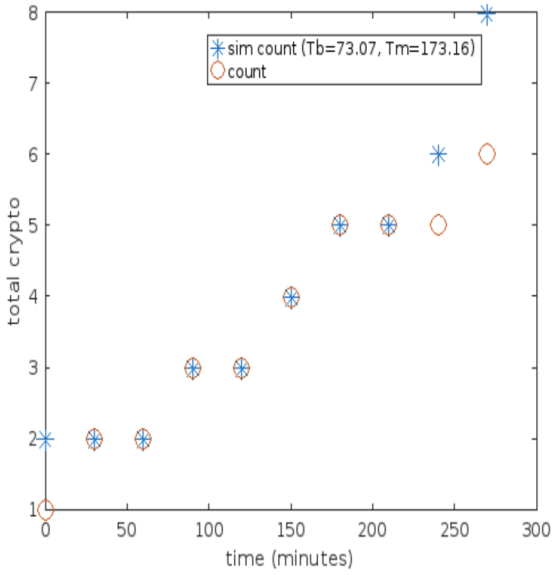

Figure 11: *Initial Model Results Compared to Empirical Data*

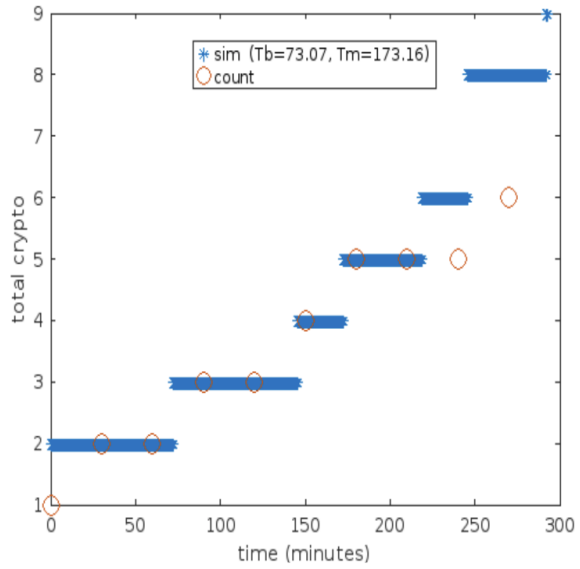

Figure 12: *Initial Model Results Extended Through Time*

CHAPTER 7

CONCLUSION

In conclusion, using the data provided by our biological collaborators, we created a model for describing the population growth of *Cryptococcus neoformans* in macrophages. The population growth model we developed is exactly solvable via the method of characteristic equations. This population growth model was coupled with a probabilistic model to account for errors in the counting process. This process allowed us to fit the model to the data by maximizing the likelihood of the data given the model parameters.

After fitting the parameters to the data associated with one J774 macrophage, we were able to generate initial results that fit closely with the given data. In the future, we aim to fit this model to multiple J774 macrophages and also to FLAM macrophages in order to compare the results for the two contrasting macrophages.

Despite these positive results, it is important to note that our modeling process does not account for the full complexity of the biological system. As discussed in Section 2.2.1, additional cell fates including cell death and non-lytic expulsion can sometimes occur. These events are not captured by our model. Although the counting rules are aimed at eliminating cells that experience these events from the counts, we cannot be completely sure that such events did not occur due to the limited temporal resolution of the data. Hence, adding additional cell fates to the model would eliminate one potential source of error and allow us to collect larger, more inclusive, data sets.

There are also additional sources of error that are not included in our model. For example, if a bud emerges near the top or bottom of a mother cell, a visible cross section of the bud could be mistaken for a cross section of the mother. In this case, the bud might not be counted even though it is visible. We did not account for this potential error in the current version of our model, but we plan on adding it in the

future.

There are some changes to the imaging process that could be adopted to reduce error. For example, as mentioned above, the current microscopy settings and growth rate of small buds means that a cell is unreliably visible for up to 47 minutes, or two observation times. Adjusting the width between z-slices to 1 μm would reduce this time to just over 19 minutes, so that a bud is unreliably visible for at most one observation time. We can also increase the accuracy of the parameter estimation by increasing the number of observations. This can be accomplished by either increasing the number of cells observed or decreasing the intervals between observations. While these two strategies are distinct, we expect them to yield similar improvements in parameter estimation; because the cells are not synchronized, increasing the number of cells observed effectively samples different points of the cell cycle. Nevertheless, observing additional cells is likely somewhat less effective than decreasing the time between observations because the distribution of ages within the population is unlikely to be uniform, meaning some ages are unlikely to be sampled. However, given that reducing the time between observations requires us to reduce either the number of z-slices or the number of fields of view, and that frequent observations can induce cell death via phototoxicity, increasing the number of cells observed is probably the most effective means of improving the accuracy of the parameter estimates.

In future work, we will fit the model to more individual J774 macrophages and analyze population statistics. Then, we will fit the model to multiple FLAMs and analyze population statistics. Comparing these population-level statistics will help us determine if intracellular replication of *Cryptococcus neoformans* differs between these two macrophage types. We also plan to fit both populations to an exponential growth model and compare the results to the model of asymmetrical growth. This will help us quantify the improvement that is achieved by accounting for the population's age and stage structure.

It will also be interesting to consider identifiability of the model parameters given

the data. While distinct choices of the model parameters generate distinct trajectories, it is possible for two choices of the parameters to generate the same visible cell counts at the discrete observation times. However, the probabilistic model of error in the count recorded means that parameterizations which generate the same visible cell counts can nevertheless have different likelihoods. Thus, the probabilistic error model improves parameter identifiability. As the number of observations goes up, it becomes increasingly unlikely that two distinct parameterizations have the same likelihood, however, we have yet to quantify conditions which guarantee parameter identifiability. Hence, this is another direction for future research.

Finally, in future work it would be interesting to adapt the model to include additional cell fates and intrinsic stochasticity in cell growth or proliferation.

BIBLIOGRAPHY

[1] C. Bogdan, U. Schleicher, *Production of interferon-γ by myeloid cells – fact or fancy?*, Trends in Immunology, Vol. 27 (2006), p 282-290, [https://doi.org/10.1016/j.it.2006.04.004.]

[2] C. Coelho et. al. "Macrophage mitochondrial and stress response to ingestion of [Cryptococcus neoformans]" J Immunol(2015) [https://pubmed.ncbi.nlm.nih.gov/25646306/]

[3] Comprehensive Cancer Information - NCI. (n.d.).*NCI Dictionary of Cancer* terms.https://www.cancer.gov/publications/dictionaries/cancer-terms/def/lysis

[4] L. C. Evans. *Partial Differential Equations*. American Mathematical Society, 2022.

[5] R. de S. A. Glauber, C. de L. Alcantara , R. Noêmia , W. de Souza, B. Pontes, Frases Susana *Ultrastructural Study of Cryptococcus neoformans Surface During Budding Events*. Frontiers in Microbiology, vol 12 (2021) [https://www.frontiersin.org/journals/microbiology/articles/10.3389/fmicb.2021.609244]

[6] P.K. Mada, R.T. Jamil, M. U. Alam. *Cryptococcus*. StatPearls Publishing, 2025.

[7] Metz and Diekman (Ed.) "The Dynamics of Physiologically Structured Populations" Lecture Notes in Biomathematics, vol 68 (1986). Springer-Verlang.

[8] T. Parvin,K. Ahmed, A. Alatwi, A. Rashed, "Differential optical absorption spectroscopy-based refractive index sensor for cancer cell detection". Vol 28 (2021).

[9] D. Sage, L. Donati, F. Soulez, D. Fortun, G. Schmit, A. Seitz, R. Guiet, C. Vonesch, M. Unser, *DeconvolutionLab2: An open-source software for deconvolution microscopy, Methods*, Vol. 115 (2017), p 28-41.

[10] J. C. Waters; "Accuracy and precision in quantitative fluorescence microscopy." , J Cell Biol, 7(2009), p.185.

[11] A. Zilman, V. Ganusov, and A. Perelson, "Stochastic models of lymphocyte proliferation and death," PLoS One, vol. 5, p. e12775, 2010.

[12] Gabriel, Pierre and Garbett, Shawn P and Quaranta, Vito and Tyson, Darren R and Webb, Glenn F, '"The contribution of age structure to cell population responses to targeted therapeutics," Journal of Theoretical Biology, vol. 311, p. 19–27, 2012.

[13] S. Talia, J. Skotheim, J. Bean, et al. "The effects of molecular noise and size control on variability in the budding yeast cell cycle". Nature 448, 947–951 (2007). [https://doi.org/10.1038/nature06072]

[14] S. Thomas,L. Ankley,K. Conner, A. Rapp,A. McGee, F. LeSage, C. Tanner, T. Vielma,E. Scheeres, J. Obar,A. Olive. "TGFβ primes alveolar-like macrophages to induce type I IFN following TLR2 activation". bioRxiv [Preprint]. (2024) [https://pmc.ncbi.nlm.nih.gov/articles/PMC11398362/]

[15] S. Tucker and A. Casadevall. "Replication of Cryptococcus Neoformans in Macrophages Is Accompanied by Phagosomal Permeabilization and Accumulation of Vesicles Containing Polysaccharide in the Cytoplasm." Proceedings of the National Academy of Sciences of the United States of America, vol. 99 (2002), pp. 3165–70.