Journal of Mathematical Biology manuscript No.
(will be inserted by the editor)

Vijay G. Subramanian · Ken R. Duffy · Marian L. Turner · Philip D. Hodgkin

Determining the expected variability of immune responses using the Cyton Model

12th September 2007

Keywords Immune response; expected variability; continuous time branching processes; time dependent offspring distributions

Mathematics Subject Classification (2000) 60J85, 92D25

Abstract During an adaptive immune response, lymphocytes proliferate for five to twenty cell divisions, then stop and die over a period of weeks. The cyton model for regulation of lymphocyte proliferation and survival was introduced by Hawkins *et al.* [17] to provide a framework for understanding this response and its regulation. The model assumes stochastic values for division and survival times for each cell in a responding population. Experimental evidence indicates that the choice of times is drawn from a skewed distribution such as the lognormal, with the fate of individual cells being potentially highly variable. For this reason we calculate the higher moments of the model so that the expected variability can be determined. To do this we formulate a new analytic framework for the cyton model by introducing a generalization to the Bellman-Harris branching process.

We use this framework to introduce two distinct approaches to predicting variability in the immune response to a mitogenic signal. The first method enables explicit calculations for certain distributions and qualitatively exhibits the full range of observed immune responses. The second approach does not facilitate analytic solutions, but allows simple numerical schemes for distributions for which there is little prospect of analytic formulae.

We compare the predictions derived from the second method to experimentally observed lymphocyte population sizes from *in vivo* and *in vitro* experiments. The model predictions for both data sets are remarkably accurate. The important biological conclusion is that there is limited variation around the expected value of the population size irrespective of whether the response is mediated by small numbers of cells undergoing many divisions or for many cells pursuing a small number of divisions. Therefore, we conclude the immune response is robust and predictable despite the potential for great variability in the experience of each individual cell.

Vijay Subramanian and Ken Duffy, Hamilton Institute, National University of Ireland, Maynooth, Ireland
E-mail: Vijay.Subramanian@nuim.ie · Marian Turner and Philip Hodgkin, Walter & Eliza Hall Institute of Medical Research, Parkville, Melbourne, Australia

1 Introduction

When T and B lymphocytes are exposed to a stimulus such as an antigen, they respond by increasing their population size through a series of cell divisions. Typically this expansion does not proceed indefinitely and once the response is complete, the population size decreases through apoptosis with a small proportion of the maximum cell number retained as a memory population [4][18]. Understanding and predicting the strength of this response and its regulation by extrinsic signals is a key goal in immunology.

Advances in flow cytometry and the introduction of techniques for following cell division by labeling with the fluorescent dye carboxyfluorescein succinimidyl ester (CFSE) have enabled the collection of detailed experimental data on the kinetics of lymphocyte division progression and cell survival that form the basis for developing quantitative models. For example, a series of recent papers [15][29][11][17] has reported on extensive *in vitro* experiments on the behavior of purified naïve lymphocytes exposed to a mitogenic stimulus. These studies indicate that the time taken for the first cell division is typically longer than subsequent divisions and varies considerably between cells within a homogeneous population. As a consequence of this variability, asynchrony in division progression develops within the population. The appropriate distribution for entry to division closely follows a lognormal for both T and B lymphocyte stimulation [11][29][17]. Subsequent divisions are typically less variable in time, although this also varies with stimulation conditions. Strongly stimulated T cell proliferation contributes little additional variation to the population structure and can be modeled as a deterministic division time [15]. However, stimulation with lower concentrations of growth supporting cytokines reveals the variance in time of passage through subsequent divisions increases and now contributes significantly to the developing division asynchrony of the population [11].

Regulation of cell death is also a feature of lymphocyte responses. When cells are placed in culture, in the absence of a mitogenic stimulus, they gradually die by apoptosis [15][11][17]. The addition of a mitogenic stimulus does not prevent cells dying in this manner in the time period prior to entry to division [15][11][17]. Once cells are dividing the survival properties of cells remain regulable. T cells receiving strong stimulus divide with little death, whereas weaker levels of growth stimulus lead to progressively more cells being lost with each generation [11][13][19]. Finally, it is clear that the number of divisions cells undergo can be regulated. *In vivo* studies following T cell responses to a viral infection suggest at least fifteen to twenty divisions, following which cells die by apoptosis [10]. This behavior on a smaller scale can be observed *in vitro* by removing the stimulus. When the stimulus is removed, some cells continue to divide for a period of time, but then ultimately all cells die by apoptosis [17]. Thus, the challenge for modeling these apparently complex processes is to find appropriate descriptions for the simultaneous operation of cell death and cell division, the inheritance of division and death parameters and the cessation of proliferation and subsequent death of the population.

A number of recent papers propose mathematical models for lymphocyte proliferation that capture some of the above features and include a generational structure to enable fitting to CFSE division tracking data. A common approach is to use Ordinary Differential Equations (ODEs) for each generation incorporating exponential waiting times for part of the cell cycle as the source of divisional variation. These models have also included simultaneous death rates modeled with ODEs, therefore leading to exponential cell loss functions [31][14][9].

An alternative to the simple ODE approach was proposed by León *et al.* [22] who provide a general framework for calculating the mean population per generation as a function of time, assuming a stochastic time to cell division. The analysis suggests using skewed distributions for time between cell divisions. This framework, however, does not incorporate terms for cell death.

The stochastic Cyton Model introduced by Hawkins *et al.* [17] addresses the requirement for randomness in each division by postulating independent control of times to divide and die in each cell. The combination of the two time-based controllers in each cell was called the cyton. With this structure the actual distribution for division and death times can be varied and, in particular, parameter values obtained from experiments can be used. They also introduced the concept of "division destiny" to the modeling framework. A cell's destiny is considered as the number of times it will divide before stopping. In [17] the cyton model is used to predict the average lymphocyte population size as a function of time, based on a small number of measured parameters. They demonstrate the validity of their model hypotheses through its ability to match experimental observations for cell growth both *in vitro* and *in vivo*.

The question of variability in the overall immune response is an important outstanding issue for the stochastic cyton model that was not addressed by Hawkins *et al.* [17]. Here we provide an advanced mathematical analysis of the cyton model. For example, our new results enable the calculation of higher moments that provide a refinement to the law of large numbers and yield strong prediction regarding variability in the immune response. Our important biological conclusion is that despite the large number of different experiences of individual cells, the immune response is predicted to be strongly concentrated around the average behavior.

2 Overview

As the lymphocyte population ultimately dies out because of division destiny, we are interested in studying the transient behavior of the population size rather than its behavior at large times. We provide two new methodologies for studying the cyton model. The first is based on a modification of traditional branching process techniques, while the latter is a hybrid of those techniques and a generation-based approach. For an introduction to branching processes, see one of the standard texts, e.g. [16][5][21]. Both methods are based on the deduction of the probability generating function for the population size as a function of time. As well as being able to determine the mean population size at any given time, we determine the higher moments, such as variance, of the population distribution. Once the probability generating function is known, the population distribution can be recovered using numerical inversion techniques such as those described in [2][3][8][20] and Section 1.2.4 of [1].

Our reasons for presenting two approaches are that the first allows an exact solution for certain distributions. This gives expressions from which the range of qualitative behaviors can be observed in an explicit fashion. The second approach lends itself to simple numerics for general distributions, but is unlikely to lead to explicit expressions. In particular, this is the case for the lognormal distributions that provide an excellent fit to available time to division data for T and B lymphocytes [11][29][17]. For numerical solutions the second approach is particularly well suited to the setting where there is a minimum time between cell divisions or division destiny ensures a bounded number of divisions, as a series of integrals in the framework converges after a finite number of iterations. A minimum time between cell divisions has been reported since the 1960s and '70s [26][24][28]. It is thought to correspond to the minimum time cells take to progress through the S and G_2/M phases of the cell cycle. We provide an example that treats these distributions using the second methodology and compare them with data taken from *in vitro* and *in vivo* experiments.

The rest of this paper is organized as follows. In Section 3 we introduce the assumptions of the cyton model and the mathematical framework to study it. In Section 4 we introduce an analysis based on total cell population, where division destiny is treated as a *random time*. We analyse it using a modification of the traditional branching processes approach. In Section 5 we present an example for

which the mean and variance of the population size can be determined explicitly and illustrate its features. In Section 6 we give a second analysis based on the number of cells that are present which are the product of a given number of divisions. Division destiny here is treated as a *random number of divisions*. We analyse this system using an iterative scheme. In Section 7 we use the proposed numerical scheme for lognormal distributions. We match the model to *in vitro* experimental data presented here for the first time and *in vivo* data taken from [18], presenting a new contribution made possible by our approach: we show that, despite the nature of the lognormal distribution, standard deviation error bars are small. That is, despite the inherent variability of an individual cell's behavior, the overall immune response is highly predictable.

3 The Cyton Model

Experimental observations of immune response have culminated in an article by Hawkins *et al.* [17] that draws conclusions about the nature of the mechanisms at work. Here we summarize some of their deductions.

- Each cell has a personal time to die and each mitogenic stimulus activates the cells and imposes a time to divide (Hawkins emphet al. propose the word *cyton* for this unit; a word we adopt). These times vary from cell to cell within the population. It appears that the two timed processes within each cell are unaware of each other and whichever outcome is reached first determines the fate of that cell.
- The cytons in distinct cells operate independently.
- The cyton in each initial cell has different distributions for time to die and time to division than their progeny. After the first division, cells have the same cyton parameters.
- There is no evidence for a significant inheritance of division times [7][12][27].
- Each initial cell has a division destiny. It will divide at most a given number of times (or until a given time)[30], after which its progeny can no longer divide. This division destiny can differ for each initial cell.

Based on these observations Hawkins *et al.* suggest the following working hypotheses for a stochastic model of the behavior of a lymphocyte population exposed to a mitogenic stimulus.

- Each cell's cyton is stochastically independent.
- The operation of the regulable cyton controlling division and survival, seen leading up to the first division, is repeated through subsequent divisions.
- Individual cells will, upon division, "erase" the values of their parents time to divide and time to die, and adopt new values drawn from the appropriate distributions.
- There is a division destiny for each initial cell in the starting population that is drawn from a stochastic distribution.

For the purpose of the modeling in this paper, we abstract the cyton mechanism as two stochastic clocks. Notionally, a clock is a timer that once started will trigger after a random time. We are not suggesting that there is necessarily equivalent time sensing machinery operating within the cell.

Matching the hypotheses, we assume that every cell sets two stochastically independent clocks: a time to division clock represented by the random variable T_B and a time to death clock represented by the random variable T_D. Each cell exists until the minimum of T_B and T_D. If the time to death clock expires first, the cell dies and does not divide. If the time to division clock expires first, then the cell dies, but progeny are born. Mathematically, we can readily treat a general distribution for the number

of progeny, but as this is not relevant for cell division we restrict cells to either two progeny or none. We assume that progeny set independent clocks at the time of their division and follow the same process, albeit with potentially different distributions. If both clocks take the same value, $T_B = T_D$, we assume that death occurs. To avoid pathologies, throughout we assume $P(T_B = 0) = 0$.

In terms familiar to the literature in continuous time Branching processes, T_B and T_D define a pair of random variables

$$L = \min(T_B, T_D) \qquad \text{and} \qquad \zeta = \begin{cases} 2 \text{ if } T_B < T_D \\ 0 \text{ if } T_B \geq T_D, \end{cases}$$

where L is the life-time and ζ is the number of offspring. For this to lead to a Bellman-Harris process, L and ζ should be independent. However, apart from exceptional cases, independence of T_B and T_D does not lead to independence of L and ζ. For example, assume that T_B is uniformly distributed on $[0, 2]$ and that T_D is uniformly distributed on $[0, 1]$. Then $P(\zeta = 2) = P(T_B < T_D) = 1/4$, $P(\zeta = 0) = P(T_D \leq T_B) = 3/4$ and $P(L \leq t) = 1 - P(T_B > t, T_D > t) = 1 - P(T_B > t)P(T_D > t) = 1 - (1 - t)(2 - t)$ for $t \in [0, 1]$. However, $P(\zeta = 2, L \leq t) = P(T_B < T_D, T_B \leq t) = \int_0^t P(T_D > r) dP(T_B = r) = 1/2 - t^2/4 \neq P(\zeta = 2)P(L \leq t)$. That is, L and ζ are not independent. Thus the cyton model leads to a generalization of the Bellman-Harris process in which life-time and number of progeny are not necessarily independent.

Let $Z(t)$ denote the total number of cells alive at time t. We consider the probability generating function $F(s, t) := E(s^{Z(t)})$ and its generation based equivalents. If, for given t, this is finite for s in any interval (s^-, s^+) that contains 1, then $F(s, t)$ completely specifies the distribution of $Z(t)$ (e.g., pg. 278 of [6]). Once $F(s, t)$ is known, the moments of $Z(t)$ can then be determined by taking partial derivatives of $F(s, t)$ with respect to s. For example, the mean population size at time t is

$$m(t) := E(Z(t)) = \frac{\partial F(s, t)}{\partial s}\Big|_{s=1}$$

and, defining

$$v(t) := E(Z(t)^2) - E(Z(t)) = \frac{\partial^2 F(s, t)}{\partial s^2}\Big|_{s=1},$$

the variance of the population size at time t is

$$\text{var}(t) := E(Z(t)^2) - E(Z(t))^2 = v(t) + m(t) - m(t)^2.$$

Our first approach will determine $F(s, t)$. We define 0 to be the generation of the initial cells. For any $k > 0$, we define the k^{th} generation to be cells that appear as a consequence of division from the $(k - 1)^{\text{th}}$ generation. The second approach involves first determining the equivalent of $F(s, t)$ for the number of cells alive in each generation as a function of time.

Assume that initially there are d cells, $Z(0) = d$. Let $X_i(t)$ denote the number of descendents of cell i that are alive at time t. The number of cells alive in the population is the sum of the number of progeny alive from each of the $i = \{1, \ldots, d\}$ initial cells: $Z(t) = \sum_{i=1}^d X_i(t)$. As we assume probabilistic independence in the clock times of each cell, the probability generating function satisfies $E(s^{Z(t)}) = \prod_{i=1}^d E(s^{X_i(t)})$. Moreover, as the clocks are identically distributed for different initial cells, $E(s^{Z(t)}) = E(s^{X_1(t)})^d$. Thus we need only to consider the probability generating function for the progeny of an individual cell, as the probability generating function for an ensemble of d cells is just its d^{th} power. With an initial population of d cells, the mean number of cells at time t is $d\, m(t)$ and the standard deviation is $\sqrt{d\, \text{var}(t)}$.

4 An analytical framework for the cyton model

For certain clock distributions the following analytic framework allows us to explicitly determine the mean and variance of the number of lymphocytes alive as a function of time. This enables us to understand the contribution of each part of the cyton model mechanism and qualitatively determine the range of possible immune responses. However, it is not possible to give explicit expressions for the lognormal distributions that appear experimentally and therefore we do not have an analytic formula to compare with observations. In Section 6 we will introduce a numerical framework that can deal with any class of distributions, and in particular lognormal distributions. It enables comparison with experiment, but the contribution of each part of the cyton mechanism is not as transparent.

We set up and solve the model successively in three steps, each refining the previous one. In Section 4.1 we assume that the time to division random variable is independent and identically distributed for each cell and that the time to death clock is independent and identically distributed for each cell. Furthermore, we assume there is no division destiny. This enables us to deduce the behavior of the cells from the first generation after the initial cells until the division destiny is reached. In Section 4.2 we then show how to take into account the observation that the initial batch of cells appear to have different clock distributions. Finally, in Section 4.3 we include division destiny. We do this by assuming that each initial cell has a *random time* after which its progeny can no longer divide. At this time, each living descendent of a given initial cell picks a new random time to die. Selecting this final time to die from yet another distribution causes no extra mathematical difficulty.

4.1 Homogeneous clock distributions, from the first generation to division destiny

We use the superscript H to indicate we are assuming homogeneous clock distributions. Consider a single cell at time 0 whose time to division, T_B^H, and time to death, T_D^H, clocks have just been initiated. In general we use the following integral equation to specify the generating function of the population size distribution:

$$F^H(s,t) = sP(T_B^H > t, T_D^H > t) + P(T_D^H \leq t, T_B^H \geq T_D^H) + \int_0^t F^H(s,t-r)^2 dP(T_D^H > r, T_B^H = r) \tag{1}$$

$$= sP(T_B^H > t)P(T_D^H > t) + P(T_D^H \leq t, T_B^H \geq T_D^H) + \int_0^t F^H(s,t-r)^2 P(T_D^H > r) dP(T_B^H = r), \tag{2}$$

where the second line follows from independence of birth and death clocks. This arises from three parts:

1. if the first cell lives at time t (which happens if both T_B^H and T_D^H are greater than t) then $Z(t) = 1$ and contribution to the expectation $E(s^{Z(t)})$ is s;
2. if the first cell dies at or before t (which happens if $T_D^H \leq t$ and $T_B^H \geq T_D^H$), then $Z(t) = 0$ and contribution to the expectation is 1;
3. finally, if the first cell divides at time r in $[0,t]$, then we have two cells for whom the same calculation needs to be performed for the remainder of time $t-r$. Due to probabilistic independence of progeny and the identical distributions of their clocks, we get the $F^H(s,t-r)^2$ term.

It is most natural to write equation (2) as a nonlinear Volterra equation (see, for example, Polyanin and Manzhirov [25]):

$$F^H(s,t) = sP(T_B^H > t)P(T_D^H > t) + P(T_D^H \le t, T_B^H \ge T_D^H) \\ - \int_0^t F^H(s,r)^2 P(T_D^H > t-r) dP(T_B^H = t-r). \tag{3}$$

We will return to the solution of equation (3) for $F^H(s,t)$ in Section 5.

Comment: note that even to determine the average number of cells alive at time t, we would also have to solve an equation of the same form as equation (3):

$$m^H(t) = \frac{\partial}{\partial s} F^H(s,t)|_{s=1} = P(T_B^H > t)P(T_D^H > t) - 2\int_0^t m^H(r) P(T_D^H > t-r) dP(T_B^H = t-r).$$

An integral equation with similar structure holds for any higher moment.

4.2 Adding different distributions for initial cells

If the initial cell has different random variables (distributions) for time to division T_B^M and time to death T_D^M to later cells (whose random variables are T_B^H and T_D^H), then we first solve equation (3) to determine $F^H(s,t)$. The probability generating function $F^M(s,t)$ for the total population at time t, without division destiny, is then determined by the following integral equation, which follows the same logic as equation (2):

$$F^M(s,t) = sP(T_B^M > t)P(T_D^M > t) + P(T_D^M \le t, T_B^M \ge T_D^M) \\ - \int_0^t F^H(s,r)^2 P(T_D^M > t-r) dP(T_B^M = t-r). \tag{4}$$

Having determined $F^H(s,t)$, then $F^M(s,t)$ is determined by an integral. Let $m^M(t)$ and $\mathrm{var}^M(t)$ denote the mean and variance of the cell population when the initial cells have this different distribution. We also have the following direct integral representations for the mean

$$m^M(t) = P(T_B^M > t)P(T_D^M > t) - 2\int_0^t m^H(r) P(T_D^M > t-r) dP(T_B^M = t-r) \tag{5}$$

and for the variance

$$\mathrm{var}^M(t) = m^M(t) - m^M(t)^2 - 2\int_0^t (\mathrm{var}^H(r) - m^H(r) + 2m^H(r)^2) P(T_D^M > t-r) dP(T_B^M = t-r). \tag{6}$$

4.3 Incorporating division destiny

The final complication we must consider is division destiny. As this varies within the population of initial cells, we model it as a stochastic quantity. At the random time T^* we hypothesize that all cells in the population that are descendent from a given initial cell draw a new time of death from a given distribution T_D^{final} independently of what has happened to them in the past; their new time to division is set to be $+\infty$. After T^*, they can only die.

In this case, $F^M(s,t)$ obeys the integral equation (4) for all $t \leq T^*$. As the death clocks are drawn independently from a new distribution at time T^*, we have that, conditioned on $T^* = t^*$, for all $t \geq t^*$:

$$\begin{aligned} E\left(s^{Z(t)}|T^*=t^*\right) &= E\left(E\left(s^{Z(t)}|Z(t^*),T^*=t^*\right)\right) \\ &= E\left(\left(sP(T_D^{\text{final}} > t-t^*) + P(T_D^{\text{final}} \leq t-t^*)\right)^{Z(t^*)}|T^*=t^*\right). \end{aligned}$$

Thus conditioned on $T^* = t^*$, we have

$$E\left(s^{Z(t)}|T^*=t^*\right) = \begin{cases} F^M(s,t) & \text{if } t \leq t^*, \\ F^M(sP(T_D^{\text{final}} > t-t^*) + P(T_D^{\text{final}} \leq t-t^*), t^*) & \text{if } t \geq t^*. \end{cases}$$

To get the probability generating function $F(s,t)$ for the total population alive at time t (including stimulus removal and division destiny), we must take the expectation over the distribution of T^* values:

$$\begin{aligned} F(s,t) &= \int_0^\infty E\left(s^{Z(t)}|T^*=t^*\right) dP(T^*=t^*) \\ &= \int_0^\infty F^M(s,t) 1_{\{t\leq t^*\}}\, dP(T^*=t^*) \\ &\quad + \int_0^\infty F^M\left(sP(T_D^{\text{final}} > t-t^*) + P(T_D^{\text{final}} \leq t-t^*), t^*\right) 1_{\{t\geq t^*\}} dP(T^*=t^*). \end{aligned}$$

That is, the probability generating function for the population size (including division destiny) is

$$\begin{aligned} F(s,t) &= F^M(s,t)P(T^* \geq t) \\ &\quad + \int_0^t F^M\left(1+(s-1)P(T_D^{\text{final}} > t-t^*), t^*\right) dP(T^*=t^*). \end{aligned} \tag{7}$$

Taking partial derivatives with respect to s one gets that the average number of cells alive at time t is also given by an integral equation:

$$m(t) = m^M(t)P(T^* \geq t) + \int_0^t m^M(t^*)P(T_D^{\text{final}} > t-t^*) dP(T^*=t^*) \tag{8}$$

and the variance at time t is

$$\begin{aligned} \text{var}(t) &= m(t)(1-m(t)) \\ &\quad + (\text{var}^M(t) - m^M(t)(1-m^M(t)))P(T^* \geq t) \\ &\quad + \int_0^t (\text{var}^M(t^*) - m^M(t^*)(1-m^M(t^*)))P(T_D^{\text{final}} > t-t^*)^2 dP(T^*=t^*). \end{aligned} \tag{9}$$

4.4 Analytic cyton model framework summary

At time $t = 0$ we have d cells whose time to division and time to die random variables are T_B^M and T_D^M. For each initial cell a random division destiny time is selected according to the random variable T^*. After T^* any living progeny of the cell cannot further divide; they select new time to die T_D^{final}. Any cells alive after the initial batch and before T^* have clocks distributed as T_B^H and T_D^H.

The probability generating function for the total population is $E(s^{Z(t)}) = F(s,t)^d$, where $F(s,t)$ is determined by first solving equation (3) for $F^H(s,t)$, then the integral (4) for $F^M(s,t)$ prior to T^* and finally the integral (7).

5 Analytic example

We consider a particular example of the modeling approach in Section 4 where explicit calculations are possible. As described in the framework, we tackle the example by first considering the homogeneous population, then adding the effect of different clock distributions for the initial cells and finally adding division destiny.

The final analytic equations are (18), (19) and (20) in section 5.3.

5.1 The homogeneous population; solving equation (3)

Assume that the time to division is exponentially distributed, $P(T_B^H > t) = e^{-\lambda_B t}$, and the time to death is exponentially distributed $P(T_D^H > t) = e^{-\lambda_D t}$. In this case each cell's life time random variable L, the time to whichever of division or death occurs first, has distribution $P(L \leq t) = 1 - P(T_B^H > t, T_D^H > t) = 1 - \exp(-(\lambda_B + \lambda_D)t)$. The number of children ζ at the end of its lifetime satisfies $P(\zeta = 0) = \lambda_D/(\lambda_B + \lambda_D)$ and $P(\zeta = 2) = \lambda_B/(\lambda_B + \lambda_D)$. As $P(\zeta = 0, L \leq t) = P(T_D^H < T_B^H, T_D^H \leq t) = \int_0^t P(T_B^H > r) dP(T_D^H = r) = \lambda_D/(\lambda_B + \lambda_D)(1 - \exp(-(\lambda_B + \lambda_D)t)) = P(\zeta = 0)P(L \leq t)$, the random variables ζ and L are independent. Thus this forms a classic Bellman-Harris process and the homogeneous part of the population falls within the standard branching process methodology.

We have that $P(T_D^H \leq t, T_B^H \geq T_D^H) = \lambda_D/(\lambda_B + \lambda_D)(1 - e^{-(\lambda_B+\lambda_D)t})$ and $P(T_B^H < t, T_D^H > T_B^H) = \lambda_B/(\lambda_B + \lambda_D)(1 - e^{-(\lambda_B+\lambda_D)t})$. Thus equation (3) becomes the following nonlinear integral equation

$$F^H(s,t) = \left(s - \frac{\lambda_D}{\lambda_B + \lambda_D}\right) e^{-(\lambda_B+\lambda_D)t} + \frac{\lambda_D}{\lambda_B + \lambda_D} + \lambda_B \int_0^t F^H(s,r)^2 e^{-(\lambda_D+\lambda_B)(t-r)} dr. \qquad (10)$$

For fixed s, equation (10) is of the form

$$y(t) = g(t) - \int_0^t e^{\lambda(t-r)} f(r, y(r)) dr.$$

It is known (e.g. [25]) that its solution is given by the solution of the following first-order ODE:

$$\frac{dg(t)}{dt} = \frac{dy(t)}{dt} + f(t, y(t)) - \lambda y(t) + \lambda g(t).$$

Here the equation that results is

$$y(0) = F(s,0) = E(s^{Z(0)}) = s; \qquad \frac{dy(t)}{dt} = (\lambda_B y(t) - \lambda_D)(y(t) - 1).$$

Solving this differential equation gives us $F^H(s,t) = y(t)$. If $\lambda_B \neq \lambda_D$, then the generating function in equation (3) is

$$F^H(s,t) = \frac{\lambda_D(s-1)e^{(\lambda_B-\lambda_D)t} - s\lambda_B + \lambda_D}{\lambda_B(s-1)e^{(\lambda_B-\lambda_D)t} - s\lambda_B + \lambda_D}. \tag{11}$$

From this we can deduce that in the absence of division destiny, the mean population as a function of time is

$$m^H(t) = e^{(\lambda_B-\lambda_D)t} \tag{12}$$

and that the variance of the of the population at time t is

$$\mathrm{var}^H(t) = \left(e^{2(\lambda_B-\lambda_D)t} - e^{(\lambda_B-\lambda_D)t}\right)\left(\frac{2\lambda_B(\lambda_B-\lambda_D)}{(\lambda_B-\lambda_D)^2} - 1\right). \tag{13}$$

When $\lambda_B > \lambda_D$, so that on average the division clock expires before the death clock, there is exponential growth in both the average population size and the variance in population size as a function of time.

When $\lambda_B < \lambda_D$, the mean population decays exponentially, but there can be transient behavior in the variance where it first increases, before decreasing.

In experiments, it is typically the case that there is an initial growth in the population size followed by population decay through cell death. This initial behavior corresponds to the mean division time being smaller than the mean death time, but in the absence of a mechanism such as division destiny, this would lead to an undamped exponential growth in the population size.

When $\lambda_B = \lambda_D$, so that the time to divide and time to death clocks have the same statistics, equation (3) gives

$$F^H(s,t) = 1 - \frac{1-s}{(1-s)\lambda_B t + 1}, \tag{14}$$

the mean population is $m^H(t) = 1$, but the variance grows as $\mathrm{var}^H(t) = 2\lambda_B t$.

5.2 Adding different initial distributions; integrating equation (4)

In [17] it is reported that a proportion of cells in the initial culture do not respond to the stimulus. This can be explained as a consequence of trauma incurred in their harvesting and/or intrinsic insensitivities to the stimulus. Hawkins *et al.* introduce a factor pF_0 that represents the fraction of initial cells that do respond to the stimulus.

Here we assume that independently with probability $1 - p \in [0,1]$ each cell in the initial culture cannot divide. Thus our parameter p corresponds to pF_0. For cells that can divide, their probability generating function is $F^H(s,t)$ prior to division destiny. This defined in equation (11) if $\lambda_B \neq \lambda_D$ and in equation (14) if $\lambda_B = \lambda_D$.

Hence if $\lambda_B \neq \lambda_D$ the mean is

$$m^M(t) = (1-p)e^{-\lambda_D t} + pe^{(\lambda_B - \lambda_D)t} \tag{15}$$

and the variance is

$$\begin{aligned}\mathrm{var}^M(t) = p\left(e^{2(\lambda_B-\lambda_D)t} - e^{(\lambda_B-\lambda_D)t}\right)\left(\frac{2\lambda_B(\lambda_B-\lambda_D)}{(\lambda_B-\lambda_D)^2}\right) \\ + pe^{(\lambda_B-\lambda_D)t} - p^2e^{2(\lambda_B-\lambda_D)t} + (1-p)e^{-\lambda_D t}\left(1-(1-p)e^{-\lambda_D t} - 2pe^{(\lambda_B-\lambda_D)t}\right).\end{aligned} \tag{16}$$

If $\lambda_B = \lambda_D$, then $m^M(t) = (1-p)\exp(-\lambda_B t) + p$ and $\mathrm{var}^M(t) = 2p\lambda_B t + m^M(t) - m^M(t)^2$.

5.3 Including division destiny time; solving equation (7)

We consider two division time distributions: (i) exponential, $P(T^* > t) = \exp(-\lambda t)$; and (ii) Weibull, $P(T^* > t) = \exp(-t^2)$. With the former, we will show that for certain parameterizations it is possible that division destiny and apoptosis *cannot* control the rate of cell division. Indeed it is possible, for example, that the mean population grows to infinity or that the mean tends to a constant while the variance tends to infinity. Neither of these is observed experimentally. This suggests using a division destiny distribution whose tail decays faster than exponentially. Based on experiments it has been suggested that division destiny is similar to the Normal distribution [17]. We will use a Weibull distribution that has the same tail behavior as the Normal distribution, but yields elegant formulae. In this case, the population distribution is well behaved as time becomes large for all parameterizations of the division and death clocks, with division destiny ultimately ensuring the population dies.

Exponentially distributed division destiny. Let T^* be exponentially distributed with rate $\lambda \neq \lambda_B$ and assume that T_D^{final} is equal in distribution to T_D^H. Equation (7) gives

$$\begin{aligned}F(s,t) = p\left(F^H(s,t)e^{-\lambda t} + \int_0^t F^H\left((s-1)e^{-\lambda_D(t-t^*)}+1, t^*\right)\lambda e^{-\lambda t^*}dt^*\right) \\ + (1-p)\left(1+(s-1)e^{-\lambda_D t}\right),\end{aligned}$$

but the second term cannot be explicitly integrated. However, using equations (8) and (9) in conjunction with equations (15) and (16) we can determine the mean, $m(t)$, and variance, $\mathrm{var}(t)$, explicitly.

For $\lambda_B \neq \lambda_D$,

$$m(t) = \begin{cases}(1-p)e^{-\lambda_D t} + \frac{pe^{-\lambda_D t}}{\lambda-\lambda_B}\left(\lambda - \lambda_B e^{(\lambda_B-\lambda)t}\right) & \text{if } \lambda_B \neq \lambda, \\ e^{-\lambda_D t}(1+p\lambda t) & \text{if } \lambda_B = \lambda.\end{cases} \tag{17}$$

At large times one of three things happens: (i) if $\lambda + \lambda_D > \lambda_B$, then a combination of natural deaths and division destiny controls the average population which decreases to zero as t becomes large; (ii) if $\lambda + \lambda_D = \lambda_B$, there is perfect balance, on average, between division, death and division destiny so that the average population tends to a constant $m(t) \to p\lambda_B/(\lambda_B - \lambda)$; and (iii) if $\lambda + \lambda_D < \lambda_B$, death and division destiny does not control the division rate and on average the population grows to infinity, $m(t) \to \infty$.

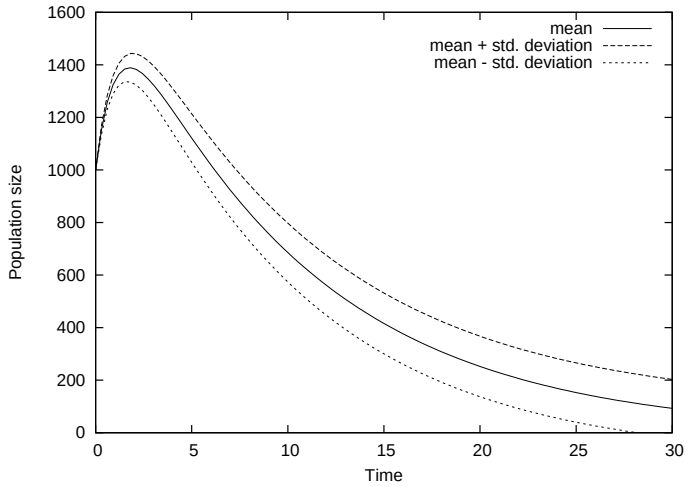

Fig. 1 One thousand lymphocytes in initial culture, with first cell survival probability $p = 0.7$, homogeneous clocks and exponential division destiny parameters $\lambda_B = 1$, $\lambda_D = 0.1$ and $\lambda = 1.81$. Analytic predictions.

When $\lambda_B \neq \lambda_D$ the variance is:

$$\begin{aligned}\operatorname{var}(t) &= m(t)(1-m(t)) \\ &\quad + pCe^{-\lambda t}\left(e^{2(\lambda_B-\lambda_D)t} - e^{(\lambda_B-\lambda_D)t}\right) + pC\lambda e^{-2\lambda_D t}\left(\frac{1-e^{(2\lambda_B-\lambda)t}}{\lambda-2\lambda_B} - \frac{1-e^{(\lambda_B+\lambda_D-\lambda)t}}{\lambda-\lambda_B-\lambda_D}\right),\end{aligned}$$

where $m(t)$ is defined in equation (17),

$$C := \frac{2\lambda_B(\lambda_B-\lambda_D)}{(\lambda_B-\lambda_D)^2}.$$

To determine $\operatorname{var}(t)$ when $2\lambda_B = \lambda$ or $\lambda_B + \lambda_D = \lambda$, it suffices to take limits.

The variance's realms of divergence are different to those for $m(t)$. Again, one of three things happens: (i) if $\lambda + 2\lambda_D > 2\lambda_B$, then the combination of death and division destiny controls the variance which tends 0 as $t \to \infty$; (ii) if $\lambda + 2\lambda_D = 2\lambda_B$, $\operatorname{var}(t) \to pC$; and (iii) if $\lambda + 2\lambda_D < 2\lambda_B$, then the variance diverges in time $\operatorname{var}(t) \to \infty$. Thus as $m(t)$ tends to a constant when $\lambda + \lambda_D = \lambda_B$, the variance is diverging to infinity for this parameterization.

Figure 1 plots $d\,m(t) \pm \sqrt{d\,\operatorname{var}(t)}$ for a set of parameters such that there is an initial explosion in the average population, but division destiny ultimately controls growth and the population ultimately dies out.

When division and death rates are balanced, $\lambda_B = \lambda_D$, with exponentially distributed division destiny, we have

$$m(t) = pe^{-\lambda t} + (1-p)e^{-(\lambda_B+\lambda)t} + (1-p)(1-e^{-\lambda t})e^{-\lambda_B t} + \frac{\lambda p e^{-\lambda_B t}(1-e^{-(\lambda-\lambda_B)t})}{\lambda-\lambda_B}$$

and if $\lambda \neq 2\lambda_B$

$$\operatorname{var}(t) = m(t)(1-m(t)) + 2p\lambda_B e^{-2\lambda_B t} + \frac{2p\lambda\lambda_B}{(\lambda-2\lambda_B)^2}\left(e^{-\lambda t}\left((2\lambda_B-\lambda)t-1\right) + e^{-2\lambda_B}\right).$$

For any $\lambda \in (0, \infty)$, regardless of the values of λ_B, both $m(t)$ and $\mathrm{var}(t)$ tend to 0 as t becomes large, as division and death are already balanced without division destiny.

Weibull distributed destiny division time. As an exponentially distributed division destiny cannot guarantee control of the average population size at large times for all clock parameterizations, we consider a division destiny whose tail decays faster than exponential; this is observed experimentally.

If T^* is distributed as $P(T^* > t) = \exp(-t^2)$, which is a Weibull distribution, then again it is not possible to give a closed form for $F(s,t)$, but for any value of λ_B and λ_D we can use equations (8) and (9) to determine the mean and variance:

$$m(t) = (1-p)e^{-\lambda_D t} + pe^{-\lambda_D t} + pe^{-\lambda_D t} e^{\lambda_B^2/4} \frac{\lambda_B}{2}\sqrt{\pi}\left(\mathrm{erf}\left(\frac{\lambda_B}{2}\right) - \mathrm{erf}\left(\frac{\lambda_B}{2} - t\right)\right), \tag{18}$$

where erf is the error function. Here, irrespective of the value of λ_B, division destiny controls the mean population size so that $m(t) \to 0$. When $\lambda_B \neq \lambda_D$ the variance is

$$\begin{aligned}\mathrm{var}(t) = m(t)(1-m(t)) + e^{-2\lambda_D t}\frac{2p\lambda_B(\lambda_D-\lambda_B)\sqrt{\pi}}{(\lambda_D-\lambda_B)^2} \\ \left(\lambda_B e^{\lambda_B^2}\left(\mathrm{erf}(\lambda_B) - \mathrm{erf}(\lambda_B - t)\right) - \frac{\lambda_B+\lambda_D}{2} e^{\frac{(\lambda_B+\lambda_D)^2}{4}}\left(\mathrm{erf}\left(\frac{\lambda_B+\lambda_D}{2}\right) - \mathrm{erf}\left(\frac{\lambda_B+\lambda_D}{2} - t\right)\right)\right).\end{aligned} \tag{19}$$

and when $\lambda_B = \lambda_D$

$$\begin{aligned}\mathrm{var}(t) = m(t)(1-m(t)) + 2p\lambda_B e^{-t^2}\left(1 - e^{2\lambda_B t} + \lambda_B e^{-2\lambda_B t}(1 - e^{-2\lambda_B t})\right) \\ + p\lambda_B e^{-2\lambda_B t}(1+\lambda_B^2)\sqrt{\pi}\left(\mathrm{erf}(\lambda_B)e^{-t^2} - \mathrm{erf}(\lambda_B - t)\right).\end{aligned} \tag{20}$$

With an initial population of $d = 10,000$ lymphocytes, Figure 2 shows $d\,m(t)$ and $d\,m(t) \pm \sqrt{d\mathrm{var}(t)}$ with $p = 0.7$, $\lambda_B = 4$ and $\lambda_D = 0.2$. The shape shown here is qualitatively close to that observed experimentally [15][29][11][17], including the dramatic initial ramp-up in population size, followed by a smooth transition to apoptosis for all cells. Using equations (19) and (20) it can be seen how the stochastic components of death, division and destiny combine to effect the predicted immune response.

6 A numerical framework for the cyton model

The framework in Section 4 is a time-driven view of the overall lymphocyte population distribution and is capable of giving closed form solutions for certain classes of distributions. We now develop an alternate view that emphasizes the contribution of different generations. The framework presented here is designed for numerical methods. We will use the model in Section 7 with lognormal clock distributions that have been proposed as providing an excellent fit to observations [11][29][17], for which there is little prospect of obtaining analytic formulae.

The framework is particularly well suited to the setting where division destiny ensures there can only be a finite number of generations or when the time to division distribution is bounded from below. This time is usually taken to be the time all cells must take to traverse the S and G_2/M phases of the cell cycle [26][24][28]. It is not yet known if there is a minimum time spent in the earlier $G1$ phase. If this minimum time is known, it can be used as a model input.

In this framework we allow different division and death clock distributions from generation to generation. As in the analytic model, we first consider the model without division destiny and then add

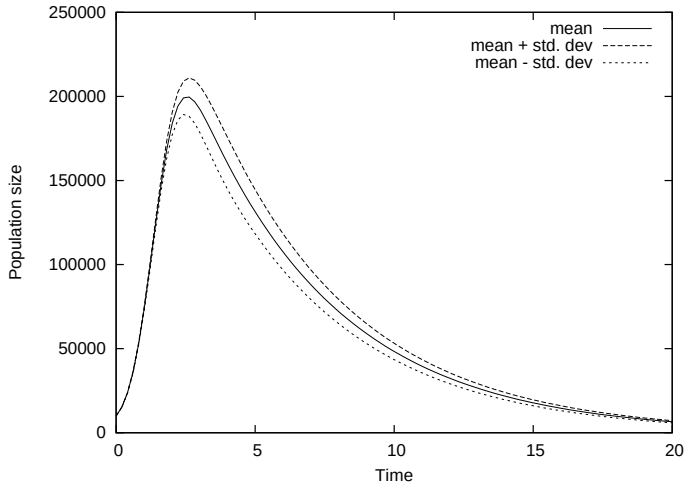

Fig. 2 Ten thousand lymphocytes in initial culture, with first cell survival probability $p = 0.7$. Weibull division destiny. Analytic predictions.

it. In Section 6.2 we show how in principle $F^H(s,t)$ in Section 4.1 can be recovered from the method presented here. In Section 6.3 we present an alternate method for deriving the equations presented in [17] that govern the mean population size evolution. In Section 6.4 we take into account division destiny. We do this by assuming that each initial cell's division destiny is specified as a *random number* of generations; its progeny reach at most a given (random) generation and then cannot divide further. After the final division is reached, the living descendent picks a new random time to die.

6.1 The model

Let $Z^k(t)$ denote the number of cells of the k^{th} generation alive at time t, where k takes values in the non-negative integers. As the clocks from generation to generation are independent, we consider the population generated by one cell of the k^{th} generation. We define the probability generating function over all of the generations from k to l at time t starting out with just one cell of the k^{th} generation at time 0:

$$F_k^l(s,t) := E\left(s^{\sum_{i=k}^{l} Z^i(t)} | Z^k(0) = 1,\ Z^m(0) = 0 \text{ for all } m \neq k\right).$$

If $k > l$, then we define $F_k^l(s,t) := 1$. In the model of Section 4 we were interested in $Z(t) = \sum_{i=0}^{\infty} Z^i(t)$, where $Z^0(0) := 1$. Thus $F^H(s,t) = F_0^{\infty}(s,t)$ under the assumption of homogeneous distributions. We will return to this point in the next section.

Let (T_D^k, T_B^k) denote the time to death and time to division random variables for the k^{th} generation. We have the following relationships

$$\begin{aligned}
F_{l+1}^l(s,t) &:= 1 \\
F_l^l(s,t) &= sP(T_D^l > t)P(T_B^l > t) + 1 - P(T_D^l > t)P(T_B^l > t) \\
F_{l-1}^l(s,t) &= sP(T_D^{l-1} > t)P(T_B^{l-1} > t) + P(T_D^{l-1} \le t, T_D^{l-1} \le T_B) \\
&\quad + \int_0^t \left(1 + (s-1)P(T_D^l > t-r, T_B^l > t-r)\right)^2 P(T_D^{l-1} > r) dP(T_B^{l-1} = r) \\
&\vdots \quad \vdots
\end{aligned}$$

These yields the following iteration scheme: for all $k \in \{0, \ldots, l\}$,

$$F_{l+1}^l(s,t) = 1 \tag{21}$$

$$\begin{aligned}
F_k^l(s,t) &= sP(T_D^k > t)P(T_B^k > t) + P(T_D^k \le t, T_D^k \le T_B^k) \\
&\quad + \int_0^t \left(F_{k+1}^l(s, t-r)\right)^2 P(T_D^k > r) dP(T_B^k = r).
\end{aligned} \tag{22}$$

This is explained as follows:

1. if the k^{th} generation cell lives at time t, then the contribution is s;
2. if the k^{th} generation cell dies at or before time t, then the contribution is 1;
3. and if the k^{th} generation cell divides into 2 cells of the $(k+1)^{\text{th}}$ generation at some time $r \in [0,t]$, then we consider the contribution of the populations generated by these 2 cells in the remaining time $t-r$;
4. we do not count contributions from any generation beyond l.

6.2 Recovering $F^H(s,t)$

When the clock distributions are the same for all generations, the iterative procedure in equations (21)–(22) gives an alternate method to calculate $F^H(s,t)$ in equation (3): $F^H(s,t) = \lim_{k\to\infty} F_0^k(s,t)$. Using this characterization we can reverse the labeling to write a simpler procedure for computing $F^H(s,t)$ in this homogeneous case. To calculate $F^H(s,t)$, it is sufficient to execute the following iteration scheme:

$$\begin{aligned}
\tilde{F}_{-1}(s,t) &= 1 \\
\tilde{F}_0(s,t) &= 1 + (s-1)P(T_D^H > t)P(T_B^H > t) \\
\tilde{F}_k(s,t) &= sP(T_D^H > t)P(T_B^H > t) + P(T_D^H \le t, T_D^H \le T_B^H) + \int_0^t \tilde{F}_{k-1}^2(s, t-r)P(T_D^H > r) dP(T_B^H = r),
\end{aligned}$$

where $k \ge 1$. Then $F^H(s,t) = \lim_{k\to\infty} \tilde{F}_k(s,t)$.

6.3 An alternate mean population size analysis

Define $m_k^l(t)$ to be the mean size of the population encompassing generations k to l descended from one cell of the k^{th} generation starting at time 0:

$$m_k^l(t) = \frac{\partial}{\partial s} F_k^l(s,t)|_{s=1}.$$

Taking partial derivatives in s and then setting $s = 1$ in (21)–(22) the iterative procedure for the mean is, for all $k \in \{0, \ldots, l\}$,

$$\begin{aligned} m_{l+1}^l(s,t) &= 0 \\ m_k^l(t) &= P(T_D^k > t)P(T_B^k > t) + 2\int_0^t m_{k+1}^l(t-r)P(T_D^k > r)dP(T_B^k = r), \end{aligned}$$

where $m_k^l(t) := 0$ if $k > l$. Using the fact that we start out at time 0 with only one cell of the 0^{th} generation we have that the mean number of generation l cells alive at time t is $E(Z^l(t)) = m_0^l(t) - m_0^{l-1}(t)$. This gives an alternate formulation of the direct mean-based procedure in [17].

6.4 Incorporating division destiny

It remains debatable whether the limits set on division number are controlled by internal time sensing or through division counting [30]. In the framework of Section 4, division destiny for the progeny of a given cell is determined by a random time T^* after which the cells can no longer divide. Here, instead, we follow Hawkins *et al.* [17] and hypothesize that a given cell's division destiny is based on generations. The cell has a random generation number $K^* \in \{0, 1, 2, \ldots\}$ after which its progeny can no longer divide. Taking an expectation over the distribution of K^* gives the probability generating function for the total population with division destiny distributed as K^*:

$$F(s,t) = \sum_{k=0}^{\infty} F_0^k(s,t)P(K^* = k). \tag{23}$$

It is well known that there exists a minimum time between cell divisions [28], so there exists some $b > 0$ such that $T_B^M > b$ and $T_B^H > b$. Thus for a given time t at most members of the first $[t/b]$ (the greatest integer smaller than t/b) generations can be present. Let $K_{\min} = \min\{k : P(K^* = k) > 0\}$. Then

$$F(s,t) = \begin{cases} F_0^{[t/b]}(s,k) = F_0^{K_{\min}}(s,k) & \text{if } [t/b] \le K_{\min}, \\ \sum_{k=K_{\min}}^{[t/b]} F_0^k(s,t)P(K^* = k) + F_0^{[t/b]}(s,t)P(K^* > [t/b]), & \text{if } [t/b] > K_{\min}. \end{cases}$$

This greatly simplifies numerical techniques, as our iterative scheme terminates after a finite number of iterations. Alternatively, if K^* has bounded support ($\sup\{k^* : P(K^* = k^*) > 0\} < \infty$), this also suffices to ensure the scheme terminates after a finite number of iterations.

6.5 Numerical cyton model framework summary

At time $t = 0$ we have d cells whose time to division and time to die random variables are T_B^M and T_D^M. For each $i \in \{1, \ldots, d\}$ a random division destiny generation is selected according to the random variable K^*. After generation K^* any living progeny of cell i cannot further divide, but can only die. Any cells alive after the initial batch and before K^* have clocks distributed as T_B^H and T_D^H.

The probability generating function for the total population is $E(s^{Z(t)}) = F(s,t)^d$, where $F(s,t)$ is determined by first solving the integral recursions (21)–(22) for $\{F_k^l(s,t)\}$ and finally evaluating the weighted sum in equation (23).

7 Numerical model examples compared with experimental data

We expound on the methodology of Section 6 by first giving an algorithmic description of its implementation. In Section 7.3, we compare the models predictions to data taken from an *in vitro* experiment and an *in vivo* experiment. The *in vitro* data is shown here for the first time. Details of the experimental setup can be found in [30]. The *in vivo* data is taken from Homann, Teyton and Oldstone [18]. Lognormal distributions have been proposed as appropriate for clock distributions [17], so we present numerical results for that case. We come to the significant biological conclusion that despite the high degree of variability in the experience of any individual cell and its progeny, the overall immune response is highly predictable.

We give the initial cell a division destiny number k^* in the support of the distribution of K^*. We can think of cells in the first generation (immediate descendents of the initial cell) as having division destiny $k^* - 1$ and their descendents as having division destiny $k^* - 2$, and so forth. For example, if the initial cell has $k^* = 3$, then any of its descendents born into the third generation cannot divide any further. Ultimately we will sum over the likelihood that the division destiny K^* takes the value k^* to determine the overall ensemble's probabilistic behavior.

With fixed $K^* = k^*$, we augment our definition of the probability generating functions to define

$$F_k^{l,k^*}(s,t) := E\left(s^{\sum_{i=k}^{l} Z^i(t)} | Z^k(0) = 1,\ Z^m(0) = 0 \text{ for all } m \neq k, K^* = k^*\right).$$

If $l < k$, then we define $F_k^{l,k^*}(s,t) := 1$ irrespective of k^*. If $k > k^*$, then we define $F_k^{l,k^*}(s,t) = 1$ for every l. A key point that follows from the definition of k^* is that

$$F_k^{l,k^*}(s,t) = F_k^{k^*,k^*}(s,t) \quad \text{for all } l \geq k^*.$$

We will use this when constructing an algorithmic realization.

Matching with our initial hypotheses, we assume homogeneity. Every cell with a non-zero division destiny counter has the same set of distributions for time to division and time to death. We allow the initial cell and the final generation cells to have different distributions. For the boundary case of $k^* = 0$, i.e., if the initial generation is the final generation, we choose the final generation distribution for all the parameters.

For the initial cell we define: the distribution of time to live, $g_M(t) := P(T_D^M > t)P(T_B^M > t)$; the probability of dying before dividing, $h_M(t) := P(T_D^M \leq t, T_B^M \geq T_D^M) = \int_0^t P(T_B^M \geq r)dP(T_D^M = r)$; and the probability density function of division before death, $f_M(t) := P(T_D^M > t)dP(T_B^M > t)/dt$. For the final definition we assume that a density exists, at least using generalized functions such as

the Dirac delta function. For intermediate generations with random clocks T_D^H and T_B^H we have the analogous definitions of $g_H(t), h_H(t)$ and $f_H(t)$. For the final generation with T_D^{final} and T_B^{final}, we use the subscript F.

Assuming that for the final generation there is no possibility of cell division, we set T_B^{final} to be $+\infty$; algorithmically, we need only set T_B^{final} to be a time beyond the period of our interest in the population. Thus for the final generation we have: $g_F(t) = P(T_D^{\text{final}} > t)$, $h_F(t) = 1 - g_F(t)$ and $f_F(t) = 0$.

The homogeneity assumption imposes further constraints on the set of functions $F_k^{l,k^*}(s,t)$ for $k \geq 1$:

$$F_k^{l,k^*}(s,t) = F_{k+1}^{l+1,k^*+1}(s,t).$$

Let $b > 0$ be a lower bound on the time between cell divisions. That is $0 = P(T_B^M < b) = P(T_B^H < b)$. A lower bound of this sort is often seen experimentally and immensely simplifies numerical calculations: at time $t \geq 0$ one can have up to $[t/b]$ generations present. This discreteness imposes constraints on the set of functions $F_k^{l,j}(s,t)$:

$$F_k^{l,j}(s,t) = F_k^{[t/b],j}(s,t) \quad \forall\, l \geq [t/b].$$

These constraints are in addition to the machinery of the integral equations (21) and (22).

If we are interested in predicting the populations behavior up to a time $T_{\max}$, then we (numerically) perform the following iterations (**Schema A**) up to $K_{\max} = [T_{\max}/b]$.

1. Set $\hat{F}_0(s,t) := 1 + (s-1)g_F(t)$ and $F_0^{0,0}(s,t) = \hat{F}_0(s,t)$.
2. Calculate $\hat{F}_k(s,t) := sg_H(t) + h_H(t) + \int_0^t \hat{F}_{k-1}(s,t-r)^2 f_H(r)dr$ for $1 \leq k \leq K_{\max}$. Note that the integration operation is, in fact, a convolution.
3. Calculate $F_0^{k,k}(s,t) := sg_M(t) + h_M(t) + \int_0^t \hat{F}_{k-1}(s,t-r)^2 f_M(r)dr$ for $1 \leq k \leq K_{\max}+1$.

The intermediate term $\hat{F}_k(s,t)$ calculates the probability generating function of a population for which all generations including the initial generation have the same distribution for the time to divide and the time to die clocks and the division destiny variable $K^* = k$.

Define $K_{\min} \in \{0, 1, \ldots\}$ be the greatest lower bound on K^* that has positive probability, i.e., $K_{\min} = \min\{k : P(K^* = k^*) > 0\}$. Incorporating these constraints, we have

$$F(s,t) = \begin{cases} F_0^{K_{\min},K_{\min}}(s,t) & \text{if } [t/b] < K_{\min}, \\ \displaystyle\sum_{k=K_{\min}}^{[t/b]} P(K^* = k^*) F_0^{k^*,k^*}(s,t) + P(K^* > [t/b]) F_0^{[t/b]+1,[t/b]+1}(s,t) & \text{otherwise.} \end{cases}$$

Having determined the probability generating function, the complete distribution can be recovered using numerical inversion techniques such as those developed by queueing theorists. See, for example, [2][3][8][20] and Section 1.2.4 of [1].

7.1 Explicit iteration schemes for means and variances

If we evaluate $F(s,t)$ numerically and wish to determine the mean and variance of $Z(t)$, we must numerically take derivatives. Alternatively, if we are only interested in a finite collection of moments

of the distribution of $Z(t)$, then we can design an explicit iterative scheme for them directly based on the same ideas as for $F(s,t)$ by taking derivatives of the iteration scheme given above. We illustrate this here, giving new direct iteration schemes for the mean and variance.

It will only be necessary to consider division destiny K^* equal to our dummy variable k. Define the following quantities.

$$\hat{m}_k(t) := \frac{\partial \hat{F}_k(s,t)}{\partial s}\Big|_{s=1},\ m_0^{k,k}(t) := \frac{\partial F_0^{k,k}(s,t)}{\partial s}\Big|_{s=1},$$

$$\hat{v}_k(t) := \frac{\partial^2 \hat{F}_k(s,t)}{\partial s^2}\Big|_{s=1} \text{ and } v_0^{k,k}(t) := \frac{\partial^2 F_0^{k,k}(s,t)}{\partial s^2}\Big|_{s=1}.$$

By taking derivatives of the $F(s,t)$ scheme, we have the following explicit iteration scheme (**Schema B**) to determine means and variances.

1. Set $\hat{m}_0(t) = g_F(t)$, $m_0^{0,0}(t) = \hat{m}_0(t)$ and $\hat{v}_0(t) = v_0^{0,0}(t) = 0$.
2. Calculate

$$\hat{m}_k(t) = g_H(t) + 2\int_0^t \hat{m}_{k-1}(t-r) f_H(r) dr; \text{ and}$$
$$\hat{v}_k(t) = 2\int_0^t \left(\hat{v}_{k-1}(t-r) + \hat{m}_{k-1}^2(t-r)\right) f_H(r) d(r) \quad 1 \le k \le K_{\max}.$$

3. Calculate

$$m_0^{k,k}(t) = g_M(t) + 2\int_0^t \hat{m}_{k-1}(s, t-r) f_M(r) dr; \text{ and}$$
$$v_0^{k,k}(t) = 2\int_0^t \left(\hat{v}_{k-1}(t-r) + \hat{m}_{k-1}^2(t-r)\right) f_M(r) d(r) \quad 1 \le k \le K_{\max}+1.$$

As in schema A, the intermediate terms $\hat{m}_k(t)$ and $\hat{v}_k(t)$ compute the contributions of a population for which all generations including the initial generation have the same distribution for the time to divide and the time to die clocks and the division destiny variable K^* is k.

Finally we get:

$$m(t) = \begin{cases} m_0^{K_{\min},K_{\min}}(t) & \text{if } [t/b] < K_{\min} \\ \sum_{k=K_{\min}}^{[t/b]} P(K^*=k)\, m_0^{k,k}(t) + P(K^* > [t/b])\, m_0^{[t/b]+1,[t/b]+1}(t) & \text{otherwise} \end{cases}$$
$$v(t) = \begin{cases} v_0^{K_{\min},K_{\min}}(t) & \text{if } [t/b] < K_{\min} \\ \sum_{k=K_{\min}}^{[t/b]} P(K^*=k)\, v_0^{k,k}(t) + P(K^* > [t/b])\, v_0^{[t/b]+1,[t/b]+1}(t) & \text{otherwise} \end{cases}$$

and $\text{var}(t) = v(t) + m(t)(1-m(t))$.

7.2 Isolating each generation's contribution

With a little additional effort we can isolate how the mean of every generation changes with time. The extra effort is necessary to account for cells that have not reached their division destiny. This approach, however, does not generalize to higher moments as the population sizes across generations are not stochastically independent.

We are interested in the probability generating function defined by

$$F_0^{k,K^*}(s,t) = \sum_{k^*=0}^{\infty} P(K^*=k^*) E\left(s^{\sum_{j=0}^{k} Z^k(t)} \mid Z^0(0)=1,\ Z^i(0)=0 \text{ for all } i\geq 1, K^*=k^*\right),$$

where (with an abuse of notation) we wish to explicitly follow the progress of the population up to generation k while accounting for K^*. Using the earlier defined quantities we write

$$\begin{aligned} F_0^{k,K^*}(s,t) &= \sum_{j=0}^{+\infty} P(K^*=j) F_0^{k,j}(s,t) \\ &= \sum_{j=0}^{\min([t/b],k)} P(K^*=j) F_0^{j,j}(s,t) + P(K^*>\min([t/b],k)) F_0^{\min([t/b],k),\min([t/b],k)+1}(s,t) \\ &= \begin{cases} F_0^{\min([t/b],k),\min([t/b],k)+1}(s,t) & \text{if } \min([t/b],k) < K_{\min}, \\ \sum_{j=K_{\min}}^{\min([t/b],k)} P(K^*=j) F_0^{j,j}(s,t) & \\ +P(K^*>\min([t/b],k)) F_0^{\min([t/b],k),\min([t/b],k)+1}(s,t) & \text{otherwise.} \end{cases} \end{aligned} \tag{24}$$

From this it directly follows that the mean of the sum of up to k generations is

$$m_0^{k,K^*}(t) = \begin{cases} m_0^{\min([t/b],k),\min([t/b],k)+1}(t) & \text{if } \min([t/b],k) < K_{\min}, \\ \sum_{j=K_{\min}}^{\min([t/b],k)} P(K^*=j) m_0^{j,j}(t) & \\ +P(K^*>\min([t/b],k)) m_0^{\min([t/b],k),\min([t/b],k)+1}(t) & \text{otherwise.} \end{cases} \tag{25}$$

Thereafter, we use the linearity of the mean to assert that

$$\sum_{k^*=0}^{\infty} P(K^*=k^*) E(Z^k(t)|K^*=k^*) = m_0^{k,K^*}(t) - 1_{\{k>0\}} m_0^{k-1,K^*}(t), \text{ where } 1_{\{k>0\}} := \begin{cases} 1 \text{ if } k>0, \\ 0 \text{ if } k\leq 0. \end{cases}$$

To evaluate (24) and (25) we need to determine $F_0^{k,k}(s,t)$, $m_0^{k,k}(t)$, $F_0^{k,k+1}(s,t)$ and $m_0^{k,k+1}(t)$ for different values of k. These correspond to the case when $K^*=k$ and $K^*>k$, respectively. Following schema **A** and schema **B** we have already obtained $F_0^{k,k}(s,t)$ and $m_0^{k,k}(t)$. Thus we only need to specify a procedure to compute $F_0^{k,k+1}(s,t)$ and $m_0^{k,k+1}(t)$. For this purpose we have the following explicit iteration scheme (**Schema C**).

1. Set $\bar{m}_0(t) := g_H(t)$, $\bar{F}_0(s,t) := 1+(s-1)g_H(t)$, $\tilde{m}_0^0(t) := g_M(t)$ and $\tilde{F}_0^k(s,t) := \bar{F}_0(s,t)$.
2. Calculate

$$\begin{aligned} \bar{m}_k(t) &= g_H(t) + 2\int_0^t \bar{m}_{k-1}(t-r) f_H(r) dr \\ \bar{F}_k(s,t) &= s g_H(t) + h_H(t) + \int_0^t \bar{F}_{k-1}(s,t-r)^2 f_H(r) dr \quad 1\leq k\leq K_{\max}. \end{aligned}$$

3. Calculate

$$\breve{m}_k(t) = g_M(t) + 2\int_0^t \bar{m}_{k-1}(s, t-r) f_M(r) dr$$
$$\breve{F}_0^k(s,t) = sg_M(t) + h_M(t) + \int_0^t \bar{F}_{k-1}(s,t-r)^2 f_M(r) dr \quad 1 \le k \le K_{\max}.$$

The intermediate terms $\bar{m}_k(t)$ and $\bar{F}_k(s,t)$ compute the mean and probability generating function of the population encompassing generations 0 to k, where all generations including the initial generation follow the same distributions for the time to divide and time to die clocks, and where $K^* = +\infty$, i.e., division destiny never sets in. The terms $\breve{m}_k(t)$ and $\breve{F}_k(s,t)$ compute the mean and probability generating function of the total population from generation 0 to k under the assumption that the initial generation is different from the rest of the generations, but once again with no division destiny.

As we only account for generations 0 through k, it is clear that

$$F_0^{k,k+1}(s,t) = \breve{F}_0^k(s,t) \qquad \text{and} \qquad m_0^{k,k+1}(t) = \breve{m}_k(t),$$

even though on the right side of both equations there is no division destiny. Thus all terms are available to substitute into equations (24) and (25).

This procedure yields the higher moments of the population up to generation k, but the inability to isolate higher moments of a specific generation now becomes clear due to inter-generational dependencies. Exact values of moments other than the mean cannot be calculated, nevertheless it is possible to bound them using quantities that can be calculated. We demonstrate this with a bound on the variance of any given generation. For $l \ge 0$ and $k \ge l$ define $\mathrm{var}_l^{k,K^*}(t) := E\left(\left(\sum_{j=k}^l Z^j(t) - m_l^{k,K^*}(t)\right)^2\right)$. The bound we use is developed as follows:

$$\left(\sum_{j=l}^k Z^j(t)\right)^2 = Z^k(t)^2 + \left(\sum_{j=l}^{k-1} Z^j(t)\right)^2 + 2Z^k(t)\left(\sum_{j=l}^{k-1} Z^j(t)\right) \ge Z^k(t)^2 + \left(\sum_{j=l}^{k-1} Z^j(t)\right)^2.$$

For all $k \ge l+1$. Using the above relation and

$$E\left(\left(\sum_{j=l}^k Z^j(t)\right)^2\right) = \mathrm{var}_l^{k,K^*}(t) + m_l^{k,K^*}(t)^2$$

we have

$$\mathrm{var}_l^{k,K^*}(t) + m_l^{k,K^*}(t)^2 \ge \mathrm{var}_k^{k,K^*}(t) + m_k^{k,K^*}(t)^2 + \mathrm{var}_l^{k-1,K^*}(t) + m_l^{k-1,K^*}(t)^2.$$

After a little rearrangement, this yields

$$\mathrm{var}_k^{k,K^*}(t) \le \mathrm{var}_l^{k,K^*}(t) - \mathrm{var}_l^{k-1,K^*}(t) + 2m_k^{k,K^*}(t)m_l^{k-1,K^*}(t). \tag{26}$$

In terms of calculable quantities, the bound in equation (26) gives

$$\mathrm{var}_k^{k,K^*}(t) \le \mathrm{var}_0^{k,K^*}(t) - \mathrm{var}_0^{k-1,K^*}(t) + 2m_k^{k,K^*}(t)m_0^{k-1,K^*}(t)$$

for all $k \ge 1$. As before the iterations to obtain $\mathrm{var}_0^{k,K^*}(t)$ can be derived from (24) after differentiating twice.

Generation Number	0	1	2	3	4	5	6	7	8
Probability	0.0200	0.0308	0.1342	0.2894	0.3103	0.1655	0.0438	0.0057	0.0004

Table 1 *In vitro* division destiny probability mass function.

7.3 Comparison of the numerical model's predictions and experimental data

In this section we compare the second model's predictions with observations from two sets of data. Firstly with *in vitro* experimental data that follows B lymphocytes stimulated by 3 μM CpG DNA. This method of stimulating B lymphocytes induces limited division rounds and cells follow closely the rules underlying the cyton model [30] with proliferation, cessation and death completed within 10 days. Note this system provides an excellent evaluation of variation as cells are not disrupted to remove stimulation as was necessary to explore division destiny in Hawkins *et al.* [17]. Cyton parameter values for the CpG data set were calculated as previously described [17]. Secondly, we compare the model predictions with data from an *in vivo* experiment reported on in [18], with cyton parameterization given in [17].

The *in vitro* experiment has two parts. Firstly, approximately fifteen thousand CFSE-labeled purified B-cells were exposed to the mitogenic stimulus. Roughly 10% of the initial cells do not respond to the stimulus. The overall cell population size, and the per-generation cell population, were recorded from the introduction of the stimulus to beyond division destiny by flow cytometric analysis [17]. This gives data with which to compare model predictions.

Secondly, it is reported that in independent experiments, the empirical clock distributions appear to be well matched by members of the family of lognormal distributions [11][17]. For example, $P(T_D > t) \approx P(N(\mu, \sigma^2) > \log t)$, where $N(\mu, \sigma^2)$ is a Normally distributed random variable with mean μ and variance σ^2. They fit the mean and variance for the lognormal distributions of the random clocks T_D^M, T_B^M, T_D^H and T_B^H to the experimental observations. With this information it is possible to run the model and make predictions that can be compared with the data.

Cyton model parameters were determined using the methodology described in [17], mildly modified to introduce a minimum time of 1 hour between cell division while leaving the mean and variance unchanged. No minimum value is introduced for the time to death random variables. The division destiny, K^*, has the probability mass function given in Table 1. The division destiny distribution is such that a maximum of 8 generations can occur. The complete list of clock parameters are:

- Time to division of initial cells: mean 34.86 hours (including 1 hour minimum); standard deviation 4.9 hours. Time to death of initial cells: mean 151.27 hours; standard deviation 19.75 hours.
- Time to division of cells in each subsequent generation: mean 7.22 hours (including 1 hour minimum); standard deviation 1.31 hour. Time to death of cells in each subsequent generation: mean 82.25 hours; standard deviation 111.28 hours.

With this parameterization, we make predictions regarding the mean and standard deviation of the population size as a function of time. We compare this with the experimentally observed values.

Figure 3 reports the model predicted mean and mean plus/minus five standard deviations. Thus, with extremely high likelihood, the immune response will progress within these bounds. There is also a scatter plot of observed data. The model predicts that despite the high degree of variability experienced by individual cells, the overall immune response is highly predictable; that is, there are small error bars around mean behavior. This is significant as it suggests that even though each cell is responding independently to the mitogenic stimulus and with a large amount of variation, the

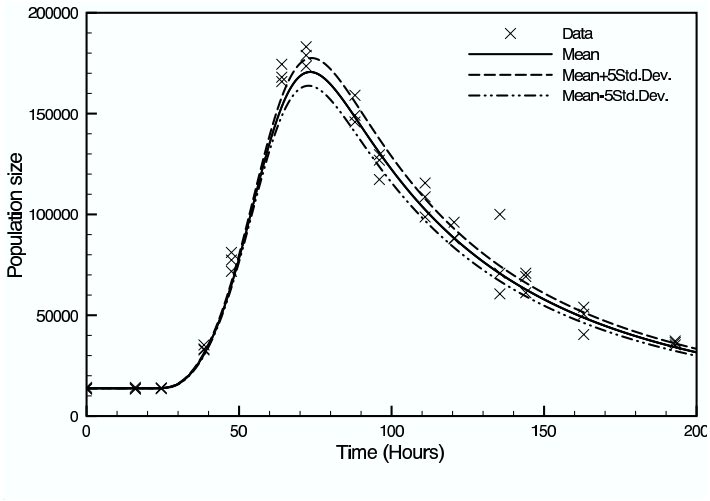

Fig. 3 Total cell population size vs. time. Comparison of model and *in vitro* experimental data. Resting B cells purified from spleens of C57BL mice were stimulated in 200 μl culture with 3 μM CpG (sequence - 5'-TCCATGACGTTCCTGATGCT-3'). Triplicate cultures were harvested at each time point.

Generation	12	13	14	15	16	17	18	19	20	21	22
Probability	0.0001	0.0017	0.0165	0.0826	0.2206	0.3151	0.2408	0.0984	0.0215	0.0025	0.0002

Table 2 *In vivo* division destiny probability mass function.

overall behavior is predictable. Mathematically, the existence of this robustness in the overall immune system comes about as a "law of large numbers". That is, the vast majority of likely microscopic states (governed by each individual cell's death/division experience) give rise to the same macroscopic immune response.

The predictions match well with the experimental observations, but with a few outliers. Due to the complexity involved in the experimentation, outliers are expected. Clearly the bulk of the variation observed in the data derives from experimental errors.

Figure 4 reports on the time evolution of means for each generation, from the initial cells to those that are the consequence of eight divisions. For each individual generation we cannot determine its variance explicitly. However, we provide an upper bound on the standard deviation using the methods described in Section 7.2. As even these bounds predict little variability in the overall response, it is clear that the number of cells of each generation that are alive at each time is highly predictable. The shapes observed for each generation are qualitatively similar, but quantitatively quite different. The model matches excellently with the observations.

In the *in vivo* experiment, from Homann, Teyton and Oldstone [18], CD8+ T lymphocytes specific to a single peptide/MHC epitope were followed after infection with lymphocytic choriomeningitis virus. The overall cell population size was recorded from the introduction of the virus to beyond division destiny and on to the period where remaining cell numbers are retained at a homeostatic level. For a detailed description of the experimental setup and methodology, we refer the reader to [18]. The true starting cell number is unknown, although it is estimated to be approximately one hundred [10][17]. It should be noted too that the starting value of one hundred is a minimal estimate for a virus response,

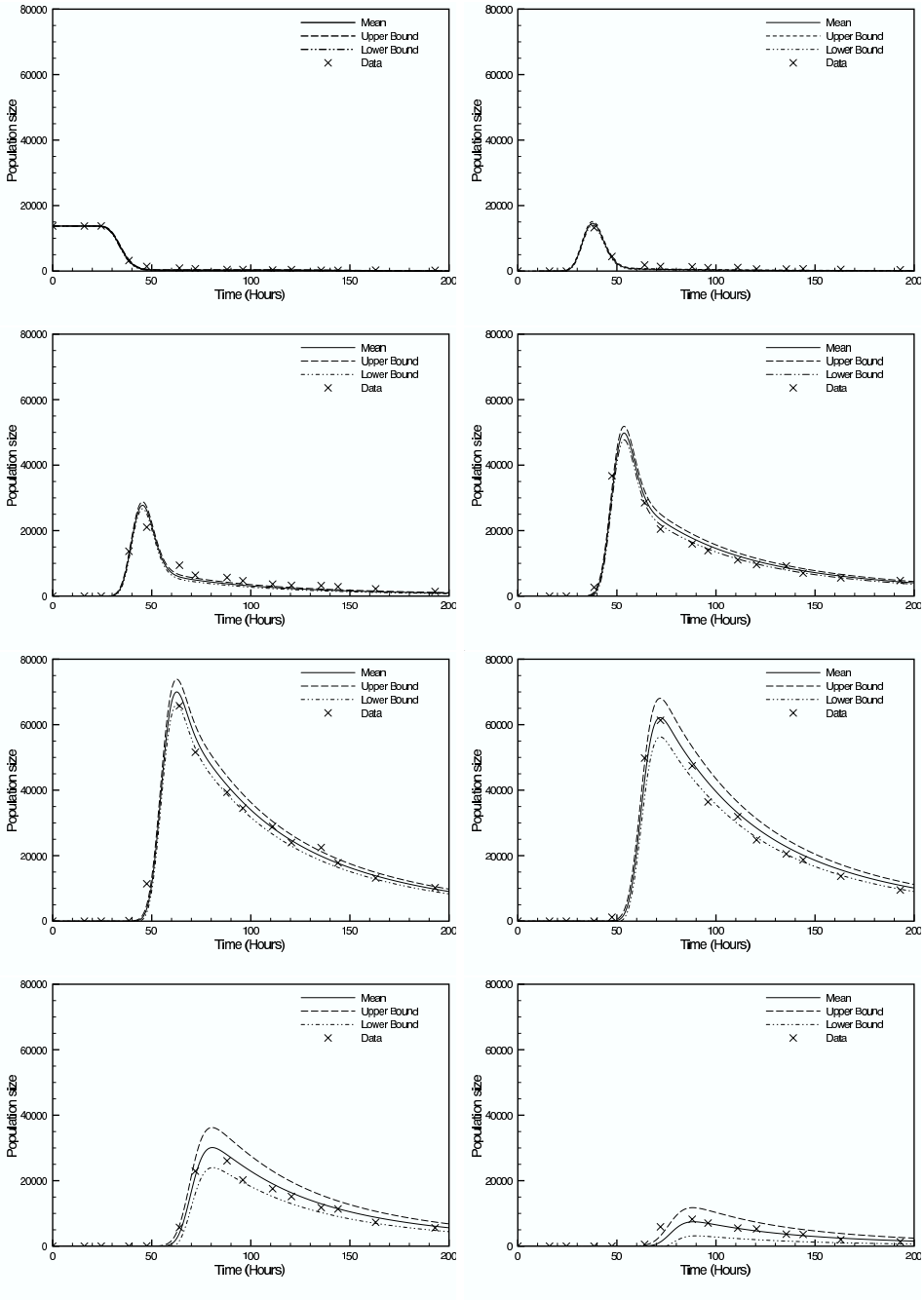

Fig. 4 Population of 0^{th} through 7^{st} generations as a function of time. Model bounds from Section 7.2. Comparison of model and *in vitro* experimental data.

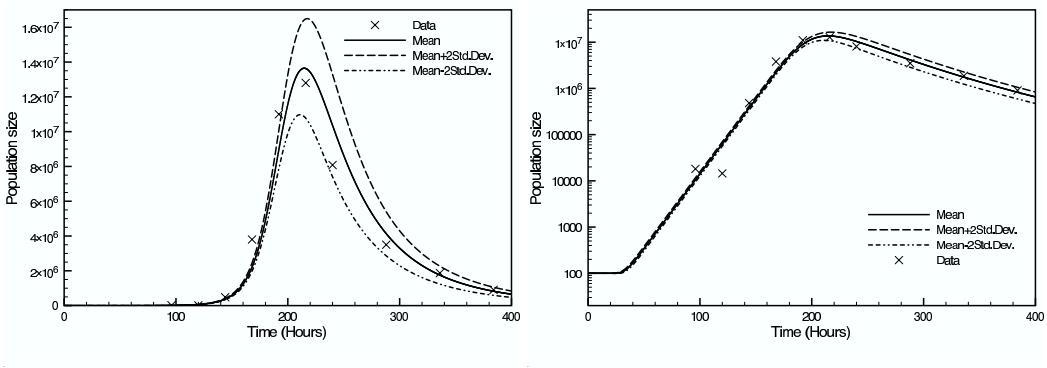

Fig. 5 Total cell population size vs. time (normal and log-scale). Comparison of model and *in vivo* experimental data. Data shows population response to viral epitope NP64.

as it represents a single epitope and is representative of similar data that could be collected on at least four other epitopes, bringing the population response to an initial number over five hundred [18][10].

We adopted the Hawkins *et al.* [17] parameterization, but again mildly modified to introduce a minimum time of 1 hour between cell division while leaving the mean and variance unchanged. No minimum value is introduced for the time to death random variables. Division destiny, K^*, has the probability mass function given in Table 2. The division destiny distribution is such that a minimum of 12 and maximum of 22 generations can occur. The complete list of clock parameters are:

- Time to division of initial cells: mean 40.44 hours (including 1 hour minimum); standard deviation 7.55 hours. Time to death of initial cells: mean 400.50 hours; standard deviation 20.04 hours.
- Time to division of cells in each subsequent division: mean 9.17 hours (including 1 hour minimum); standard deviation 0.73 hour. Time to death cells in each subsequent division: mean 70.23 hours; standard deviation 62.31 hours.

Figure 5 reports on the mean derived from the cyton model plus/minus two standard deviations and is overlaid with the observed data. Despite the small number of initial cells, it can be seen that the immune response is highly predictable.

8 Discussion

Based on cell-level stochastic assumptions on the dynamics of immune response that were driven by experimental observations, we have introduced two distinct approaches to predicting the immune response to a mitogenic signal. One method enables explicit calculations for certain distributions and qualitatively exhibits the full range of observed immune responses. The second approach allows simple numerical schemes for distributions for which there is little prospect of analytic results. We compared the predictions of the second method to experimentally observed lymphocyte population changes over time. The model predictions are remarkably accurate, but the important advance over the previous calculation method [17] is that the higher moments can now be calculated.

The calculation of higher moments is interesting for a number of important reasons relating to features of the immune response. First, the cyton model highlights the randomness associated with stimulation

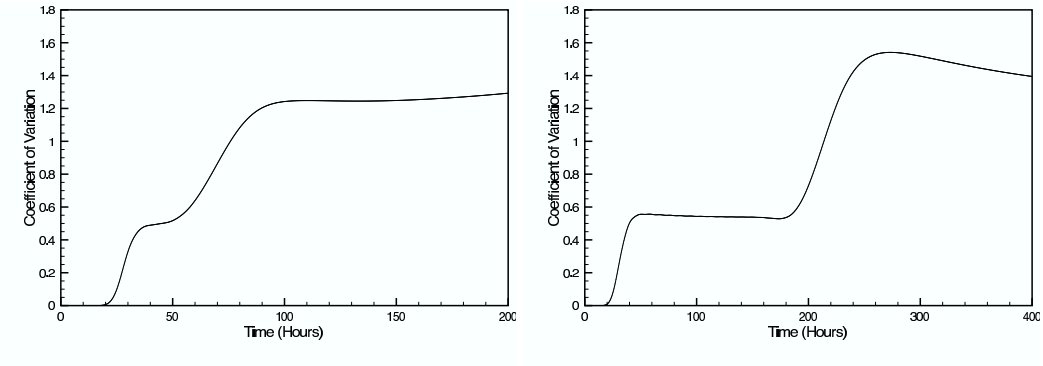

Fig. 6 Coefficient of variation for the progeny of a single cell as a function of time. *In vitro* parameterization on the left and *in vivo* parameterization on the right.

and progression of proliferating cells as a feature of the adaptive immune response. This inherent randomness ensures that, in effect, every participating cell will follow a different course over time. Thus, there is potential for responses simply not to occur due to stochastic effects, or occasionally to proceed so weakly that the individual would not be protected. However, we demonstrate here that these scenarios are highly unlikely.

As the cyton model treats each cell as being stochastically independent, we can readily deduce the impact of initial cell number on the variability in immune response. The usual measure of the dispersion of a probability distribution is the coefficient of variation: the standard deviation divided by the mean. For a given cyton parameterization, let $m(t)$ denote the mean number of progeny of a single cell at time t and $\mathrm{var}(t)$ denote its variance. Then with d initial cells, the coefficient of variation is $\sqrt{\mathrm{var}(t)}/(m(t)\sqrt{d})$. Thus, once the coefficient of variation is known for a single cell, it can be immediately determined for the ensemble. For example, Figure 6 plots the coefficient of variation, $\sqrt{\mathrm{var}(t)}/m(t)$, for a single cell with the parameterizations of both the *in vivo* and *in vitro* experiments reported on in Section 7.3. The coefficient of variation increases at two distinct time-scales: the time of first division; and during division destiny. The largest observed coefficient of variation is approximately 1.3 and 1.6 respectively. With one hundred initial cells, the coefficient of variation is one tenth of these numbers and with ten thousand initial cells, it is one hundredth, making the dispersion in the distribution extremely small.

The method developed here offers an alternative way to calculate the expected variation attributed to the selection and stochastic variation in cell division and death lifetimes to that used by Milutinović and De Boer [23]. These authors referred to this source of variation as "process variation" to distinguish it from that contributed by experimental error. They assume a triphasic model where the immune response breaks into three distinct phases: (1) an initial period of non-proliferation; (2) an exponential increase due to cell division; and (3) after some time, an exponential cell loss due to cell death. Their aim is to identify ordinary differential equation (ODE) parameters that match experimental data, but - in an advance on ODE parameter fitting - they take into account the stochastic nature of the data through a Gaussian approximation. Their approach assumes that the immune response will be of a given form, while the cyton model predicts this form from lower-level stochastic dynamical hypotheses. We also note that their assumption that the coefficient of variation does not change with time clashes with the cyton model prediction. They conclude that this process error is significant and therefore differences in response of individual mice (or humans) need not reflect parameter differences in the underlying response. For the parameter values we analyzed, our analysis also provides a method for

calculating the expected variance between animals, and suggests that in general the differences would be small.

It is important to note that not all immune responses are as strong as that observed against a virus. It is conceivable that some stimuli will react with less than one hundred starting cells and initiate suboptimal stimulation conditions that lead to greater variation in times to divide and die, as occurs for weaker stimulation protocols *in vitro* [11][13][19]. Under these conditions the expected variation between otherwise identical individuals could be significant.

Our approach here has been to formulate the mechanical and kinetic axioms of the cyton model in terms of the time-evolution of a probability generating function, adapting branching process ideas. This allows the higher moments to be calculated. A distinct branching process model to help describe the underlying asynchrony in CFSE data is reported by Yates *et al.* [32]. They propose a cell population size model based on a discrete time approximation to a continuous time branching process, which enables them to account for the minimum time between cell division, and focus on parameter estimation issues. In their Appendix 1, they comment that a more general approach would involve continuous time models where lifetimes are not restricted to be exponential, but that these processes are harder to analyse. Our second method treats this more general analysis and thus our approach to the cyton model can serve as alternative platform for CFSE analysis performed by Yates *et al.*

We have analysed the model proposed in [17], but the approach taken here can also treat alternate hypotheses. For example, if new experimental data suggests that birth and death random variables are not independent, then if the joint distribution of birth and death clocks is known, a similar analysis can be performed (in both approaches) starting from equation (1). In addition, if new experimental observations suggest an alternative hypothesis for inheritance of parameters from generation to generation, then this can be incorporated into the analysis by including conditional information at the time of cell birth.

In summary, we expect the method presented here should prove useful for calculating expected variance for immune responses at many different initial conditions dictated by such variables as starting cell number, receptor affinity and availability of costimulatory and growth factor signals.

Acknowledgments: This work was carried out while Philip Hodgkin was visiting the Hamilton Institute, NUI Maynooth, Ireland as a Walton Fellow, supported by Science Foundation Ireland. We thank Peter Wellstead for helping to facilitate this collaborative work at the Hamilton Institute. We thank Wilhelm Huisinga and Cameron Wellard for helpful comments on an earlier draft of this paper. We also thank Dirk Homann for providing the *in vivo* data.

References

1. Joseph Abate, Gagan L. Choudhury, and Ward Whitt. *Computational Probability*, chapter An Introduction to Numerical Transform Inversion and its Application to Probability Models, W. Grassman ed., pages 257–323. Kluwer, Boston, 1999.
2. Joseph Abate and Ward Whitt. The Fourier-series method for inverting transforms of probability distributions. *Queueing Systems Theory Appl.*, 10(1-2):5–87, 1992.
3. Joseph Abate and Ward Whitt. Numerical inversion of probability generating functions. *Oper. Res. Lett.*, 12(4):245–251, 1992.
4. R. Ahmed and D. Gray. Immunological memory and protective immunity: understanding their relation. *Science*, 272(5238):54–60, 1996.
5. K. B. Athreya and P. E. Ney. *Branching processes*. Dover Publications Inc., Mineola, NY, 2004.
6. P. Billingsley. *Probability and Measure*. John Willey & Sons, 1995.
7. D.A. Cantrell and K. A. Smith. The interleukin-2 T-cell system: a new cell growth model. *Science*, 224(4655):1292–1361, 1984.

8. Gagan L. Choudhury, David M. Lucantoni, and Ward Whitt. Multidimensional transform inversion with applications to the transient $M/G/1$ queue. *Ann. Appl. Probab.*, 4(3):719–740, 1994.
9. R. J. De Boer and Alan S. Perelson. Estimating division and death rates from CFSE data. *J. Comput. Appl. Math.*, 184(1):140–164, 2005.
10. R.J. De Boer, D. Homann, and AS. Perelson. Different dynamics of CD4+ and CD8+ T cell responses during and after acute lymphocytic choriomeningitis virus infection. *J Immunol.*, 171(8):3928–3935, 2003.
11. Elissa K. Deenick, Amanda V. Gett, and Philip D. Hodgkin. Stochastic model of T cell proliferation: A calculus revealing IL-2 regulation of precursor frequencies, cell cycle time, and survival. *Journal of Immunology*, 170(10):4963–4972, 2003.
12. G. Froese. The distribution and interdependence of generation times of HeLa cells. *Experimental Cell Research*, 35(2):415–419, 1964.
13. V.V. Ganusov, D. Milutinović, and R.J. De Boer. IL-2 regulates expansion of CD4 T cell populations by affecting cell death: insights from modeling CFSE data. *Journal of Immunology*, In press, 2007.
14. V.V. Ganusov, S.S. Pilyugin, R.J. De Boer, K. Murali-Krishna, R. Ahmed, and R. Antia. Quantifying cell turnover using CFSE data. *J Immunol Methods*, Mar.(1-2):183–200, 2005.
15. Amanda V. Gett and Philip D. Hodgkin. A cellular calculus for signal integration by T cells. *Nature Immunology*, 1(4):239–244, 2000.
16. Theodore E. Harris. *The theory of branching processes*. Dover Phoenix Editions. Dover Publications Inc., Mineola, NY, 2002.
17. E. D. Hawkins, M. L. Turner, M. R. Dowling, C. van Gend, and P. D. Hodgkin. A model of immune regulation as a consequence of randomized lymphocyte division and death times. *Proceedings of the National Academy of Sciences of the United States of America*, 104:5032–5037, 2007.
18. D. Homann, L. Teyton, and M.B.A. Oldstone. Differential regulation of antiviral T-cell immunity results in stable CD8+ but declining CD4+ T-cell memory. *Nature Medicine*, 7:913–919, 2001.
19. M. Hommel and R.J. De Boer. TCR affinity promotes CD8+ T cell expansion by regulation survival. *Journal of Immunology*, In press, 2007.
20. E. P. C. Kao. *An Introduction to Stochastic Processes*. Duxbury Press, NY, 1997.
21. Marek Kimmel and David E. Axelrod. *Branching processes in biology*, volume 19 of *Interdisciplinary Applied Mathematics*. Springer-Verlag, New York, 2002.
22. K. León, J. Faro, and J. Carneiro. A general mathematical framework to model generation structure in a population of asynchronously dividing cells. *J. Theoret. Biol.*, 229(4):455–476, 2004.
23. D. Milutinović and R. J. De Boer. Process noicse: an explanation for the fluctiations in the immune response during acute viral infection. *Biophysical journal*, 92:3358–3367, 2007.
24. D. S. Nachtwey and I. L. Cameron. *Methods in Cell Physiology*, volume III, pages 213–257. Academic Press, New York, 1968.
25. Andrei D. Polyanin and Alexander V. Manzhirov. *Handbook of integral equations*. CRC Press, Boca Raton, FL, 1998.
26. David M. Prescott. Regulation of cell reproduction. *Cancer Research*, 28(9):1815–1820, 1968.
27. R. Shields. Transition probability and the origin of variation in the cell cycle. *Nature*, 267(5613):704–707, 1977.
28. J.A. Smith and L. Martin. Do cells cycle? *Proceedings of the National Academy of Sciences of the United States of America*, 70(4):1263–1267, 1973.
29. Stuart G. Tangye, Danielle T. Avery, Elissa K. Deenick, and Philip D. Hodgkin. Intrinsic differences in the proliferation of naive and memory human B cells as a mechanism for enhanced secondary immune responses. *Journal of Immunology*, 170(2):686–694, 2003.
30. M. Turner, E. Hawkins, and P.D. Hodgkin. Manuscript in preparation, 2007.
31. Henrique Veiga-Fernandes, Ulrich Walter, Christine Bourgeois, Angela McLean, and Benedita Rocha. Response of naïve and memory CD8+ t cells to antigen stimulation in vivo. *Nature Immunology*, 1:47–53, 2000.
32. A. Yates, C. Chan, J. Strid, S. Moon, R. Callard, A.J.T. George, and J. Stark. Reconstruction of cell population dynamics using CFSE. *BMC Bioinformatics*, In press, 2007.