

Computational modeling

Quantifying cell turnover using CFSE data

Vitaly V. Ganusov^{a,1,2,*}, Sergei S. Pilyugin^{b,1}, Rob J. de Boer^c, Kaja Murali-Krishna^d,
Rafi Ahmed^c, Rustom Antia^a

^aDepartment of Biology, Emory University, Atlanta, GA 30322, United States

^bDepartment of Mathematics, University of Florida, Gainesville, FL 32611, United States

^cTheoretical Biology, Utrecht University, Padualaan 8, 3584 CH, Utrecht, The Netherlands

^dDepartment of Immunology, University of Washington, Seattle, WA 98195, United States

^eDepartment of Microbiology and Immunology, Emory University, Atlanta, GA 30322, United States

Received 9 April 2004; received in revised form 23 December 2004; accepted 28 January 2005

Abstract

The CFSE dye dilution assay is widely used to determine the number of divisions a given CFSE labelled cell has undergone in vitro and in vivo. In this paper, we consider how the data obtained with the use of CFSE (CFSE data) can be used to estimate the parameters determining cell division and death. For a homogeneous cell population (i.e., a population with the parameters for cell division and death being independent of time and the number of divisions cells have undergone), we consider a specific biologically based “Smith–Martin” model of cell turnover and analyze three different techniques for estimation of its parameters: direct fitting, indirect fitting and rescaling method. We find that using only CFSE data, the duration of the division phase (i.e., approximately the S+G₂+M phase of the cell cycle) can be estimated with the use of either technique. In some cases, the average division or cell cycle time can be estimated using the direct fitting of the model solution to the data or by using the Gett–Hodgkin method [Gett A. and Hodgkin, P. 2000. A cellular calculus for signal integration by T cells. *Nat. Immunol.* 1:239-244]. Estimation of the death rates during commitment to division (i.e., approximately the G₁ phase of the cell cycle) and during the division phase may not be feasible with the use of only CFSE data. We propose that measuring an additional parameter, the fraction of cells in division, may allow estimation of all model parameters including the death rates during different stages of the cell cycle.

© 2005 Elsevier B.V. All rights reserved.

Keywords: CFSE; Cell turnover; Cell cycle; Smith–Martin model; Parameter estimation; Nonlinear regression; Linear regression; Gett–Hodgkin method; Average division time; Death rate

* Corresponding author.

E-mail addresses: vitaly.ganusov@alum.emory.edu (V.V. Ganusov), pilyugin@math.ufl.edu (S.S. Pilyugin).

¹ These authors contributed equally to this work.

² On the leave of absence from Institute of Biophysics, Akademgorodok, Krasnoyarsk, Russia, 660036.

1. Introduction

The dynamic nature of immune responses requires the development of appropriate experimental and

theoretical tools to quantitatively understand the division and death processes which determine the turnover of immune cells. For example, we need to obtain a reliable quantitative method for estimating the division and death rates of immune cells with or without antigenic stimulation; to understand how these rates differ for naive and memory cells, antigen-specific B and T lymphocytes, and other types of cells.

The problem of estimating the turnover rate of immune cells is not new and has been approached with the use of BrdU and D-glucose labels allowing to distinguish between divided and undivided cells (for a review see [Asquith and Bangham, 2003](#); [De Boer et al., 2003](#)). In the past few years, BrdU assay has been supplemented by the CarboxyFluorescein diacetate Succinimidyl Ester (CFSE) label which provides a more detailed information on the proliferation history of labelled cells ([Lyons and Parish, 1994](#)). Specifically, CFSE dye is loaded into cells in vitro and CFSE fluorescence in a given cell is monitored over time. Upon division, CFSE segregates equally between daughter cells so that the intensity of cellular fluorescence decreases twofold with each successive generation. This property of CFSE allows accurate tracking of the number of divisions that a given cell has undergone either in vitro or following transfer in vivo ([Weston and Parish, 1990](#); [Lyons and Parish, 1994](#); [Parish, 1999](#); [Lyons, 2000](#)).

Our work extends some of the earlier studies that used CFSE data to estimate the rates of cell turnover ([Nordon et al., 1999](#); [Gett and Hodgkin, 2000](#); [Veiga-Fernandes et al., 2000](#); [Revy et al., 2001](#); [Bernard et al., 2003](#); [Pilyugin et al., 2003](#); [De Boer and Perelson, 2005](#)). In this paper, we discuss the quantitative information on the cell turnover that is contained within the CFSE data. We analyze several methods that may be used to process this information. We also describe the limitations of the CFSE data that may preclude a more detailed understanding of the immune cell turnover.

In this paper, we restrict our analysis to the case of a homogeneous cell population, i.e., a population with the parameters determining cell division and death being identical for all cells and independent of time or the number of divisions cells have undergone. (In Section 8 we review some of the limitations of this

assumption.) There are at least two major approaches that could be used to quantify the turnover of immune cells.

- One can formulate a general model of cell division and death that does not provide the specific description for division and death, and estimate the invariant parameters related to cell turnover ([Pilyugin et al., 2003](#)).
- Alternatively, one can construct a specific model of cell division and death, and then estimate its parameters.

In this paper, we focus on the second approach. We first briefly describe a random “birth–death” model commonly used to describe turnover of immune cells and discuss its biological limitations (Section 3). We then formulate a more biologically realistic “Smith–Martin” (SM) model for cell division and death that lacks the limitations of the earlier random birth–death model (Section 4). We suggest three methods that may be used to estimate the parameters of the SM model (Section 5), discuss their limitations, and importantly propose additional experimental measures (such as the fraction of cells in division) that may improve the parameter estimates obtained. We then investigate the conditions under which the average division time of CFSE labelled cells can be estimated using the method suggested by [Gett and Hodgkin \(2000, Section 6\)](#). For comparison and illustration purposes, we apply all these different methods to a set of CFSE data, describing proliferation of naive CD8 T cells transferred into irradiated hosts (Section 7). All mathematical derivations are provided in the Appendix.

2. Description of CFSE data

A given set of CFSE data contains samples of the total number of cells $X(t)$ and the CFSE division profile at several time points. The CFSE division profile includes the fraction of cells $f_n(t)$ that have undergone n divisions by time t . The actual number of cells that have undergone n divisions by time t equals $X_n(t) = X(t)f_n(t)$. The mean number of divisions and the variance in the

number of divisions of labelled cells at time t , $\mu(t)$ and $\sigma^2(t)$, are given by

$$\mu(t) = \sum_{n=0}^{\infty} n f_n(t), \quad \sigma^2(t) = \sum_{n=0}^{\infty} n^2 f_n(t) - \mu^2(t). \tag{1}$$

After n divisions, each cell has given rise to 2^n progeny. Therefore, the number of precursors $p(t)$ that would have generated the current cell population in the absence of death (Gett and Hodgkin, 2000) is given by

$$p(t) = \sum_{n=0}^{\infty} X_n(t) 2^{-n}. \tag{2}$$

3. Quantifying cell turnover using random birth–death model

One particular model has been often used to quantify cell turnover in vitro and in vivo (Veiga-Fernandes et al., 2000; Bonhoeffer et al., 2000; Revy et al., 2001; Mohri et al., 2001; Asquith et al., 2002). In this model, cell division and death are modeled as independent random processes, and in this paper we refer to this model as the random birth–death (RBD) model. The model is based on the following assumptions: (i) the process of division is stochastic with probability rates of cell death d and commitment to cell division λ being constant; (ii) division occurs instantaneously (i.e., after a cell has divided, there is a non-zero chance for this cell to divide again in the next short period of time); (iii) there is no death during the division (Table 1).

In this model, changes in the number of cells that have undergone n divisions by time t are given by equations

$$\frac{dX_n(t)}{dt} = 2\lambda X_{n-1}(t) - (\lambda + d)X_n(t), \tag{3}$$

$n = 0, 1, 2, \dots, \infty,$

where $X_{-1}(t)=0$. Assuming that all cells at $t=0$ are undivided (i.e., $X_0(0)=X_0$ and $X_n(0)=0$ for $n>0$),

Table 1

A summary of definitions and parameters used in this paper

$X_n(t)$	The number of cells undergone n divisions by time t
$X(t) = \sum_{n=0}^{\infty} X_n(t)$	The total number of cells at time t
$f_n(t)=X_n(t)/X(t)$	The fraction of cell undergone n divisions by time t
$\mu(t) = \sum_{n=0}^{\infty} n f_n(t)$	The mean number of divisions cells have undergone at time t
$\sigma^2(t) = \sum_{n=0}^{\infty} n^2 f_n(t) - \mu^2(t)$	The variance of the number of divisions cells have undergone at time t
$p(t) = \sum_{n=0}^{\infty} X_n(t) 2^{-n}$	The number of cells that would give rise to the current population if there were no death (precursor number)
$\mu_2(t) = \sum_{n=0}^{\infty} n X_n(t) 2^{-n} / p(t)$	The mean number of divisions of cells in a normalized (by a factor 2^{-n}) population
λ	The rate of commitment to division (from the A-state to the B-phase)
Δ	The duration of division (B-phase)
d_A	The death rate in the A-state
d_B	The death rate in the B-phase
r	The rate of increase/decrease of the total population size $X(t)$
d	The rate of decrease of the precursor number $p(t)$
$T = \Delta + \lambda^{-1}$	The average division time in the SM model
T_1	The duration of the first division
$\delta = 1 - \lambda e^{-d_B \Delta} / (\lambda + d_A)$	The probability of cell's death during the cell cycle
$k = r + \lambda + d_A$	An accessory parameter
P	Proliferating fraction (the fraction of cells in division)
$r(a)$	The rate of increase of the rescaled (by a factor a) population

we find that the solution of the system (3) is given by

$$X_n(t) = \underbrace{\frac{(2\lambda t)^n}{n!}}_{\text{distribution}} e^{-2\lambda t} \underbrace{\left[X_0 e^{(\lambda-d)t} \right]}_{\text{total \#}}. \tag{4}$$

Importantly, in this model both the mean and variance of the number of divisions cells have undergone (calculated in accord with Eq. (1)) increase linearly with time with identical slopes equal to 2λ . This prediction is unlikely to be met for rapidly dividing cell populations for which variance increases slower than the mean (Gett and Hodgkin, 2000; Deenick et al., 2003, Section 7). Therefore, strictly speaking, the RBD model should

not be used to quantify the turnover of rapidly dividing cell populations.

We expect that for slowly dividing cell populations the mean and variance of the number of divisions will increase approximately at similar rates as predicted by the RBD model. However, an alternative model in which death can occur only during the division makes a similar prediction (Pilyugin et al., 2003). Such division-linked-death (DLD) model also assumes that the probability rate of commitment to cell division λ is constant and division is instantaneous. In contrast to the RBD model, the DLD model assumes that cells die only during the division; the probability of cell death during the division is α . Changes in the number of cells that have undergone n divisions by time t in this alternative model are given by equations (Pilyugin et al., 2003)

$$\frac{dX_n(t)}{dt} = 2\lambda(1 - \alpha)X_{n-1}(t) - \lambda X_n(t),$$

$$n = 0, 1, 2 \dots \infty, \tag{5}$$

where again $X_{-1}(t)=0$. Using the same initial conditions as for the RBD model, we find that the solution of the system (5) is given by

$$X_n(t) = \underbrace{\frac{(2\lambda(1 - \alpha)t)^n}{n!}}_{\text{distribution}} \underbrace{e^{-2\lambda(1-\alpha)t} \left[X_0 e^{\lambda(1-2\alpha)t} \right]}_{\text{total \#}}. \tag{6}$$

Similarly to the RBD model, in the DLD model the mean and variance of the number of divisions cells have undergone increase linearly with time with equal slopes. However, while in the RBD model this rate of increase equals 2λ and depends only on the rate of commitment to division λ , in the DLD model this rate equals $2\lambda(1 - \alpha)$ and depends on both the rate λ and probability of cell death during division α . Thus, because of these differences in predictions of the RBD and DLD models, we expect that in general, fitting the two models to the CFSE data will lead to different estimates of biologically relevant parameters such as the average cell division time (which in both models equals λ^{-1}). For example, for a turning over cell population of a constant size (i.e., when $\lambda=d$ in the RBD model and $\alpha=1/2$ in the DLD model), there will be a twofold difference in the average division time estimated in two models (Pilyugin et al., 2003). Since a priori it is not generally known if death occurs only either during the recruitment into division (RBD

model) or during the division (DLD model), neither model can be used with confidence to quantify cell turnover for slowly dividing cell populations. In the next section, we describe a more general Smith–Martin model of cell division and death that lacks these limitations of the earlier RBD model.

4. The Smith–Martin model

An example of a biologically reasonable specific model of cell division is given by the Smith–Martin model. Based on their quantitative study of the FLM (fraction of labelled mitoses) curves of dividing cell populations in vitro, Smith and Martin (1973) formulated an elegant quantitative description for the process of cell division that in fact is similar to the model developed by Burns and Tannock (1970) (Fig. 1). In this (Smith–Martin or SM) model, the progression of cells through the cell cycle involves a stochastic recruitment of cells from an A-state (corresponding approximately to the G_1 phase of the cell cycle) into the dividing B-phase (approximately equivalent to the S, G_2 , and M phases of the cycle). The B-phase has a fixed duration Δ . The recruitment of cells from the A-state into the B-phase occurs at the fixed rate λ (the waiting time in the A-state is exponentially distributed with the parameter λ). The two parameters λ and Δ provide a complete description of cell division (Cain and Chao, 1997a,b; Nordon

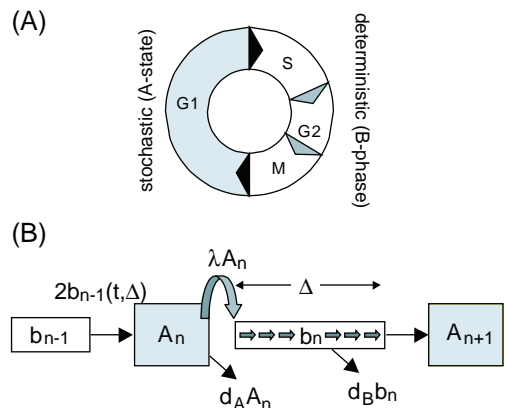


Fig. 1. A simplified representation of the Smith–Martin model (panel A), and generalized model accounting for the number of divisions cells have undergone (panel B). See main text for more details.

et al., 1999). The SM model is a reasonable first approximation for the process of cell division (for some critical considerations of the FLM data, see Castor, 1980; Brooks et al., 1980; Grasman, 1990).

In contrast with cell division, our understanding of processes that regulate cell death in different phases of the cell cycle is much poorer. We do not know if the death rate is constant during the entire cell cycle, or if the death rates are different for the A-state and the B-phase, or even if the death events are restricted to discrete checkpoints within the A-state and the B-phase of the cell cycle. Nevertheless, as a first approximation, we assume that the death rates in the A-state and the B-phase are given by two constants d_A and d_B , respectively (Fig. 1B).

With these assumptions we can describe the rates of change of cell numbers in the A-state and the B-phase having undergone different numbers of divisions by a set of differential equations. These equations are a simple extension of the model by Cain and Chao (1997a) to keep track of the number of divisions a cell has undergone and is similar to the model studied by Bernard et al. (2003):

$$A_n'(t) = 2b_{n-1}(t, \Delta) - (\lambda + d_A)A_n(t),$$

$$\frac{\partial b_n(t, s)}{\partial t} + \frac{\partial b_n(t, s)}{\partial s} = -d_B b_n(t, s), \quad (7)$$

where $A_n(t)$ and $B_n(t) = \int_0^\Delta b_n(t, s) ds$ are the number of cells having undergone n divisions in the A-state and the B-phase, respectively, and $b_n(t, s)$ is the time density of cells that have spent time s in the B-phase

($0 \leq s \leq \Delta$), $b_n(t, 0) = \lambda A_n(t)$, and $b_{-1}(t, s) = 0$. As noted previously, we assume the time- and division-independence of the parameters λ , Δ , d_A and d_B .

Analysis of the SM model, found in the Appendix (Section A.2), shows that asymptotically

1. The total number of cells, $X(t)$, changes exponentially (at the rate r),
2. The number of precursors, $p(t)$, declines exponentially (at the rate d),
3. The number of undivided cells, $X_0(t)$, declines exponentially (at the rate $\lambda + d_A$), and
4. Both the mean number of divisions, $\mu(t)$, and the variance in the number of divisions, $\sigma^2(t)$, increase linearly over time at rates a and b , respectively (Fig. 2).

The rates of change of the total population size, precursor number, the mean number of divisions and its variance are given by (see Sections A.2.1 and A.2.4 in the Appendix for derivations)

$$r = 2\lambda e^{-(d_B+r)\Delta} - (\lambda + d_A), \quad (8)$$

$$d = (\lambda + d_A) - \lambda e^{-(d_B-d)\Delta}, \quad (9)$$

$$\mu' = \frac{k}{1 + k\Delta}, \quad (10)$$

$$(\sigma^2)' = \frac{k}{(1 + k\Delta)^3}, \quad (11)$$

where $k = r + \lambda + d_A = 2\lambda e^{-(r+d_B)\Delta}$. Note that if $\Delta = 0$, then $\mu' = (\sigma^2)' = k = 2\lambda$. Thus, difference in the rates of

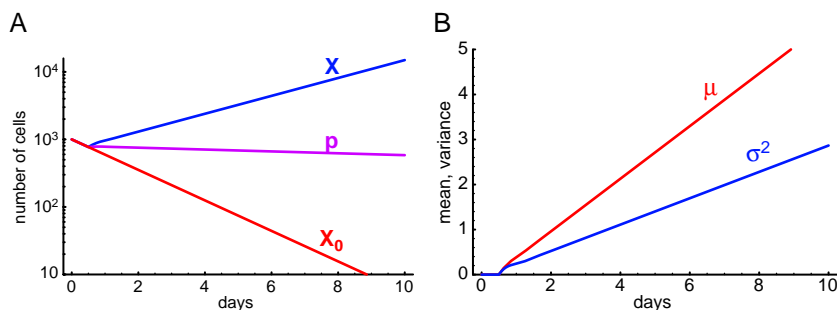


Fig. 2. The dynamics of several characteristics of the Smith–Martin model according to its numerical solution. Panel A shows the changes in the total number of cells X , the precursor number p , and the number of undivided cells X_0 with time. Panel B shows the changes in the mean μ and the variance σ^2 of the division number with time. Parameters used for the numerical solution are: $\lambda = 0.50 \text{ day}^{-1}$, $\Delta = 0.50 \text{ day}$, $d_A = 0.02 \text{ day}^{-1}$, $d_B = 0.08 \text{ day}^{-1}$, $A_0(0) = 10^3$, $b_0(s, 0) = 0$ (we assumed that at $t = 0$ all cells are in the A_0 -state). The average division time is $T = \Delta + \lambda^{-1} = 2.5 \text{ days}$. The rates given in Eqs. (8), (9), (10), and (11) are: $r = 0.31 \text{ day}^{-1}$, $d = 0.03 \text{ day}^{-1}$, $\mu' = 0.58 \text{ day}^{-1}$, and $(\sigma^2)' = 0.29 \text{ day}^{-1}$.

increase of the mean and variance of the number of divisions is a relative measure of the duration of the B-phase. Also note that r and d are not independent in the SM model, and that the decline rate of the number of undivided cells, $\lambda + d_A = k - r$, does not represent an independent parameter.

5. Estimating parameters of the SM model

The SM model involves four parameters: λ , Δ , d_A , and d_B that we would like to estimate. In this section, we suggest three methods that can potentially be used for such estimation.

5.1. Direct fitting

This approach is an attempt to solve system (7) and fit the solution to the CFSE data to infer the model parameters. The first step in this approach is to choose the correct initial conditions. Although it is natural to assume that all cells are initially undivided, the initial condition must also include the number of cells $A_0(0)$ in the A-state and the initial distribution of cells $b_0(s, 0)$ in the B-phase. While the total number of undivided cells at $t=0$, $X_0(0) = A_0(0) + \int_0^\Delta b_0(s, 0)ds$, is generally known, the distribution of cells in the B-phase is not. There are at least two experimental solutions to this problem. One solution is to synchronize cells before the experiment in the A-state. Then $b_0(s, 0) = 0$, and the analytical solution of the SM model in this case is given in the Appendix by Eq. (31). Another solution is to adoptively transfer cells, taken from one animal and labelled in vitro, into another animal assuming that cells are in the exponential phase of growth/decay and that such transfer does not change the cell turnover parameters. As we show in the Appendix (Section A.2.2), the initial conditions for the dynamics of CFSE labelled cells in the second animal are given by

$$A_0(0) = cX_0, \quad b_0(0, s) = \lambda c X_0 e^{(-d_B+r)s}, \quad (12)$$

where $c = \frac{r+d_B}{r+d_B+\lambda(1-e^{-(r+d_B)\Delta})}$ and X_0 is the total number of labelled cells at $t=0$. These initial conditions can be used to solve the SM model numerically.

The second step is to obtain reasonable parameter estimates and confidence intervals. While it might be possible to do for the first type of initial conditions

(when all cells are in the A-state at $t=0$) our previous analysis suggests that it is unlikely to be done using the second type of initial conditions (adoptive transfer) because of the degeneracy of the SM model during the exponential phase of growth. Indeed in the Appendix (Section A.3) we use the rescaling method to show that during the exponential growth phase only three parameter combinations and not the parameters themselves determine the dynamics of the SM model indicating degeneracy of the model during this phase.

5.2. Indirect fitting

Instead of trying to choose the right initial conditions, one can neglect the initial transients and simply assume that the population has reached the phase of exponential growth/decay. It is then important to determine when the population has reached this phase. As De Boer and Perelson (2005) pointed out for the case of the RBD model (described in Section 3), this phase is approximately reached when the mean number of divisions increases linearly with time. Indeed, our analysis suggests that during the phase of exponential growth/decay, in the SM model both mean and variance of the division number should increase linearly with time, while the total number of cells and the precursor number should change exponentially with time. Therefore, for the fitting procedure only data points that lie on a straight line should be used. For example, in Fig. 2 this phase is reached after approximately 12 h.

Importantly, however, our analysis suggests that during the exponential growth/decay phase, the SM model becomes degenerate since only three parameter combinations determine its dynamics. Thus, given the CFSE data alone, using Eqs. (10) and (11) one can only estimate the duration of the B-phase Δ and the accessory parameter k :

$$\Delta = \frac{1}{a} \left(1 - \sqrt{\frac{b}{a}} \right), \quad k = a \sqrt{\frac{a}{b}}, \quad (13)$$

where $a = \mu'$ and $b = (\sigma^2)'$. While death rates d_A and d_B cannot be determined using only CFSE data, we can nevertheless calculate the probability of cell death per cell cycle δ . The probability that a cell commits to division (i.e., leaves A-state and does not die) is $\lambda/$

$(\lambda + d_A)$ and the probability that a cell successfully divides and does not die is $e^{-d_B \Delta}$. Then the probability of cell death during a cell cycle is given by

$$\delta = 1 - \frac{\lambda e^{-d_B \Delta}}{\lambda + d_A} = 1 - \frac{k e^{r \Delta}}{2(k - r)}, \quad (14)$$

where k , Δ , and r are calculated from the CFSE data. Furthermore, by assuming that death during the cell cycle occurs only in the A-state (i.e., $d_B = 0$) or the B-phase (i.e., $d_A = 0$), using Eqs. (8), (10), and (11) one can also estimate the range for the parameters λ , d_A , d_B and the average division time $T = \Delta + \lambda^{-1}$ where all parameters are positive:

$$\lambda \in \left(\frac{k e^{r \Delta}}{2}, k - r \right), \quad (15)$$

$$d_A \in \left(0, k - r - \frac{k e^{r \Delta}}{2} \right), \quad (16)$$

$$d_B \in \left(0, \frac{\ln[2(1 - r/k)]}{\Delta} - r \right), \quad (17)$$

$$T \in \left(\Delta + \frac{1}{k - r}, \Delta + \frac{2e^{-r \Delta}}{k} \right). \quad (18)$$

Importantly, all four parameters can be estimated if an additional experimental measurement is available, such as the fraction of cells in the B-phase of the cell cycle. Approximately, this corresponds to the fraction of cells in the S+G₂+M phase of the cell cycle. The proliferation fraction can be estimated either by measuring the DNA content per cell (with cells in division having more DNA than resting cells) or by a two parameter, RNA–DNA content analysis. In essence, commitment to division is associated with upregulation of RNA expression followed by DNA synthesis (Pera et al., 1977; Bohmer, 1979; Kubbies et al., 1992; Stout and Suttles, 1992; Veiga-Fernandes et al., 2000). As we show in the Appendix (Section A.2.3), asymptotically the proliferation fraction P is given by

$$P = \frac{k(e^{(d_B+r)\Delta} - 1)}{2(d_B + r) + k(e^{(d_B+r)\Delta} - 1)}. \quad (19)$$

If the proliferation fraction P is measured, then all parameters but the death rate d_B are known in

Eq. (19). By solving this equation (graphically or numerically), we estimate d_B ; the remaining parameters of the SM model can be calculated by substituting the estimated parameters into the expressions:

$$\lambda = \frac{k}{2} e^{(d_B+r)\Delta}, d_A = k - \lambda - r. \quad (20)$$

5.3. Rescaling method

Pilyugin et al. (2003) have suggested a rescaling method that can be used to estimate parameters of any specific model. In this method, the number of cells undergone n divisions at time t is multiplied by a factor a^n where the rescaling parameter a is constant. For this “normalized” population the total number of cells $X(t, a) = \sum_{n=0}^{\infty} X_n(t) a^n$ is calculated. For example, for the data presented in Table 3, at time $t=3$ days we rescaled the number of cells that have undergone n divisions by $a=1/2$. The rescaled numbers are $1.77 \cdot 10^4 \cdot (2 \cdot 1)^{-0}$, $6.10 \cdot 10^4 \cdot (2 \cdot 1)^{-1}$, $6.58 \cdot 10^4 \cdot (2 \cdot 1)^{-2}$, and $1.28 \cdot 10^4 \cdot (2 \cdot 1)^{-3}$ for cells undergone 0, 1, 2, and 3 divisions, respectively. The total number of cells of this normalized population at $t=3$ days is then $X(t=3, a=1/2) = 1.77 \cdot 10^4 + 6.10 \cdot 10^4/2 + 6.58 \cdot 10^4/4 + 1.28 \cdot 10^4/8 \approx 6.63 \cdot 10^4$. Calculating similarly the total number of cells for other times while keeping a fixed (in the example at 1), we then can determine the rate of exponential increase of the total number of cells $X(t, a)$ for a given a . The rate of increase is calculated by linear regression of the natural logarithm of the total population size $\ln X(t, a)$ versus time (see Fig. 4A). Doing similarly for several values of a (for example in the range from 0 to 2 with the step 0.01), we calculate the experimental dependence of the rate of increase of the total population size $\mathbf{r}(a)$ on the rescaling factor a .

As we show in the Appendix (Section A.3), by rescaling the SM model, we find an analytical expression for the rate of increase $\mathbf{r}(a)$ as the function of the rescaling factor a

$$\mathbf{r}(a) = 2a\lambda e^{-\mathbf{r}(a)+d_B\Delta} - (\lambda + d_A). \quad (21)$$

Now one could fit this formula to the experimentally obtained dependence $\mathbf{r}(a)$ to estimate the parameters of the SM model. However, similarly to the indirect fitting, only the duration of the B-phase Δ

and the range for the other parameters including the average division time can be estimated from the CFSE data alone (see Section A.3 in the Appendix).

Measuring an additional parameter, the fraction of cells in the B-phase of the cell cycle, will allow estimation of all four parameters of the SM model. For this we rewrite Eq. (21) using the known value for the proliferating fraction P (see Section A.3 in the Appendix for derivations):

$$\mathbf{r}(a) = 2a \cdot \frac{P(r + d_B)e^{-(r(a)+d_B)\Delta}}{(1 - P)(1 - e^{-(r+d_B)\Delta})} - \left(d_A + \frac{P(r + d_B)}{(1 - P)(1 - e^{-(r+d_B)\Delta})} \right), \quad (22)$$

where $r = \mathbf{r}(1)$ is the rate of exponential increase of the total population size $X(t)$. This formula can be fitted to the experimentally obtained dependence $\mathbf{r}(a)$ to obtain estimates for the parameters Δ , d_A and d_B . Then using Eq. (55), the remaining parameter λ can be also estimated:

$$\lambda = \frac{P(r + d_B)}{(1 - P)(1 - e^{-(r+d_B)\Delta})}. \quad (23)$$

6. Estimating the average division time: the Gett–Hodgkin method

Instead of trying to estimate parameters of specific models, Gett and Hodgkin (2000) proposed an elegant and simple way to estimate the average division time as well as the time of first division for CFSE labelled cells. Shortly, in this method the number of cells $X_n(t)$ that have undergone n divisions by time t is divided by 2^n (Gett and Hodgkin, 2000). For this “normalized” population, the average number of divisions cells have undergone is calculated using the formula (De Boer and Perelson, 2005):

$$\mu_2(t) = \frac{\sum_{n=0}^{\infty} nX_n(t)2^{-n}}{\sum_{n=0}^{\infty} X_n(t)2^{-n}}. \quad (24)$$

Then, the average division time T of dividing cells is estimated as the reciprocal of the slope of $\mu_2(t)$ and the time of first division T_1 is calculated from the equation $\mu_2(T_1) = 1$ (Fig. 3).

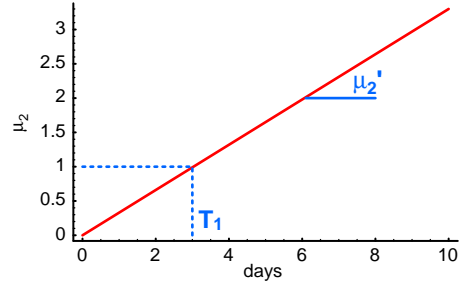


Fig. 3. A cartoon of the change of the average number of divisions $\mu_2(t)$ in the normalized by 2^{-n} population (see main text for details). An estimate of the average division time is $1/(\mu_2')$ and an estimate of the time to first division T_1 is found in the equation $\mu_2(T_1) = 1$. In the cartoon, $\mu_2(t) = 0.33t$; then $(\mu_2') = 0.33 \text{ day}^{-1}$, $T_1 = 3$ day.

Although this method of estimation of the average division time is intuitively appealing, we have sought to determine whether the inverse of the slope for $\mu_2(t)$ actually corresponds to the average division time $T = \Delta + \lambda^{-1}$ when we use the SM model for cell division and death. In the Appendix (Section A.2.4) we have shown that after initial transient μ_2 increases linearly with time (see De Boer and Perelson (2005) on how the length of this transient can be calculated for some simple models). The asymptotic rate of increase $(\mu_2')^{-1}$ is given by

$$(\mu_2')^{-1} = \Delta + \frac{1}{\lambda + d_A - d}, \quad (25)$$

where d is the decline rate of precursors given in Eq. (9). Given that all parameters are positive, we immediately see that $(\mu_2')^{-1}$ corresponds to the true average division time only when $d = d_A$ which happens only when the death rates in the A-state and the B-phase are equal, i.e. at $d_A = d_B$. As we show in Table 2, for a population of constant size (i.e., $r = 0$), the error in estimation of T can be as large as twofold if cells spend most of their time in the A-state (i.e., if $\lambda^{-1} \gg \Delta$) but may be relatively small if cells spend most of their time in the B-phase (i.e., if $\lambda^{-1} \ll \Delta$). The errors become more pronounced for declining populations ($r < 0$) and less pronounced for growing populations ($r > 0$, not shown). Since for rapidly dividing cells, the G_1 phase of the cell cycle is relatively short (i.e., $\Delta > \lambda^{-1}$), we expect the Gett–Hodgkin method to give reasonable

Table 2

The estimates of the average division time obtained by the Gett–Hodgkin method for different parameter combinations of the SM model given in Eq. (25)

(μ_2^{-1})	$d_A=0$	$d_A=d_B$	$d_B=0$
$\lambda^{-1} \gg \Delta$	3.80	2.00	1.91
$\lambda^{-1} = \Delta$	2.46	2.00	1.64
$\lambda^{-1} \ll \Delta$	2.00	2.00	1.95

Three parameter combinations are considered: The average recruitment time is much greater than the division time ($\lambda=0.53 \text{ day}^{-1}$, $\Delta=0.1 \text{ day}$), recruitment time is equal to the division time ($\lambda^{-1}=\Delta=1 \text{ day}$), and recruitment time is much shorter than the division time ($\lambda=10 \text{ day}^{-1}$, $\Delta=1.9 \text{ days}$). In all cases, the average division time $T=\Delta+\lambda^{-1}=2 \text{ days}$. We also assumed the population of a constant size (i.e., with $r=0$), implying the following equality for the parameters of the SM model: $\lambda+d_A=2\lambda e^{-d_B\Delta}$.

estimates for the average division time of rapidly proliferating cell populations.

7. Example: proliferation of naive CD8 T cells in irradiated hosts

In this section, we apply the methods for parameter estimation to a set of experimental data. In doing so we pursue two major goals. Firstly, we demonstrate how these methods can be used to estimate parameters of the SM model. Secondly, by using experimental data which are “noisy” we also illustrate the limitations of these methods. It is also important to note that at present this procedure of parameter estimation is rather illustrative since some of the assumptions for these methods such as homogeneity of the cell population, cannot be validated (see Discussion).

The particular dataset that we use here has been obtained by Murali-Krishna and Ahmed (2000) in a study of the dynamics of CFSE labelled P-14 transgenic naive CD8 T cells after adoptive transfer into irradiated hosts (Table 3). In this experiment, naive T cells were collected from one animal, labelled with the CFSE in vitro, and then transferred into irradiated animals lacking T cells (T cells had been destroyed by irradiation). Following the transfer, the labelled T cells began to divide and their total number increased with time. In Fig. 4 we plot the changes in the total number of cells, $X(t)$, the precursor number $p(t)$, the mean

number of divisions, $\mu(t)$, and the variance in the number of divisions, $\sigma^2(t)$, over time.

It is generally believed that in normal conditions naive T cells are resting nondividing cells in the G_0/G_1 phase of the cell cycle. For direct fitting we therefore assumed that all labelled cells were synchronized in the A-state at $t=0$ (i.e., cells are approximately in the G_1 phase of the cell cycle). For the fitting, we used the analytical solution of the SM model found for this type of initial conditions (given in the Appendix in Eq. (31)) to CFSE data. Since in the adoptive transfer experiments the actual number of T cells engrafted by the host is not generally known, we estimated this number by calculating the intercept of the linear regression of the logarithm of the total size vs. time (see Fig. 4A). The estimated value $A_0(0)=6.1 \cdot 10^4$ is approximately equal to 10% of the transferred CD8 T cells that agrees well with estimates on engraftment of naive CD8 T cells into normal hosts (Blattman et al., 2002).

The estimation of parameters was done by minimization of the residual sum of squares using the Levenberg–Marquardt algorithm implemented in Mathematica (Wolfram, 1990). The 67% confidence intervals (CIs) were determined by a bootstrap method with 1000 resamplings of the residuals (Efron and Tibshirani, 1993). As shown in Table 4, for this particular dataset, we could obtain biologically reasonable estimates and confidence intervals for the commitment rate λ , the duration of the B-phase Δ and the death rate in the A-state d_A . The estimate of the death rate in the B-phase d_B had large CIs. We suggest two explanations of this result. First, this dataset might be too small to give reasonably good estimates for all four parameters. Second, the model might be

Table 3

The CFSE profile of P-14 Tg naive CD8 T cells after adoptive transfer into irradiated hosts (Murali-Krishna and Ahmed, 2000)

	t (days)				
	$X_n(t)$	0.5	1.25	3	8
n (divisions)	0	7.38	7.07	1.77	0.0
	1	0.0	0.64	6.10	0.29
	2	0.0	0.0	6.58	5.71
	3	0.0	0.0	1.28	19.97
	4	0.0	0.0	0.0	18.83
	5	0.0	0.0	0.0	7.99
	6+	0.0	0.0	0.0	4.00

The numbers above equal the number of cells per spleen divided by 10^4 .

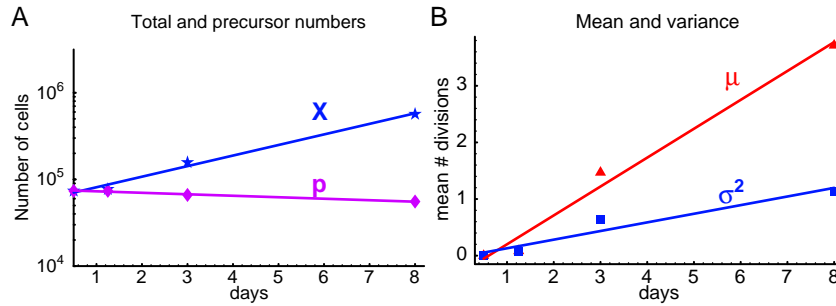


Fig. 4. Changes in several characteristics of P-14 Tg naive CD8 T cells after adoptive transfer into irradiated hosts (Murali-Krishna and Ahmed, 2000). Panel A shows the change in the total number of Tg cells X (stars) and precursors p (diamonds). Panel B shows the change in the mean μ (triangles) and the variance σ^2 (boxes) of the number of divisions cells have undergone. Solid lines represent the best regression lines, the slopes of the regressions are $r=0.28 \text{ day}^{-1}$, $d=0.04 \text{ day}^{-1}$, $a=\mu'=0.51 \text{ day}^{-1}$, $b=(\sigma^2)'=0.15 \text{ day}^{-1}$.

overdefined with respect to the given CFSE data and the form of the initial conditions that we used, i.e., several different parameter combinations may fit the data reasonably well. While we could not directly test either of these explanations, in support of the second explanation, we found that fitting Eq. (31) to the data generated from it (i.e., from Eq. (31)) with added noise, may lead to incorrect and imprecise parameter estimates obtained for both death rates d_A and d_B even if the introduced errors are relatively small ($\sim 5\text{--}10\%$, not shown). Importantly, however, the estimates of the duration of the B-phase and the rate of commitment to division were generally close to the expected values and had relatively small CIs.

It is also worthwhile to note that Bernard et al. (2003) used the direct fitting of a similar analytical

solution to estimate only three parameters of the SM model while assuming some small ad hoc value for the death rate in the A-state (Bernard et al., 2003). Indeed, we also found that reducing the number of parameters of the SM model (by assuming $d_A=0$, $d_B=0$, or $d_A=d_B$) may allow estimation of the remaining parameters with relatively small confidence intervals (not shown). How making these assumptions on the death rates affects estimates for other parameters will be investigated elsewhere. Finally, using the obtained estimates for the parameters using Eq. (14) we could also estimate the probability of cell's death per cell cycle and CIs for it (see Table 4).

Using either indirect fitting (Method 2) or the rescaling method (Method 3) one can only estimate the duration of the B-phase (Table 4). For other

Table 4

The estimates for the parameters of the SM model obtained by direct fitting of the analytical solution of the model (Method 1), indirect fitting (Method 2), and the rescaling method (Method 3)

Parameters	Direct fitting		Indirect fitting		Rescaling method	
	Method 1		Method 2		Method 3	
	Mean	67% CIs	Mean	67% CIs/range	Mean	67% CIs/range
λ , day^{-1}	1.57	(0.62, 1.49)	–	(0.60–0.65)*	–	(0.67, 0.74)*
A , day	1.13	(0.92, 1.25)	0.89	(0.63, 1.20)	0.97	(0.67, 1.46)
d_A , day^{-1}	0.00	(0.0, 0.0)	–	(0–0.05)*	–	(0–0.07)*
d_B , day^{-1}	0.09	(0.02, 0.16)	–	(0–0.10)*	–	(0–0.10)*
T , day	2.31	(2.03, 2.58)	–	(2.42–2.56)*	–	(2.33–2.47)*
δ	0.09	(0.01, 0.15)	0.08	–	0.09	(0.08, 0.11)

For the Method 1, confidence intervals were calculated using a bootstrap method with 1000 resamplings of the residuals. For the Method 2 the confidence intervals are calculated using standard methods for linear regressions. For the Method 3, confidence intervals are calculated using standard methods for linear regressions where the errors in the estimated values for the rate of increase $r(a)$ are also taken into account. The 67% confidence intervals are shown in brackets with a comma. The ranges for parameters found by indirect fitting or rescaling method are shown in brackets with a dash. The estimate of the average division time using the Gett–Hodgkin method is $(\mu'_2)^{-1}=2.39 \text{ days}$ (2.22, 2.60).

* Ranges for the relevant parameters.

parameters, however, we can find the ranges where all the model parameters are positive. Interestingly, for this particular data set, these ranges are relatively small and are very similar to the estimates obtained with direct fitting. Also, the estimates for the average division time T obtained by all three methods are also very close to the estimate obtained using the Gett–Hodgkin technique which gives $(\mu_2^{-1}) = 2.39$ days with 67% confidence intervals (2.22, 2.60). Since the decline rate of the precursor number is small ($dT \approx 0.1 \ll 1$), the Gett–Hodgkin method gives a reasonably good estimate of the average division time T (see Section 6). Finally, using both indirect fitting and rescaling method, we estimate the probability of cell death during the cell cycle $\delta \approx 0.09$; only the rescaling method gave relatively small CIs for this parameter.

Thus, we found that naive transgenic CD8 T cells transferred into irradiated hosts proliferate with the average division time of 2.4 days with approximately 1 day spent on actual division (i.e., in S+G₂+M phase of the cell cycle). We found that approximately 10% of cells are dying per each division with death predominantly occurring during the division phase. Additional data, such as the fraction of cells in division, is required to verify the last observation.

8. Discussion

8.1. Quantifying cell turnover

The CFSE label is widely used in immunology to quantify the number of divisions immune cells have undergone (Lyons, 2000). In this paper, we have analyzed how the data obtained using the CFSE label, can be used to estimate the rates of cell division and death. There are two basic approaches in addressing this problem.

The first is to formulate a general model of cell division and death where the probabilities of cell division and cell death are general functions dependent only on time since last division. Using such a model, we have previously shown how parameters that are independent of specific functions describing cell division and death, can be estimated from the CFSE data (Pilyugin et al., 2003). The main problem with this approach, however, is that the invariant

parameters provide limited insight into the specifics of cell division and death.

This problem can be resolved by the second approach in which one can formulate a specific model for cell division and death and then estimate its parameters using CFSE data. In this paper, we extended the model of cell division pioneered by Smith and Martin (1973) and Burns and Tannock (1970) to include cell death during the cell cycle and to track the number of division cells have undergone.

We focused our analysis on the case of turnover of homogeneous cell populations, i.e., populations with division and death rates independent of time and the numbers of divisions cells have undergone. However, even in this simplest case, the problem of estimation of parameters determining cell division and death appears to be much harder than was previously thought.

Given CFSE data alone, we found no general robust method that allows estimation of all parameters of the SM model. It is, however, possible to estimate the duration of division Δ , probability of cell death during the cell cycle δ , and in some cases the average division time T (Table 5). Measuring an additional parameter, the fraction of cells in division P (i.e., fraction of cells that are in the S+G₂+M phase of the cell cycle) may allow estimation of all parameters of the SM model (Table 6). The fraction of cells in the S+G₂+M phase of the cell cycle can be estimated by looking at either DNA content per cell (for example, using PI staining) or by a two parameter, RNA–DNA content analysis (Pera et al., 1977; Bohmer, 1979; Kubbies et al., 1992; Stout and Suttles, 1992; Veiga-Fernandes et al., 2000).

8.2. Caveats

8.2.1. Inhomogeneity of cell populations

Importantly, the problem of estimation of the division and death rates of CFSE labelled cells may become even harder for inhomogeneous cell populations. In general, two types of inhomogeneity of cell populations may occur: when the parameters determining cell division and death are dependent on (i) time or (ii) the number of divisions a given cell has undergone.

The dependence of parameters determining cell division and death on time may arise when factors extrinsic to CFSE labelled cells affect their rates of

Table 5

Recommendation for the estimation of parameters of the SM model when only CFSE data are available

-
- (1) If the initial conditions for the population are known, use the analytical or numerical solutions of the SM model (7) to estimate its parameters by fitting the model to the data. Note that this approach may give large confidence intervals for some parameters for “noisy” data.
 - (2) Otherwise, only the duration of the B-phase Δ , the probability of cell death per cell cycle δ and at some conditions (see below) the average division time T can be determined. To do so,
 - estimate the instantaneous rate of increase of the total population size r , the decline rate of the precursor number d , and the rate of increase in the mean number of divisions and its variance $\mu'(t)=a$ and $(\sigma^2)'(t)=b$, respectively (see Fig. 4A and B). From the slopes a and b , and the rate of increase r estimate the duration of the B-phase Δ , the probability of cell death per cell cycle δ , and the range for the remaining parameters λ , d_A , d_B and the average division time T given in Eqs. (13)–(18).
 - if $\Delta \approx T$ or there is little death occurring (i.e., $dT \ll 1$), then use the Gett–Hodgkin method to estimate the average division time T from the rate of increase of the mean number of divisions of the normalized CFSE data using Eq. (24).
 - (3) In all other cases, use the rescaling method to estimate invariant parameters τ and δ , where τ is the average division time of surviving cells and δ is the probability of cell death during the cell cycle. In this case, τ will be an underestimate of T (see Pilyugin et al. (2003) for more details).
-

division and death in a time-dependent fashion. For example, adoptive transfer of CFSE labelled cells into naive hosts generally involves some degree of local damage that may result in temporary changes in levels of some cytokines (or other extracellular proteins) leading to time-dependent changes of the parameters for cell division and death.

The dependence of parameters determining cell division and death on the number of divisions cells have undergone may arise when factors intrinsic to CFSE labelled cells affect their rates of division and death. For example, it is well established that some cell properties such as abundance of different molecules on the cell surface change as cell undergo more divisions (Hodgkin et al., 1996; Lyons, 2000; Ma et al., 2004; Tangye and Hodgkin, 2004). It is possible that changes in the expression of some genes (Fas, for example) may well lead to changes in the rates of cell division and death depending on how many divisions cells have undergone.

One particular type of inhomogeneity has received more attention. It was noted, that in some experiments

where proliferation of CFSE labelled cells is followed after stimulation in vitro, time required to complete the first division appears to be larger than the time of subsequent divisions (Gett and Hodgkin, 2000). Gett and Hodgkin (2000) proposed a method of estimating the average division time of proliferating cells and the time of first division; some of the limitations of this method for the estimation of both parameters have also been investigated (De Boer and Perelson, 2005, Section 6).

However, to our knowledge, little work has been done in developing general methods that allow detection of general types of inhomogeneity. The properties of the SM and a more general model suggest that in a homogeneous cell population mean and variance of the number of divisions cells have undergone, after initial transient should increase linearly with time (Pilyugin S.S., unpublished). Observing departures from such a linear increase would imply inhomogeneity of the cell population.

Experimentally, inhomogeneity of dividing cells can be detected using adoptive transfer of cells undergone n division at time t (where both division number and time should vary) into new hosts and detecting CFSE profiles of transferred cells. Significant differences in CFSE profiles generated from cells undergone, for example, 1 and 5 divisions would imply that population is indeed inhomogeneous. This method, however, is clearly time and resource consuming and may not be available to perform in most situations. Developing more robust theoretical

Table 6

Recommendation for the estimation of parameters of the SM model when both CFSE data and the proliferating fraction P are available

-
- (1) Estimate the instantaneous rate of increase of the total population size r and the rate of increase in the mean and variance of the number of divisions $\mu'(t)=a$ and $(\sigma^2)'(t)=b$, respectively (see Fig. 4). From the slopes a and b , estimate the duration of the B-phase Δ and the parameter k given in Eq. (13). Given the known value for the fraction of cells in division P , using Eqs. (19) and (20), estimate the remaining parameters of the SM model.
 - (2) Alternatively, use the rescaling method as outlined in Section 5. For this, using CFSE data, calculate the rate of increase $\mathbf{r}(a)$ of the total population size $X(t)$ rescaled by a parameter a where a varies from 0 to 2 with some step (for example, 0.01). Then, use the analytical expression for the rate of increase $\mathbf{r}(a)$ given in Eq. (22) to estimate Δ , d_A , d_B , and then use Eq. (23) to estimate λ .
-

and experimental methods of detecting inhomogeneity of cell populations is an important area that should be addressed in future studies.

8.2.2. Limitations of the SM model

In our analysis we assumed that division of lymphocytes can be well approximated by the SM model. While it might be so to a first approximation (Smith and Martin, 1973; Shields and Smith, 1977), additional studies have suggested that changing some of the assumptions of the SM model is required to describe new data (Castor, 1980; Brooks et al., 1980; Tyson and Hannsgen, 1985; Brooks and Riddle, 1980; Cain and Chao, 1997a). For example, several studies have suggested that two transition probabilities are required to describe cell division, especially for slowly dividing cells (Brooks et al., 1980; Brooks and Riddle, 1980). Other studies have proposed that the deterministic phase should include another stochastic state to describe variation in the duration of the division phase (Cain and Chao, 1997a). The SM model also assumes that death rates are uniformly distributed in the A-state and B-phase. While the assumption of the constant death rate may be correct for the A-state, it may be not so for the B-phase where death events may be restricted to particular checkpoints. How changing these assumptions may affect the parameter estimates obtained needs to be investigated.

8.3. Previous work

There has been a number of theoretical studies estimating the rates of cell turnover using the CFSE label (Nordon et al., 1999; Gett and Hodgkin, 2000; Veiga-Fernandes et al., 2000; Revy et al., 2001; Bernard et al., 2003). While these papers employed different models and methods for parameter estimation, there was little emphasis on the problems associated with this task. For example, Veiga-Fernandes et al. (2000) and Revy et al. (2001) used a random birth–death model to analyze survival and proliferation of T cells after stimulation with specific antigens. We have shown that this simple model may give incorrect estimates for parameters determining cell division and death (Pilyugin et al., 2003, Section 3). Nordon et al. (1999) have analyzed a particular case of the SM model with no death during division (i.e., with $d_B=0$) that allowed the authors to estimate

three remaining parameters using a similar to ours indirect fitting method. Finally, Bernard et al. (2003) have studied a general SM model and obtained estimates for three model parameters by direct fitting of an analytical solution to the data assuming a fixed (small) death rate in the A-state. We have also observed that reducing the number of parameters of the SM model may allow estimation of the remaining parameters with relatively small confidence intervals (not shown). However, since a priori we cannot make such simplifications (without invoking additional data), estimation of all parameters of the SM model using CFSE data alone appears to be problematic.

8.4. Concluding remarks

Although CFSE labelling technique provides the information on the number of divisions a given cell has undergone, in some biologically relevant cases it is sufficient to estimate the division but not the death rates of CFSE labelled cells. The rescaling method can be used to infer the invariant parameters of the cell cycle (such as the probability of cell death during the cell cycle and the average division time of surviving cells) but these parameters tell us little about the nature of cell division and death. Thus, additional data such as the fraction of cells in division are required if we are to estimate both the rates of cell division and death. This requirement should eventually stimulate future collaborations between theoreticians and experimentalists.

Acknowledgements

This work was supported by National Institutes of Health grant R01-AI-49334 to RA. We thank Alan Perelson and Roland Regoes for many helpful discussions.

Appendix A

A.1. Analytical solutions of the SM model

In this section, we present the explicit solution of the Smith–Martin model (7) with special initial conditions. Specifically, we assume that at $t=0$ all

cells are in the A-state of the 0-th generation, that is, $A_0(0)=X_0>0$ and $b_0(s,0)=0$ ($A_n(0)=b_n(s,0)=0$ for $n>0$). Biologically this may occur when all labelled cells are synchronized in the A-state at the beginning of the experiment.

For each generation number $n \geq 0$, we use method of characteristics and solve for $B_n(t,s)$ in terms of $A_n(t)$ to obtain

$$b_n(t,s) = \lambda A_n(t-s)\exp(-d_B s), \quad 0 \leq s \leq \Delta. \quad (26)$$

Substituting Eq. (26) into Eq. (7), we obtain the equations for the dynamics of cells in the A-state

$$A_n'(t) = 2\lambda \exp(-d_B \Delta) A_{n-1}(t-\Delta) - (\lambda + d_A) A_n(t), \quad (27)$$

Given the initial conditions, the Laplace transforms of $A_n(t)$ must therefore satisfy

$$\mathcal{L}\{A_n(t)\} = \frac{X_0 2^n \lambda^n e^{-(s+d_B)n\Delta}}{(s + \lambda + d_A)^{n+1}}, \quad n \geq 0. \quad (28)$$

Inverting the Laplace transform in Eq. (28), we obtain the explicit solutions for the number of cells in the A-state for each generation

$$A_n(t) = X_0 \left(2\lambda e^{(\lambda+d_A-d_B)\Delta} \right)^n e^{-(\lambda+d_A)t} \frac{(t-n\Delta)^n}{n!} \times H(t-n\Delta). \quad (29)$$

where $H(x)$ is the Heaviside step function (i.e., $H(x)=0$ if $x<0$ and $H(x)=1$ otherwise). The number of cells in the B-phase in each generation is then:

$$B_n(t) = X_0 \frac{2^n \lambda^{n+1} e^{-d_B t}}{n!(\lambda + d_A - d_B)^{n+1}} \cdot \Gamma(n+1, \max(0, t-(n+1)\Delta) \cdot (\lambda + d_A - d_B), \max(0, t-n\Delta) \cdot (\lambda + d_A - d_B)), \quad (30)$$

where $\Gamma(n+1, a, b) = \int_a^b z^n e^{-z} dz$ is incomplete generalized gamma function. The total number of cells $X_n(t)$ in the n -th generation at time t can be determined by substituting Eqs. (29) and (30) into

$$X_n(t) = A_n(t) + B_n(t). \quad (31)$$

Note that this solution is similar to one obtained by Bernard et al. (2003). Having the general solution

(31), we can now easily calculate the dynamics of undivided cells $X_0(t)$:

$$X_0(t) = \begin{cases} X_0 e^{-(\lambda+d_A)t} + \frac{\lambda A_0}{\lambda+d_A-d_B} e^{-(\lambda+d_A)t} (e^{(\lambda+d_A-d_B)t} - 1), & \text{if } t < \Delta \\ X_0 e^{-(\lambda+d_A)t} + \frac{\lambda A_0}{\lambda+d_A-d_B} e^{-(\lambda+d_A)t} (e^{(\lambda+d_A-d_B)\Delta} - 1), & \text{otherwise} \end{cases} \quad (32)$$

A.2. Properties of the SM model

A.2.1. Exponential rate of increase r and decline rate of precursors d

Summing over n in the model (7) we derive the following equation for the change of the total number of cells in the A-state:

$$A'(t) = 2\lambda e^{-d_B \Delta} A(t-\Delta) - (\lambda + d_A) A(t). \quad (33)$$

Since the time-delay is fixed and Eq. (33) is linear, we conclude that any solution of Eq. (33) is asymptotic to an exponential function $C \exp(rt)$ where r is the principal eigenvalue and C is a constant determined by the initial conditions (Bellman and Cooke, 1963). The principal eigenvalue r is the unique real root of the characteristic equation

$$r = 2\lambda e^{-(d_B+r)\Delta} - (\lambda + d_A). \quad (34)$$

Similarly, we can normalize model (7) by dividing the number of cells that have undergone n divisions by 2^n . Summing over n , we obtain the following equation for $A_N(t) = \sum_{n=0}^{\infty} A_n(t) 2^{-n}$:

$$A_N'(t) = \lambda e^{-d_B \Delta} A_N(t-\Delta) - (\lambda + d_A) A_N(t). \quad (35)$$

Similarly to the previous case, the solution to this equations is given by $C' \exp(-dt)$, where d is the unique real root of the characteristic equation:

$$d = (\lambda + d_A) - \lambda e^{-(d_B-d)\Delta}. \quad (36)$$

A.2.2. Asymptotic steady state distribution of cells in the A-state and the B-phase

Using Eq. (26) we find that the time density of cells that spent time s in the B-phase $b(t,s) = \sum_n b_n(t,s)$ is

$$b(t,s) = \lambda A(t-s)\exp(-d_B s), \quad 0 \leq s \leq \Delta. \quad (37)$$

Asymptotically $A(t)$ enjoys the exponential growth pattern, $A(t) = C \exp(rt)$, and therefore

$$b(t, s) = \lambda A(t) \exp(- (d_B + r)s). \tag{38}$$

Now if X_0 cells are taken from the population in the exponential phase of growth, are labelled and adoptively transferred into a new host then

$$\begin{aligned} X_0 &= A(0) + \int_0^{\Delta} b(0, s) ds \\ &= C \left(1 + \frac{\lambda(1 - e^{-(d_B+r)\Delta})}{d_B + r} \right), \end{aligned} \tag{39}$$

and the number of cells in the A-state and the time density of cells in the B-phase in the new host is

$$\begin{aligned} A_0(0) &= \frac{X_0(r + d_B)}{r + d_B + \lambda(1 - e^{-(r+d_B)\Delta})}, \\ b_0(0, s) &= \frac{\lambda X_0(r + d_B)}{r + d_B + \lambda(1 - e^{-(r+d_B)\Delta})} e^{-(d_B+r)s}. \end{aligned} \tag{40}$$

A.2.3. Proliferating fraction P

Since asymptotically $A(t)$ enjoys the exponential growth pattern, that is, $A(t) = C \exp(rt)$, the total number of cells in the B-phase can also be calculated:

$$B(t) = \int_0^{\Delta} \lambda A(t - \tau) \exp(- d_B \tau) d\tau = A(t) \int_0^{\Delta} F(\tau) d\tau,$$

where $F(\tau) = \lambda e^{-(r+d_B)\tau}$. Since the total population size, $X(t) = A(t) + B(t)$, is also increasing exponentially, we can calculate the constant fraction of cells in the B-phase (i.e., in division):

$$P = \frac{\int_0^{\Delta} F(\tau) d\tau}{1 + \int_0^{\Delta} F(\tau) d\tau} = \frac{\lambda(1 - e^{-(r+d_B)\Delta})}{r + d_B + \lambda(1 - e^{-(r+d_B)\Delta})}, \tag{41}$$

and by substituting $\lambda = \frac{k}{2} e^{-(r+d_B)\Delta}$ we obtain Eq. (19).

A.2.4. Increase of the mean number of division

We define $f_n^A(t) = (A_n(t)/A(t))$, and let $\mu^A(t)$ and $(\sigma^2)^A(t)$ be the mean and the variance of the

distribution generated by $f_n^A(t)$. In general, the rate of change of $f_n^A(t)$ is given by

$$\begin{aligned} (f_n^A)'(t) &= \left(\frac{A_n(t)}{A(t)} \right)' \\ &= 2\lambda e^{-d_B \Delta} \left(\frac{A_{n-1}(t - \Delta)}{A(t)} - \frac{A_n(t)A(t - \Delta)}{A^2(t)} \right), \end{aligned} \tag{42}$$

and using exponential growth assumption $A(t) = C \exp(rt)$, we obtain

$$(f_n^A)'(t) = 2\lambda \exp(- (d_B + r)\Delta) (f_{n-1}^A(t - \Delta) - f_n^A(t)). \tag{43}$$

From Eq. (43), we obtain the equation for $\mu^A = \sum_n n f_n^A$:

$$(\mu^A)'(t) = k[\mu^A(t - \Delta) - \mu^A(t)] + k. \tag{44}$$

Eq. (44) is a neutral non-homogeneous linear equation since its principal eigenvalue is zero, and variation of constants yields an asymptotical solution as a linear function of time. Letting $\mu^A(t) = at + a_1$, we substitute this expression into Eq. (44) and solve for a ,

$$(\mu^A)'(t) = a = \frac{k}{1 + k\Delta}.$$

A similar argument shows that $(\sigma^2)^A(t)$ also grows linearly but at a different rate

$$\left((\sigma^2)^A \right)'(t) = b = a(1 - \Delta a)^2 = \frac{k}{(1 + k\Delta)^3},$$

and therefore $(\sigma^2)^A(t) = a(1 - \Delta a)^2 t + a_2$. Using these results it is easy to demonstrate that the rates of increase for the mean number of divisions and its variance of the total cell population μ^{A+B} and $(\sigma^2)^{A+B}$ are the same, i.e., $\mu'(t) = a$ and $(\sigma^2)'(t) = b$.

It can be shown that since a similar to Eq. (44) equation holds for the average number of divisions of the normalized population $\mu_2^A(t)$, the rate of increase of $\mu^2(t)$ is given by

$$(\mu_2)'(t) = \frac{k_N}{1 + k_N \Delta}, \tag{45}$$

where $k_N = \lambda + d_A - d$.

A.3. Rescaling technique

In (Pilyugin et al., 2003), we have formulated a general model where the probability rates of cell division and death are dependent only on the time since the last division and described the rescaling method for estimating the general, invariant parameters of the cell cycle. The rescaling method can be also used to estimate the parameters of particular models. To do so we need to calculate the rate of exponential increase of the total population size $\mathbf{r}(a)$ for the specific model given that a mother cell produces $2a$ daughter cells (see Pilyugin et al. (2003) for more details). Then by generating the experimental function $\mathbf{r}(a)$ from the CFSE data and by fitting the theoretical $\mathbf{r}(a)$ to the experimental one, the model parameters for cell division and death can be estimated.

For example, from a previously calculated rate of exponential change of the total population size in the SM model after rescaling we obtain:

$$\mathbf{r}(a) = 2a\lambda e^{-(r(a)+d_B)\Delta} - (\lambda + d_A). \quad (46)$$

Importantly, using experimentally obtained $\mathbf{r}(a)$ one can only estimate three parameter combinations in Eq. (46). These combinations can be written as invariant model-independent parameters:

$$\begin{aligned} \delta &= 1 - \frac{\lambda e^{-d_B\Delta}}{\lambda + d_A}, \quad \tau = \Delta + \frac{1}{\lambda + d_A}, \\ \sigma_\tau^2 &= \frac{1}{(\lambda + d_A)^2}, \end{aligned} \quad (47)$$

where $\delta = 1 - (2a^*)^{-1}$ (the probability that, in a single cell cycle, the death event occurs before division), $\tau = (a^* \mathbf{r}'(a^*))^{-1}$ (the mean generation time of surviving cells), $\sigma_\tau^2 = \tau^2(1 + (a^*)^2 \mathbf{r}''(a^*))\tau$ (the variance of the generation time of surviving cells), a^* is a solution of the equation $\mathbf{r}(a^*) = 0$, $\mathbf{r}'(a^*)$ and $\mathbf{r}''(a^*)$ are the first and second derivatives of \mathbf{r} with respect to a evaluated at $a = a^*$ (Pilyugin et al., 2003). Rewriting Eq. (46) in terms of invariant parameters and Δ , we find

$$\mathbf{r}(a) = \frac{2ae^{-r(a)\Delta}(1-\delta) - 1}{\tau - \Delta}. \quad (48)$$

By fitting Eq. (48) to the experimentally obtained $\mathbf{r}(a)$, we can estimate δ , τ and Δ . Note that only the duration of the B-phase $\Delta = \tau - \sigma_\tau$ can be estimated from the CFSE data alone. Assuming that death is

restricted only to the A-state ($d_B=0$) or the B-phase ($d_A=0$), we find the range for the remaining parameters λ , d_A , d_B and the average division time T :

$$\lambda \in \left(\frac{1}{\tau - \Delta}, \frac{1 - \delta}{\tau - \Delta} \right), \quad (49)$$

$$d_A \in \left(0, \frac{\delta}{\tau - \Delta} \right), \quad (50)$$

$$d_B \in \left(0, -\frac{\ln(1 - \delta)}{\Delta} \right), \quad (51)$$

$$T \in \left(\tau, \tau + \frac{\delta}{(1 - \delta)(\tau - \Delta)} \right). \quad (52)$$

Finally, if the proliferating fraction P is known, all the four parameters of the SM model can be estimated. This can be done in two ways. One is to rewrite Eq. (19) in terms of the invariant parameters and Δ

$$r + d_B = e^{-r\Delta} \frac{1 - \delta}{\tau - \Delta} (1 - P)(e^{d_B\Delta} - 1), \quad (53)$$

where $r = \mathbf{r}(1)$ (the rate of exponential increase of the total population size given in Eq. (8)) and d_B is unknown. Solving Eq. (53) numerically, we can estimate d_B . Using the estimate of d_B , we can then estimate λ and d_A :

$$\lambda = \frac{1 - \delta}{\tau - \Delta} e^{d_B\Delta}, \quad d_A = \frac{1 - (1 - \delta)e^{d_B\Delta}}{\tau - \Delta}. \quad (54)$$

Alternatively, we can express one of the parameters of the model as the function of other parameters and the proliferating fraction. Expressing the rate of commitment to division λ we find

$$\lambda = \frac{P(r + d_B)}{(1 - P)(1 - e^{-(r+d_B)\Delta})}. \quad (55)$$

Replacing λ given in Eq. (46) with the expression given Eq. (55) we obtain

$$\begin{aligned} \mathbf{r}(a) &= 2a \cdot \frac{P(r + d_B)e^{-(r(a)+d_B)\Delta}}{(1 - P)(1 - e^{-(r+d_B)\Delta})} \\ &\quad - \left(d_A + \frac{P(r + d_B)}{(1 - P)(1 - e^{-(r+d_B)\Delta})} \right). \end{aligned} \quad (56)$$

where r and P are known. By fitting this function to the experimentally obtained dependence $\mathbf{r}(a)$ we can estimate parameters Δ , d_A , d_B , and after substituting the obtained values in Eq. (55) we can also estimate λ .

References

- Asquith, B., Bangham, C., 2003. An introduction to lymphocyte and viral dynamics: the power and limitations of mathematical analysis. *Proc. R. Soc. Lond., B Biol. Sci.* 270, 1651.
- Asquith, B., Debacq, C., Macallan, D., Willems, L., Bangham, C., 2002. Lymphocyte kinetics: the interpretation of labelling data. *Trends Immunol.* 23, 596.
- Bellman, R., Cooke, K.L., 1963. *Differential–Difference Equations*. Academic Press.
- Bernard, S., Pujo-Menjouet, L., Mackey, M., 2003. Analysis of cell kinetics using a cell division marker: mathematical modeling of experimental data. *Biophys. J.* 84, 3414.
- Blattman, J.N., Antia, R., Sourdive, D.J., Wang, X., Kaech, S.M., Murali-Krishna, K., Altman, J.D., Ahmed, R., 2002. Estimating the precursor frequency of naive antigen-specific CD8 T cells. *J. Exp. Med.* 195, 657.
- Bohmer, R.M., 1979. Flow cytometric cell cycle analysis using the quenching of 33258 Hoechst fluorescence by bromodeoxyuridine incorporation. *Cell Tissue Kinet.* 12, 101.
- Bonhoeffer, S., Mohri, H., Ho, D., Perelson, A.S., 2000. Quantification of cell turnover kinetics using 5-bromo-2'-deoxyuridine. *J. Immunol.* 164, 5049.
- Brooks, R.F., Riddle, P.N., 1980. The 3T3 cell cycle at low proliferation rates. *J. Cell Sci.* 90, 601.
- Brooks, R.F., Bennett, D.C., Smith, J.A., 1980. Mammalian cell cycles need two random transitions. *Cell* 19, 493.
- Burns, F.J., Tannock, I.F., 1970. On the existence of a G0-phase in the cell cycle. *Cell Tissue Kinet.* 3, 321.
- Cain, S.J., Chao, P.S., 1997a. Transition probability cell cycle model: 1. Balanced growth. *J. Theor. Biol.* 185, 55.
- Cain, S.J., Chao, P.S., 1997b. Transition probability cell cycle model: 2. Non-balanced growth. *J. Theor. Biol.* 185, 69.
- Castor, L.N., 1980. A G1 rate model accounts for cell-cycle kinetics attributed to “transition probability”. *Nature* 287, 857.
- De Boer, R.J., Perelson, A.S., in press. Estimating division and death rates from CFSE data. *J. Comput. Appl. Math.*
- De Boer, R., Mohri, H., Ho, D., Perelson, A., 2003. Turnover rates of B cells, T cells, and NK cells in simian immunodeficiency virus-infected and uninfected rhesus macaques. *J. Immunol.* 170, 2479.
- Deenick, E., Gett, A., Hodgkin, P., 2003. Stochastic model of T cell proliferation: a calculus revealing IL-2 regulation of precursor frequencies, cell cycle time, and survival. *J. Immunol.* 170, 4963.
- Efron, B., Tibshirani, R., 1993. *An Introduction to the Bootstrap*. Chapman and Hall, New York.
- Gett, A., Hodgkin, P., 2000. A cellular calculus for signal integration by T cells. *Nat. Immunol.* 1, 239.
- Grasman, J., 1990. A deterministic model of the cell cycle. *Bull. Math. Biol.* 52, 535.
- Hodgkin, P., Lee, J., Lyons, A., 1996. B cell differentiation and isotype switching is related to division cycle number. *J. Exp. Med.* 184, 277.
- Kubbies, M., Goller, B., Van Bockstaele, D.R., 1992. Improved BrdUrd-Hoechst bivariate cell kinetic analysis by helium–cadmium single laser excitation. *Cytometry* 13, 782.
- Lyons, A.B., 2000. Analysing cell division in vivo and in vitro using flow cytometric measurement of CFSE dye dilution. *J. Immunol. Methods* 243, 147.
- Lyons, A., Parish, C.R., 1994. Determination of lymphocyte division by flow cytometry. *J. Immunol. Methods* 171, 131.
- Ma, C., Hodgkin, P., Tangye, S., 2004. Automatic generation of lymphocyte hetero-geneity: division-dependent changes in the expression of CD27, CCR7 and CD45 by activated human naive CD4+ T cells are independently regulated. *Immunol. Cell Biol.* 82, 67.
- Mohri, H., Perelson, A.S., Tung, K., Ribeiro, R.M., Ramratnam, B., Markowitz, M., Kost, R., Hurley, A., Weinberger, L., Cesar, D., Hellerstein, M.K., Ho, D.D., 2001. Increased turnover of T lymphocytes in HIV-1 infection and its reduction by antiretroviral therapy. *J. Exp. Med.* 194, 1277.
- Murali-Krishna, K., Ahmed, R., 2000. Cutting edge: naive T cells masquerading as memory cells. *J. Immunol.* 165, 1733.
- Nordon, R.E., Nakamura, M., Ramirez, C., Odell, R., 1999. Analysis of growth kinetics by division tracking. *Immunol. Cell Biol.* 77, 523.
- Parish, C.R., 1999. Fluorescent dyes for lymphocyte migration and proliferation studies. *Immunol. Cell Biol.* 77, 499.
- Pera, F., Mattias, P., Detzer, K., 1977. Methods for determining the proliferation kinetics of cells by means of 5-bromodeoxyuridine. *Cell Tissue Kinet.* 10, 255.
- Pilyugin, S.S., Ganusov, V.V., Murali-Krishna, K., Ahmed, R., Antia, R., 2003. The rescaling method for quantifying the turnover of cell populations. *J. Theor. Biol.* 225, 275.
- Revy, P., Sospedra, M., Barbour, B., Trautmann, A., 2001. Functional antigen-independent synapses formed between T cells and dendritic cells. *Nat. Immunol.* 2, 925.
- Shields, R., Smith, J.A., 1977. Cells regulate their proliferation through alterations in transition probability. *J. Cell. Physiol.* 91, 345.
- Smith, J.A., Martin, L., 1973. Do cells cycle? *Proc. Natl. Acad. Sci. U. S. A.* 85, 411.
- Stout, R.D., Suttles, J., 1992. T cells bearing the CD44hi “memory” phenotype display characteristics of activated cells in G1 stage of cell cycle. *Cell. Immunol.* 141, 433.
- Tangye, S., Hodgkin, P., 2004. Divide and conquer: the importance of cell division in regulating B-cell responses. *Immunology* 112, 509.
- Tyson, J.J., Hannsgen, K.B., 1985. The distributions of cell size and generation time in a model of the cell cycle incorporating size control and random transitions. *J. Theor. Biol.* 113, 29.

- Veiga-Fernandes, H., Walter, U., Bourgeois, C., McLean, A., Rocha, B., 2000. Response of naive and memory CD8+ T cells to antigen stimulation in vivo. *Nat. Immunol.* 1, 47.
- Weston, S.A., Parish, C.R., 1990. New fluorescent dyes for lymphocyte migration studies. *Analysis by flow cytometry and fluorescence microscopy.* *J. Immunol. Methods* 133, 87.
- Wolfram, S., 1990. *Mathematica: A System for Doing Mathematics by Computer.* 2nd edition. Addison-Wesley Pub. Co., Redwood City, CA.