Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Computational and Applied Mathematics 184 (2005) 140–164
JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS
www.elsevier.com/locate/cam

Estimating division and death rates from CFSE data

Rob J. De Boer[a,*], Alan S. Perelson[b]

[a]*Theoretical Biology, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands*
[b]*Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA*

Received 21 April 2004

Abstract

The division tracking dye, carboxyfluorescin diacetate succinimidyl ester (CFSE) is currently the most informative labeling technique for characterizing the division history of cells in the immune system. Gett and Hodgkin (Nat. Immunol. 1 (2000) 239–244) have proposed to normalize CFSE data by the 2-fold expansion that is associated with each division, and have argued that the mean of the normalized data increases linearly with time, t, with a slope reflecting the division rate p. We develop a number of mathematical models for the clonal expansion of quiescent cells after stimulation and show, within the context of these models, under which conditions this approach is valid. We compare three means of the distribution of cells over the CFSE profile at time t: the mean, $\mu(t)$, the mean of the normalized distribution, $\mu_2(t)$, and the mean of the normalized distribution excluding nondivided cells, $\widehat{\mu_2}(t)$.

In the simplest models, which deal with homogeneous populations of cells with constant division and death rates, the normalized frequency distribution of the cells over the respective division numbers is a Poisson distribution with mean $\mu_2(t) = pt$, where p is the division rate. The fact that in the data these distributions seem Gaussian is therefore insufficient to establish that the times at which cells are recruited into the first division have a Gaussian variation because the Poisson distribution approaches the Gaussian distribution for large pt. Excluding nondivided cells complicates the data analysis because $\widehat{\mu_2}(t) \neq pt$, and only approaches a slope p after an initial transient.

In models where the first division of the quiescent cells takes longer than later divisions, all three means have an initial transient before they approach an asymptotic regime, which is the expected $\mu(t) = 2pt$ and $\mu_2(t) = \widehat{\mu_2}(t) = pt$. Such a transient markedly complicates the data analysis. After the same initial transients, the normalized cell numbers tend to decrease at a rate e^{-dt}, where d is the death rate.

Nonlinear parameter fitting of CFSE data obtained from Gett and Hodgkin to ordinary differential equation (ODE) models with first-order terms for cell proliferation and death gave poor fits to the data. The Smith–Martin model with an explicit time delay for the deterministic phase of the cell cycle performed much better. Nevertheless, the insights gained from analysis of the ODEs proved useful as we showed by generating virtual CFSE data with

* Corresponding author. Tel.: +31 30 253 7560; fax: +31 30 251 3655.
E-mail address: R.J.DeBoer@bio.uu.nl (R.J. De Boer).

0377-0427/$ - see front matter
doi:10.1016/j.cam.2004.08.020

a simulation model, where cell cycle times were drawn from various distributions, and then computing the various mean division numbers.

Keywords: CFSE; Cell kinetics; Immunology; Cell division; Proliferation; Smith–Martin model

1. Introduction

Carboxyfluorescin diacetate succinimidyl ester (CFSE) is an intracellular fluorescent dye that dilutes 2-fold when a cell divides [7]. The labeling of cells with CFSE is done in vitro, and labeled cells can be followed thereafter both in vitro and in vivo. Harvesting the cells and sorting them by the CFSE intensity generates profiles with maximally 7 or 8 peaks, each reflecting the number of divisions the cells have undergone. The limit of 7 or 8 divisions is caused by the dilution of the dye: after 8 divisions the CFSE intensity is 2^8 fold lower than the original intensity. CFSE labeling is currently the most informative technique for characterizing the kinetics of cell types in the immune system.

Gett and Hodgkin [5] proposed a very elegant and intuitive approach to analyzing CFSE data. First, they integrate the CFSE intensity under each peak to measure the number of cells that have divided a given number of times. They then normalize this data by the 2-fold expansion that is associated with each division. This has the immediate advantage of allowing one to see the distribution of the cells from the original starting population in the various division numbers not confounded by their clonal expansion. CFSE data have frequently been misinterpreted by arguing that the large number of cells in the highest division number means that *most cells* have proceeded through this many divisions over the time of the experiment [1,6,11]. Because each division involves a 2-fold expansion one expects more cells for the higher division numbers. Normalization reveals that the cells in the highest division number typically represents a tiny fraction of the starting population. Gett and Hodgkin [5] analyzed the normalized frequency distributions by fitting a normal distribution to the data, excluding the data points corresponding to the division number zero, since some cells in this class may never divide. Using the mean of the fitted normal distribution, μ_{GH}, they showed that this mean increased linearly in time, and argued that the slope represented the division rate. This seems a powerful notion, because intuitively one indeed expects the increase of the mean division number to reflect the division rate, *whatever the precise mechanism by which cells divide*. Because the fitted normal distribution in the Gett and Hodgkin [5] method includes cells in division number zero, the method implicitly estimates the number of cells in division number zero that will ultimately divide. Thus, their mean division number, μ_{GH}, is not the same as the mean division number estimated from all the data.

The Gett and Hodgkin [5] method is elegant and simple, and it would be very important if the rate of increase of the mean division number could indeed be used to estimate the cell cycle time. However, a recent paper developing a mathematical model for analyzing CFSE data demonstrates that the Gett and Hodgkin [5] method is not always valid [8]. These authors consider the “homogeneous” case where the division and death rates are independent of the division number, and confirm that the mean division number, μ, increases linearly in time (after a short initial transient). Their important result is that the rate at which the mean increases also depends on the distribution of death rates among the phases of the cell cycle. For instance, cell death could occur mainly in the S, G2 and M phase of the cell cycle because of the various check points during these phases, and death could be rare during the quiescent G1 (or G0)

phase. If this is the case, and if the period to complete the S, G2 and M phases of the cell cycle is a small fraction of the total time between divisions, the slope of the linear increase of the mean division number is a composite of the cell cycle time and the death rate of the cells [8]. If the cell cycle time were estimated from this slope, it would maximally be overestimated 2-fold when the population is at steady state, less than 2-fold when it expands, and more than 2-fold if it contracts [8]. Because this casts serious doubts on the validity of the Gett and Hodgkin [5] approach, the authors propose an elegant novel scaling technique that generically allows one to estimate (a) the average cycle time of the cells successfully completing division, and (b) the fraction of cells surviving the cell cycle [8]. However, most biologists seek estimates for the average cell cycle time and the average cell death rate, where the average is taken over all cells.

In this paper, we address another problem associated with estimating cell division times from CFSE data. Even under the circumstances where the rate of linear increase of the mean division number should approximately reflect the cell cycle time, one needs to consider the initial transient before the increase has become linear. In experiments where quiescent cells are activated to proliferate one expects the first division to take longer than the later divisions [5]. Additionally, quiescent nondivided cells are expected to have a longer life-span than activated cells. This makes the problem nonhomogeneous because the nondivided cells have different death and division rates than the cells having completed at least one division. Additionally, strong activation of quiescent cells typically triggers a rapid cascade of divisions resulting in growth of the population. Under such circumstances, i.e., rapid population growth with a short G1 phase for all divisions but the first, the rate of increase of the mean division number tends to reflect the cell cycle time in the homogeneous model [4,8].

Gett and Hodgkin [5] have circumvented most of these problems by arguing that the first division involves a stochastic recruitment into a proliferative state and that subsequent divisions are deterministic. Their main argument was that the standard deviation of the division number distribution hardly increases over time, and thus seemed due to the variability in the time at which cells are recruited into the first division. Further, they developed a computer-based analysis tool for fitting CFSE data based on these assumption [2]. Our analysis will cast doubt on the assumption that after the first division the subsequent divisions are deterministic.

A standard mathematical model for analyzing the kinetics of the cell cycle is the Smith–Martin [10] model, which allows for two phases of the cell cycle. Cells in the “A” state are approximately in the G1 (or G0) phase of the cell cycle, and are recruited stochastically to divide, i.e., to enter the so-called “B” phase. The B-phase approximately corresponds to the S, G2, and M phases of the cell cycle and has a fixed total duration Δ. After completing the “deterministic” B-phase, a cell delivers two daughter cells in the “stochastic” A-state in which the cells may be recruited for another round of division [10]. Importantly, cells differing in the length of the cell cycle typically differ in the length of the A state. Once triggered to divide, cells typically proceed through the deterministic B phase at approximately the same speed [10]. Because cells cannot divide arbitrarily fast due to the finite length of the B phase, the Smith and Martin [10] model does a better job at describing the dynamics of subpopulations of cells that have divided a given number of times than systems of ordinary differential equations (ODEs). In ODE models the probability of cell division per unit time is effectively distributed exponentially, which implies that cells are likely to divide immediately after their previous division. Without a minimum division time cell populations proceed too rapidly through a cascade of divisions. Thus, while the mean behavior of a population of cells can be modeled well by ODEs, modeling the number of divisions a cell subpopulation has gone through may not be accurately represented. ODEs have been used before to analyse CFSE data [9,12] and we will show that under some circumstances their use is appropriate, while under other conditions they may give

rise to misleading conclusions. Moreover, one should bear in mind that the simple ODE models that we develop here, and that were used before [9,12] to analyse CFSE data, do not distinguish between the A and B phase of the Smith–Martin model. This is important because Pilyugin et al. [8] and Ganusov et al. [4] have shown that if the death rates during the two phases are not equal, this will influence the increase of the mean division number, and one will have to estimate two death rates from the data. Nevertheless, we will employ the simplicity of the ODE models to illustrate a number of basic points, including the Gett and Hodgkin [5] method.

2. Homogeneous case

Revy et al. [9] proposed the following simple system of ODEs to analyse their CFSE data:

$$\begin{aligned}\frac{\mathrm{d}N_0}{\mathrm{d}t} &= -(p+d)N_0,\\ \frac{\mathrm{d}N_i}{\mathrm{d}t} &= 2pN_{i-1} - (p+d)N_i, \quad \text{for } i = 1,\ldots,\infty.\end{aligned} \tag{1}$$

where N_i be the number of cells that have completed i divisions, and p and d are the cell proliferation and death rates, respectively. The division number or division index is i. The solution of this system, with the initial condition $N_0(0) > 0$, $N_i(0) = 0$ for $i = 1,\ldots,\infty$, is [9]

$$\begin{aligned}N_0(t) &= N_0(0)\mathrm{e}^{-(p+d)t},\\ N_i(t) &= \frac{(2pt)^i}{i!}N_0(t), \quad \text{for } i = 1,\ldots,\infty.\end{aligned} \tag{2}$$

The total number of cells, $N(t) = \sum_0^\infty N_i(t) = N_0(t)\mathrm{e}^{2pt} = N_0(0)\mathrm{e}^{(p-d)t}$, increases exponentially with the net growth rate of the population $p - d$.

The frequency distribution of cells over the division numbers, $F_i(t) \equiv N_i(t)/N(t)$, is given by a Poisson distribution

$$F_i(t) = \frac{(2pt)^i}{i!}\mathrm{e}^{-2pt}, \quad \text{for } i = 0,\ldots,\infty, \text{ with mean } \mu(t) \equiv \sum_0^\infty i F_i(t) = 2pt. \tag{3}$$

Because this distribution only depends on the proliferation rate, p, it was used to estimate p from CFSE data collected at various times after cells were placed in culture [9].

One can easily extend this model by incorporating a time delay to allow for the extra time nondivided cells may require to complete the first division. Defining the extra time required for the first division as τ, one can define $N_0(\tau) = N(\tau) = P(0)\mathrm{e}^{-d\tau}$ as the number of (undivided) cells that have survived the time period τ before proliferation starts, where $P(0)$ is the original number of precursor cells. Substituting this into the model one obtains

$$\begin{cases} N_0(t) = P(0)\mathrm{e}^{-dt},\ N_i(t) = 0 \text{ and } \mu(t) = 0 & \text{when } 0 \leqslant t \leqslant \tau \text{ and}\\ N_0(t) = N_0(\tau)\mathrm{e}^{-(p+d)t'},\ N_i(t) = \dfrac{(2pt')^i}{i!}N_0(t) \text{ and } \mu(t) = 2pt' & \text{otherwise,}\end{cases} \tag{4}$$

for $i = 1,\ldots,\infty$ and where $t' \equiv t - \tau$. For $t > \tau$ it makes no qualitative difference to allow for the extra time to first division since the only difference between Eq. (2) and Eq. (4) is a time shift, τ, and the

Table 1
The asymptotic regime of the ODE models and the length of the transient

Method	Model	Transient	Slope cells	Intersect time[a]	Slope mean
μ	Eq. (4)	τ	$e^{(p-d)t}$	$\tau+[2p]^{-1}$	$2pt$
μ_2	Eq. (7)	τ	e^{-dt}	$\tau+p^{-1}$	pt
$\widehat{\mu_2}$	Eq. (9)	$\tau+p^{-1}$	e^{-dt}	$\tau+p^{-1}$	pt
μ_2	Eq. (14)	τ	e^{-dt}	$\tau+(\phi p)^{-1}$	ϕpt
$\widehat{\mu_2}$	Eq. (9)	$\tau+p^{-1}$	e^{-dt}	$\tau+p^{-1}$	pt
μ	Eq. (23)	$[p+p_0-(d-d_0)]^{-1}$	$e^{(p-d)t}$	$[p+p_0-d+d_0]^{-1}$	$2pt$
μ_2	Eq. (28)	$[p_0+d_0-d]^{-1}$	e^{-dt}	$[p_0+d_0-d]^{-1}$	pt
$\widehat{\mu_2}$	Eq. (32)	$[p_0+d_0-d]^{-1}$	e^{-dt}	$[p_0+d_0-d]^{-1}$	pt

[a]The intersect time is the time at which the asymptote of the mean division index equals one.

initial condition in the solution of the undivided cells, i.e., $N_0(\tau)$ versus $N_0(0)$. Note that solving $\mu(t)=2$ from Eq. (4) yields the time to complete the first division, i.e., $\tau+p^{-1}$, and that solving $\mu(t)=1$ yields $t=\tau+[2p]^{-1}$ (see Table 1).

The Gett and Hodgkin [5] method can be illustrated with this model by normalizing the number of cells in each division class i by the corresponding degree of clonal expansion 2^i. This normalization is extremely useful and gives the distribution of cells from the starting population $P(0)$ over the various division classes. This is what Gett and Hodgkin [5] call the *precursor cohort distribution*. When $n_i(t)\equiv N_i(t)/2^i$ we obtain

$$\begin{cases} n_0(t)=N_0(t)=P(0)e^{-dt} \text{ and } n_i(t)=0 & \text{when } 0\leqslant t\leqslant\tau \text{ and} \\ n_0(t)=N_0(t)=N_0(\tau)e^{-(p+d)t'} \text{ and } n_i(t)=\dfrac{(pt')^i}{i!}N_0(t) & \text{otherwise,} \end{cases} \tag{5}$$

for $i=1,\ldots,\infty$ and where $t'\equiv t-\tau$. Thus, the total normalized cell number

$$\begin{cases} n(t)=N_0(t)=P(0)e^{-dt} & \text{when } 0\leqslant t\leqslant\tau \text{ and} \\ n(t)\equiv\sum_0^\infty n_i(t)=N_0(\tau)e^{-dt'}=P(0)e^{-dt} & \text{otherwise,} \end{cases} \tag{6}$$

declines constantly with the death rate d. For $t>\tau$, the frequency distribution is

$$f_i(t)\equiv n_i(t)/n(t)=\frac{(pt')^i}{i!}e^{-pt'}, \tag{7}$$

i.e., Poisson with mean and variance

$$\mu_2(t)\equiv\sum_{i=0}^{\infty} i f_i(t)=pt' \quad \text{and} \quad \sigma^2(t)=\sum_{i=0}^{\infty} i^2 f_i(t)-\mu_2(t)^2=pt'. \tag{8}$$

For $0\leqslant t\leqslant\tau$, $f_0=1$, $f_i=0$ for $i=1,\ldots,\infty$, and $\mu_2(t)=\sigma^2(t)=0$.

These results allow one to estimate three parameters from normalized CFSE profiles (see Table 1). First, the death rate d can be estimated from the slope e^{-dt} by which the total normalized cell number decreases. Second, the division rate is given by the slope of the linear increase of $\mu_2(t)$ when $t\geqslant\tau$. Third,

the time at which $\mu_2(t) = 1$ gives the time to first division $\tau + p^{-1}$. Although we have not ignored the nondivided cells, and we have computed the mean directly from the Poisson distribution, rather than from a fitted normal distribution, the latter two conclusions seem to confirm the estimation procedure proposed by Gett and Hodgkin [5]. We have added the estimation of the death rate to the procedure.

Gett and Hodgkin [5], however, did not compute the arithmetic mean of the normalized distribution, $\mu_2(t)$, but rather the conditional mean, i.e., $\widehat{\mu}_2(t) \equiv \sum_{i=1}^{\infty} i f_i(t) / \sum_{i=1}^{\infty} f_i$, in which only cells that have divided at least once were considered. For the homogeneous normalized model of Eq. (7) this conditional mean is

$$\begin{cases} \widehat{\mu}_2(t) = 0 & \text{when } 0 \leqslant t < \tau \text{ and} \\ \widehat{\mu}_2(t) = \dfrac{pt'}{1 - \mathrm{e}^{-pt'}} & \text{otherwise,} \end{cases} \tag{9}$$

which jumps between $\widehat{\mu}_2(t) = 0$ for $0 \leqslant t < \tau$ to $\widehat{\mu}_2(\tau) = 1$, i.e., the limit for $t' \to 0$ of $pt'/(1 - \mathrm{e}^{-pt'})$ is one, and then continues increasing for $t > \tau$. This jump itself is not important but the behavior after the jump seems to contradict the two Gett and Hodgkin [5] interpretations because (1) $\widehat{\mu}_2(t)$ is not increasing linearly in time with slope p (see Fig. 2), and (2) estimating the time to first division by solving $\widehat{\mu}_2(t) = 1$ gives $t = \tau$, which is not equal to the time required to complete the first division $(\tau + p^{-1})$. However, after an initial transient exceeding τ plus several cell cycle times, the term $[1 - \mathrm{e}^{-pt'}] \to 1$, and $\widehat{\mu}_2(t)$ will approach $\mu_2(t)$. For instance, the slope

$$\frac{\mathrm{d}\widehat{\mu}_2(t)}{\mathrm{d}t} = \frac{\mathrm{d}\widehat{\mu}_2(t')}{\mathrm{d}t'} \frac{p(1 - [1 + pt']\mathrm{e}^{-pt'}}{(1 - \mathrm{e}^{-pt'})^2}, \tag{10}$$

approaches $0.5p$ when $t \to \tau$, equals $0.66p$ at $t = \tau + p^{-1}$, and is $0.79p$ at $t = \tau + 2/p$. Because the initial slope of $\widehat{\mu}_2(t)$ is only 50% of the true division rate, and because this slope approaches the division rate p only after several more cell cycles, one tends to underestimate the division rate whenever most data points fall within this initial transient.

Surprisingly, because Gett and Hodgkin [5] estimated the mean division number from a normal distribution that was fitted to the data, rather than estimating the mean $\widehat{\mu}_2(t)$ from the data, their mean division number tends to be lower than $\widehat{\mu}_2(t)$ at early time points, and approach $\widehat{\mu}_2(t)$ at late time points. The reason for this is that the fitted normal distribution was not truncated, and thus allowed for an estimated contribution of nondivided cells, and possibly even of cells with a negative division number. Their method is interesting because it decreases the negative impact that the transient of $\widehat{\mu}_2(t)$ has in estimating the division rate. Because the data analyzed by Gett and Hodgkin [5] adhered very well to a normal distribution, the method was build upon the notion that normalized precursor cohort profiles at different time points are Gaussians that just shift with time. Basically, we have generalized the model to systems where the distribution need not be Gaussian, at the expense of having to correct for a transient.

The normalized number of dividing cells is only defined for times exceeding τ and is given by $\mathrm{d}\widehat{n}/\mathrm{d}t = -d\widehat{n} + pn_0$, where $n_0(t) = N_0(\tau)\mathrm{e}^{-(p+d)t'}$. For $t' > 0$ this has the solution

$$\widehat{n}(t) = N_0(\tau)\mathrm{e}^{-dt'}[1 - \mathrm{e}^{-pt'}], \tag{11}$$

which approaches an exponential with slope $-d$ after the same initial transient. Thus, to reliably estimate cell cycle times from CFSE data one should convince oneself (1) that $\widehat{\mu}_2(t)$ has approached its linear regime, and (2) that $\widehat{\mu}_2(t) \gg 1$ for most of the time. For this simple homogeneous model one would be

tempted to conclude that it is better to study the arithmetic mean, $\mu_2(t)$, of the data because it lacks the initial nonlinear transient. However, the next extension of the model will show that this is only true for this simple homogeneous model, in more complicated situations $\widehat{\mu}_2(t)$ is more likely to reflect the cell cycle time.

Gett and Hodgkin [5] observed that the frequency distribution of the normalized cell numbers over the division classes has a Gaussian shape. They then interpreted this as evidence for a Gaussian distribution of the times at which precursor cells are recruited into their first division. The Poisson distribution of Eq. (7) demonstrates that Gaussian shaped frequency distributions are also expected in this simple ODE model where recruitment times are distributed exponentially because the Poisson distribution approaches the Gaussian distribution for large values of pt. Thus, without further tests like the one performed by Deenick et al. [2], one cannot conclude that the observed Gaussian frequency distribution implies a Gaussian distribution of recruitment times. Basically, by the central limit theorem, most models with a chain of cell divisions would deliver gamma distributions that also look Gaussian. Unfortunately, this casts doubt on the most essential assumption of the Gett and Hodgkin [5] model, i.e., that the variance of the normalized frequency distribution is mostly due to a distribution of recruitment times into the first division.

2.1. Cells that never divide

One important factor is still missing from the model: a fraction of precursors cells may fail to divide. Assuming that all cells have the same death rate, d, an experiment starting with $P(0)$ precursor cells at time zero would have $N_0(\tau) = \phi P(0)\mathrm{e}^{-d\tau}$ cells ready to divide, and a total cell number $N(\tau) = P(0)\mathrm{e}^{-d\tau}$ when proliferation starts, where ϕ is the fraction of precursor cells that will divide. For this new definition of $N_0(\tau)$, the solution of this extended model remains to be given by Eq. (5). For $t > \tau$, the total normalized cell number

$$n(t) \equiv \sum_{i=0}^{\infty} n_i(t) + (1-\phi)P(0)\mathrm{e}^{-dt} = N_0(0)\mathrm{e}^{-dt'} + (1-\phi)P(0)\mathrm{e}^{-dt} = P(0)\mathrm{e}^{-dt} \tag{12}$$

remains unchanged. The frequency distribution $f_0(t) \equiv (n_0(t) + (1-\phi)P(0)\mathrm{e}^{-dt})/n(t)$ and $f_i(t) \equiv n_i(t)/n(t)$, for $i = 1, \ldots, \infty$, becomes

$$f_0(t) = \phi\mathrm{e}^{-pt'} + 1 - \phi \quad \text{and} \quad f_i(t) = \phi\,\frac{(pt')^i}{i!}\mathrm{e}^{-pt'} \tag{13}$$

with mean

$$\mu_2(t) \equiv \sum_{i=0}^{\infty} i f_i(t') = \phi p t'. \tag{14}$$

For $0 < t \leqslant \tau$, $f_0(t) = 1$ and $\mu_2(t) = 0$.

Thus, for $t > \tau$, the mean division number increases linearly in time but with slope ϕp, which would argue that one cannot simply estimate the division time from the slope. This also means that the time at which $\mu_2(t) = 1$ fails to provide the time to complete the first division (see Table 1).

Importantly, the original Gett and Hodgkin [5] concept of the mean division number of cells that have divided, $\widehat{\mu}_2(t)$, behaves quite differently. Because ϕ cancels from Eq. (13) when one computes $\widehat{\mu}_2(t) \equiv \sum_{i=1}^{\infty} i f_i(t)/\sum_{i=1}^{\infty} f_i$, one obtains the same mean $\widehat{\mu}_2(t)$ as in the previous model (see Eq. (14)

and Table 1). An example, where half of the cells fails to divide is depicted in Fig. 2b. To estimate the time to first division, one could employ linear regression to estimate the asymptote $\widehat{\mu}_{2\infty}$, i.e., $y = pt'$, which is expected to cross $y = 1$ at $t = \tau + p^{-1}$ (see Fig. 2b). Thus, for data sets where an unknown fraction, ϕ, of the cells fails to divide, it seems better to use $\widehat{\mu}_2(t)$ to estimate the cell cycle time, after one has convinced oneself that $\widehat{\mu}_2(t)$ has approached its linear regime. Finally, from the difference between $\mu_2(t)$ and $\widehat{\mu}_2(t)$ one should be able estimate the fraction,ϕ, of precursor cells that never divide. Summarizing, to properly analyze CFSE data one should depict both normalized means to estimate ϕ, and estimate the cell cycle time from the Gett and Hodgkin [5] mean $\widehat{\mu}_2(t)$ if the slopes μ_2' and $\widehat{\mu}_2{}'$ are different.

3. Data

The data from Gett and Hodgkin [5] suggests that the various mean division numbers, $\mu(t)$, $\mu_2(t)$, and $\widehat{\mu}_2(t)$, approach the regime where they increase linearly with time (see Fig. 1). CSFE data were taken from cells 60 to 96 h after they were triggered to divide by polyclonal stimulation. Fitting the three means by linear regression gives the lines $y = 0.095t - 4.24$, $y = 0.053t - 2.45$, and $y = 0.053t - 1.71$, for $\mu(t)$, $\mu_2(t)$, and $\widehat{\mu}_2(t)$, respectively. All three slopes suggest a division rate of $p \simeq 0.05\,\mathrm{h}^{-1}$ (i.e., a cell cycle time of 20 h). From the observation that the slopes, μ_2' and $\widehat{\mu}_2{}'$, are similar one would conclude that almost all cells divide, i.e., that $\phi \simeq 1$. From the regression lines one would estimate that $\mu(t) = 2$ at $t = 66\,\mathrm{h}$,

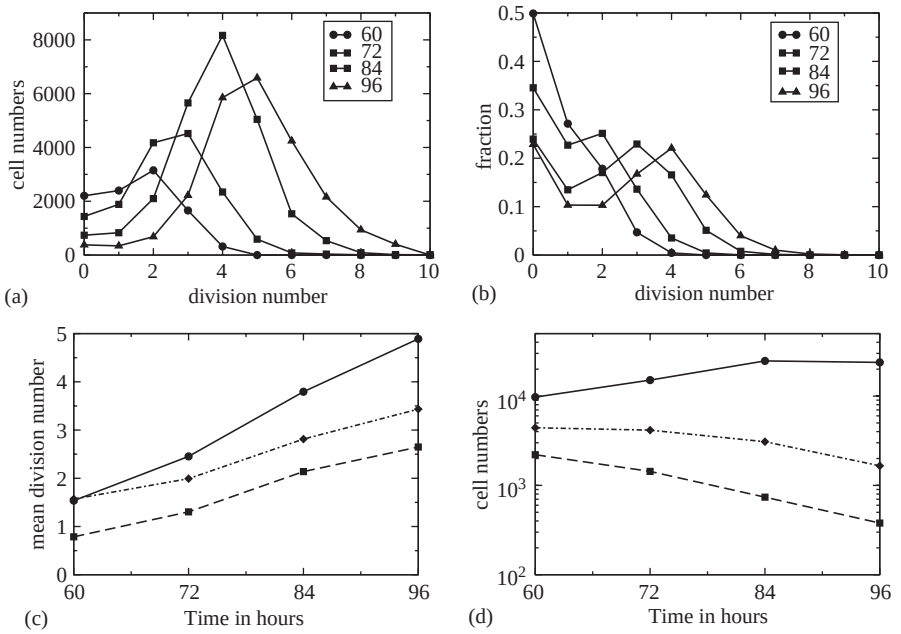

Fig. 1. The Gett and Hodgkin [5] data for CD4$^+$ T cells stimulated with anti-CD3 under saturating IL-2 concentrations (their Fig. 4). (a) the data represented as cell number, N_i, versus the division number, i, at different times given in hours. (b) the data normalized by dividing by the clonal expansion factor 2^i, and then plotted as a frequency distribution (f_i). (c) the three means: μ (circles connected by solid lines), μ_2 (squares connected by dashed lines), and $\widehat{\mu}_2$ (diamonds connected by dash-dotted lines). (d) the total number of cells (N, circles connected by solid lines), the total number of normalized cells (n, diamonds connected by dash-dotted lines), and the number of nondivided cells (N_0, squares connected by dashed lines).

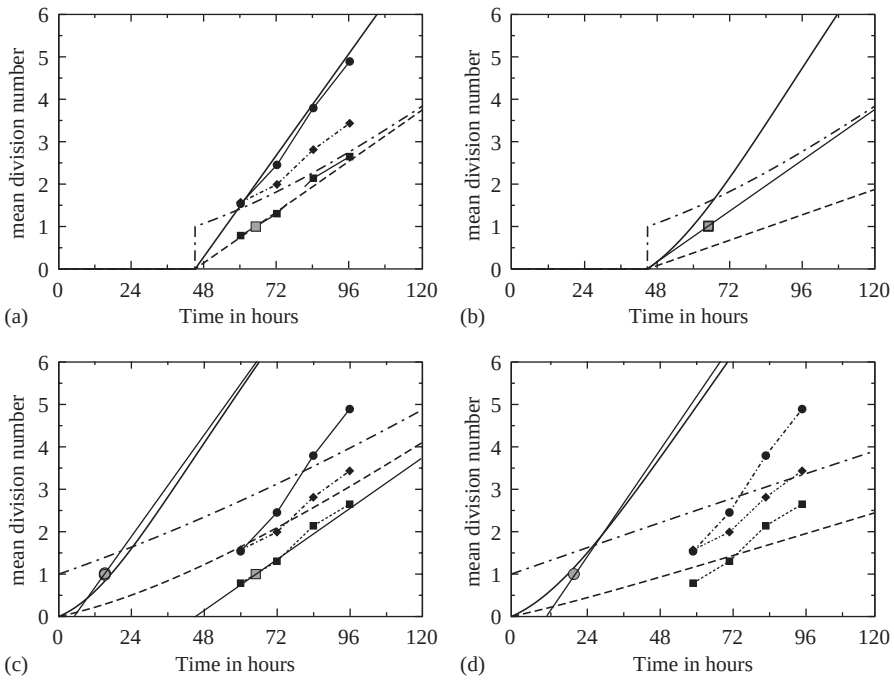

Fig. 2. The behavior of the three mean division numbers for the homogeneous ((a) and (b)) and heterogeneous models ((c) and (d)) plotted alongside the means computed from the data in Fig. 1 (symbols). The figure shows the mean $\mu(t)$ (solid line theory; filled circle data); the mean of the normalized distribution μ_2 (dashed line theory; squares data), the Gett and Hodgkin [5] mean $\widehat{\mu_2}$ (dash-dotted line theory; diamonds data). Parameters in (a): $\tau = 45\,\mathrm{h}$, $p_0 = p = 0.05\,\mathrm{h}^{-1}$, $d_0 = d = 0.027\,\mathrm{h}^{-1}$, and $\phi = 1$; in (b) we show the case where 50% of the precursors cells fail to divide ever ($\phi = 0.5$). In panels (a) and (b) the large grey square depicts the time where $\widehat{\mu_2}_\infty(t) = 1$, i.e., the point $(\tau + 1/p, 1)$, which reflects the time to complete the first division. The light solid line in (b) is the asymptote of $\widehat{\mu_2}$. In the heterogeneous model all cells will ultimately divide ($\phi = 1$), but this may take long because $p_0 \ll p$. Parameters in (c): $p_0 = 1/65 \simeq 0.015\,\mathrm{h}^{-1}$, $p = 0.05\,\mathrm{h}^{-1}$, $d_0 = d = 0.027\,\mathrm{h}^{-1}$. In (d) we study the effect of different death rates by assuming that nondivided cells have a much longer expected life-span, i.e., $d_0 = 0.01\,\mathrm{h}^{-1}$ and $d = 0.027\,\mathrm{h}^{-1}$. In (c) the grey square denotes the time where $\mu_{2\infty}(t) = 1$, i.e., the point $(1/p_0 = 65, 1)$, which reflects the time to complete the first division because $d = d_0$. The grey circle in panels (c) and (d) depict the time points where $\mu_\infty(t) = 1$, i.e., $(c^{-1} = [p + p_0]^{-1} = 15.3, 1)$ in (c), and $(1/c = 1/[p + p_0 - (d - d_0)] = 20.67, 1)$ in (d). The light solid lines are the asymptotes of Eq. (23) and Eq. (28) or Eq. (32). Although our parameter choice $d_0 \ll d$ seems very reasonable, $\gamma^{-1} = [p_0 + d_0 - d]^{-1} = -619\,\mathrm{h}^{-1}$, which implies that at $t = 0$ one obtains that $\mu_2(0) = 32$ and that the asymptote Eq. (28) is not approached at reasonable times.

$\mu_2(t) = 1$ at $t = 65\,\mathrm{h}$, and that $\widehat{\mu_2}(t) = 1$ at $t = 51\,\mathrm{h}$. The first two suggest that $\tau + p^{-1} \simeq 65$, which with the estimate $p^{-1} \simeq 20\,\mathrm{h}$ suggests that $\tau \simeq 45\,\mathrm{h}$. Because the asymptote of the latter, i.e., $\widehat{\mu_2}_\infty(t)$, also equals one at $t = \tau + p^{-1}$ (see Table 1), and the regression line equals one at time $t = 51\,\mathrm{h}$, one would have to conclude that $\widehat{\mu_2}(t)$ is still in its initial transient, and has not yet approached the asymptote $\widehat{\mu_2}_\infty$. This is in reasonable agreement with the latter, which suggests that $\tau \simeq 51\,\mathrm{h}$. The exact behavior of the three means is plotted for these parameter values in Fig. 2a.

The slopes of the three lines in Fig. 1d representing $N(t)$, $n(t)$, and $N_0(t)$, can also be used for parameter estimation. Fitting the natural log of the data by linear regression yields the slopes 0.026, -0.027, and -0.05, and for $N(t)$, $n(t)$, and $N_0(t)$ curves, respectively. Since $n(t)$, by Eq. (6), decays exponentially

at rate d, this procedure suggests a death rate of $d = 0.027\,\mathrm{h}^{-1}$. By Eq. (2) for $N(t)$, the net growth rate of the population is approximately $r = p - d = 0.026\,\mathrm{h}^{-1}$. After the initial transient due to the slow first division, the growth rate is expected to approach the difference between the division and the death rates, i.e., $\mathrm{d}N/\mathrm{d}t = rN = (p-d)N$. From the slope based estimates of d and r one would predict that $p = r + d = 0.053\,\mathrm{h}^{-1}$, which is in good agreement with the division rate estimated from the increase of the mean division number. However, the fact that $\mathrm{d}N_0/\mathrm{d}t = -(p+d)N_0 \simeq -0.05N_0$, cannot be reconciled with these estimates, since $p + d \simeq 0.072\,\mathrm{h}^{-1}$. Rather this suggests that the nondivided cells have a slower division rate (see below). Summarizing, it seems that one can obtain rough estimates of several parameters of the model by examining the rates at which the total number of cells, the total number of normalized cells, and the total number of nondivided cells change with time. Below we show that this procedure is unreliable because the net growth rate r depends strongly on the underlying mechanism of cell division. Moreover, Ganusov et al. [4], Pilyugin et al. [8] have shown that this procedure will give misleading results when the death rates in the G0 or G1 phase of the cell cycle differs from that in the other phases.

4. Heterogeneous case

Instead of assuming that a fraction of the precursor cells will never divide, one could also argue that the first division may just take very long. The loss of the nondivided cells in Fig. 1c already suggests this because they are declining somewhat faster than the other cells at an estimated rate of $0.05\,\mathrm{h}^{-1}$. If the average time to first division is indeed about 65 h, one would expect a loss rate, $p + d$, of $1/65 + 0.027 = 0.042\,\mathrm{h}^{-1}$, which is not far from the observed loss rate of $0.05\,\mathrm{h}^{-1}$. Moreover, the experimental observation of Gett and Hodgkin [5] that the standard deviation of the division index was not increasing with time was taken as evidence to suggest that the time to complete the first division was not only slow but also more variable than that of subsequent divisions [5]. Using this as motivation, we therefore also analyse a "heterogeneous" model where the rate of the first division is slower than that of later divisions. Because this new model has two division rates (or probabilities) it is not expected to yield the same linear increase of the mean division index with time that we found when the underlying distribution was Poisson, Eq. (7).

Consider the model analysed above, but with division rate p_0 and death rate d_0 for the nondivided cells, i.e.,

$$\begin{aligned}
\frac{\mathrm{d}N_0}{\mathrm{d}t} &= -(p_0 + d_0)N_0,\\
\frac{\mathrm{d}N_1}{\mathrm{d}t} &= 2p_0N_0 - (p+d)N_1,\\
\frac{\mathrm{d}N_i}{\mathrm{d}t} &= 2pN_{i-1} - (p+d)N_i, \quad \text{for } i = 2,\ldots,\infty.
\end{aligned} \tag{15}$$

This has solutions

$$N_0(t) = N_0(0)\mathrm{e}^{-(p_0+d_0)t},$$

$$N_i(t) = N_0(0)\mathrm{e}^{-(p+d)t}\ \frac{p_0}{p}\left(\frac{2p}{a}\right)^i\left[\mathrm{e}^{at} - \sum_{j=0}^{i-1}\frac{(at)^j}{j!}\right]$$
$$= N_0(0)\mathrm{e}^{-(p+d)t}\ \frac{p_0}{p}\left(\frac{2p}{a}\right)^i\sum_{j=i}^{\infty}\frac{(at)^j}{j!} \quad \text{for } i = 1,\ldots,\infty, \tag{16}$$

where $a \equiv p - p_0 + d - d_0 \geqslant 0$.

Summing the differential equations, Eq. (15), one finds that the total cell number obeys $\mathrm{d}N/\mathrm{d}t = (p - d)N + (p_0 - p + d - d_0)N_0$, with $N(0) = N_0(0)$. This has the solution

$$N(t) = \frac{N_0(0)\mathrm{e}^{(p-d)t}}{c}[2p_0 + b\mathrm{e}^{-ct}], \tag{17}$$

where $b = p - p_0 - (d - d_0)$ and where $c = p + p_0 - (d - d_0) > 0$ because $p > d$. For the homogeneous case, i.e., $p = p_0$ and $d = d_0$, one obtains $a = b = 0$ and $c = 2p$. In general, the frequency distribution, $F_i(t) \equiv N_i(t)/N(t)$, is

$$F_0(t) = \frac{c\mathrm{e}^{-ct}}{2p_0 + b\mathrm{e}^{-ct}}, \tag{18}$$

$$F_i(t) = \frac{c\mathrm{e}^{-2pt}}{2p_0 + b\mathrm{e}^{-ct}}\ \frac{p_0}{p}\left(\frac{2p}{a}\right)^i\sum_{j=i}^{\infty}\frac{(at)^j}{j!} \quad \text{for } i = 1,\ldots,\infty. \tag{19}$$

The mean, $\mu(t) \equiv \sum_{i=0}^{\infty} i F_i$, of this frequency distribution is

$$\mu(t) = \left[\frac{c\mathrm{e}^{-2pt}}{2p_0 + b\mathrm{e}^{-ct}}\ \frac{p_0}{p}\right] \times \sum_{i=1}^{\infty}\sum_{j=i}^{\infty} i\left(\frac{2p}{a}\right)^i\frac{(at)^j}{j!}, \tag{20}$$

where the double sum can be rewritten as

$$\sum_{i=1}^{\infty}\sum_{j=i}^{\infty} i\left(\frac{2p}{a}\right)^i\frac{(at)^j}{j!} = \sum_{j=0}^{\infty}\frac{(at)^j}{j!}\sum_{i=0}^{j} i\left(\frac{2p}{a}\right)^i$$
$$= \sum_{j=0}^{\infty}\frac{(at)^j}{j!}\ \frac{2ap}{(2p-a)^2}[1 - (2p/a)^j - j(2p/a)^j + j(2p/a)^{j+1}]$$
$$= \frac{2ap}{(2p-a)^2}\left[\sum_{j=0}^{\infty}\frac{(at)^j}{j!} - \sum_{j=0}^{\infty}\frac{(2pt)^j}{j!} + \frac{2p-a}{a}\sum_{j=0}^{\infty} j\frac{(2pt)^j}{j!}\right]$$
$$= \frac{2ap}{(2p-a)^2}\,\mathrm{e}^{2pt}\left[\mathrm{e}^{(a-2p)t} - 1 + \frac{2p-a}{a}2pt\right]. \tag{21}$$

Because $2p - a = c > 0$ one obtains for the mean

$$\mu(t) = \left[\frac{c\mathrm{e}^{-2pt}}{2p_0 + b\mathrm{e}^{-ct}}\,\frac{p_0}{p}\right] \times \frac{2ap}{c^2}\,\mathrm{e}^{2pt}\left[\mathrm{e}^{(a-2p)t} - 1 + \frac{c}{a}\,2pt\right]$$
$$= \frac{2p_0[a(\mathrm{e}^{-ct} - 1) + 2pct]}{c[2p_0 + b\mathrm{e}^{-ct}]}. \tag{22}$$

Substituting $d_0 = d$ and $p_0 = p$ (i.e., $a = b = 0$ and $c = 2p$) gives $\mu(t) = 2pt$, as obtained above with the homogeneous model. For times much larger than $1/c$, the mean Eq. (22) will approach the asymptote

$$\mu_\infty(t) = 2pt - \frac{a}{c} = 2pt - \frac{p - p_0 + d - d_0}{p + p_0 - (d - d_0)}, \tag{23}$$

which increases with the expected slope $\mu_\infty(t)' = 2p$. Solving $\mu_\infty(t) = 1$ gives $t = 1/(p + p_0 - d + d_0) = 1/c$. Thus, the time it takes to breach $\mu = 1$ provides an indication of the length of the transient. In other words, data before and around $\mu = 1$ are not expected to fit a straight line with slope $\mu(t)' = 2p$. For the Gett and Hodgkin data, $\mu(t)$ was fitted to the regression line $y = 0.095t - 4.24$ in Fig. 1c. Thus, $\mu(t) \simeq 1$ around $t = 1/c \simeq 55$ h. This suggests that the last data point at 120 h should still be in the transient, i.e., the e^{-ct} terms in Eq. (22) are $\mathrm{e}^{-1} = 0.37$ around $t = c^{-1}$ and $\mathrm{e}^{-2} = 0.13$ around $t = 2/c \simeq 110$ h. If the ODE model were correct, this would cast doubt on the parameter estimates made from the Gett and Hodgkin data with the homogeneous model: if most data points fall within an initial transient the proliferation rate is expected to be an underestimate.

The behavior of $\mu(t)$ is plotted in Fig. 2c and d for $p_0^{-1} = 65$ h, $p^{-1} = 20$ h, and $d_0 = d = 0.027\,\mathrm{h}^{-1}$ in Fig. 2c, and for $d_0 \ll d$ in Fig. 2d. Although the parameters in Fig. 2c are identical to those estimated in Fig. 1, the fit of $\mu(t)$ in Fig. 2c is very poor compared to that of the homogeneous $\mu(t)$ in Fig. 2a. The slope of $\mu(t)$ rapidly approaches the expected value of $2p$, see the asymptote of Eq. (23) that runs parallel to the data, but the mean breaches $\mu(t) = 1$ far too early. In retrospect this makes sense because in the ODE model cells proceed too fast through the division cascade. Indeed, from the observation that $c^{-1} \simeq 55$ h, with the earlier estimate that $p = 0.05\,\mathrm{h}^{-1}$ and assuming that $d_0 = d$, one obtains from the equation defining c a negative value for p_0 (i.e., $p_0 \simeq -0.03$). Similarly, for $d_0 < d$, i.e., Fig. 2d, one obtains $p_0 \simeq -0.00487 - d_0$. Thus, although the increase in the mean division number can be described by this ODE model, it fails to correctly capture the time to first division. Rather the time to $\mu(t) = 1$ is an indication of the length of the initial transient: observe that the slope of the solid line approaches that of the asymptote beyond the large grey circles in Fig. 2a and d. The grey circles in Fig. 2a and d also showthat the length of the initial transient increases somewhat when $d_0 < d$. Thus, the longer the nondivided cells live, the longer they influence the increase of $\mu(t)$. We have computed expressions for the variance, but because of space limitations we proceed by studying the behavior of $\mu_2(t)$ and $\widehat{\mu_2}(t)$ in these ODE models.

4.1. Normalization

To compute the frequency distribution after the normalization by 2^i, we remove the factors two from Eqs. (15) and (16). Summing the normalized differential equations one finds that $\mathrm{d}n/\mathrm{d}t = -dn + (d - d_0)n_0$

with initial condition $n(0) = N_0(0)$, which has the solution

$$n(t) = \frac{N_0(0)\mathrm{e}^{-dt}}{\gamma}[p_0 + (d_0 - d)\mathrm{e}^{-\gamma t}], \tag{24}$$

where $\gamma \equiv p_0 + d_0 - d = p - a$. In the homogeneous case, $d = d_0$, $\gamma = p = p_0$, and hence $n(t) = N_0(0)\mathrm{e}^{-dt}$. When $d \neq d_0$, one can only estimate the death rate, d, from the slope of ln n versus t for $t \gg 1/\gamma$, so that $\mathrm{e}^{-\gamma t} \to 0$.

Using the solutions from Eq. (15), the normalized frequency distribution, $f_i(t) = n_i(t)/n(t)$, for the heterogeneous case becomes

$$\begin{aligned} f_0(t) &= \frac{\gamma \mathrm{e}^{-\gamma t}}{p_0 + (d_0 - d)\mathrm{e}^{-\gamma t}}, \\ f_i(t) &= \frac{\gamma \mathrm{e}^{-pt}}{p_0 + (d_0 - d)\mathrm{e}^{-\gamma t}} \frac{p_0}{p}\left(\frac{p}{a}\right)^i \sum_{j=i}^{\infty} \frac{(at)^j}{j!}, \quad \text{for } i = 1, \ldots, \infty. \end{aligned} \tag{25}$$

The mean $\mu_2(t) \equiv \sum_{i=0}^{\infty} i f_i$ of the normalized frequency distribution is

$$\mu_2(t) = \left[\frac{\gamma \mathrm{e}^{-pt}}{p_0 + (d_0 - d)\mathrm{e}^{-\gamma t}} \frac{p_0}{p}\right] \times \sum_{i=1}^{\infty} \sum_{j=i}^{\infty} i\left(\frac{p}{a}\right)^i \frac{(at)^j}{j!}, \tag{26}$$

which after rewriting the double sum becomes

$$\begin{aligned} \mu_2(t) &= \left[\frac{\gamma \mathrm{e}^{-pt}}{p_0 + (d_0 - d)\mathrm{e}^{-\gamma t}} \frac{p_0}{p}\right] \times \frac{ap}{(p-a)^2} \mathrm{e}^{pt}\left[\mathrm{e}^{(a-p)t} - 1 + \frac{p-a}{a}\, pt\right] \\ &= \frac{p_0}{\gamma} \frac{\gamma pt + a(\mathrm{e}^{-\gamma t} - 1)}{p_0 + (d_0 - d)\mathrm{e}^{-\gamma t}}, \end{aligned} \tag{27}$$

using $\gamma = p - a$.

Substituting $d_0 = d$ and $p_0 = p$ (i.e., setting $\gamma = p$) gives the $\mu_2(t) = pt$ obtained above with the homogeneous model. To study the increase in the mean division number in the full model we let $t \gg 1/\gamma$ in Eq. (27) and obtain that

$$\mu_{2\infty}(t) = pt - \frac{a}{\gamma}, \tag{28}$$

which increases in time with slope p. This asymptotic regime is approached at times greater than $1/\gamma$. Noting that $\gamma = p - a$, one sees from Eq. (28) that the time at which $\mu_{2\infty}(t) = 1$ is again $1/\gamma$. Since $\mu_2(t)$ was fitted to the regression line $y = 0.053t - 2.45$, the estimate for $1/\gamma = 65$ h. For $d_0 = d$ (i.e., for $1/\gamma = 1/(p_0 + d_0 - d) = 1/p_0$) this indeed delivers the estimated time to complete the first division (see Table 1). For $d_0 < d$ this is no longer the case and γ^{-1} only provides an estimate for the length of the initial transient.

The behavior of μ_2 is depicted as the dashed line in Fig. 2c and d. When the nondivided cells have the same death rate as the other cells (i.e., for $d_0 = d$ in panel c), the data (depicted by the black squares) are much better described by the asymptote $\mu_{2\infty}(t)$ (see Eq. (28)) than by μ_2 itself. The slope $\mu_2(t)'$ only begins to approach the slope $\mu_{2\infty}(t)' = p$ of the asymptote at times larger than γ^{-1} (as indicated by the

large grey square). For the quite reasonable parameter choice $d_0 \ll d$, i.e., $d_0 = 0.01\,\mathrm{h}^{-1}$, this is no longer true since $\gamma^{-1} \simeq -619\,\mathrm{h}$. The slope $\mu_2(t)'$ in the data fails to approach that of the asymptote $\mu_{2\infty}(t)'$ over the entire range of the observations (see Fig. 2d). Thus, if cells were to obey this ODE model, and have unequal death rates, one would not be able to estimate division rates from the increase in the mean of the normalized distribution.

4.2. Mean division number of dividing cells: $\widehat{\mu_2}(t)$

In this heterogeneous model with different proliferation and death rates the growth of the total number of normalized divided cell numbers is defined as $\mathrm{d}\widehat{n}/\mathrm{d}t = p_0 n_0 - d\widehat{n}$, with initial condition $\widehat{n} = 0$, and where $n_0(t) = N_0(0)\mathrm{e}^{-(p_0+d_0)t}$. The solution is

$$\widehat{n}(t) = \frac{p_0 N_0(0)\mathrm{e}^{-dt}}{\gamma}[1 - \mathrm{e}^{-\gamma t}], \tag{29}$$

where $\gamma \equiv p_0 + d_0 - d = p - a$. Thus, again, only after a transient, so that $\mathrm{e}^{-\gamma t} \to 0$, will the slope of a plot of the log of the normalized cell numbers versus time reflect the death rate.

Using the solution from Eq. (15), the normalized frequency distribution, $\widehat{f_i}(t) = n_i(t)/\widehat{n}(t)$, becomes

$$\widehat{f_i}(t) = \frac{\gamma \mathrm{e}^{-pt}}{p_0[1 - \mathrm{e}^{-\gamma t}]} \frac{p_0}{p}\left(\frac{p}{a}\right)^i \sum_{j=i}^{\infty} \frac{(at)^j}{j!}, \quad \text{for } i = 1, \ldots, \infty, \tag{30}$$

which after rewriting the double sum becomes

$$\begin{aligned}\widehat{\mu_2}(t) &= \left[\frac{\gamma \mathrm{e}^{-pt}}{p_0[1 - \mathrm{e}^{-\gamma t}]} \frac{p_0}{p}\right] \times \frac{ap}{(p-a)^2}\, \mathrm{e}^{pt}\left[\mathrm{e}^{(a-p)t} - 1 + \frac{p-a}{a}\, pt\right] \\ &= \frac{\gamma pt + a(\mathrm{e}^{-\gamma t} - 1)}{\gamma[1 - \mathrm{e}^{-\gamma t}]},\end{aligned} \tag{31}$$

Thus, for $t \gg 1/\gamma$ the mean of the divided cells approaches

$$\widehat{\mu_{2\infty}}(t) = pt - \frac{a}{\gamma} = pt - \frac{p - p_0 + d - d_0}{p_0 + d_0 - d}, \tag{32}$$

which is identical to the asymptotic behavior of μ_2 (see Eq. (28)). The same conclusions therefore apply: wherever $d_0 = d$ the slope of $\widehat{\mu_2}$ is similar to the that of the asymptote (Eq. (32)), and the time to first division can be estimated by solving $\mu_{2\infty}(t) = 1$ from Eq. (28). However, whenever $d_0 \ll d$ the model's mean $\widehat{\mu_2}$ fails to approach the mean $\widehat{\mu_2}$ of the data (see Fig. 2d).

5. Fitting the distributions

Instead of estimating the cell cycle time from the increase in the mean division number, one could fit the full solution of Eq. (16) to the actual cell numbers observed per division number depicted in Fig. 1a. Since this is using much more information, one could expect a much better fit. Doing so, we obtain the fit depicted in Fig. 3a with the parameter estimates $p_0 = 0.022\,\mathrm{h}^{-1}$, $p = 0.025\,\mathrm{h}^{-1}$, and $N(0) = 1.5 \times 10^4$

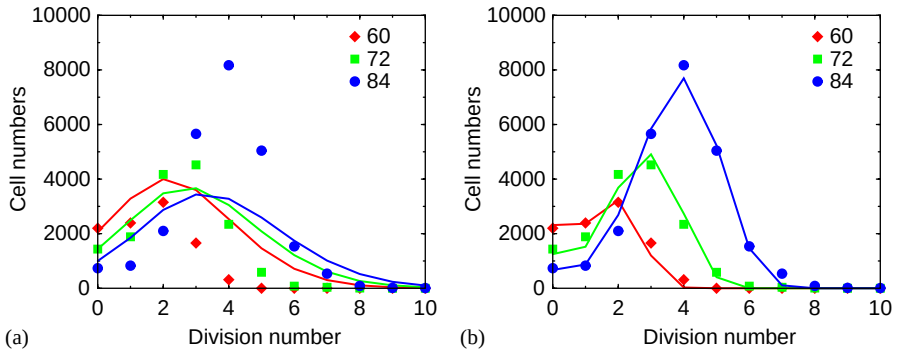

Fig. 3. (a) Fit of the heterogeneous ODE model, Eq. (16), to the data in Fig. 1c. The data are represented by symbols and the fits are represented by lines. Because $N_0(0)$ and d turned out to be dependent parameters both could not be estimated. Thus, we fixed $d = 0.025\,\mathrm{h}^{-1}$. Estimates: $p = 0.025\,\mathrm{h}^{-1}$, $p_0 = 0.022\,\mathrm{h}^{-1}$, $d_0 = 0.01\,\mathrm{h}^{-1}$, and $N_0(0) = 1.5 \times 10^4$ cells. For comparison we also show the best fit of the Smith and Martin [10] model to the same data in panel (b). For parameter estimates see the left column of Table 2.

cells, i.e., cell cycle times of 45 and 40 h, for the first and the later divisions, respectively. The best fit to the data is very poor because the distributions predicted by the model are much broader than those in the data (see Fig. 3). The fitting procedure therefore moves the distributions at the early time points to the right, and those of the late time points to the left. By estimating a slow proliferation rate the fitting procedure reduces the tail at the right with cells having undergone many divisions. This poor fit suggests that the exponential distribution of cell cycle times implicit in the ODE model is incorrect, and leads to overestimates of the cell cycle time in order to slow down the too rapid successive proliferation events in the ODE model. The exponential distribution allows too many cells to be recruited too early into the first division, and these cells then proceed too fast through the subsequent divisions. This also explains why adding a time delay before the first division (i.e., Eq. (7)) provides a better fit of the mean division number than the ODE model (compare Fig. 2a with c), and casts doubt on earlier work estimating cell cycle times by fitting ODE models to CFSE data [9,12]. Proper fitting CFSE data therefore requires models with an explicit time delay corresponding to the minimal length of the cell cycle.

5.1. Smith–Martin model

The Smith–Martin model [10] prevents too rapid progression through the cell cycle by introducing the equivalent of a time delay, i.e., a fixed length for the S, G2 and M phases of the cell cycle. The Smith–Martin model allows for two phases of the cell cycle: cells in the "A" state are randomly activated to divide, and dividing cells in the "B" phase remain in this phase for a fixed time Δ, after which they yield two daughter cells in the A phase. Cells in the A and B phase have death rates d_A and d_B, respectively. Cells in the A phase are activated at a rate λ. For experiments where quiescent cells are activated to proliferate in a programmed cascade, we use different, i.e., slower, parameters for the nondivided cells. Call them λ_0 and Δ_0, respectively. One expects $\lambda_0 < \lambda$ and $\Delta_0 > \Delta$. Because the Smith–Martin model distinguishes the two phases of the cell cycle we will be able to accommodate the findings of Pilyugin et al. [8] and Ganusov et al. [4] showing that the results obtained above with the much simpler ODE models may only be relevant when the death rates during these two phases are sufficiently similar.

Table 2
Parameter estimates of the Smith and Martin [10] model

		$d_A = 0$		$d_A = d_B$		$d_B = 0$	
d_A	h^{-1}	0	—	0.012	0.007–0.017	0.022	0.012–0.032
d_B	h^{-1}	0.027	0.016–0.039	0.012	0.007–0.017	0	—
$1/\lambda + \Delta$	h	13.0	12.2–14.1	13.4	12.6–14.4	13.8	12.9–14.9
Δ	h	5.2	4.6–5.8	5.3	4.7–5.9	5.4	4.8–6.0
$1/\lambda_0 + \Delta_0$	h	60.0	57.0–63.5	69.5	66.3–73.4	86.5	75.9–104
Δ_0	h	40.4	37.7–42.9	40.4	37.8–42.	40.3	37.2–42.7

The ranges are 95% confidence limits as estimated by bootstrapping with 500 simulations [3]. Because we know little about the death rates during the A and B phases of the cell cycle, we have fitted the data for $d_A = 0$ (left columns), $d_A = d_B$ (middle columns), and $d_B = 0$ (right columns). The quality of these three fits is very similar, as measured by the MNSQ, or by visual inspection, but decreases slightly going from left to right in the table. We have fixed $\phi = 1$ because allowing $\phi < 1$ failed to improve the quality of the fit (not shown). The latter confirms the observation that almost all cells ultimately divide. Fitting under the constraint that $\lambda = \lambda_0$ decreased the quality of the fit (partial F-test: $F_{[1,27]} = 46.6$, $p < 10^{-5}$).

The dynamics of cells moving through successive divisions where the Smith–Martin model is used to describe the cell cycle, can be formulated as a set of delay differential equations for the cells in the A phase of the various division indexes. The model is

$$\begin{aligned}
\frac{dA_0(t)}{dt} &= -(\lambda_0 + d_0)A_0(t), \quad A_0(0) > 0,\\
\frac{dA_1(t)}{dt} &= 2\lambda_0 A_0(t-\Delta_0)e^{-d_B\Delta_0} - (\lambda + d_A)A_1(t), \quad A_1(0) = 0,\\
\frac{dA_i(t)}{dt} &= 2\lambda A_{i-1}(t-\Delta)e^{-d_B\Delta} - (\lambda + d_A)A_i(t), \quad A_i(0) = 0,
\end{aligned} \tag{33}$$

for $i = 2, \ldots, \infty$.

Since the total number of cells per division number is measured, the data are fitted to $T_i(t) \equiv A_i(t) + B_i(t)$, where the number of cells in the B-phase of each division class are

$$\begin{aligned}
B_0(t) &= \lambda_0 \int_0^{\Delta_0} A_0(t-\tau)e^{-d_B\tau}\,d\tau,\\
B_i(t) &= \lambda \int_0^{\Delta} A_i(t-\tau)e^{-d_B\tau}\,d\tau, \quad \text{for} \quad i = 1, \ldots, \infty.
\end{aligned} \tag{34}$$

The solution of this model is given in the Appendix, and the best fit of this solution, with $d_0 = d_A$, to the data is depicted in Fig. 3b.

The fit of the Smith–Martin model to the Gett and Hodgkin data is much better than that of the ODE model (see Fig. 3 and Table 2). However, as explained in the Section 1, the main problem with fitting with the Smith–Martin model is that additional parameters are introduced. For example, little is known about the relative rates of death in the two phases of the cell cycle. Different assumptions about whether death occurs in the A state, the B phase or both can make a big difference in the parameter estimates obtained, and CFSE profiles by themselves contain too little information to estimate these parameters [8]. We therefore fit the data for three different assumptions about the death rates (see Table 2): $d_A = 0$,

$d_A = d_B$, and $d_B = 0$. The $d_A = 0$ fit shown in Fig. 3b is slightly better than the other two, but basically the quality of the fit that is obtained under these three assumptions is very similar (not shown). Fortunately, the parameter estimates remain similar for the three different assumptions about the death rates (see Table 2). The time to complete the first division varies between 60 and 80 h, and involves a time delay of approximately 40 h, both are reasonably similar to the parameters estimated by the simple time delay model of Eq. (7) (see Table 1). The cell cycle time of the subsequent divisions is now estimated to be 13–14 h, which is faster than the 20 h cell cycle time estimated from the increase in the mean division number. In retrospect this is understandable because we now know that the increase in the mean division number is not expected to have approached the asymptotic regime, and we have seen that during the transient the estimated slope could easily be 66% of the true division rate. Approximately one third, i.e., 5 h of the 13–14 h cell cycle time, is estimated to be deterministic. The death rate is slower than that estimated from the loss of the normalized cell number. If we estimate an average death rate by forcing $d_A = d_B$, we obtain a death rate of $0.012\,h^{-1}$, which is less than half of the death rate estimated above. This is at least partially due to the fact that we have fitted the Smith–Martin model to the data at 60,72, and 84 h, and have obtained the death rate from the data points at 60, 72, 84, and 96 h. In Fig. 1c the loss rate of total normalized cell number seems to increase between 84 and 96 h. Setting $d_A = 0$ or $d_B = 0$ increases the estimates of the death rates, but these estimates only pertain to part of the cell cycle. Because the total number of cells in each division number depends on the initial cell number, we also estimated $A_0(0)$. However, Gett and Hodgkin [5] did not measure this quantity and thus we do not include our estimates in the Table 2 because we have no data to compare them with.

6. Virtual data sets

Because the previous sections cast serious doubts on using ODE models to describe CFSE data we finish by showing that what we have learned about the increase in the means μ, μ_2 and $\widehat{\mu_2}$ of ODE models remains valid under more realistic dynamical regimes, like the Smith–Martin model, or a model where cell cycle times are drawn from Gaussian distributions [2]. To this end we have developed a simple stochastic simulator in which individual cells undergo the discrete events "death" and "division" with the appropriate probabilities, and stay in the B-phases for the appropriate length of time. With this computer simulation model we either study the Smith–Martin cell cycle model having an exponential and a fixed deterministic phase, or a model where we draw the length of each cell cycle from a normal distribution. The latter model is constructed from the Smith–Martin model by setting the length of the A-phase to zero, and drawing the parameter Δ from a Gaussian distribution. The cell cycle times in the two simulation models are taken from the first column in Table 2. The death rate is chosen to be uniform throughout the cell cycle, and is set to $d = d_A = d_B = 0.02\,h^{-1}$. If these death rates were unequal, the slope of μ versus t is no longer expected to reflect the cell cycle time [8]. In the "Gaussian" model the standard deviation of the cell cycle time of the first division is set to 25% (i.e., 15 h), and that of later divisions to 10% (i.e., 1.3 h), because Gett and Hodgkin [5] suggested that most of the variance is due to the first division. With an average cell cycle time of 13 h the expected slope of the normalized mean is $\mu_{2\infty}' = 0.077\,h^{-1}$ in both models.

In Fig. 4 we study three parameter regimes: (1) all cells ultimately divide ($\phi = 1$) and equal death rates ($d_A = d_B = 0.02\,h^{-1}$), (2) half of the precursors cells divide ($\phi = 0.5$) and equal death rates, and (3) $\phi = 1$ and precursors cells live longer than activated cells (i.e., $d_0 = 0.001\,h^{-1}$ and $d_A = d_B = 0.02\,h^{-1}$).

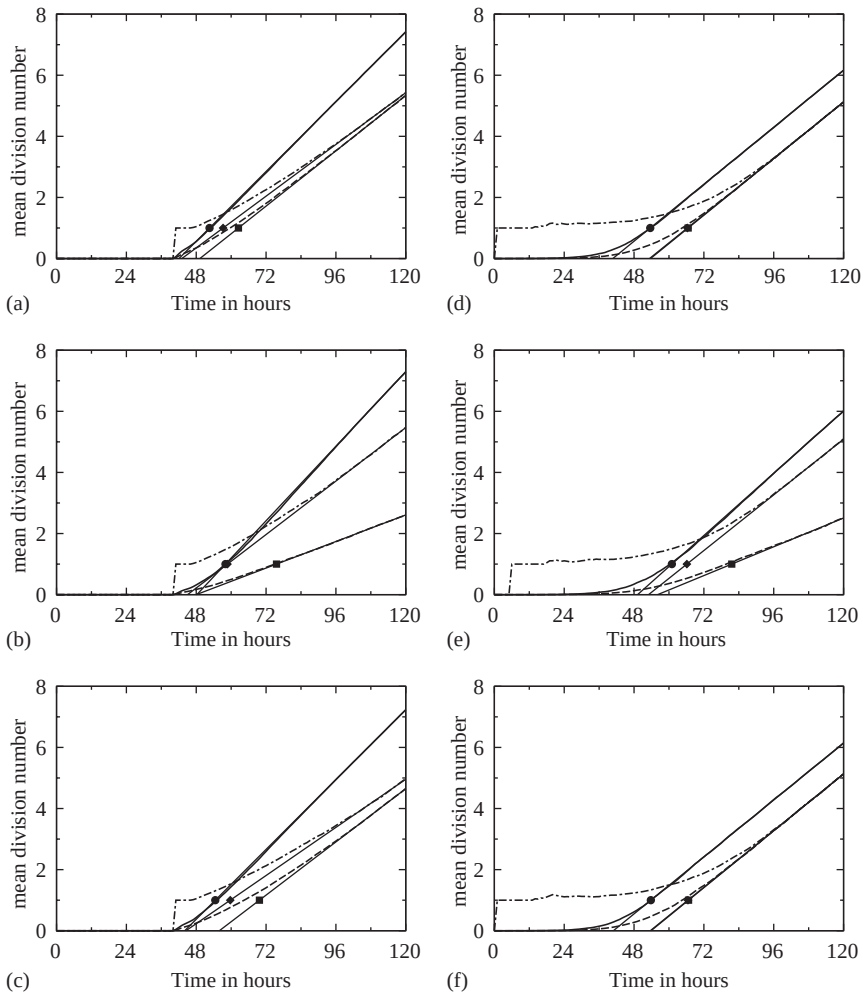

Fig. 4. To test the effect of the distribution of cell cycle times we simulate individual cells undergoing the discrete events "death" and "division" with probabilities according to the Smith–Martin model ((a)–(c)), or by drawing cell cycle times from a Gaussian distribution ((d)–(f)). In panels (a) and (d) all cells ultimately divide ($\phi = 1$) and we have equal death rates ($d_A = d_B = 0.02\,h^{-1}$), in panels (b) and (e) half of the precursors cells divide ($\phi = 0.5$) and equal death rates, and in panels (c) and (f) $\phi = 1$ and precursors cells live longer than activated cells (i.e., $d_0 = 0.001\,h^{-1}$ and $d_A = d_B = 0.02\,h^{-1}$). The proliferation parameters are taken from the first column is Table 2, i.e., the time to complete the first division is 60 h, and a subsequent cell cycle time of 13 h. In panels (c) and (f) we set the death rate of the undivided cells to $d' = 0.001\,h^{-1}$, and keep $d = d_A = d_B = 0.02\,h^{-1}$ for all other cells. The solid line represents $\mu(t)$, the dashed line $\mu_2(t)$, and the dash-dotted line $\widehat{\mu_2}(t)$. The light solid lines are the regression lines fitted through the last 20 h, and the symbols denote the time at which the regression line has a y-value of one (which is an estimate for the time to complete the first division, see Table 3).

The regression lines depicted in Fig. 4 are taken from the data over the last 20 h, and their slopes are given in Table 3. In all cases these regression lines provide a reasonable description of the asymptotic behavior of the model. Generally the results confirm our findings with the ODE models: if half of the cells fails to divide ($\phi = 0.5$) the slope of μ_2 is half of the slope of $\widehat{\mu_2}$ (see Table 3), and when nondivided

Table 3
Slopes and offsets from Figs. 4 and 5 as estimated by linear regression over the last 20 h

	$d_A = d_B$			$\phi = 0.5$			$d_0 = 0.001$		
Slopes (expected for μ_2: 0.077 h^{-1})									
	μ	$\widehat{\mu_2}$	μ_2	μ	$\widehat{\mu_2}$	μ_2	μ	$\widehat{\mu_2}$	μ_2
Smith–Martin	0.095	0.070	0.075	0.102	0.073	0.036	0.095	0.066	0.073
Gaussian	0.078	0.077	0.077	0.085	0.076	0.039	0.078	0.077	0.077
Time to mean of one (h) (expected for μ_2: 60 h)									
	μ	$\widehat{\mu_2}$	μ_2	μ	$\widehat{\mu_2}$	μ_2	μ	$\widehat{\mu_2}$	μ_2
Smith–Martin	52.6	57.2	62.5	58.0	58.7	75.6	54.5	59.7	69.7
Gaussian	53.6	66.3	66.6	61.0	66.1	81.6	53.8	66.5	66.7
Slopes of cell numbers (h^{-1})									
	r	d	d_0	r	d	d_0	r	d	d_0
Smith–Martin	0.040	−0.019	−0.070	0.039	−0.020	−0.021	0.039	−0.020	−0.053
Gaussian	0.033	−0.020	—	0.031	−0.021	−0.022	0.034	−0.020	—

Here r is the natural rate of increase, i.e., the slope of the natural logarithm of the total cell number over the last 20 h, d is the average death rate, i.e., the slope of the natural logarithm of the total normalized cell numbers, d_0 is the slope of the natural logarithm of the nondivided cells (which depends on d_0, d_B and λ_0).

cells have a longer half life than dividing cells ($d_0 < d_A = d_B$) it takes longer to approach the asymptotic behavior. However, these transients are much shorter than those in the ODE model: the corresponding value of $\gamma = 1/60 + 0.001 - 0.02 < 0$ would predict a very long transient (see Fig. 2c and d). Instead, in the Smith–Martin model and in the Gaussian model, the normalized mean breaches $\mu_2 = 1$ after 67–70 h, and shortly afterwards the asymptotic regime is approached. Thus, the long transients present in the ODE models are not confirmed by these more realistic models. However, it remains true that the time to breach $\mu_2 = 1$ fails to accurately estimate the time to complete the first division.

Interestingly, we find a large difference in the slope of μ, and not in that of μ_2, between the Gaussian and the Smith–Martin model. This would argue that the slope of μ_2 is more invariant to the underlying cell cycle model, and would hence provide a more reliable estimate of the cell cycle time than the slope μ' does. Ganusov et al. [4] and Pilyugin et al. [8] have shown for a homogeneous Smith–Martin model (in which $\lambda_0 = \lambda$ and $\Delta_0 = \Delta$) that the rate of increase in the mean division number μ approaches

$$\mu' = \frac{1}{[r + \lambda + d_A]^{-1} + \Delta}, \tag{35}$$

where r is the net rate of increase of the total cell number. This is in good agreement with our simulation results. First, for the simulations depicted in Fig. 4a where $d_A = d_B$, a rate of increase of $r = 0.04\,h^{-1}$ is approached (see Fig. 5a and Table 3). This means that $\mu' = 1/(0.188^{-1} + 5.2) = 0.095\,h^{-1}$, which is indeed found in Table 3. Thus, unlike in the ODE models, μ'_∞ fails to reflect the cell cycle time (whereas $\mu_{2\infty}'$ does because $d_A = d_B$, cf. Pilyugin et al. [8]). Moreover, in the Gaussian model where the average cell cycle time is Δ (and $\lambda \to \infty$) one obtains $\mu' = 1/13 = 0.077\,h^{-1}$, which is indeed found in the Table

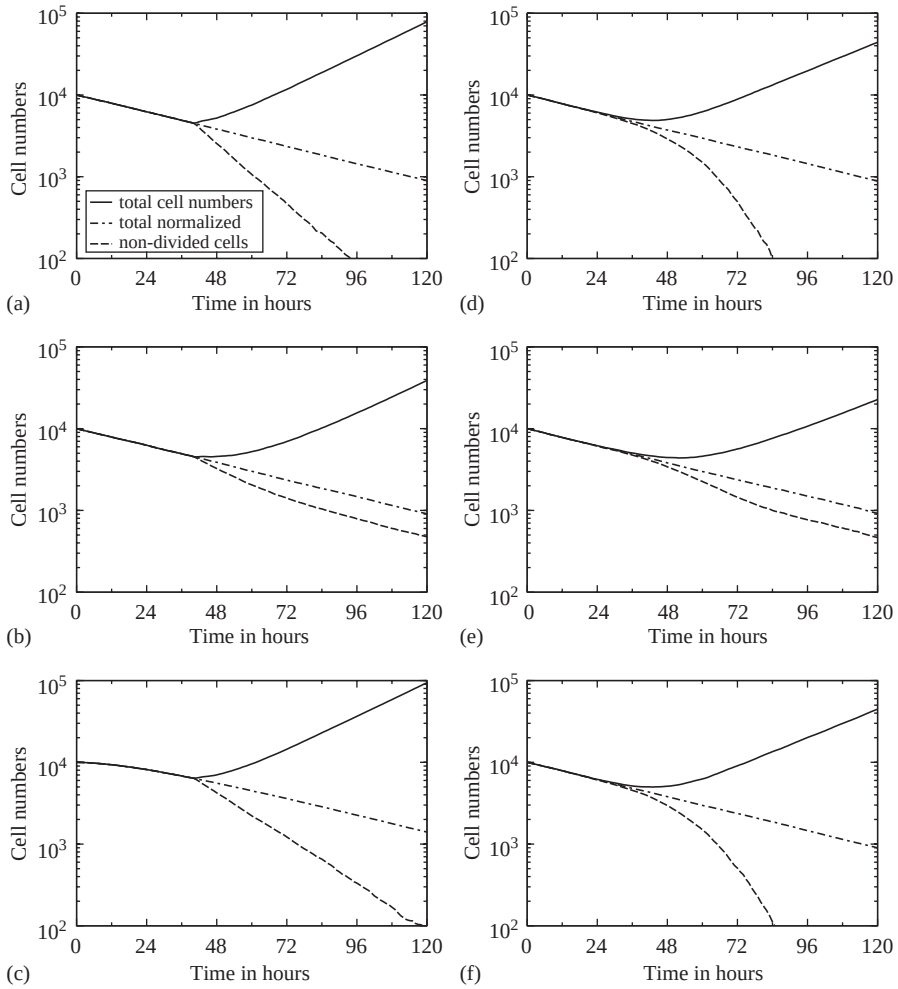

Fig. 5. The total population size (solid line), the normalized total (dashed-dotted line), and the number of nondivided cells (dashed line) of the six simulations in Fig. 4.

3, and which is similar to the slope of μ_2. Apparently, the normalization by 2^i is an appropriate scaling [8] for this model where $d_A = d_B$, and where the population is expanding. As a consequence one obtains the expected slope for the normalized mean division number whatever the underlying model. This strongly argues in favor of using the normalized mean for estimating the cell cycle time. Above we argued for the ODE model that the nonnormalized μ was more reliable because it has the shortest transient. We now see that these transients are less of an issue in cell cycle models with delays.

Above we have shown that one can use the growth rate of the total population, the death rate of the normalized total, and the rate of loss of the nondivided cells to estimate various parameters. Unfortunately, the growth rate depends strongly on the underlying model (see Table 3 and Fig. 5). Due to the exponential distribution of the triggering events in the A-phase of the Smith–Martin model, the growth rate in the Smith–Martin model is faster than that in the Gaussian model. Indeed, most cells in the

Smith–Martin model have cell cycle times faster than $1/\lambda + \Delta$ h. Because the growth rate depends so strongly on the underlying model, it is an unreliable parameter to employ for estimating other parameters. The average death rate $d = d_A = d_B$ is in all cases estimated reliably by the slope of the normalized cell numbers (see Fig. 5). The loss of the nondivided cells is more variable because it depends on the difference in the death rates, and on the fraction of cells that never divide (see Fig. 5). Interestingly the loss of the nondivided cells approaches an exponential in the Smith–Martin model, whereas in the Gaussian model undivided cells are lost much more rapidly. This could possibly be used to discrimate between the two models.

7. Discussion

We have addressed the issue of estimating proliferation and death rates from CSFE data using a variety of models. For ODE models we found that examining the rate of increase of the mean division number could be informative, but depending on details of the model, there could be transients that make interpretation difficult. For our most realistic models in which the time to complete the first division is longer than for subsequent divisions, the typical length of the initial transient is identical (or similar) to the time to complete the first division. This is a natural result because the transient is due to the fact that the first division takes longer than later divisions. The smaller the difference between the times to complete the first and the later divisions, the shorter the transient. Similarly, the fact that the transient takes longer, and that it takes longer to breach the mean division number one, when nondivided cells live longer than cells that have divided is natural. The longer the nondivided cells live, the longer they slow down the increase of the mean division number. Finally, the fact the transient is the shortest for $\mu(t)$ is natural because $\mu_\infty(t)$ is less dominated by the nondivided cells. When we compute the means μ_2 and $\widehat{\mu_2}$, due to the normalization by 2^i, all division numbers are weighed equally and the nondivided cells have a strong influence on the mean division number.

Proper interpretation of CFSE data remains a hard problem. ODE models are commonly used to describe population kinetics of cells [9,12]. However, fitting the ODE models we developed here to CFSE data gave rather poor fits (Fig. 3a), with distributions of cells over the division numbers broader than in the data. For example, at early times the best-fit distribution places more cells in the higher division numbers than the data exhibit. This appears to be caused by a lack of a lower bound in the cell cycle time, because when the Smith–Martin model is used, much more realistic fitting profiles are obtained. Said another way, in ODE models cells that have completed cell division have the same probability of dividing again as all other cells. This leads, in the model, to cells entering division more rapidly than observed, which is amplified by our assumption of exponential growth. The net effect is that the ODE models give rise to rapid division cascades, which in obtaining the best fit to the data are slowed down by overestimating the cell cycle time.

The elegantly simple model proposed by Gett and Hodgkin [5] was developed on the assumption that recruitment into the first division is stochastic and that the later divisions are deterministic. However, we have shown that the Gaussian shape of the normalized frequency distribution, that was used by Gett and Hodgkin [5] to support this assumption, is also expected to arise in ODE models when the later divisions are not deterministic but in essence occur with an exponential probability distribution. This does not prove that their assumption was wrong, but does invalidate the evidence underlying their model. Also since CFSE data can be fit with the Smith–Martin model that has a stochastic A phase, suggests that the

later divisions may not be completely deterministic. Independent of whether the first division is stochastic and later divisions deterministic, we have shown that the slope of the asymptote of the mean division number obtained from the best fit of the data of the Smith–Martin model remains to reflect the division rate, as suggested by Gett and Hodgkin [5]. Therefore, it seems that the major caveat in interpreting the slope of this asymptote remains the distribution of death over the two phases of the cell cycle [8].

The Gett and Hodgkin [5] normalization method remains very important for interpreting the distribution of cells over the various division numbers, since it reveals the average number of divisions the initial cells have completed. This prevents earlier overestimates of the number of divisions cells typically have completed over the course of the experiment based on the behavior of cells that have divided a maximal number of times [1,6,11]. Plotting the normalized means μ_2 and $\widehat{\mu_2}$ is also important because one has to check whether $\mu_2 > 1$ for most of the time, whether they approach and maintain a linear increase, and whether they have different slopes. The asymptotic slope of μ_2 gives at best an indication of the cell cycle time. This depends on the length of the transient, and on the distribution of death over the A and B phases [8]. One could use this slope as an initial guess for fitting the full Smith–Martin model to the data. The asymptotic slope of μ depends strongly on the underlying model, and is an unreliable estimator of the cell cycle time. The same is true of the growth rate of the total population. Similarly, the loss rate of the normalized population may provide an initial guess for an "average death" rate over the two phases of the cell cycle, once the normalized cell number declines linearly on a logarithmic scale. The loss of the nondivided cells may give an indication about the distribution of recruitment times into the first division. In the Smith–Martin model the number of undivided cells approaches an exponential. This is not so if recruitment into division has a Gaussian distribution (see Fig. 5).

A recipe for interpretating CFSE data would therefore be to plot the various means, μ, μ_2 and $\widehat{\mu_2}$, and the three population sizes, N, n, and N_0, and estimate their asymptotes by linear regression, as we do in Figs. 4 and 5. The slopes of these regression lines can be used as an initial guess for fitting the Smith–Martin model to the data. Moreover, a difference between the slopes of μ_2 and $\widehat{\mu_2}$ would suggest that a fraction of the precursor cells fails to divide. Because little is known about the distribution of death between the A and B phase of the cell cycle, and because when in the cell cycle cells die can make a big difference in proliferation rate estimates [8], the Smith–Martin model should be fitted to the three possibilities studied in Table 2. Alternatively, or additionally, one can use the rescaling method devised by Pilyugin et al. [8] once the system has approached its asymptotic regime, i.e., after $\mu_2 > 1$ and the increase in $\mu_2(t)$ with time is linear, to estimate the invariant "mean generation time of surviving cells" and the "fraction of cells that die in one generation". These invariant parameters remain valid even when the cells are not obeying the Smith–Martin model.

Fitting the data generated by the Gaussian model of Fig. 4d with the Smith–Martin model gives a reasonable fit with estimated cell cycle times of 67 (63–71 h) and 15 h (14.3–15.9 h), respectively (not shown). This is somewhat slower than the true average cell cycle time to compensate for the fact that most cells complete the cell cycle faster than average in the Smith–Martin model. Although the virtual cells have no A phase and only a deterministic B phase, fitting with the Smith–Martin model yielded estimates of $\Delta_0 = 43$ h (40–45 h) and $\Delta = 6$ h (5.2–6.5 h). Because a short A phase makes the data more deterministic, the Smith–Martin model fits lengths of the A and B phases that correspond best with the variance in the data. This means that an excellent fit of the Smith–Martin model to CFSE data, i.e., one with tight confidence limits on the Δ parameters fails to provide good evidence for the lengths of the B phases. Rather the tight confidence limits could be due to the fact that there is no variance on Δ in the Smith–Martin model.

Thus, the estimated lengths of the B phases are unreliable. The Smith–Martin model, however, does a reasonable job at estimating the total length of the cell cycle, and the average death rate.

Acknowledgements

We thank Jerome Percus for help with some of the sums and Sergei Pilyugin and Vitaly Ganusov for deriving the solution of the Smith–Martin model. Dan Coombs derived solutions of a version of the Smith–Martin model where the length of the A-state for the later divisions goes to zero. We thank José Borghans for critical discussions on the Gett and Hodgkin [5] method and for helpful comments on the manuscript. We thank Phil Hodgkin, Rustom Antia, Vitaly Ganusov and Segei Pilyugin for helpful discussions. Portions of the work were done under the auspices of the US Department of Energy under contract W-7405-ENG-36 and supported by NIH grants AI28433 and RR06555 to A.S.P. This work on analyzing CFSE data is supported by Research Grant RGP10/2004 from the Human Frontiers Science Program awarded to R.J.d.B, A.S.P., and Phil Hodgkin.

Appendix A.

One can use Laplace transforms to solve the delay differential equations, Eq. (33), of the Smith–Martin model. Start by writing the Laplace transforms of $A_i(t)$ for $i = 0, 1, \ldots$

$$\mathscr{L}\{A_0(t)\} = \frac{A_0(0)}{s + \lambda_0 + d_0}, \tag{36}$$

and

$$\mathscr{L}\{A_i(t)\} = \left(\frac{2\lambda \mathrm{e}^{-(s+d_\mathrm{B})\Delta}}{s + \lambda + d_\mathrm{A}}\right)^{i-1} \frac{2\lambda_0 \mathrm{e}^{-(s+d_\mathrm{B})\Delta_0}}{s + \lambda + d_\mathrm{A}} \frac{A_0(0)}{s + \lambda_0 + d_0}, \quad i \geqslant 1.$$

Rewrite the above equation as

$$\mathscr{L}\{A_i(t)\} = 2^i \lambda^{i-1} \lambda_0 A_0(0) \frac{\mathrm{e}^{-(s+d_\mathrm{B})\tilde{\Delta}}}{(s + \lambda + d_\mathrm{A})^i} \frac{1}{s + \lambda_0 + d_0}, \quad i \geqslant 1 \tag{37}$$

where $\tilde{\Delta} = (i-1)\Delta + \Delta_0$.

To obtain $A(t)$, we invert the Laplace transforms as follows,

$$\mathscr{L}^{-1}\left\{\frac{1}{s + \lambda_0 + d_0}\right\} = \mathrm{e}^{-(\lambda_0 + d_0)t}, \tag{38}$$

and

$$\mathscr{L}^{-1}\left\{\frac{\mathrm{e}^{-(s+d_\mathrm{B})\tilde{\Delta}}}{(s + \lambda + d_\mathrm{A})^i}\right\} = \frac{\mathrm{e}^{(\lambda + d_\mathrm{A} - d_\mathrm{B})\tilde{\Delta}}}{(i-1)!} \mathrm{e}^{-(\lambda + d_\mathrm{A})t} (t - \tilde{\Delta})^{i-1} H(t - \tilde{\Delta}), \tag{39}$$

where $H(t)$ is the unit step function.

Since convolution of functions in the time domain is equivalent to multiplication of their Laplace transforms, the next step is to compute the convolution of $e^{-(\lambda+d_A)t}(t-\tilde{\Delta})^{i-1}H(t-\tilde{\Delta})$ and $e^{-(\lambda_0+d_0)t}$. Thus, we compute the integral

$$\int_0^t e^{-(\lambda+d_A)z}(z-\tilde{\Delta})^{i-1}H(z-\tilde{\Delta})e^{-(\lambda_0+d_0)(t-z)}\,dz$$
$$= e^{-(\lambda_0+d_0)t}H(t-\tilde{\Delta})\int_{\tilde{\Delta}}^t e^{-(\lambda-\lambda_0)z}(z-\tilde{\Delta})^{i-1}\,dz.$$

Making the shift $z-\tilde{\Delta}\to z$ and rescaling $z(\lambda-\lambda_0)\to z$, we obtain the following expression for this convolution

$$\frac{e^{-(\lambda_0+d_0)t}e^{-(\lambda-\lambda_0)\tilde{\Delta}}}{(\lambda-\lambda_0)^i}H(t-\tilde{\Delta})\Gamma(i,0,(t-\tilde{\Delta})(\lambda-\lambda_0)),$$

where $\Gamma(a,z_0,z_1)=\int_{z_0}^{z_1}e^{-t}t^{a-1}dt$ is the generalized incomplete Gamma function.

Putting all the pieces together, we obtain that

$$A_i(t)=\frac{2^i\lambda^{i-1}\lambda_0A_0(0)e^{-(\lambda_0+d_0)t}e^{-(\lambda-\lambda_0)\tilde{\Delta}}e^{(\lambda+d_A-d_B)\tilde{\Delta}}}{(i-1)!(\lambda-\lambda_0)^i}\times H(t-\tilde{\Delta})\Gamma(i,0,(t-\tilde{\Delta})(\lambda-\lambda_0)),$$

or simply that

$$A_i(t)=\frac{2^i\lambda^{i-1}\lambda_0A_0(0)e^{(\lambda_0+d_A-d_B)\tilde{\Delta}}e^{-(\lambda_0+d_0)t}}{(i-1)!(\lambda-\lambda_0)^i}\times H(t-\tilde{\Delta})\Gamma(i,0,(t-\tilde{\Delta})(\lambda-\lambda_0)),\tag{40}$$

where $i\geqslant 1$, $\tilde{\Delta}=(i-1)\Delta+\Delta_0$, and Γ is as described above.

For $\lambda=\lambda_0$, Eq. (40) simplifies to

$$A_i(t)=\frac{2^i\lambda^iA_0(0)e^{(\lambda+d_A-d_B)\tilde{\Delta}}e^{-(\lambda+d_0)t}}{i!}\times H(t-\tilde{\Delta})(t-\tilde{\Delta})^i\ ,\quad i\geqslant 1.\tag{41}$$

From Eq. (36) and Eq. (38), $A_0(t)=A_0(0)e^{-(\lambda_0+d_0)t}$.

References

[1] V.P. Badovinac, B.B. Porter, J.T. Harty, Programmed contraction of CD8$^+$ T cells after infection, Nat. Immunol. 3 (2002) 619–626.

[2] E.K. Deenick, A.V. Gett, P.D. Hodgkin, Stochastic model of T cell proliferation: A calculus revealing IL-2 regulation of precursor frequencies, cell cycle time, and survival, J. Immunol. 170 (2003) 4963–4972.

[3] B. Efron, R. Tibshirani, Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy, Stat. Sci. 1 (1986) 54–77.

[4] V.V. Ganusov, S.S. Pilyugin, R.J. DeBoer, K. Murali-Krishna, R. Ahmed, R. Antia, Quantifying the cell turnover from CFSE data, J. Immunol. Methods, in press.

[5] A.V. Gett, P.D. Hodgkin, A cellular calculus for signal integration by T cells, Nat. Immunol. 1 (2000) 239–244.

[6] S.M. Kaech, R. Ahmed, Memory CD8$^+$ T cell differentiation: initial antigen encounter triggers a developmental program in naive cells, Nat. Immunol. 2 (2001) 415–422.

[7] A.B. Lyons, Analysing cell division in vivo and in vitro using flow cytometric measurement of CFSE dye dilution, J. Immunol. Methods 243 (2000) 147–154.
[8] S.S. Pilyugin, V.V. Ganusov, K. Murali-Krishna, R. Ahmed, R. Antia, The rescaling method for quantifying the turnover of cell populations, J. Theor. Biol. 225 (2003) 275–283.
[9] P. Revy, M. Sospedra, B. Barbour, A. Trautmann, Functional antigen-independent synapses formed between T cells and dendritic cells, Nat. Immunol. 2 (2001) 925–931.
[10] J.A. Smith, L. Martin, Do cells cycle?, Proc. Nat. Acad. Sci. USA 70 (1973) 1263–1267.
[11] M.J. Van Stipdonk, E.E. Lemmens, S.P. Schoenberger, Naive CTLs require a single brief period of antigenic stimulation for clonal expansion and differentiation, Nat. Immunol. 2 (2001) 423–429.
[12] H. Veiga-Fernandes, U. Walter, C. Bourgeois, A. McLean, B. Rocha, Response of naive and memory $CD8^+$ T cells to antigen stimulation in vivo, Nat. Immunol. 1 (2000) 47–53.