Modeling of Cell Proliferation with Flow Cytometry Data from CFSE-based Assays

H.T. Banks

Center for Quantitative Sciences in Biomedicine
Center for Research in Scientific Computation
Department of Mathematics
North Carolina State University
Raleigh, NC

Clay Thompson and Karyn Sutton
CRSC Postdocs at NCSU

Tim Schenkel
Department of Virology, Saarland University, Homburg, Germany

Jordi Argilaguet, Sandra Giest,
Cristina Peligero, Andreas Meyerhans
ICREA Infection Biology Lab, Univ. Pompeu Fabra, Barcelona, Spain

Gennady Bocharov
Institute of Numerical Mathematics, RAS, Moscow, Russia

Selected Sources and References

- H.T. Banks, Frederique Charles, Marie Doumic, Karyn L. Sutton, and W. Clayton Thompson, Label structured cell proliferation models, *Appl. Math. Letters* **23** (2010), 1412–1415; doi:10.1016/j.aml.2010.07.009

- H.T. Banks, Karyn L. Sutton, W. Clayton Thompson, Gennady Bocharov, Dirk Roose, Tim Schenkel, and Andreas Meyerhans, Estimation of cell proliferation dynamics using CFSE data, CRSC-TR09-17, North Carolina State University, August 2009; *Bull. Math. Biol.*, **70** (2011), 116–150.

- H.T. Banks, Karyn L. Sutton, W. Clayton Thompson, G. Bocharov, Marie Doumic, Tim Schenkel, Jordi Argilaguet, Sandra Giest, Cristina Peligero, and Andreas Meyerhans, A new model for the estimation of cell proliferation dynamics using CFSE data, CRSC-TR11-05, North Carolina State University, Revised July 2011; *J. Immunological Methods,* **373** (2011), 143–160; DOI:10.1016/j.jim.2011.08.014.

- H.T. Banks, W. Clayton Thompson, Jordi Argilaguet, Cristina Peligero, and Andreas Meyerhans, A division-dependent compartmental model for computing cell numbers in CFSE-based lymphocyte proliferation assays, CRSC-TR12-03, N. C. State University, Raleigh, NC, January, 2012; *Math. Biosci. Engr.*, to appear.

- H.T. Banks and W. Clayton Thompson, Mathematical models of dividing cell populations: Application to CFSE data, CRSC-TR12-10, N. C. State University, Raleigh, NC, April, 2012; *Journal on Mathematical Modelling of Natural Phenomena*, to appear.

- H.T. Banks and W. Clayton Thompson, A division-dependent compartmental model with cyton and intracellular label dynamics, CRSC-TR12-12, N. C. State University, Raleigh, NC, May, 2012; *Intl. J. of Pure and Applied Math.*, **77** (2012), 119–147.

Selected Sources and References

- H.T. Banks, Amanda Choi, Tori Huffman, John Nardini, Laura Poag and W. Clayton Thompson, Modeling CFSE label decay in flow cytometry data, CRSC-TR12-20, N. C. State University, Raleigh, NC, November, 2012; *Applied Mathematical Letters*, (2103), doi:10.1016/j.aml.2012.12.010.

- H.T. Banks, D. F. Kapraun, W. Clayton Thompson, Cristina Peligero, Jordi Argilaguet and Andreas Meyerhans, A novel statistical analysis and interpretation of flow cytometry data, CRSC-TR12-23, N. C. State University, Raleigh, NC, December, 2012; *J. Biological Dynamics*, submitted.

- A.V. Gett and P.D. Hodgkin, A cellular calculus for signal integration by T cells, *Nature Immunology* **1** (2000), 239–244.

- E.D. Hawkins, Mirja Hommel, M.L Turner, Francis Battye, J Markham and P.D Hodgkin, Measuring lymphocyte proliferation, survival and differentiation using CFSE time-series data, *Nature Protocols*, **2** (2007), 2057–2067.

- E.D. Hawkins, M.L. Turner, M.R. Dowling, C. van Gend, and P.D. Hodgkin, A model of immune regulation as a consequence of randomized lymphocyte division and death times, *Proc. Natl. Acad. Sci.*, **104** (2007), 5032–5037.

- T. Luzyanina, D. Roose, T. Schenkel, M. Sester, S. Ehl, A. Meyerhans, and G. Bocharov, Numerical modelling of label-structured cell population growth using CFSE distribution data, *Theoretical Biology and Medical Modelling* **4** (2007), Published Online.

- C. Parish, Fluorescent dyes for lymphocyte migration and proliferation studies, *Immunology and Cell Biol.* **77** (1999), 499–508.

Selected Sources and References

- B. Quah, H. Warren and C. Parish, Monitoring lymphocyte proliferation in vitro and in vivo with the intracellular fluorescent dye carboxyfluorescein diacetate succinimidyl ester, *Nature Protocols* **2**:9 (2007), 2049–2056.
- D. Schittler, J. Hasenauer, and F. Allgöwer, A generalized model for cell proliferation: Integrating division numbers and label dynamics, *Proc. Eighth International Workshop on Computational Systems Biology (WCSB 2011)*, June 2001, Zurich, Switzerland, p. 165–168.
- J. Hasenauer, D. Schittler and F. Allgöwer, Analysis and simulation of division- and label-structured population models: A new tool to analyze proliferation assays, *Bull. Math. Biol.*, **74** (2012), 2692–2732.

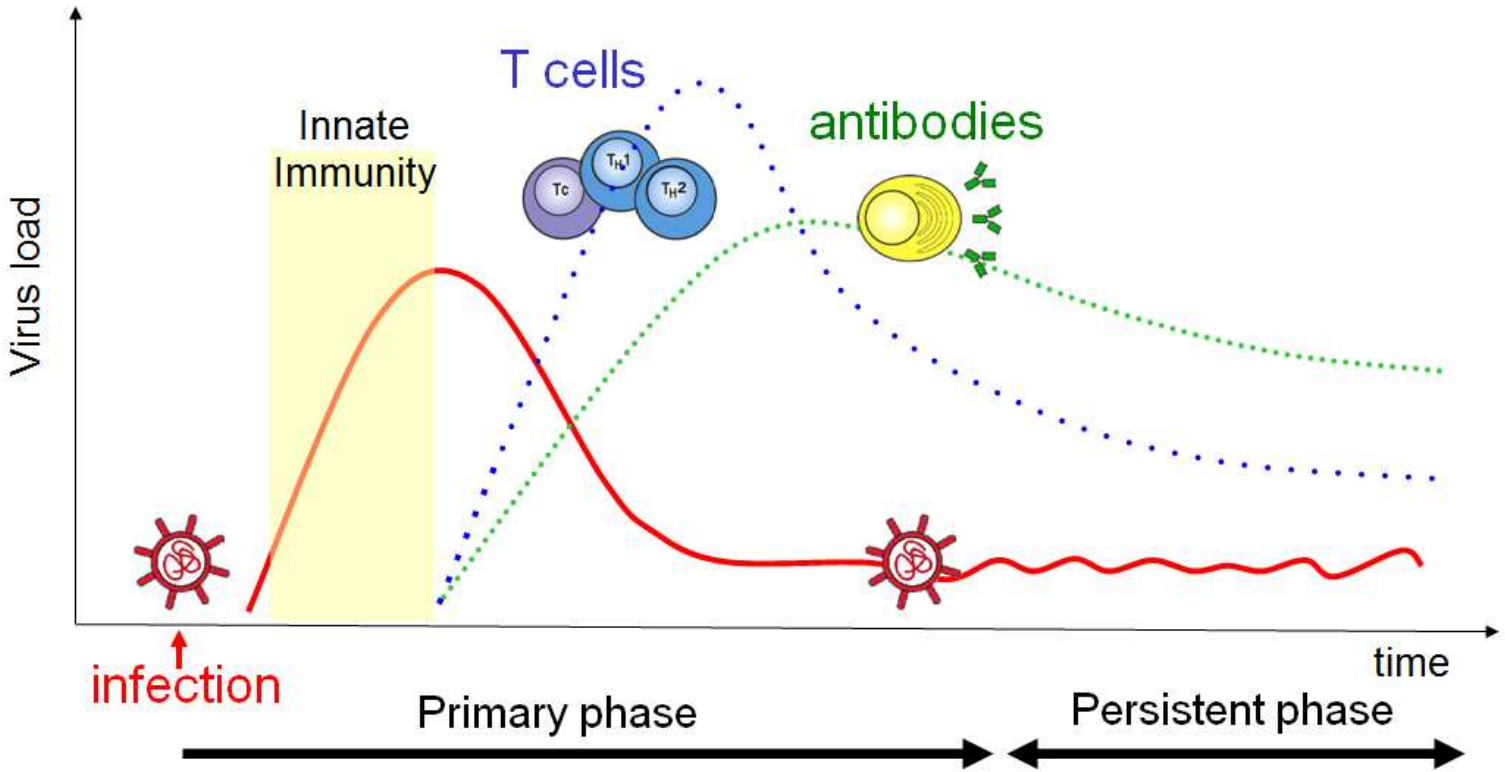
Scientific Goals

Long-term collaboration between NCSU and UPF

- Monitor cell division and differentiation
  - Assess polyfunctionality
  - Investigate immunospecific extracellular signaling pathways
  - Identify correlated and (ideally) causal relationships between immune mechanisms
- Quantitative measure of ‘dynamic responsiveness’
- Link observed (cellular) behaviors to clinical outcomes; improve clinical outcomes
- The ‘central problem in immunology’ (according to G. Bocharov): to understand the ‘cellular and molecular mechanisms that control the ability of the immune system to mount a protective response against pathogen-derived foreign antigens, but avoid a pathological response to self-antigens’

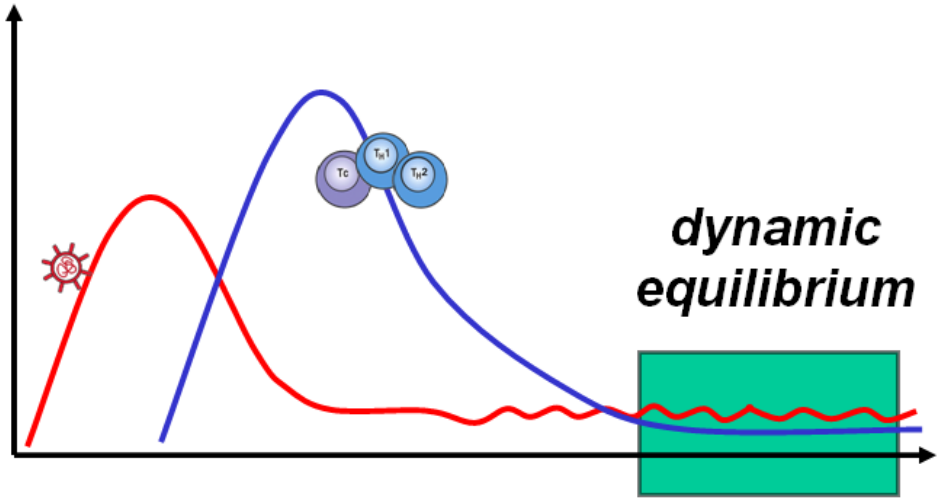
Persistent Infections

Trade-off between immune protection and immunopathy can lead to *persistent infection*



(A. Meyerhans)

Persistent Infections (cont'd)



(A. Meyerhans)

- 'Dynamic equilibrium' between host immune response and microbe expansion
- Applications in HIV, HCV, TB
- Open questions:
  - Characterization of set point
  - Outcome of perturbations, modification for host benefit

Goals

Observations:

- Clinical outcome believed to be strongly influenced by timing and magnitude of ‘clonal expansion’
- CFSE (intracellular dye) + flow cytometry = powerful new tool for tracking cell division

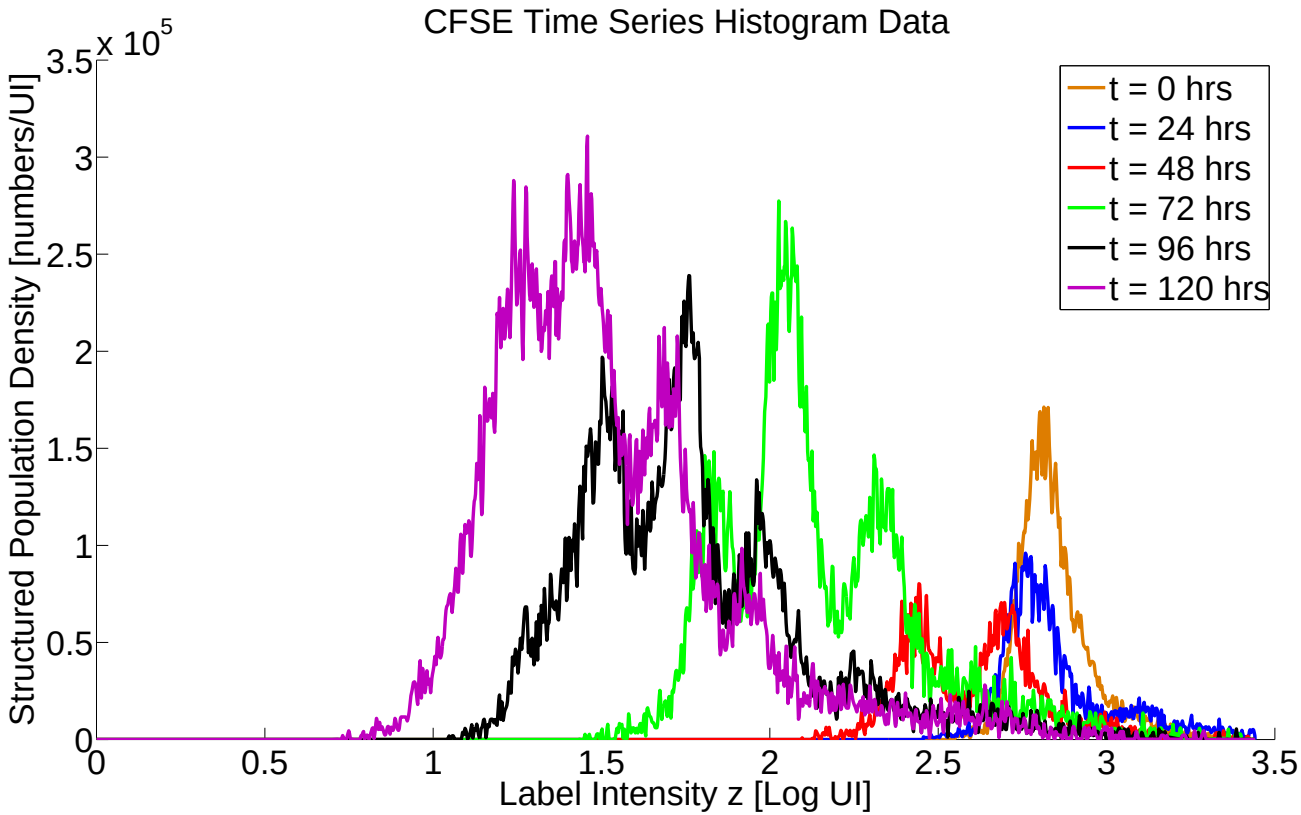
Analysis of data:

- Develop a mathematical model for CFSE data
- Link cell counts to measures of proliferation/death rates
  - Population doubling time; cell cycle time
  - Cell viability

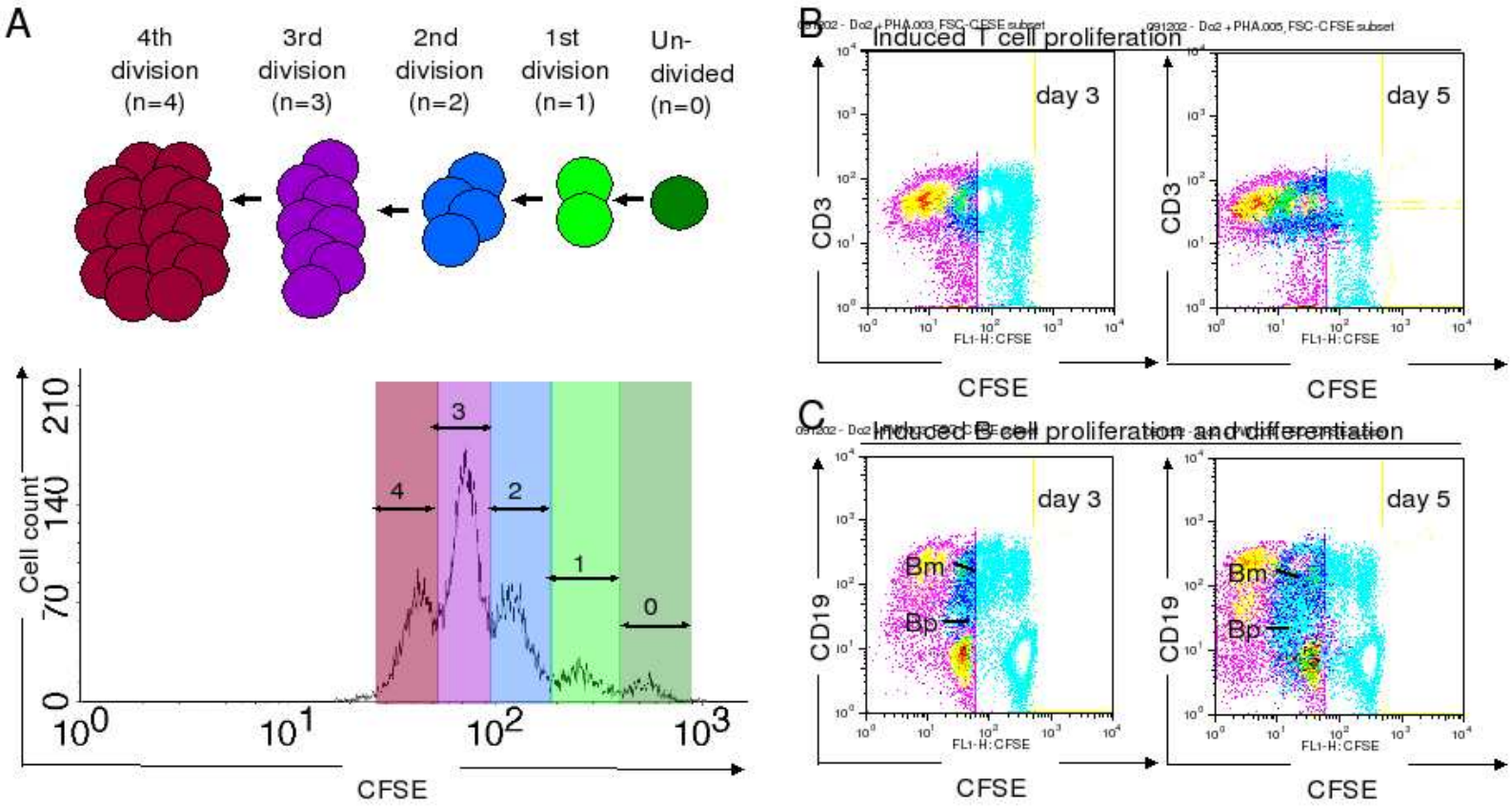
Applications:

- Analyze experimental and biological variability
- Optimal experimental design
- Lab experiments (more in a moment...)

CFSE Data Set



Data Overview

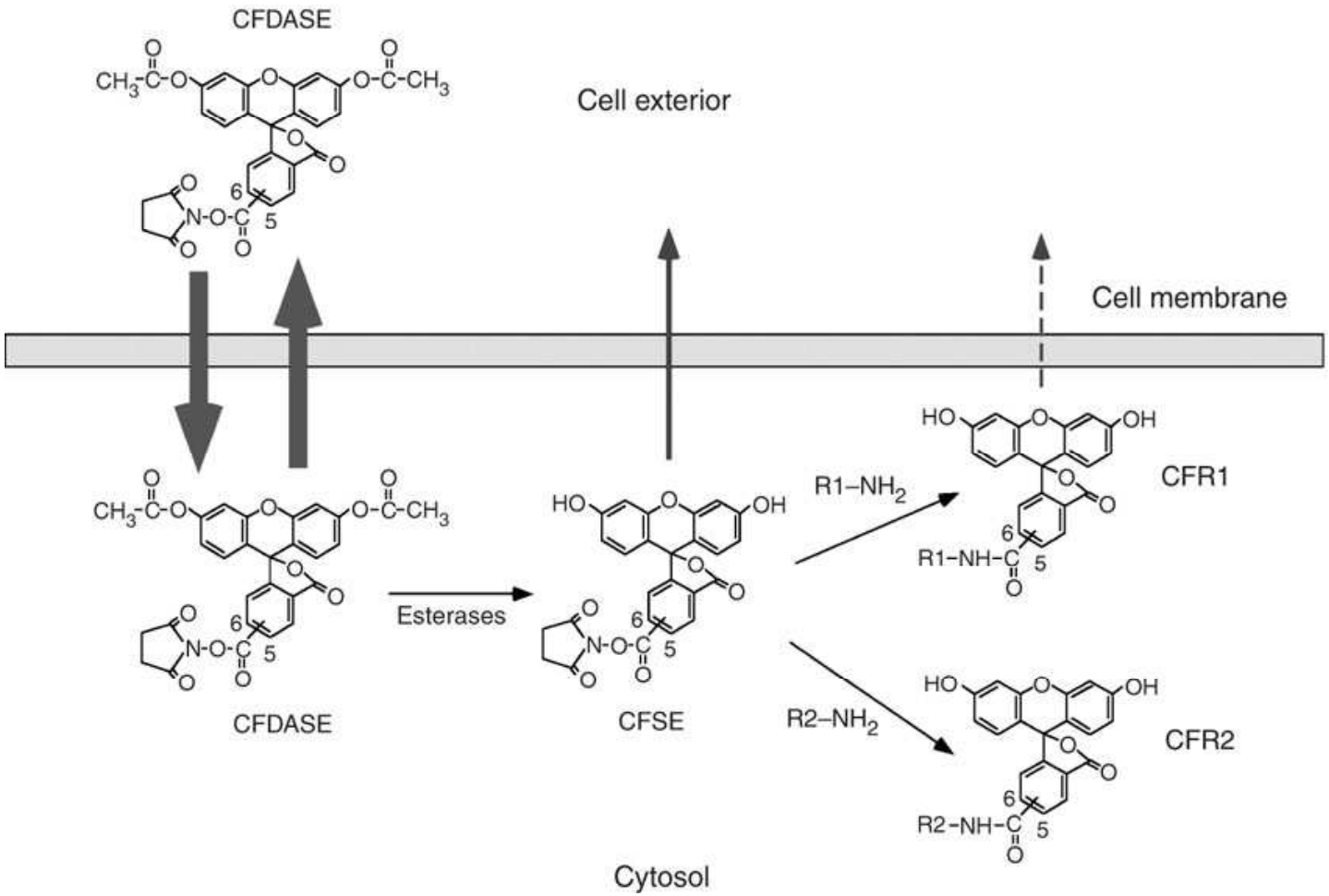


(A. Meyerhans)

CFSE Labeling (Lyons and Parish, 1994)

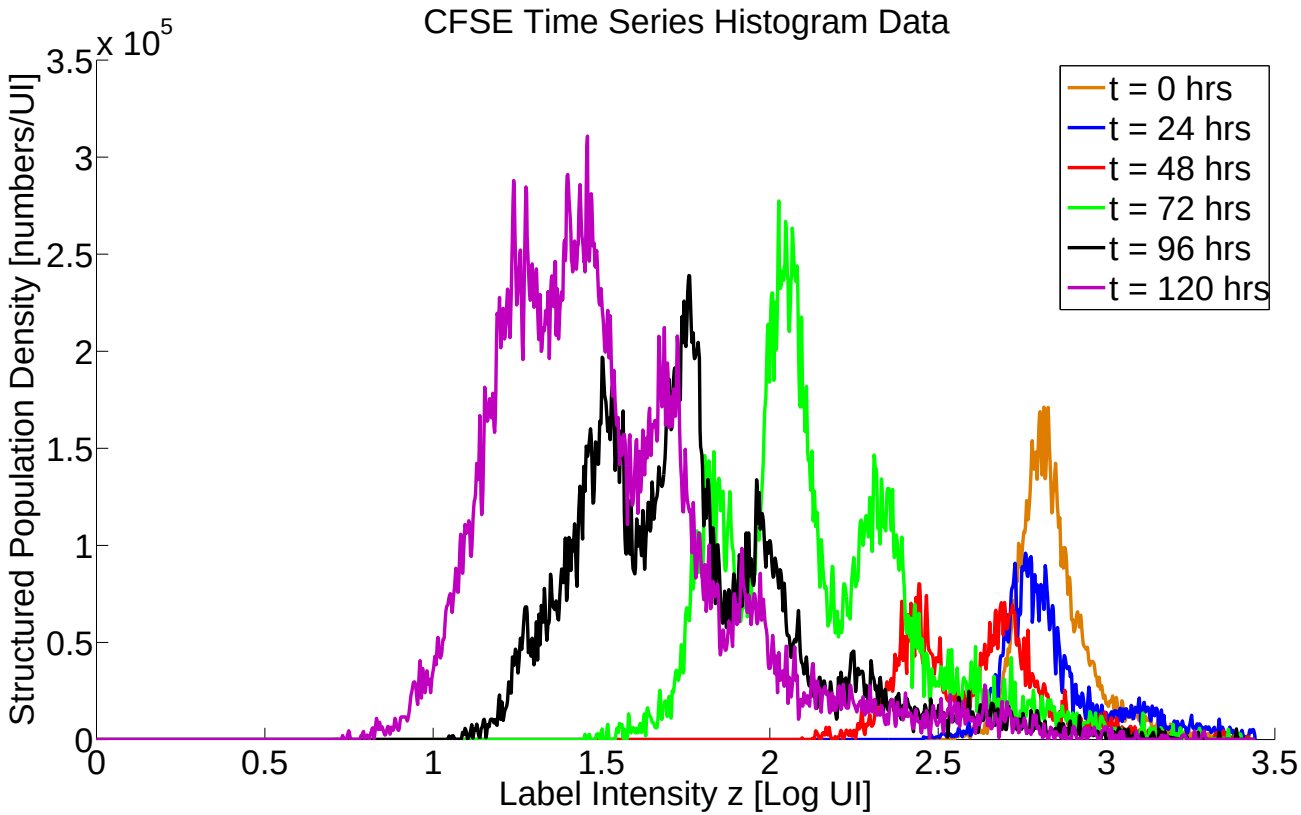
- Cells cultured with CFDA-SE (carboxyfluorescein diacetate succinimidyl ester) then washed
- CFDA-SE becomes protein-bound and fluorescent CFSE (the fluorescent dye carboxyfluorescein succinimidyl ester)
- Dye split between daughter cells at division
- Dye naturally turns over/degrades (very slowly)
- Fluorescence Intensity (FI) of CFSE measured via flow cytometry
- FI linear with dye concentration $\Rightarrow$ FI $\propto$ mass
- Several advantages over other dyes/techniques

CFSE Labeling (Lyons and Parish, 1994)



(C. Parish, Fluorescent dyes for lymphocyte migration and proliferation studies, *Immunology and Cell Biol.* **77** (1999), 499–508.)

CFSE Data Set

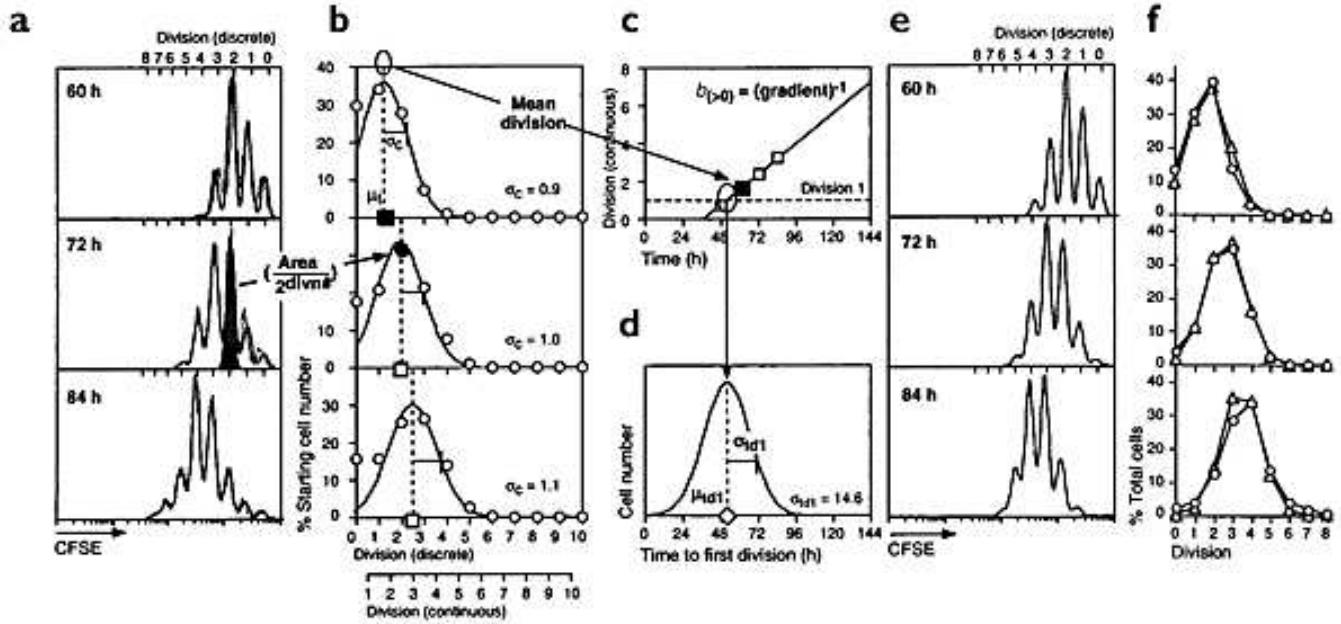


Goals of Modeling

- Cellular ‘Dynamic Responsiveness’
- Link cell counts with proliferation/death rates
  - Population doubling time
  - Cell viability
  - Biological descriptors (cell cycle time, etc.)
- Uncertainty Quantification...
  - ... in the experimental procedure
  - ... for estimated rates/etc
- Analyze cell differentiation and division-linked changes
- Investigate immunospecific extracellular signaling pathways
- Comparison among donors/cell types/disease progression

Traditional Approach (Gett and Hodgkin)

- Traditional ‘semi-quantitative analysis’ pioneered by Gett and Hodgkin et al. (2000)
- Fit data with gaussian curves to determine approximate cells per generation



(A.V. Gett and P.D. Hodgkin, A cellular calculus for signal integration by T cells, *Nature Immunology* **1** (2000), 239–244.)

Traditional Approach (cont'd)

- Gett-Hodgkin method quick, easy to implement, useful comparisons between data sets (e.g. stimulation conditions)
- Compatible with ODE, DDE models; 'indirect fitting' for parameter estimation
- Generalizations, extensions, and various other modeling efforts
  - Smith-Martin model (with generalizations)
  - Cyton model
  - Branching process models

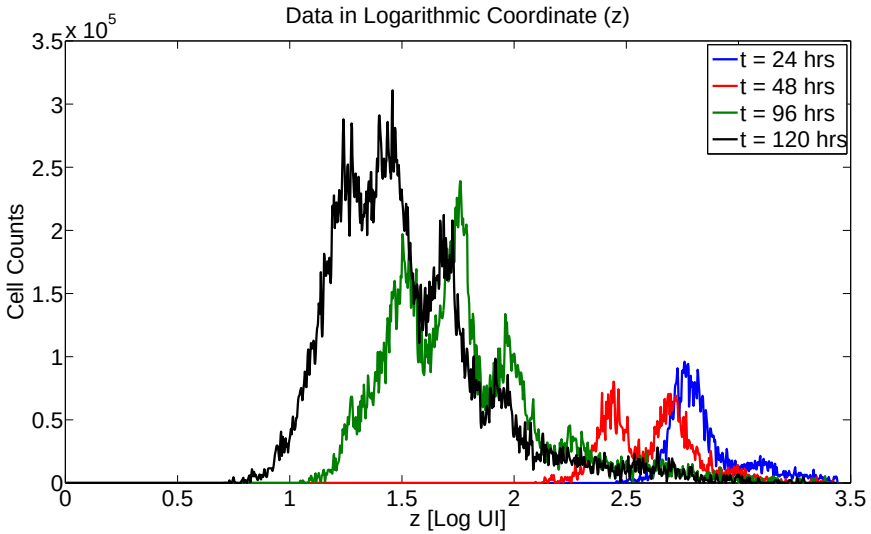
Label-Structured Model

- All previous work with *cell numbers* determined by deconvolution
- Alternatively, we propose to fit the CFSE histogram data directly
  - Capture full behavior of the population density
  - No assumption on the shape of CFSE uptake/distribution
- Histogram presentation of cytometry data makes structured population models a natural choice (as in age, size, etc) except here structure label is "CFSE content or intensity"
  - Key ideas first formulated by Luzyanina, Bocharov, et al., 2007
  - FI (or log FI) $\Leftrightarrow$ Division Number

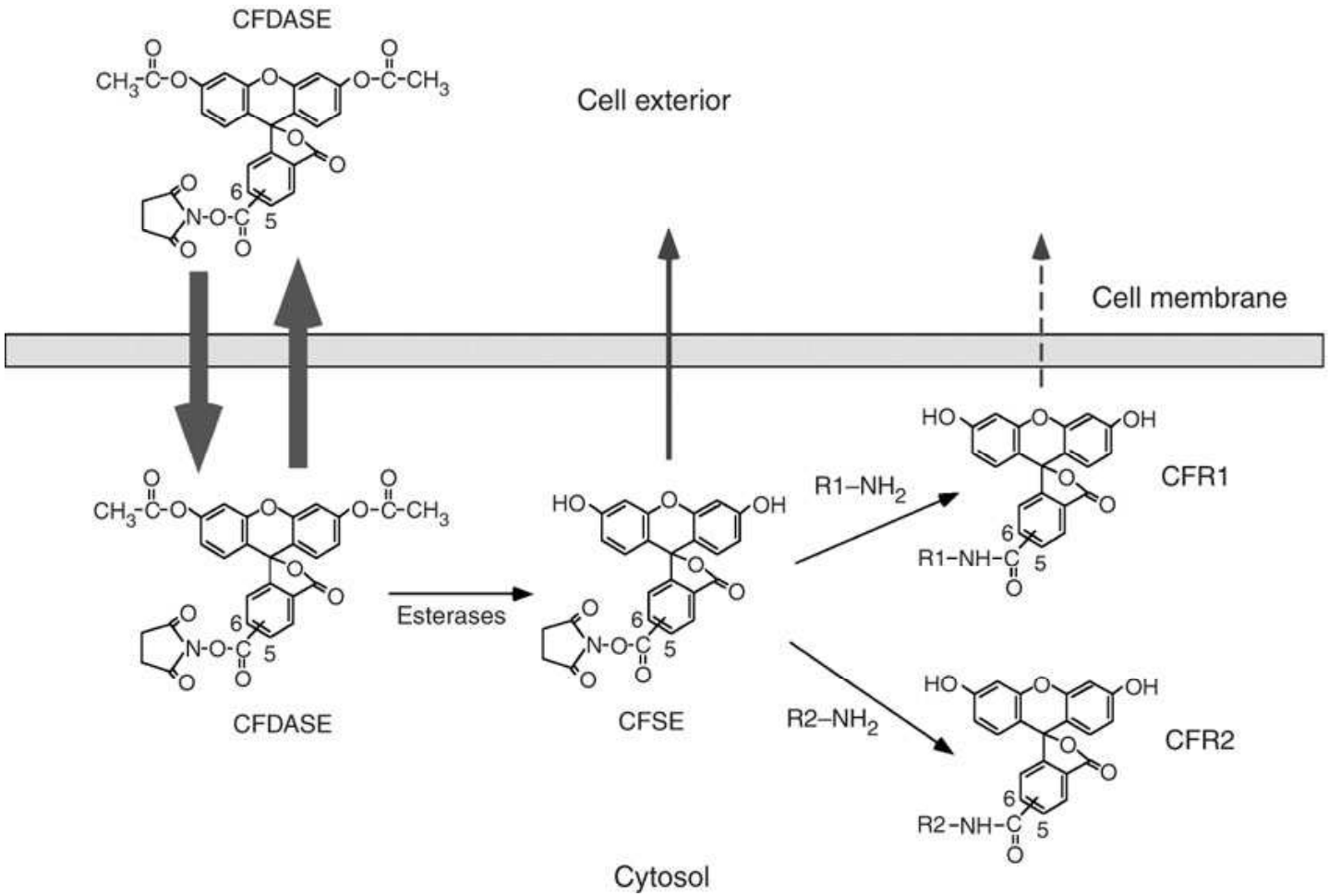
Label-Structured Model (cont'd)

This model must account for (Luzyanina et al., 2007):

- Slow decay of CFSE FI over time
- Dilution of CFSE as cells divide
- Asynchronous division times

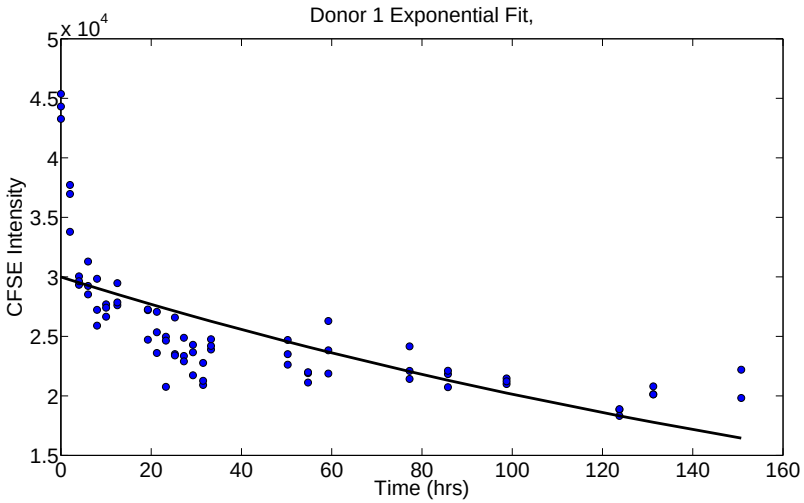


Rate of Label Decay



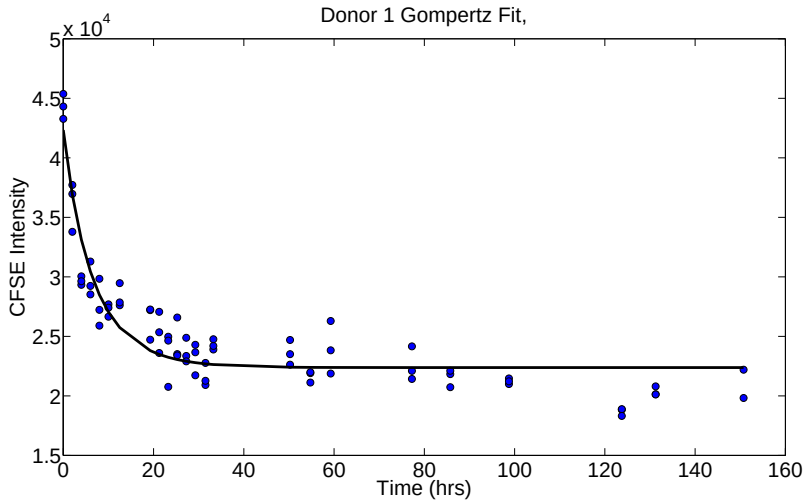
(C. Parish, Fluorescent dyes for lymphocyte migration and proliferation studies, *Immunology and Cell Biol.* **77** (1999), 499–508.)

‘Biphasic Decay’



$$\frac{dx}{dt} = v(x) = c(x - x_a)$$

Exponential



$$\frac{dx}{dt} = v(t,x) = c(x - x_a)e^{-kt}$$

Gompertz

'Translated Coordinate' (Banks et al, 2010)

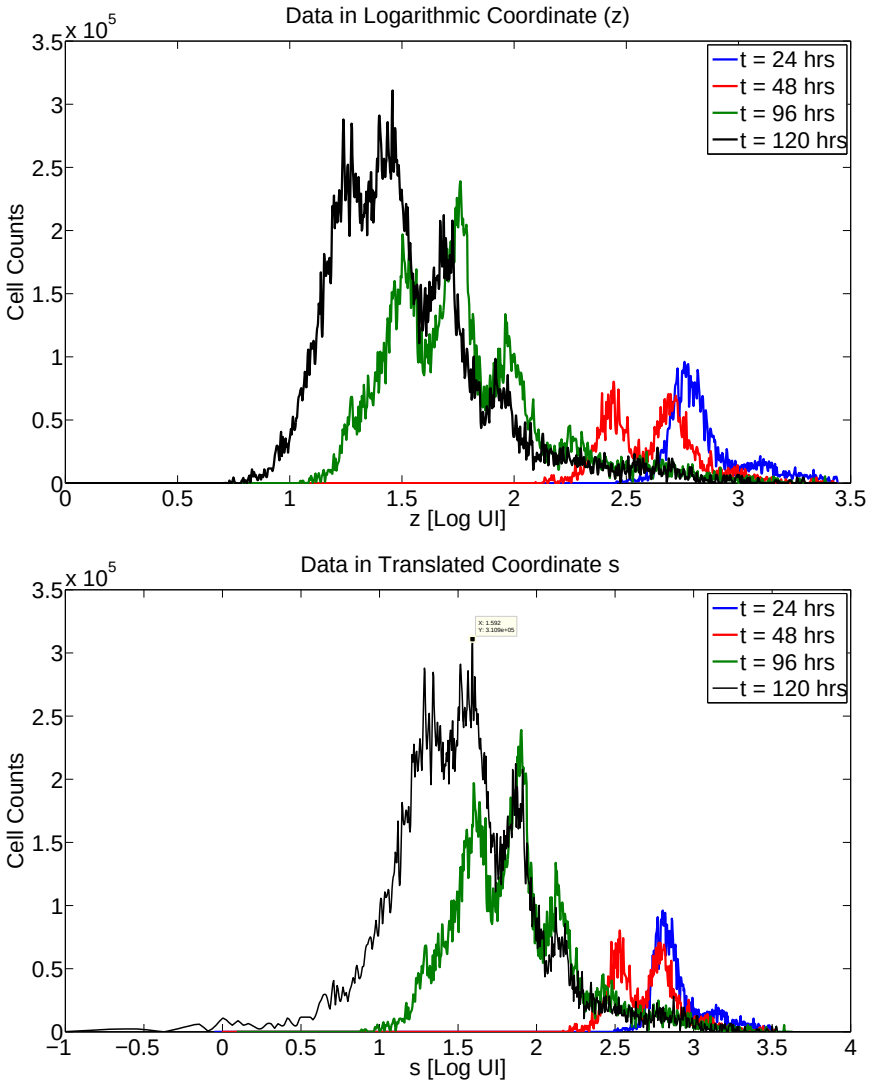
While the structure variable z does correlate with division number, the 'translated variable'

$$s(t,z) = z - \frac{c}{k \log 10}\left(e^{-kt} - 1\right)$$

has an even stronger correlation.

- Follows from the Method of Characteristics
- Intuitively, s represents the FI of a cell with a hypothetical label which does not decay
- Change of reference: *Eulerian* vs *Lagrangian* coordinates

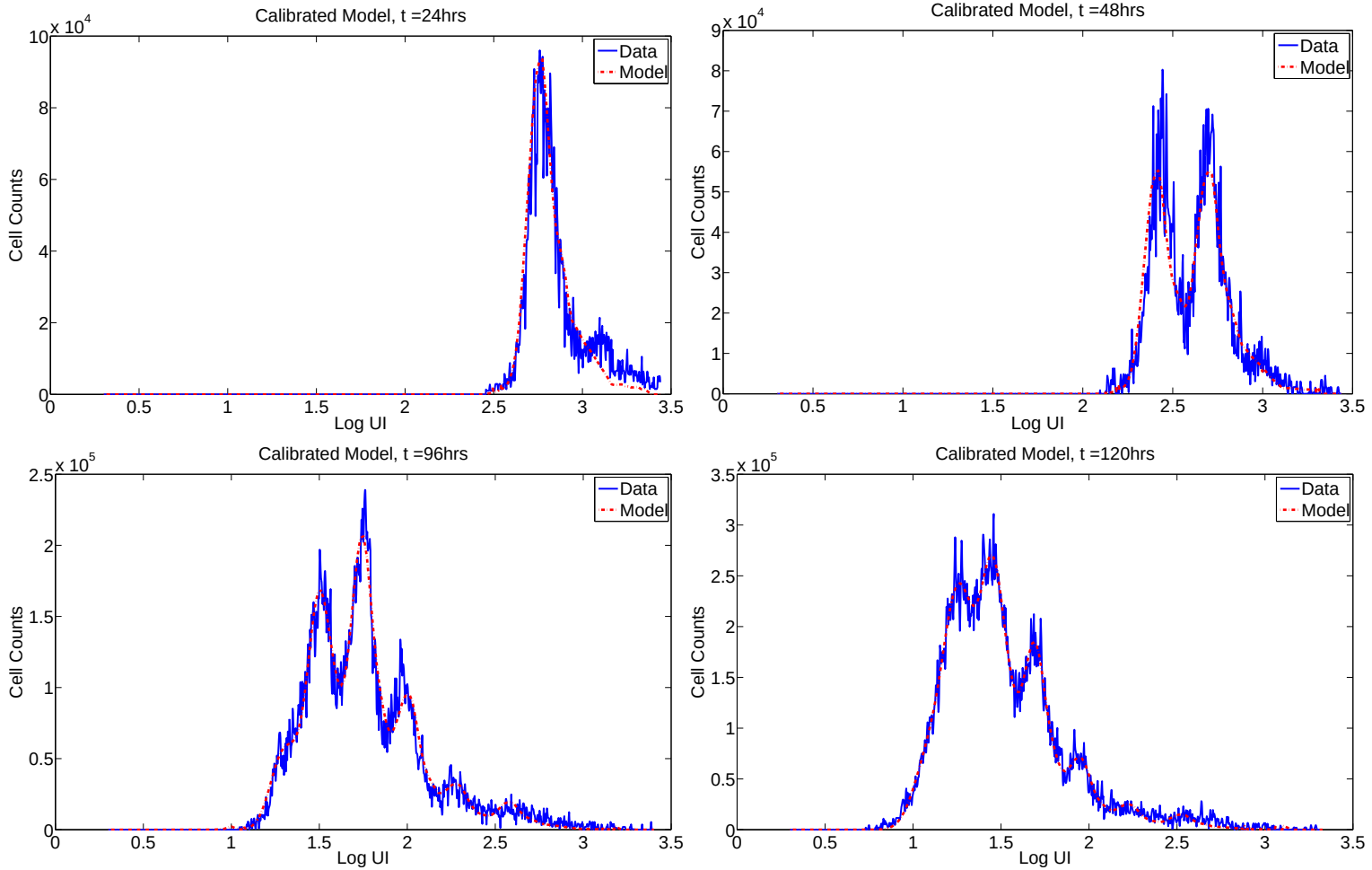
‘Translated Coordinate’ (cont’d)



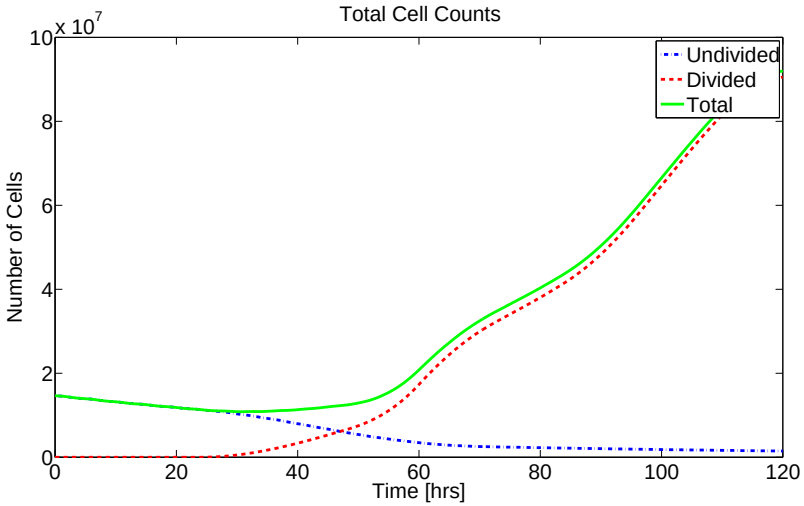
Fragmentation Model Summary

- Model is capable of precisely fitting the observed data
- c, k, x_a estimated consistently (as α and β nodes change), though subject to high experimental variability
- 'Translated coordinate' very strongly correlated with division number
- Time-dependence of the proliferation rate is an essential feature of the model
- Biologically relevant average values of proliferation and death (in terms of number of divisions undergone) are easily computable.
- But...(Aggregate Data/Aggregate Dynamics)
  - Still cannot compute cell numbers and cohort rate parameters
  - Data overlap affecting estimated rates (?)
  - Large number of parameters necessary

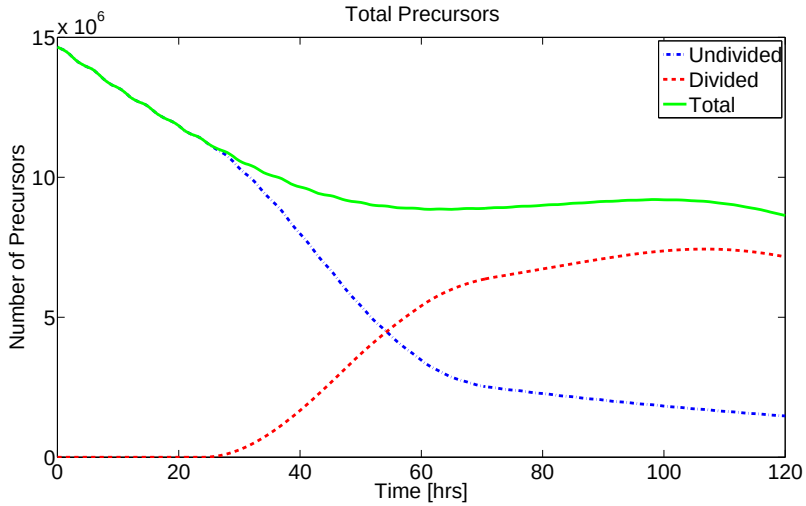
Best-fit, AIC-selected results: Model A5B5Dist



Cell Numbers



$$N_i(t) = \int n_i(t,x)$$



$$P_i(t) = N_i(t)/2^i$$

Population doubling time and precursor viability easily computable

Label Structure Fragmentation Model-Model 1

CFSE label dynamics can be described by the fragmentation equation

$$\begin{aligned}
&\frac{\partial n(t,x)}{\partial t} - ce^{-kt}\frac{\partial[(x-x_a)n(t,x)]}{\partial x} \\
&= -(\alpha(t,x)+\beta(t,x))n(t,x)\chi_{[x_a,x^*]}4\alpha(t,2x-x_a)n(t,2x-x_a) \\
&n(0,x) = \Phi(x) \\
&n(t,x_{\max}) = 0 \\
&v(t,x_a)n(t,x_a) = 0.
\end{aligned} \tag{1}$$

where $n(t,x)$ is the structured population density (cells per unit FI) at time t and measured FI x and the advection term (with parameters c and k) represents the Gompertz decay process for decrease in FI resulting from intracellular turnover of CFSE.

Compart. Models with Generation Structure-Model 2

A model can be reformulated with distinct compartments for each generation. The resulting model is a system of partial differential equations

$$\begin{aligned}
\frac{\partial n_0}{\partial t} - ce^{-kt}(x - x_a)\frac{\partial n_0}{\partial x} &= -(\alpha_0(t) + \beta_0(t) - ce^{-kt})n_0(t,x) \\
\frac{\partial n_1}{\partial t} - ce^{-kt}(x - x_a)\frac{\partial n_1}{\partial x} &= -(\alpha_1(t) + \beta_1(t) - ce^{-kt})n_1(t,x) \\
&\quad + R_1(t,x) \\
&\vdots \\
\frac{\partial n_{i_{\max}}}{\partial t} - ce^{-kt}(x - x_a)\frac{\partial n_{i_{\max}}}{\partial x} &= -(\beta_{i_{\max}}(t) - ce^{-kt})n_{i_{\max}}(t,x) \\
&\quad + R_{i_{\max}}(t,x), \qquad (2)
\end{aligned}$$

Compartmental Models-Model 2

Here $n_i(t,x)$ is structured population density for cells having undergone i divisions. The recruitment terms describe the symmetric division of CFSE upon mitosis. Given by $R_i(t,x) = 4\alpha_{i-1}(t)n_{i-1}(t,2x-x_a)$. Boundary and initial conditions same as in (1).

Observe that, because the number of divisions undergone has now been explicitly identified by the subscripted generation number, no longer necessary for division and death rates to depend upon measured fluorescence intensity.

Compartmental Models-Model 3

IDEA: Combine compartmental models with probabilistic structures (so-called Cyton Models) for cell division and death

Let $N_i(t)$, $0 \leq i \leq i_{\max}$ represent the number of cells having undergone i divisions at time t. Assume there are N_0 cells in the population at $t = 0$. In its simplest form, the cyton model relates the number of cells in the population to the number of cells which divide and die in a unit of time,

$$N_0(t) = N_0 - \int_0^t \left(n_0^{div}(s) - n_0^{die}(s) \right) ds$$

$$N_i(t) = \int_0^t \left(2n_{i-1}^{div}(s) - n_i^{div}(s) - n_i^{die}(s) \right) ds, \qquad (3)$$

where $n_i^{div}(t)$ and $n_i^{die}(t)$ indicate the numbers per unit time of cells having undergone i divisions that divide and die, respectively, at time t.

Compartmental Models-Model 3

Let $\phi_0(t)$ and $\psi_0(t)$ be probability density rate functions (in units 1/time) for the times to division and death, respectively, for an undivided cell. Let F_0, the initial precursor fraction, be the fraction of undivided cells which would hypothetically divide in the absence of any cell death. It follows that

$$n_0^{div}(t) = F_0 N_0 \left(1 - \int_0^t \psi_0(s)ds\right) \phi_0(t)$$

$$n_0^{die}(t) = N_0 \left(1 - F_0 \int_0^t \phi_0(s)ds\right) \psi_0(t). \qquad (4)$$

Compartmental Models-Model 3

Can define probability density rate functions $\phi_i(t)$ and $\psi_i(t)$ for times to division and death, respectively, for cells having undergone i divisions, as well as the progressor fractions F_i of cells which would complete the i^{th} division in the absence of cell death. Then the numbers per unit time of dividing and dying cells are computed as

$$n_i^{div}(t) = 2F_i \int_0^t n_{i-1}^{div}(s) \left(1 - \int_0^{t-s} \psi_i(\xi) d\xi\right) \phi_i(t-s) ds$$

$$n_i^{die}(t) = 2 \int_0^t n_{i-1}^{div}(s) \left(1 - F_i \int_0^{t-s} \phi_i(\xi) d\xi\right) \psi_i(t-s) ds. \quad (5)$$

Compartmental Models-Model 4

Allgöwer, et al., have proposed a model which is structured by the fluorescence intensity *resulting from the mass of CFSE within the cell*, i.e., intracellular label ignoring autofluorescence. This leads to a system for the mass of fluorescence $n(t,x)$ described by the system of equations

$$\frac{\partial n_0}{\partial t} - ce^{-kt}\frac{\partial [xn_0]}{\partial x} = -(\alpha_0(t)+\beta_0(t))n_0(t,x)$$

$$\vdots$$

$$\frac{\partial n_i}{\partial t} - ce^{-kt}\frac{\partial [xn_i]}{\partial x} = -(\alpha_i(t)+\beta_i(t))n_i(t,x) + 4\alpha_{i-1}(t)n_{i-1}(t,2x) \tag{6}$$

$$\vdots$$

with boundary and initial conditions as in (1) and (2).

Compartmental Models-Model 4

The major advantage of formulating the model in terms of mass of FI is the very simple form of the model solution. The solution to the model (6) can be written as

$$n_i(t,x) = N_i(t)\bar{n}_i(t,x) \tag{7}$$

for all i. In this representation the functions $N_i(t)$ satisfy the weakly coupled system of ordinary differential equations

$$\begin{aligned}
\frac{dN_0}{dt} &= -(\alpha_0(t) + \beta_0(t))N_0(t) \\
\frac{dN_1}{dt} &= -(\alpha_1(t) + \beta_1(t))N_1(t) + 2\alpha_{i-1}(t)N_{i-1}(t) \qquad (8) \\
&\vdots
\end{aligned}$$

with initial conditions $N_0(0) = N_0$, $N_i(0) = 0$ for all $i \geq 1$.

Compartmental Models-Model 4

The functions $\bar{n}_i(t,x)$ each satisfy the partial differential equation

$$\frac{\partial \bar{n}_i(t,x)}{\partial t} - ce^{-kt}\frac{\partial [x\bar{n}_i(t,x)]}{\partial x} = 0 \tag{9}$$

with initial condition

$$\bar{n}_i(0,x) = \frac{2^i \Phi(2^i x)}{N_0}.$$

Note that, by definition,

$$N_0 = \int_0^\infty \Phi(\xi) d\xi.$$

Compartmental Models-Model 4

Let $\tilde{n}_i(t,\tilde{x})$ be the structured density relative to the *measured fluorescence* ($\tilde{x}$ is related to x by the addition of cellular autofluorescence). To account for autofluorescence and because autofluorescence may vary from cell to cell in the population, this is most accurately treated by computing the convolution integral

$$\tilde{n}(t,\tilde{x}) = \int_0^\infty n(t,x)p(t,\tilde{x}-x)dx, \tag{10}$$

where $p(t,\tilde{x})$ is (for each fixed time t) a probability density function describing the distribution of autofluorescence in the population.

Compartmental Models-Model 4

So finally we consider the system of equations

$$\begin{aligned}
\frac{\partial n_0}{\partial t} - ce^{-kt}\frac{\partial [x n_0]}{\partial x} &= \left(n_0^{div}(t) - n_0^{die}(t)\right)\bar{n}_0(t,x) \\
\frac{\partial n_1}{\partial t} - ce^{-kt}\frac{\partial [x n_1]}{\partial x} &= \left(2n_0^{div}(t) - n_1^{div}(t) - n_1^{die}(t)\right)\bar{n}_1(t,x) \qquad (11) \\
&\vdots
\end{aligned}$$

where the definitions of $n_0^{div}(t)$, $n_0^{die}(t)$, $n_i^{div}(t)$ and $n_i^{die}(t)$ are given in Equations (4) - (5).

Compartmental Models-Model 4

This model, which is based on simple mass balance, can be solved by factorization (7) as before; the label densities $n(t,x)$ are readily computed according to Equation (9), and the cell numbers are now provided by the cyton model as discussed above. Thus the new model (11) is capable of accurately describing the evolving generation structure of a population of cells while also accounting for the manner in which the CFSE profile of the population changes in time. The model is easily relatable to biologically meaningful parameters (times to division and death) and can be solved efficiently so that it is eminently suitable for use in an inverse problem.

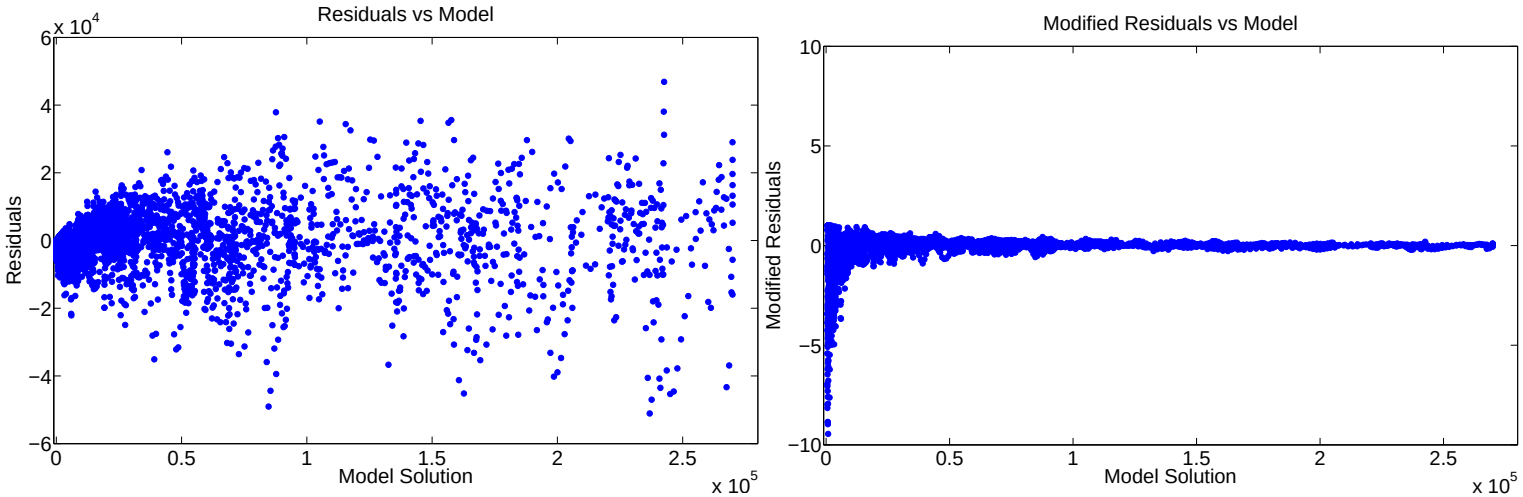
The Statistical Model

- Links the mathematical model to the data
- Implications for estimation procedure (Least squares vs. weighted least squares)

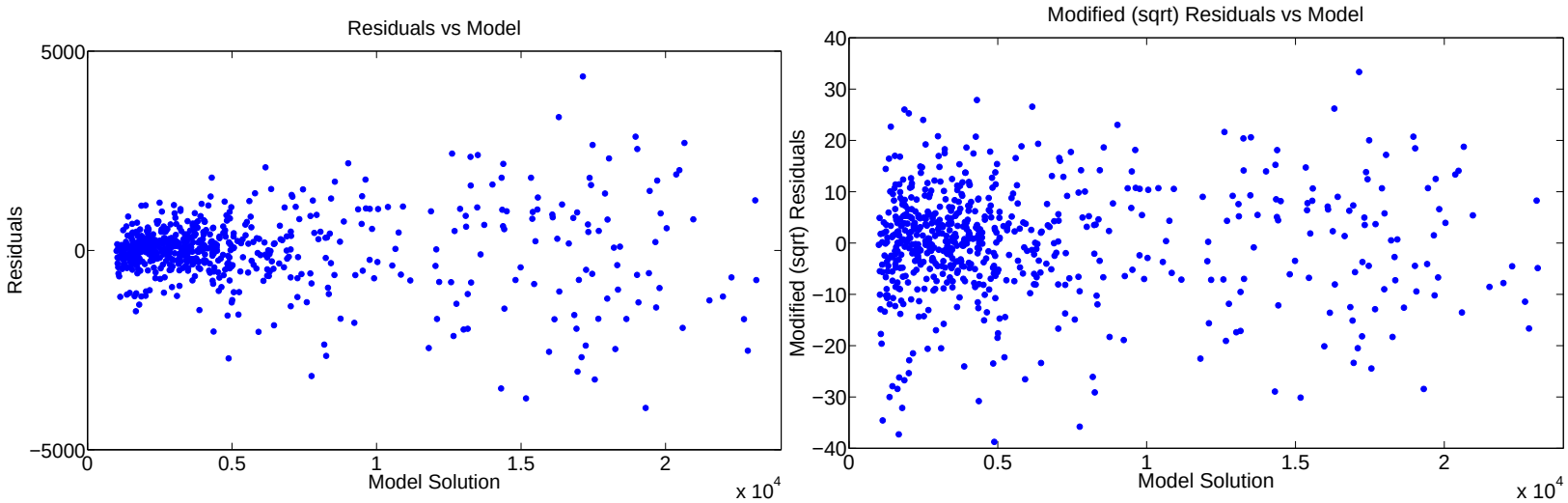
$$N_k^j = I[\tilde{n}](t_j, z_k^j; \theta_0) + (I[\tilde{n}](t_j, z_k^j; \theta_0))^\gamma \mathcal{E}_{kj}$$

- Early efforts using constant variance (CV) model, $Var(\mathcal{E}_{kj}) = \sigma_0^2$ ($\Rightarrow$ Absolute Error: $Y_{kj} = \text{model} + \mathcal{E}_{kj}$)
- Also tried constant coefficient of variance (CCV), $Var(\mathcal{E}_{kj}) = \sigma_0^2 I[\tilde{n}](t_j, z_k^j; \theta_0)^2$ ($\Rightarrow$ Relative Error:$Y_{kj} = \text{model} \times (1 + \mathcal{E}_{kj})$)
- Found in between error most closely described data, $Var(\mathcal{E}_{kj}) = \sigma_0^2 I[\tilde{n}](t_j, z_k^j; \theta_0)$ ($\Rightarrow$ SquareRoot Error: $Y_{kj} = \text{model} + \sqrt{\text{model}} \times \mathcal{E}_{kj}$)

Residual Plots $\gamma = 0$ vs. $\gamma = 1$



Residual Plots



$N_k^j = I[\tilde{n}](t_j, z_k^j; q_0) + (I[\tilde{n}](t_j, z_k^j; q_0))^\gamma \mathcal{E}_{kj}$ with $\gamma = 0$ (left) vs. $\gamma = 1/2$ (right). And $\gamma = 1/2$ works well!!

Least Squares Estimation

Define

$$q_{WLS} = \arg\min_{q\in Q} J(q;\{\lambda_j\})$$

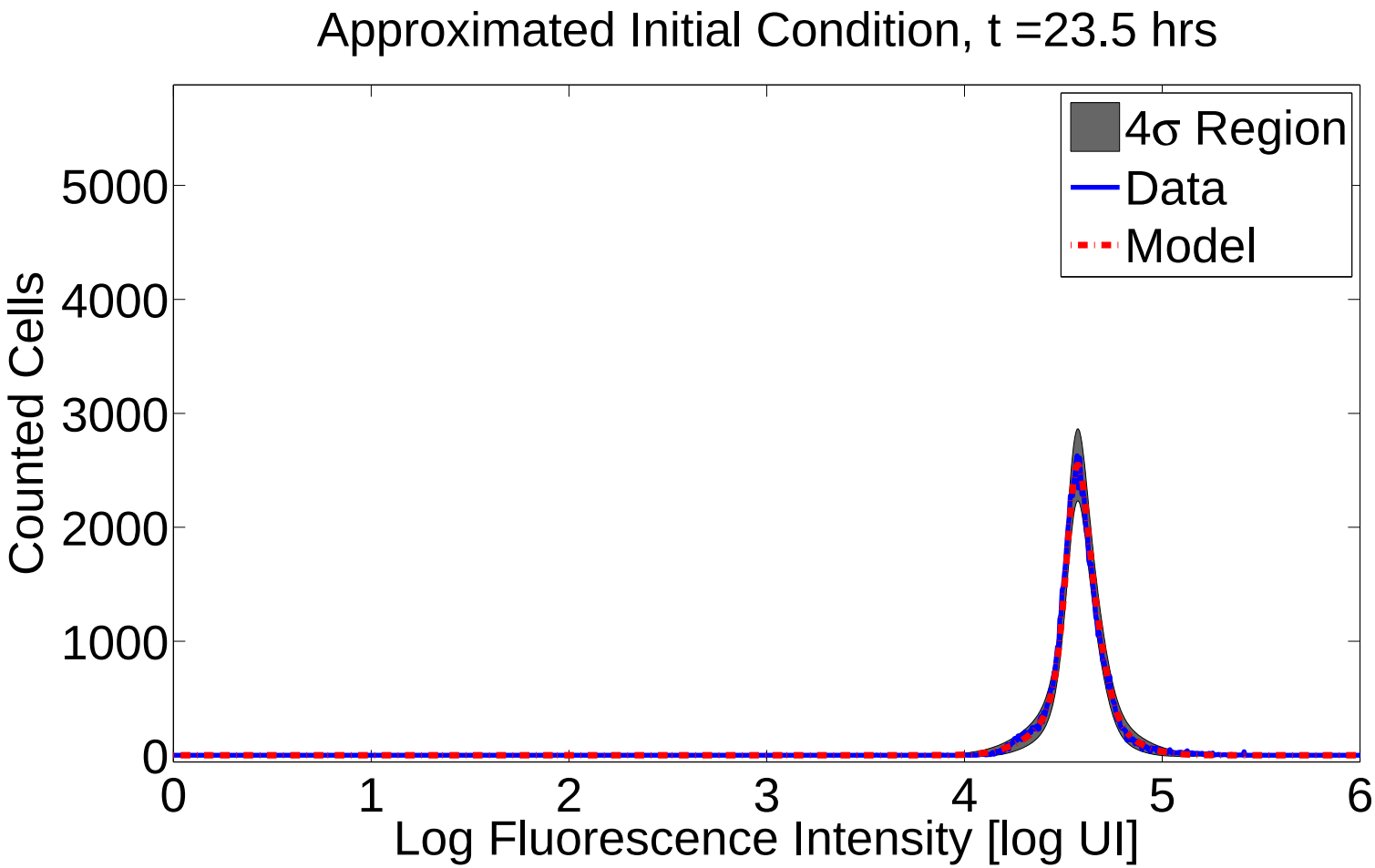
$$= \arg\min_{q\in Q} \sum_{j,k} \frac{\left(\lambda_j I[\hat{n}](t_j, z_k; q) - N_k^j\right)^2}{w_k^j}$$

where

$$w_k^j = \begin{cases} \lambda_j \frac{B}{\hat{b}_j} I[\hat{n}](t_j, z_k^j; q_0), & I[n](t_j, z_k^j; q_0) > I^* \\ \lambda_j \frac{B}{\hat{b}_j} I^*, & I[n](t_j, z_k^j; q_0) \le I^* \end{cases}.$$

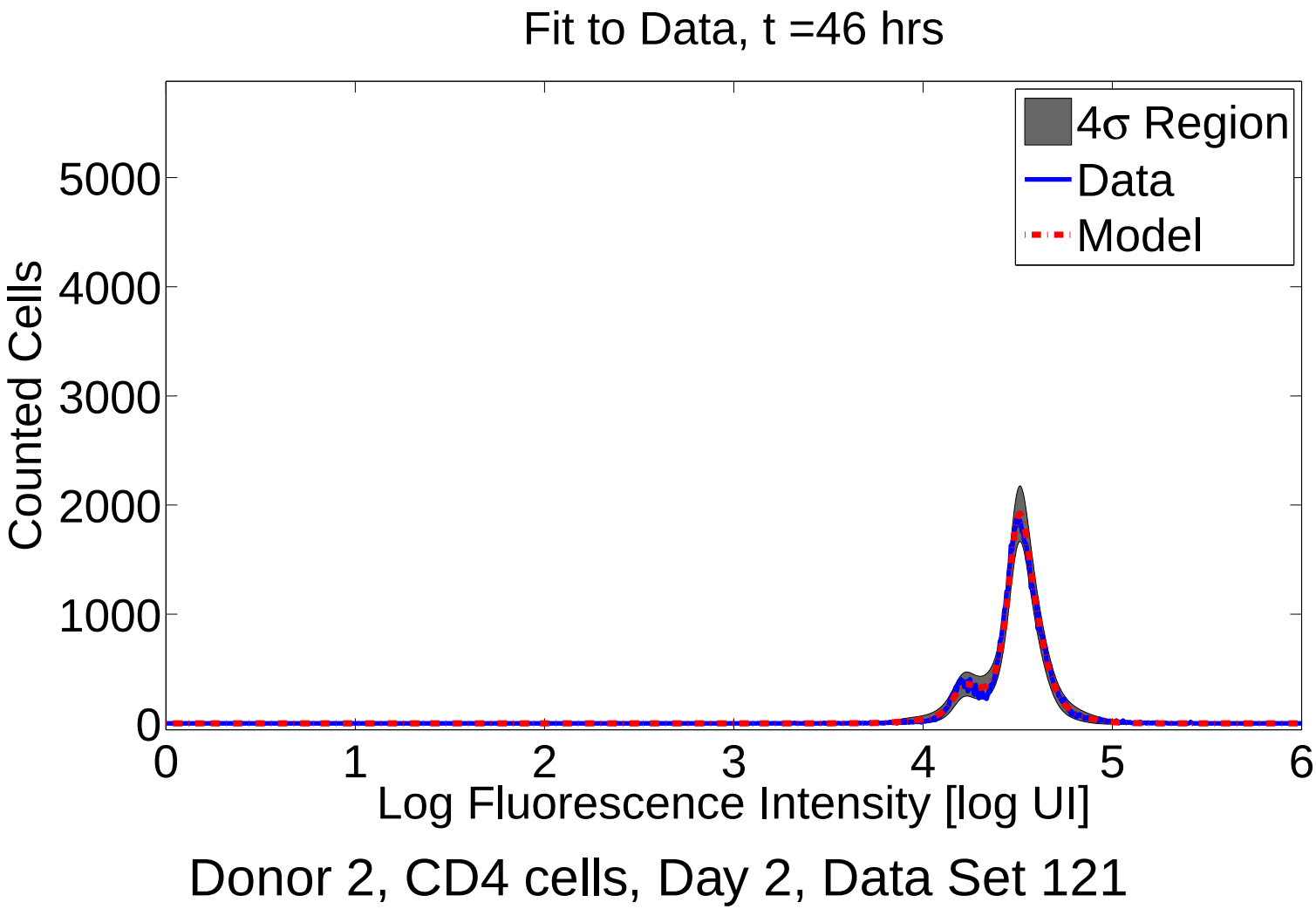
Then q_{WLS} is a consistent estimator of q_0 (Banks, Kenz, Thompson, 2012)

Fit to Data

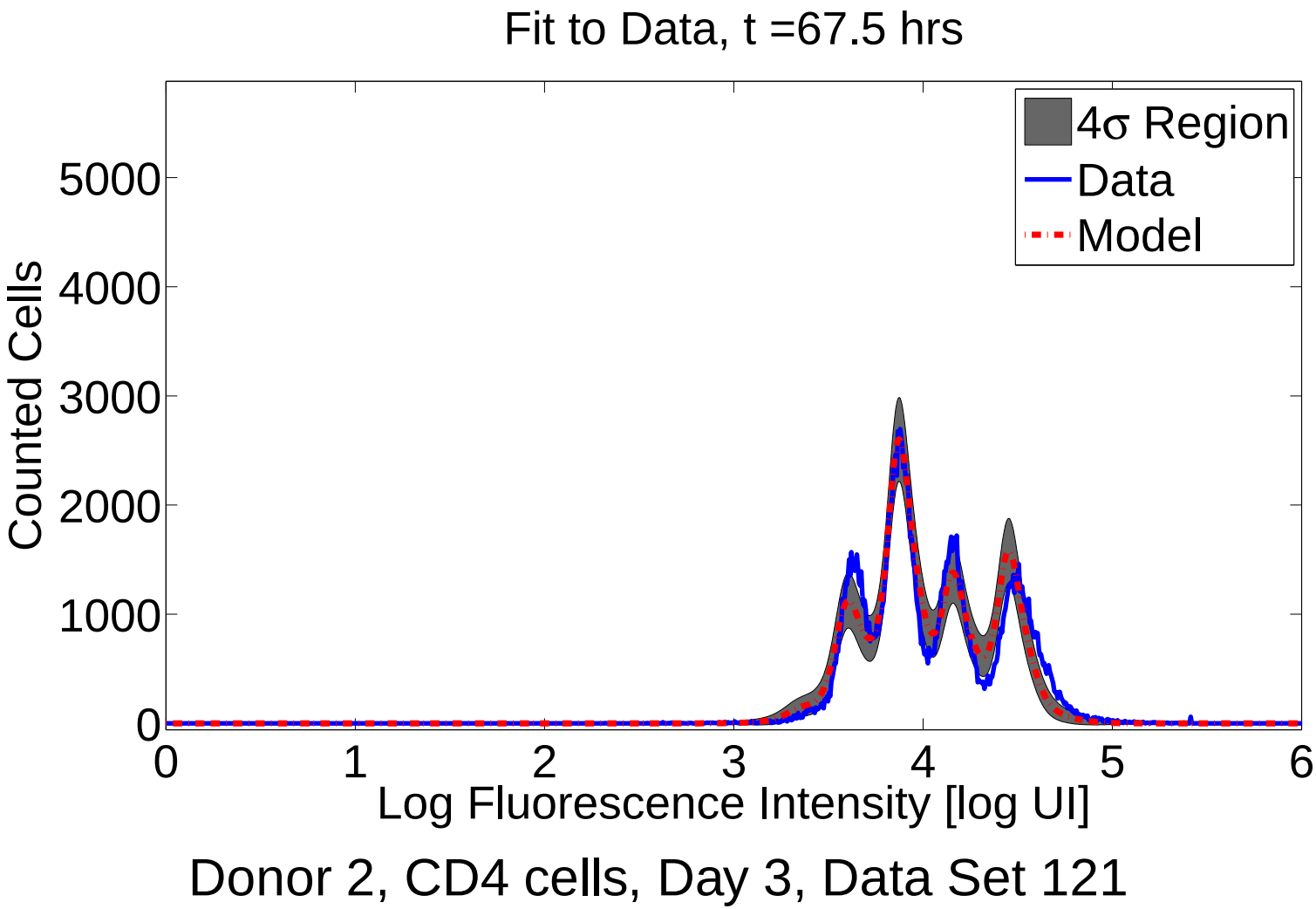


Donor 2, CD4 cells, Day 1, Data Set 121

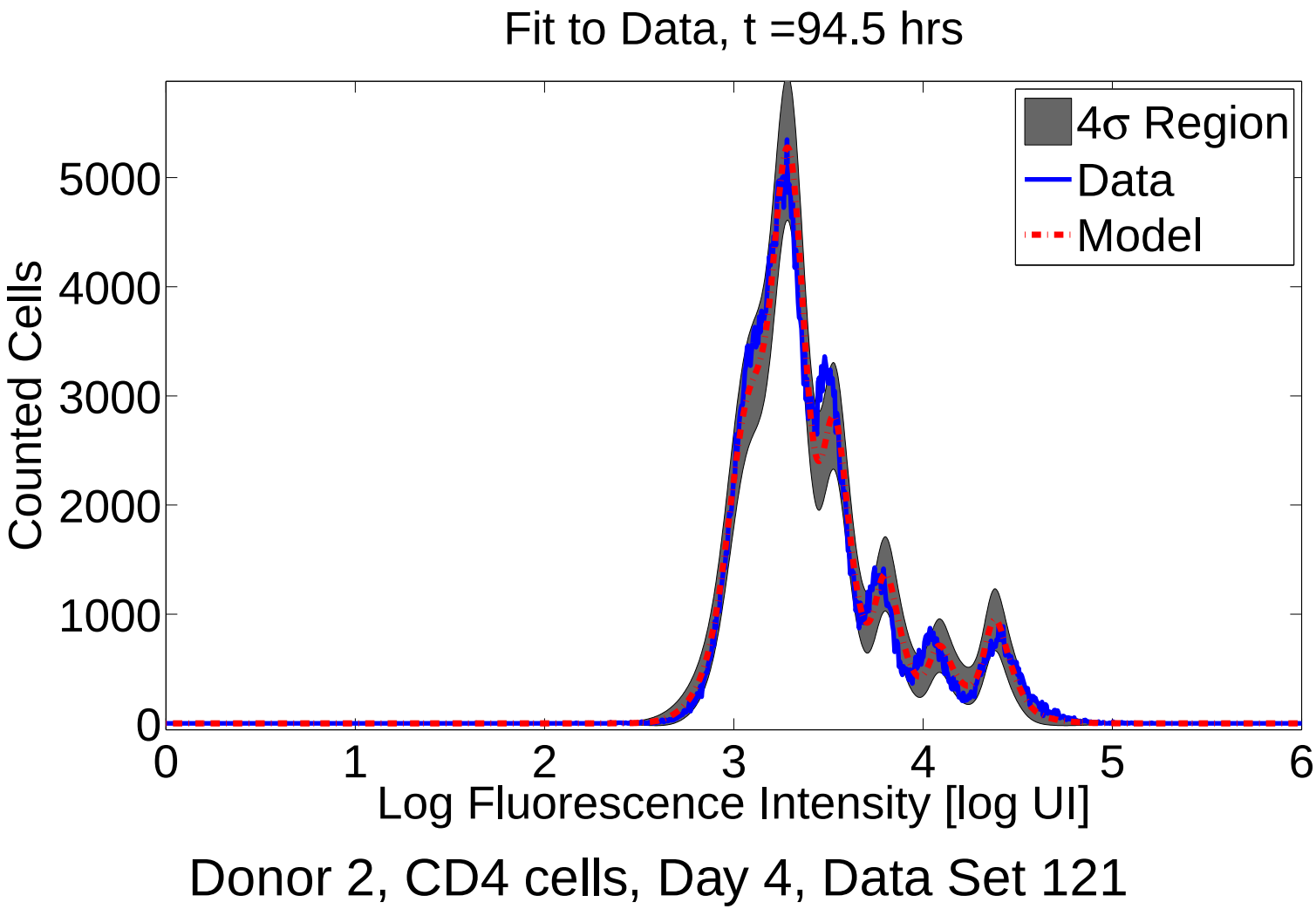
Fit to Data (cont'd)



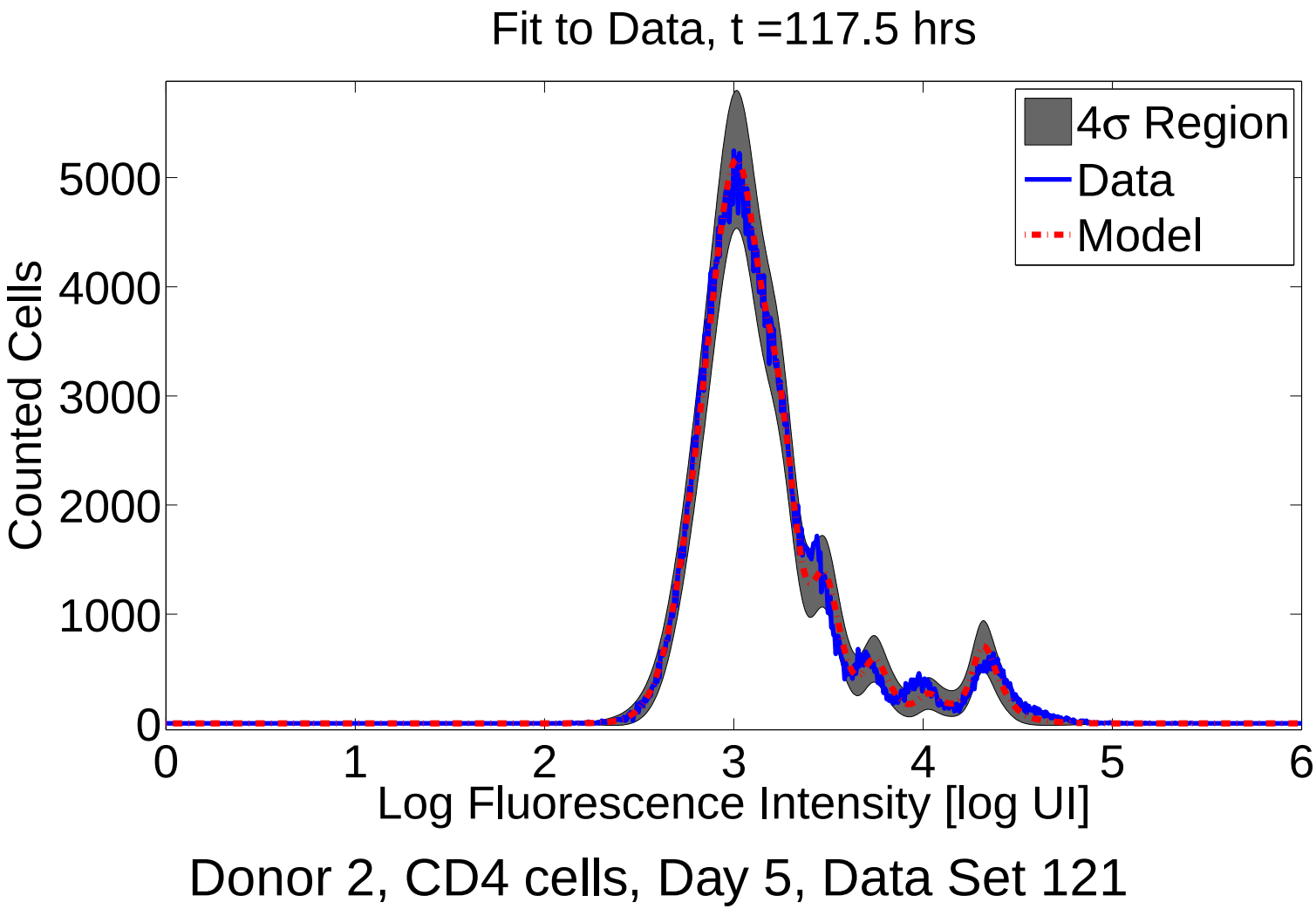
Fit to Data (cont'd)



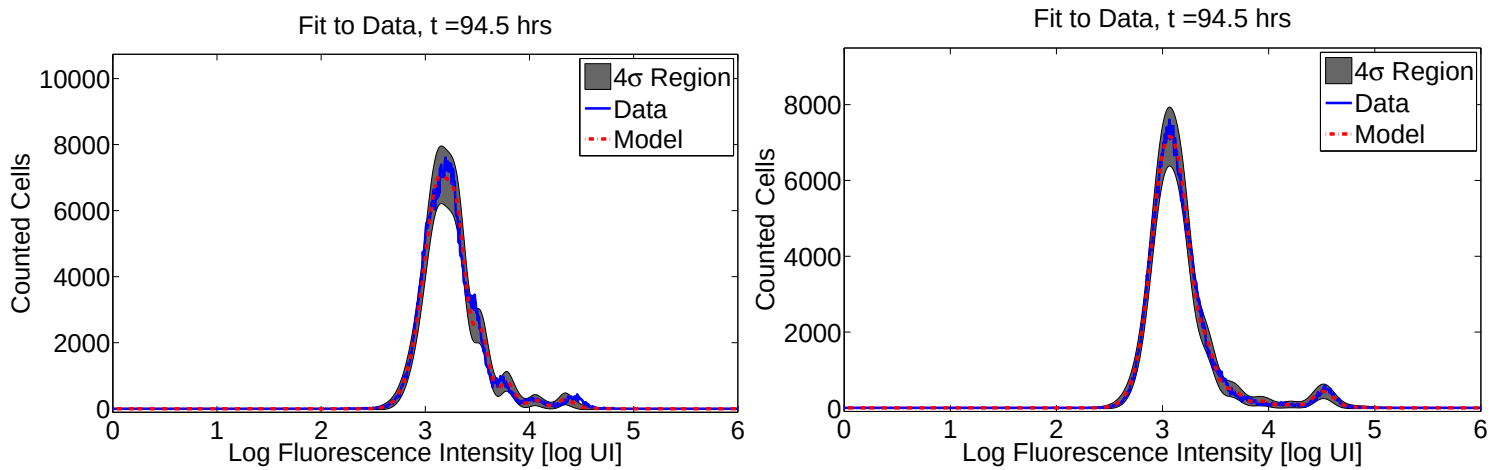
Fit to Data (cont'd)



Fit to Data (cont'd)

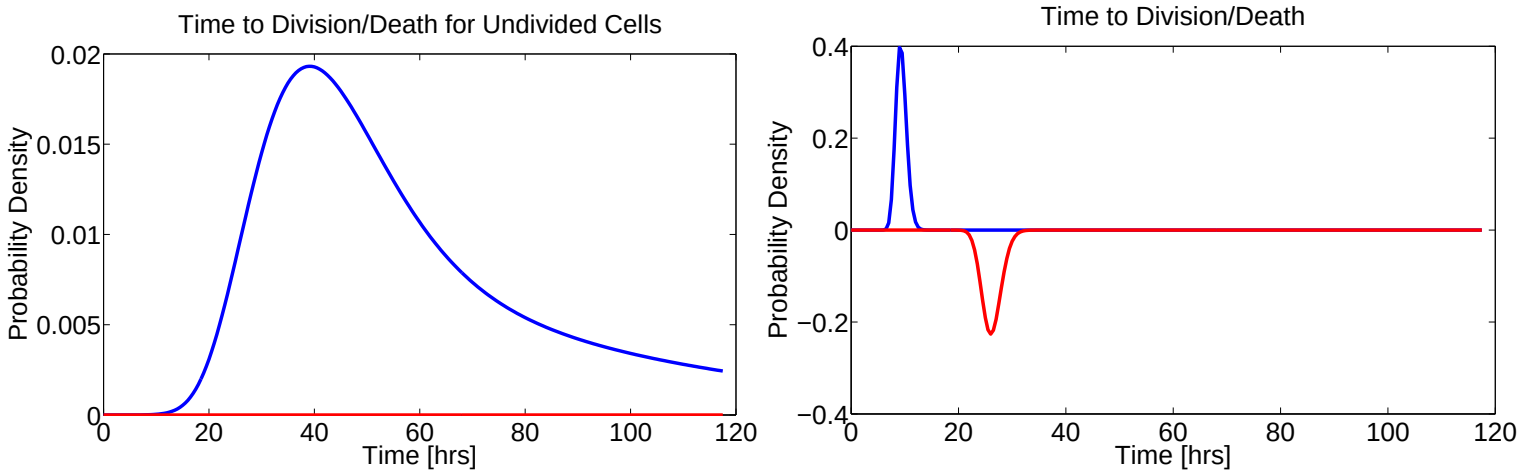


Fit to Data, CD8 Cells



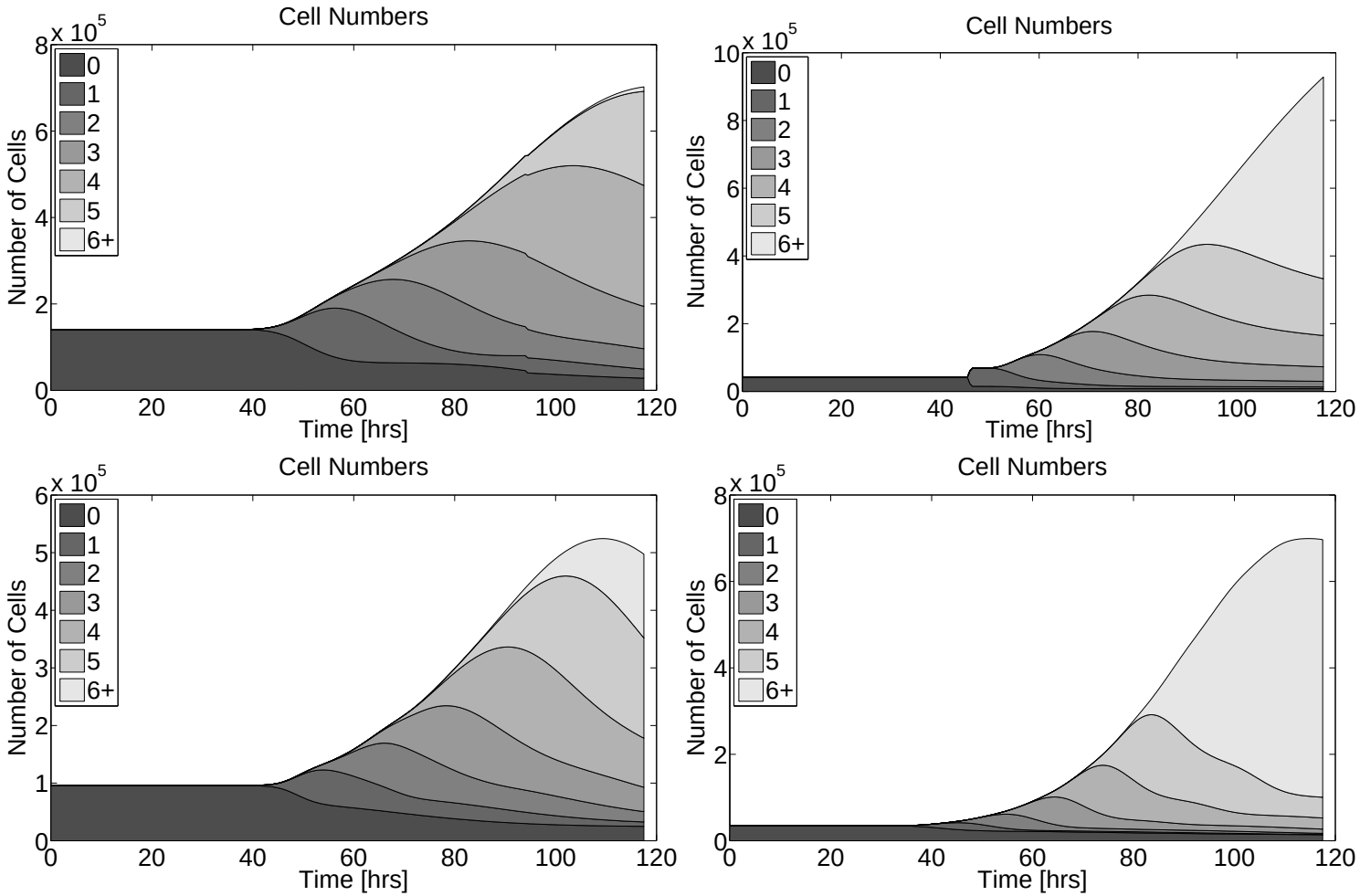
Donors 1 (left) and 2 (right), CD8 cells, Day 4, Data Set 121

Cytons



Donor 2, CD8 cells, undivided (left) and divided (right), Data Set 121

Generation Structure



Donor 1 (top) and Donor 2 (bottom), CD4 cells (left) and CD8 cells (right)

Experimental Extensions

- Account for multiple cell cultures present in PBMC culture
- Antigen-specific stimulation
- Effects of cryopreservation
- Extracellular signaling, knockout experiments
- In vitro vs in vivo differences
- Linking to immune process models
- **Analyze Proliferation in Diseased vs Healthy cells**

Acknowledgements

- National Institute of Allergy and Infectious Disease grant NIAID 9R01AI071915
- U.S. Air Force Office of Scientific Research under grant AFOSR-FA9550-09-1-0226