

A mathematical model for chronic myelogenous leukemia (CML) and T cell interaction

Helen Moore^{a,*}, Natasha K. Li^b

^a American Institute of Mathematics, 360 Portage Avenue, Palo Alto, CA 94306, USA

^b Stanford University, Stanford, CA 94305, USA

Received 19 February 2003; received in revised form 17 November 2003; accepted 18 November 2003

Abstract

In this paper, we propose and analyse a mathematical model for chronic myelogenous leukemia (CML), a cancer of the blood. We model the interaction between naive T cells, effector T cells, and CML cancer cells in the body, using a system of ordinary differential equations which gives rates of change of the three cell populations. One of the difficulties in modeling CML is the scarcity of experimental data which can be used to estimate parameters values. To compensate for the resulting uncertainties, we use Latin hypercube sampling (LHS) on large ranges of possible parameter values in our analysis. A major goal of this work is the determination of parameters which play a critical role in remission or clearance of the cancer in the model. Our analysis examines 12 parameters, and identifies two of these, the growth and death rates of CML, as critical to the outcome of the system. Our results indicate that the most promising research avenues for treatments of CML should be those that affect these two significant parameters (CML growth and death rates), while altering the other parameters should have little effect on the outcome.

© 2003 Published by Elsevier Ltd.

Keywords: Cancer; Chronic myelogenous leukemia (CML); Differential equations; Dynamics; Latin hypercube sampling (LHS); Mathematical modeling; T cells

1. Introduction

Chronic myelogenous leukemia (CML) is a cancer that affects cells circulating in the blood system. The disease occurs in more than 1 out of 100 000 people per year, and makes up approximately 15% of all leukemias in adults (Faderl, et al., 1999, p. 164). The immune system is known to play an important role in the dynamics of CML (Sawyers, 1999), which is motivation for the work in this paper.

In this paper, we propose a mathematical model for the human immune system's response to CML in a hypothetical patient. One major goal of this work is to identify promising directions for experimental research on treatments for patients with CML. Our model

consists of a system of three nonlinear ordinary differential equations. We consider rates of change for naive and effector T cell populations, as well as the CML cancer cell population. Parameter values are obtained from available experimental data and estimates, and the system is analysed for sensitivity to changes in the parameters. Our analysis reveals that changes in most parameters (such as naive and effector T cell death rates) have little effect on the long-term outcome of the system, whereas changes in two parameters, the CML cancer cell growth rate and natural death rate, result in significantly different outcomes of the system. See Section 8 below for the full set of conclusions.

There is an extensive body of work which develops models of this type for the interaction of T cells with human immunodeficiency virus (HIV). Perelson and Nelson (1999) and Nowak and May (2000) have written detailed surveys of the main ideas developed through such models. Several other diseases with similar pathogen-immune system interaction dynamics have been modeled in this way as well. These include tuberculosis (cf. Wigginton and Kirschner, 2001) and hepatitis B

*Corresponding author. Tel.: 650-845-2072.

E-mail address: moore@aimath.org (H. Moore).

¹ Supported by National Science Foundation DMS Grant 0096135. Abbreviations: AICD, activation-induced cell death; APC, antigen-presenting cell(s); CML, chronic myelogenous leukemia; HIV, human immunodeficiency virus; LHS, Latin hypercube sampling; MHC, major histocompatibility complex; 1 microliter (μ l) = 10^{-6} l = 1 cubic millimeter (mm^3).

(cf. Nowak et al., 1996). Early modeling of CML includes models of T cell precursors in the bone marrow (cf. Cronkite and Vincent, 1969; Rubinow, 1969; Rubinow and Lebowitz, 1975). A later mathematical model for CML created by Fokas et al. (1991) refined these models, and focused on the maturation and proliferation of the T cell precursors.

2. Background and assumptions for the model

The two T cell populations included in our model are defined as follows: naive T cells, which, if specific to CML, could become activated; and effector T cells specific to CML, which are fully armed and capable of immediate action. Naive T cells are activated in the lymph tissues. If a naive T cell is CML-specific, then it can bind to a peptide-MHC (major histocompatibility complex) pair on a professional antigen-presenting cell (APC). If costimulators are also present, the T cell will be retained, activated, and will proliferate. After up to 1 week of proliferation, the progeny differentiate into armed effector cells. These are then released into the blood and are capable of mounting an immune attack upon encountering CML antigen (without the need for costimulation).

These two categories (naive and effector) represent the major distinct types of behavior of T cells that we wish to capture in our model. In order to keep this model tractable, we do not consider other parts of the immune system's response, such as the populations of B lymphocyte cells or the levels of production of cytokines. Our working assumption is that these other responses correlate positively with the responses we do examine, although we do not hypothesize exactly what the correlation functions might be. Furthermore, in Kirschner and Webb's (1996, p. 370, footnote) model for T cell responses to HIV, they find no qualitative difference in outcome between a model with CD4⁺ and CD8⁺ T cells considered separately, and a model in which only the total population of T cells was considered. Therefore, we do not consider these populations separately in our model.

Our model is based in the circulatory blood system, since we wish to tie it as closely as possible to available data from blood samples. We include a source term for new T cells entering the circulatory blood from other compartments (such as the bone marrow, lymph nodes, and thymus). We also include the input of CML cells from outside the blood, in the growth term for CML. However, we assume that pre-existing cells entering and exiting the blood due to diffusion give net changes that are approximately zero. We also assume that the concentration of T cells in the circulatory blood is proportional to the concentration of T cells in other compartments. These are standard assumptions for a

system at or near equilibrium, and experimental support in the case of HIV appears in Haase et al. (1996) and Perelson et al. (1997).

In this model, the naive T cells come into contact with the APC in the lymph tissues. We assume that APC with antigen from CML cells are present in proportion to the number of CML cancer cells themselves. Therefore, in our model we use the populations of CML cells with a proportionality constant to represent the numbers of APC presenting CML antigen. The encounters of naive T cells with the APC are represented by a Michaelis–Menten saturation term, the third term on the right-hand side of Eq. (1). We use a saturating growth term to take into account the limitations of the immune response, due to the large numbers of APC presenting CML antigen which saturate the lymph, but the relatively few CML specific naive T cells (Janeway et al., 2001, pp. 295–318). Similar ideas were used in Perelson et al. (1993) and Kirschner and Panetta (1998).

We assume the effector T cells come into contact with CML cancer cells in the blood in a random fashion. Because the encounters take place in the blood, we use the “law of mass action”, which says that the total number of encounters between members of the two populations is proportional to the product of the sizes of the two populations. See Edelstein-Keshet (1988) for more discussion of the law of mass action.

Note that the use of deterministic equations means that our model is not expected to be accurate when the populations are relatively small (Freedman, 1980, pp. 3–4). Specifically, the model will not be very accurate in its predictions if the populations are so small (e.g., on the order of single cells) that the continuous rates are not good approximations for gain or loss terms.

The disease CML has three distinct phases: chronic, accelerated, and blast. The chronic phase lasts the longest, averaging 3–5 years in length. During this time, cell counts grow steadily. The accelerated and blast phases can last just a few months each, and are characterized by more rapid increases in cell counts, followed by death of the patient. Our model is intended for the chronic phase, as the rapid changes in cell counts during the accelerated and blast phases are not as well quantified by data in the later phases, and treatments for CML focus on the chronic phase. Also, the cancer cell population is large enough during the chronic phase that the law of mass action is valid for cell interactions in the blood, and yet not so large that crowding of CML in the bone marrow decreases the source rate of the naive T cells (Stryckmans et al., 1977).

3. Details and explanation of the model

Let t represent time, measured in days. We consider the following populations of cells in the circulating

blood system, measured as concentrations of cells per μl :

T_n = naive T cells,

T_e = effector T cells specific to CML,

C = chronic myelogenous leukemia (CML) cancer cells.

Each of the three cell populations is a function of time, t . The naive T cell population consists of all naive T cells, both specific and non-specific for CML. The naive T cells either have not yet been exposed to professional APC presenting CML antigen, or else are not CML-specific. The effector T cells have differentiated from naive T cells upon activation by a professional APC peptide-MHC complex and costimulators.

The system of differential equations is given below, followed by explanations of the terms:

$$\frac{dT_n}{dt} = s_n - d_n T_n - k_n T_n \left(\frac{C}{C + \eta} \right), \quad (1)$$

$$\begin{aligned} \frac{dT_e}{dt} = & \alpha_n k_n T_n \left(\frac{C}{C + \eta} \right) + \alpha_e T_e \left(\frac{C}{C + \eta} \right) \\ & - d_e T_e - \gamma_e C T_e, \end{aligned} \quad (2)$$

$$\frac{dC}{dt} = r_c C \ln \left(\frac{C_{\max}}{C} \right) - d_c C - \gamma_c C T_e. \quad (3)$$

Each equation represents the rate of change, with respect to time, of one of the cell populations. The lowercase coefficients, or parameters (e.g., s_n , α_n , and r_c) are all taken to be constants. Later in this paper we vary the parameter values, and examine the resulting changes in the system.

Fig. 1 below shows the cell population diagram for system (1)–(3).

We assume the changes in populations due to diffusion are approximately zero. That is, we assume the numbers of pre-existing cells of each population that diffuse into and out of the blood are approximately equal. Thus, there are no terms in the equations for diffusion of already-existing, mature cells into or out of the blood.

The first term on the right-hand side of Eq. (1) is a source term for *new* T cells entering the blood system. We approximate this as a constant, s_n , which is a reasonable approximation except during the last stages of CML, when crowding in the bone marrow could decrease the production of naive T cells. The second term is due to the natural attrition of T_n cells in the absence of CML. The factor d_n is the naive T cell death rate constant, which is equal to the reciprocal of the average lifespan of a T_n cell, and can be thought of as roughly approximating the fraction of the T_n population that is expected to die naturally in 1 day. The third term is a Michaelis–Menten term, and represents the change in the T_n population due to encounters with CML antigen in the lymph. We use a Michaelis–Menten term to take into account the saturation effect of CML cells

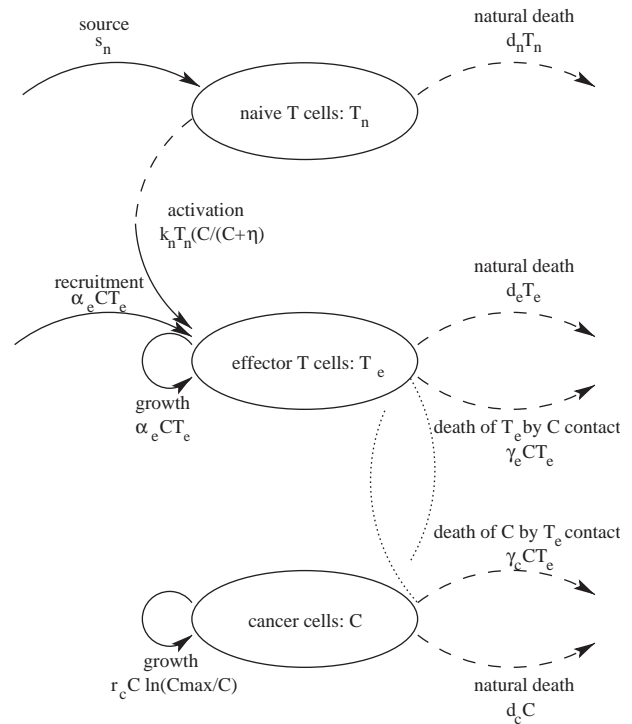


Fig. 1. Cell population diagram. In this figure, solid curves with arrows represent source terms (such as proliferation and activation), dotted curves without arrows represent interaction terms (such as contact between cells), and dashed curves with arrows represent loss of cells (such as death). Curves with arrows signify movement into or out of a population, while curves without arrows only signify interactions between populations.

in the lymph nodes, as described earlier in Section 2. We assume CML antigen is present in proportion to the number of CML cells. The constant η is the standard half-saturation concentration in a Michaelis–Menten term. The constant k_n is the rate constant which makes $k_n T_n C / (C + \eta)$ equal to the instantaneous rate of T_n loss due to encounters between naive T cells and CML antigen. While CML-specific T_n cells can become activated upon encountering CML peptides and costimulators on APC, they can also become anergic if they encounter CML antigen without costimulators. The state of anergy essentially inactivates these naive T cells, preventing later activation, even if they subsequently encounter APC with CML antigen and costimulators. Hence, we consider both activation and anergy as contributors to the loss term.

We now examine Eq. (2), which describes the rate of change for those effector T cells in the blood compartment which are specific to CML. The first term on the right-hand side is due to activation encounters between T_n and professional APC (presenting peptides from CML antigen), and comes from the naive T cells lost in Eq. (2). When a naive T cell becomes activated, it is retained in the lymph tissues and proliferates for several days. The newly-formed T cells then differentiate into

Table 1
Parameter information

Param.	Description	Value	Sampled range	Units	Reference
s_n	T_n source term	0.073	(0, 0.5)	$\frac{\text{cells}/\mu\text{l}}{\text{day}}$	Mohri et al. (2001)
d_n	T_n death rate	0.040	(0, 0.5)	day^{-1}	Mohri et al. (2001)
d_e	T_e death rate	0.06	(0, 0.5)	day^{-1}	Kuznetsov et al. (1994)
d_c	C death rate	0.2	(0, 0.8)	day^{-1}	Duvall and Perry (1968)
k_n	T_n differentiation	0.001	(0, 0.1)	day^{-1}	Essunger and Perelson (1994)
η	Michaelis–Menten	100	(0, 1000)	$\text{cells}/\mu\text{l}$	Estimated
α_n	T_e proliferation	0.41	(0, 1)		Janeway et al. (2001)
α_e	T_e recruitment	0.2	(0, 1)	day^{-1}	Estimated
C_{\max}	maximum C	3×10^5	$(1.5 \times 10^5, 4 \times 10^5)$	$\text{cells}/\mu\text{l}$	Estimated
r_c	C growth	0.03	(0, 0.5)	day^{-1}	Afenya and Calderón (2000)
γ_e	T_e loss (due to C)	0.005	(0, 0.1)	$\text{day}^{-1} \left(\frac{\text{cells}}{\mu\text{l}} \right)^{-1}$	Estimated
γ_c	C loss (due to T_e)	0.005	(0, 0.1)	$\text{day}^{-1} \left(\frac{\text{cells}}{\mu\text{l}} \right)^{-1}$	Essunger and Perelson (1994)

effector T cells, and are released into the blood (Janeway et al., 2001). Thus, the coefficient α_n incorporates the rate at which such encounters lead to activation (as opposed to anergy) of T_n , as well as the rates of proliferation and differentiation into T_e .

The second term is a recruitment term, which we again represent as Michaelis–Menten. We assume that a proportion of effector T cells will recruit other immune cells to aid in killing CML cells. In order to represent this increase in killer cells, we effectively boost the effector T cell population by a compensatory amount. The rate constant α_e incorporates this recruitment. The third term represents the loss due to the natural lifespan of T_e cells, with d_e the death rate constant for T_e . The fourth term is a law of mass action term, due to encounters between T_e and CML cancer cells. Since these encounters occur in the blood, there is no saturation effect as occurs in the lymph. Some encounters result in a loss of T_e , due to direct effect by the CML cell or due to activation-induced cell death (AICD) from excess stimulation. The coefficient γ_e is the rate constant for loss of T_e due to these encounters.

In Eq. (3), the first term on the right-hand side gives the contribution due to growth of C , in the form of a Gompertz law. This law is well-established as the best fit for cancerous tumors. (See, e.g., Xu and Ling, 1988; Bassukas and Maurer-Schultze, 1988; Steel, 1977.) Afenya and Calderón (2000) show that the Gompertz curve also provides a significantly better fit for leukemic cancer data than logistic, exponential, or polynomial curves. The Gompertz curve incorporates the almost exponential growth of cancers at first, as well as the later slowing of the growth. The constant C_{\max} is the estimate of the maximum possible concentration of CML. The second term in Eq. (2) represents the loss

of CML cells which die a natural death, where d_c is their death rate constant. The third term, with coefficient γ_c , represents the loss of C due to encounters with T_e .

The constants used in the system of differential equations (1)–(3), appear in Table 1. The table gives a brief description of the constants, the values used as initial estimates for the constants, the ranges that are used for sampling later in the paper, and the units and references for the constants.

We use the following initial values for the three populations, where $t = 0$ is the starting time for the model:

$$\begin{aligned} T_n(0) &= 1510 \text{ cells}/\mu\text{l}, \\ T_e(0) &= 20 \text{ cells}/\mu\text{l}, \\ C(0) &= 10,000 \text{ cells}/\mu\text{l}. \end{aligned}$$

4. Discussion of parameter estimates and initial population values

The value we use for s_n , the source term for the T_n cell population, is calculated using data from Mohri et al. (2001). In the data collected from healthy subjects, the average contributions from a source were found to be 0.033 CD4⁺ cells/ $\mu\text{l}/\text{day}$, and 0.040 CD8⁺ cells/ $\mu\text{l}/\text{day}$. This gives a net contribution of 0.073 cells/ $\mu\text{l}/\text{day}$, which we use as the source term for our naive T cell population, T_n . In the same paper, the average number of cell deaths in healthy subjects, both for CD4⁺ and CD8⁺ lymphocytes, was found to be 0.040/day. (This cell death number actually takes into account the very small amount of proliferation measured for the

T cells, and so should more accurately be termed the “net loss”).

The value for d_e , the rate constant for natural T_e death, is estimated to be close to the calculated value in Kuznetsov et al. (1994, p. 303). It is expected that $d_e > d_n$, although the two numbers should not be drastically different, so we choose d_e to be of the same order of magnitude as d_n . The value for d_c , the rate constant for natural C death, is taken from Duvall and Perry (1968). They report an average leukocyte half-life of 83.1 h, or roughly 3.5 days, which we use to calculate d_c . The value for k_n is based on estimates of the probability of anergy and activation of antigen-specific T_n in Janeway et al. (2001), and on the very rough estimates of encounter rates incorporated into other T cell-pathogen models, such as that of Essunger and Perelson (1994). The value for α_n , the production factor for T_e from activated T_n , is estimated using proliferation rates in Janeway et al. (2001). Janeway et al. state that an activated T cell proliferates for close to 7 days, producing a total of approximately 1000 daughter cells. The value for α_e , the recruitment constant for T_e due to encounters with C , is approximated at 0.2, because it should be lower than the rate of activation, but is not expected to be an entire order of magnitude lower. The value for r_c , the growth rate of C , is taken from Afenya and Calderón (2000). To get an estimate for γ_c , the loss rate for C due to encounters between T_e and C , we again use encounter rates that are incorporated into other T cell-pathogen models (e.g., Essunger and Perelson, 1994). Based on our γ_c , we were able to estimate the value for γ_e .

In Mohri et al. (2001), the average concentration of CD4⁺ T cells in healthy individuals was found to be approximately 1080 cells/ μ l of blood in the steady state. The average concentration of CD8⁺ measured in healthy individuals was approximately 600 cells/ μ l of blood, giving a total of 1680 T cells/ μ l of blood, which we use for our estimate of the initial concentration of naive T cells in the blood. We decrease this by roughly 10% with the assumption that at this stage of CML, there would have been a small decrease of the usual levels of naive T cells due to activation. Thus we take $T_n(0) = 1510$. Correspondingly, we assume that the initial concentration of effector T cells, $T_e(0)$, is approximately 20 cells/ μ l of blood. This represents the proliferation that occurs when naive T cells receive the signal to become activated and undergo cell division, minus a small loss of T_e due to interactions with C . Small changes in these numbers were investigated numerically in the model, with no significant change observed in the outcome of the system.

The estimate of 10 000 cells/ μ l of blood is used as an initial guess for $C(0)$; the maximum concentration of CML cells in the blood observed in living patients is reported to be approximately 300 000 cells/ μ l (Lee, 2001).

5. Analytic results

To study the behavior of the system of differential equations, we find the equilibrium solutions, and linearize about these to examine their stability. To ease the calculations, we first simplify the model by rescaling the variables.

5.1. Simplification of the model

Using standard rescaling techniques (non-dimensionalization), we reduce the number of parameters from twelve to eight. The populations and time are rescaled as follows: T_n is rescaled by multiplying by a factor of d_n/s_n , T_e is rescaled by a factor of γ_c/d_n , C is rescaled by γ_e/d_n , and t is rescaled by a factor of d_n .

The resulting system is the following:

$$\frac{dT_n}{dt} = 1 - T_n - \zeta_1 T_n \left(\frac{C}{C + \zeta_2} \right), \quad (4)$$

$$\begin{aligned} \frac{dT_e}{dt} = & \zeta_3 T_n \left(\frac{C}{C + \zeta_2} \right) + \zeta_4 T_e \left(\frac{C}{C + \zeta_2} \right) \\ & - \zeta_5 T_e - C T_e, \end{aligned} \quad (5)$$

$$\frac{dC}{dt} = \zeta_6 C \ln \left(\frac{\zeta_7}{C} \right) - \zeta_8 C - C T_e. \quad (6)$$

The new coefficients are defined as follows: $\zeta_1 = k_n/d_n$, $\zeta_2 = \gamma_e \eta/d_n$, $\zeta_3 = \alpha_n k_n s_n \gamma_c/d_n^3$, $\zeta_4 = \alpha_e/d_n$, $\zeta_5 = d_e/d_n$, $\zeta_6 = r_c/d_n$, $\zeta_7 = \gamma_e C_{max}/d_n$, and $\zeta_8 = d_c/d_n$.

5.2. Equilibrium solutions

We now set each of the three simplified differential equations equal to zero, and solve for T_n , T_e , and C . This gives the equilibrium solutions; i.e., it gives values of T_n , T_e , and C for which the system will no longer change (since all of the derivatives, or rates of change, will be zero).

We first consider the equilibrium solution(s) for which $C = 0$. In this case, Eq. (4) implies that $T_n = 1$, and Eq. (5) implies that $T_e = 0$, and there are no other equilibria for which $C = 0$. We let $P_1 = (1, 0, 0)$ represent this trivial (healthy) equilibrium solution.

Any other equilibrium values of C are given by the solution to the following equation obtained from all three Eqs. (4)–(6):

$$\begin{aligned} 0 = & \zeta_8 - \zeta_6 \ln(\zeta_7) + \zeta_6 \ln(C) \\ & + \frac{\zeta_3 C(C + \zeta_2)}{(C + \zeta_2 + \zeta_1 C)[(C + \zeta_2)(C + \zeta_5) - \zeta_4]}. \end{aligned} \quad (7)$$

The third term on the right-hand side of Eq. (7) is logarithmic in C , and hence increases as C increases. For the factors of the fourth term to be zero, C must be negative (for the clinically feasible ranges of parameters in Table 1, Section 3). Thus, this fourth term is a

rational function which decreases as C increases (when C is positive), and so there is at most one value of C which makes the right-hand side of Eq. (7) equal to zero. This means there is at most one other equilibrium solution, which we will denote by $P_2 = (\tilde{T}_n, \tilde{T}_e, \tilde{C})$. This equilibrium solution occurs when $\zeta_8 < \zeta_6 \ln(\zeta_7)$, or when $d_c/r_c < \ln(\gamma_e C_{\max}/d_n)$.

In order for this equilibrium to be physically meaningful, all three populations must be nonnegative. The resulting conditions on Eqs. (4)–(6) are satisfied trivially, so the equilibrium is a realistic one, and there is no new information obtained from the conditions.

5.3. Linearization of the system

To determine the behavior of the cell populations near each of the equilibrium solutions, we need to compute the linearization of the system, which is obtained from the Jacobian matrix of the system. See, e.g., Borrelli and Coleman (1998, p.446) for more information about such calculations. For the system of Eqs. (4)–(6), the Jacobian is the following:

$$A = \begin{pmatrix} -1 - \zeta_1 \left(\frac{C}{C + \zeta_2} \right) & 0 & -\zeta_1 T_n \left(\frac{\zeta_2}{(C + \zeta_2)^2} \right) \\ \zeta_3 \left(\frac{C}{C + \zeta_2} \right) & \zeta_4 \left(\frac{C}{C + \zeta_2} \right) - \zeta_5 - C & \zeta_3 T_n \left(\frac{\zeta_2}{(C + \zeta_2)^2} \right) + \zeta_4 T_e \left(\frac{\zeta_2}{(C + \zeta_2)^2} \right) - T_e \\ 0 & C & \zeta_6 \ln \left(\frac{\zeta_7}{C} \right) - \zeta_6 - \zeta_8 - T_e \end{pmatrix}. \quad (8)$$

If we substitute a set of equilibrium values for T_n , T_e , and C in the matrix A , then the matrix A will represent the linearization of the system of differential equations about that equilibrium solution. By the Hartman–Grobman theorem (see, Othmer et al., 1997), the behavior of the original (nonlinear) system will be approximated by the behavior of the linearized system near that equilibrium solution, as long as the equilibrium solution is hyperbolic (i.e., the real parts of the eigenvalues of the matrix are not equal to zero). We will now examine the two equilibrium solutions separately.

5.3.1. Equilibrium 1

Evaluating A at the trivial solution $P_1 = (1, 0, 0)$ gives

$$A(P_1) = \begin{pmatrix} -1 & 0 & -\frac{\zeta_1}{\zeta_2} \\ 0 & -\zeta_5 & \frac{\zeta_3}{\zeta_2} \\ 0 & 0 & -\zeta_6 - \zeta_8 \end{pmatrix}. \quad (9)$$

The eigenvalues of this matrix are $\lambda_1 = -1$, $\lambda_2 = -\zeta_5$, and $\lambda_3 = -\zeta_6 - \zeta_8$, which are all real numbers. The equilibrium solution P_1 is asymptotically stable, since λ_1 , λ_2 , and λ_3 are all negative.

5.3.2. Equilibrium 2

Here we examine the equilibrium solution $P_2 = (\tilde{T}_n, \tilde{T}_e, \tilde{C})$, when it exists. In this case, analytic techniques fail to yield information about the eigenvalues, due to the complexity of the system. Conditions such as the Routh test (Borrelli and Coleman, 1998, p. 415), for example, give transcendental equations which cannot be solved analytically. As an alternative, we calculate the eigenvalues for a wide variety of possible parameter values, by systematically sampling through the ranges listed in Table 1. To ensure representative sampling, we use a technique called Latin hypercube sampling (LHS), which is discussed in detail in Section 6, to randomly select values from the specified ranges. For the equilibrium populations, we sample from the interval (0, 5000) for \tilde{T}_n and \tilde{T}_e , and we sample from the interval (1, 400 000) for \tilde{C} . In our 3000 LHS runs, all of the eigenvalues were either negative real or were imaginary with all of the real parts bounded above by -1.000076 .

The large ranges of \tilde{T}_n , \tilde{T}_e , and \tilde{C} used to calculate the eigenvalue estimates give reasonable confidence that the

eigenvalues indeed have negative real parts for this equilibrium solution. Hence we assume the equilibrium solution considered here is asymptotically stable, and we proceed with numerical analysis of the system.

5.3.3. Summary of stability

If $d_c/r_c < \ln(\gamma_e C_{\max}/d_n)$ (which is the case for the initial parameter estimates), then the system of Eqs. (4)–(6) has two asymptotically stable equilibria, P_1 and P_2 . Otherwise, the system has only one equilibrium, the trivial (healthy) state P_1 , and this equilibrium is asymptotically stable.

6. Numerical analysis: relationships between parameters and C

In this section, we examine the relationship between C and the model parameters. We numerically solve Eqs. (4)–(6) for a variety of choices of parameter values, using the ranges listed in Table 1. For all of the numerical analysis that follows, we use code we wrote for Mathematica™ 4.2. Each time we numerically solve the system of differential equations we use a time span

of 750 days. We use the largest concentration of CML cells during the 750 days as a marker for disease progression in these solutions. Other time periods (both much larger and much smaller) yield the same qualitative results in determining critical parameters in the outcome of the disease.

We look for significance of the relationships between choices of particular parameter values and outcomes for CML by keeping track of whether or not C exceeds a set threshold value at any time during the 750 day period. We use a threshold of 22 500 cells/ μl to represent a stage at which the patient is no longer in full remission, although we also obtained the same results when larger and smaller values for the threshold were used. Partial remission is evidenced by a white blood cell count in the range of approximately 15 000–30 000 cells/ μl , but it is unclear what fraction of the total count are CML cells (Athens et al., 1965; Lee et al., 1999). Stryckmans et al. (1977) show kinetic data where the fraction of white blood cells that are leukemic increases as the disease progresses, and this fraction varies greatly between patients. However, it is likely that the number of total healthy white blood cells stays around the normal range (7500–10 000) (Gordon et al., 1999). Thus, in partial remission we expect there are fewer than 22 500 CML cells/ μl , and perhaps many fewer.

Because of the uncertainty in the initial parameter estimates, we perform Latin hypercube sampling (LHS) on all of the parameters in all of the simulations below. LHS is a well-established method for investigating systems which are too complex to be tractable analytically, and for which there is possibly large uncertainty in the parameter values used in the equations (see Blower and Dowlatabadi, 1994). The technique involves numerically solving the equations multiple times, using randomly sampled values for all of the parameters in the system. Although the standard Monte Carlo sampling varies all parameters at once, LHS differs in that it ensures that the combinations of parameter choices are well distributed. The ranges of values from which samples are chosen are specified ahead of time, and are chosen here to ensure that all clinically feasible parameter values are included. Each range is divided into a specified number of equi-probable intervals, which are then put into a random order. On any given run, a random value is selected for each parameter, from the appropriate interval. This ensures the mixing of the intervals from which the random values are chosen, and provides higher confidence in the qualitative behavior of the solutions, despite the uncertainty in the parameter values.

6.1. Graphical results of parameters vs. C

First, we examine the plots of individual parameter values versus C , where the other parameters are sampled

using LHS. We reduce the dependence of the outcome on the particular choice of other parameters by varying them simultaneously. We run LHS to give 500 sets of parameter choices. For each of the 500 sets, we determine \tilde{C} , defined to be the largest value of C over the 750 day time interval. We use \tilde{C} to examine dependence on the parameters, since this decreases the dependence of our results on the choice of the correct time interval. For each parameter, we then create a graph showing the 500 values for that parameter, plotted against \tilde{C} . In our analysis of all 12 parameters in this manner, we find that the values of the parameters d_c and r_c appear to strongly correlate with \tilde{C} , with r_c appearing to have a somewhat stronger correlation. A statistical analysis of these plots appears in the following section.

In the plot of d_c vs. \tilde{C} (Fig. 2), the dependence of \tilde{C} on d_c can be clearly seen. However, even for extremely large values of d_c , there are many combinations of other parameter values which cause \tilde{C} to be 50 000 cells/ μl or more. This contrasts with the case of r_c , whose sampled values are plotted vs. \tilde{C} in Fig. 3.

In Fig. 3, there are still a few combinations of parameter values which cause \tilde{C} to be large, even for

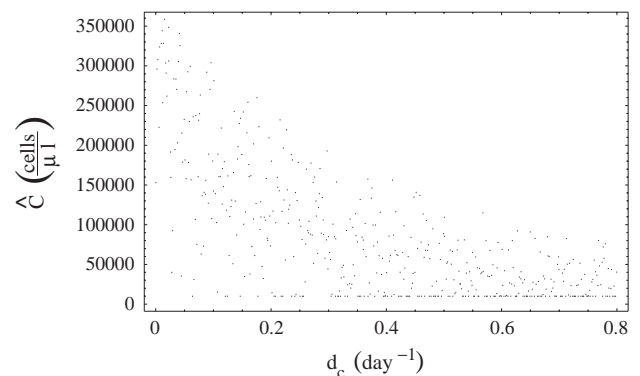


Fig. 2. A plot of 500 points with d_c vs. \tilde{C} , the maximum value of C over a 750 day time period. All other parameters are varied simultaneously, using LHS.

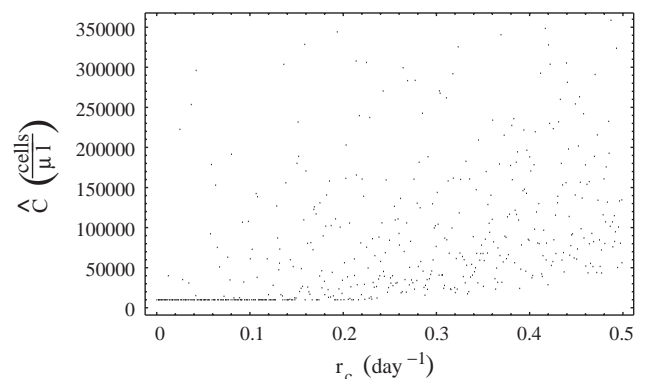


Fig. 3. A plot of 500 points with r_c vs. \tilde{C} , the maximum value of C over a 750 day time period. All other parameters are varied simultaneously, using LHS.

extremely small values of r_c . However, the large majority of the outcomes for \tilde{C} are indistinguishable from 0 cells/ μl when r_c is sufficiently small (around 0.06 cells/ $\mu\text{l}/\text{day}$), regardless of the values of other parameters.

6.2. Analysis of numerical results

In order to distinguish individual parameters that significantly affect the long-term behavior of C in the model, we record and analyse the parameter values randomly chosen during the LHS where we vary all parameters simultaneously for each of the 500 runs. The maximum value of C over the 750 day time period, \tilde{C} , is recorded for each run. We associate a value of 1 to a particular \tilde{C} if \tilde{C} is above the threshold value of 22 500 cells/ μl discussed previously. Otherwise, we associate a value of 0 to the particular value of \tilde{C} .

We use the statistical package R to fit a logistic regression to the 500 different values of a given parameter, and the associated 1's and 0's. We then perform an analysis of deviance on the regression model and a chi-squared test to determine whether or not the parameter is statistically significant. The results appear in Table 2:

Using a significance level of 0.05, there are two parameters which are clearly significant: r_c , the growth rate of CML, and d_c , the death rate of CML. It is important to note the statistical analysis is done on data which is not stochastic. Furthermore, by performing a logistic regression we must make the assumption that by focusing on a single parameter at a time, the effects of all the other parameters introduce enough noise that they can be considered “random” in some sense. These limitations occur because presently there are not better tools for exploring the behavior of complex high-dimensional differential equation models.

6.3. Additional graphical evidence examining significance

For this analysis, we look at all possible pairs of the 12 parameters. This gives 66 (“12 choose 2”) pairs to examine. For each of these pairs, we make a graph of plotted points of one parameter versus the other, with the coordinates of the points in the graph given by the sampled values for those parameters from the LHS. The points in the plot are colored light gray if the

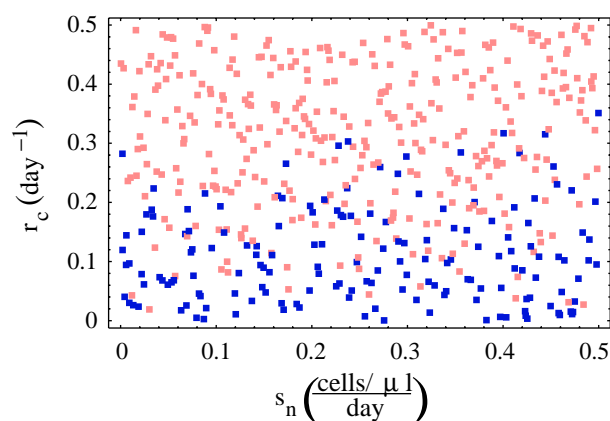


Fig. 4. Plot of the LHS sampled values for s_n vs. r_c . The light gray squares are values which caused C to go above the given threshold value of 22 500 cells/ μl , and the dark gray squares are values for which C remained below the threshold. All other parameters are varied simultaneously.

corresponding solution for C goes above our chosen threshold of 22 500 cells/ μl , and dark gray if it does not.

In the majority of the graphs (all except for r_c and d_c), there are no apparent regions of exclusion of points for which C is above the threshold value of 22 500 cells/ μl or below. However, in all of the graphs that include r_c (the effective growth rate of C), the coordinates of dark gray points are restricted to values of r_c below approximately 0.35 cells/ $\mu\text{l}/\text{day}$. The percentage of points which are dark gray increases as the values of r_c decreases below this. This indicates that no matter what the values of the other 11 parameters, the value of r_c determines the outcome of the system in the majority of cases. See the plot of s_n versus r_c given in Fig. 4.

In the graphs that include d_c (the natural per capita death rate of C) few of the dark gray points occur for values of d_c below approximately 0.2/day. The percentage of points which are dark gray increases as d_c increases, but do not achieve the percentages achieved in the case of r_c . This indicates that having $d_c \geq 0.2$ makes it more likely that C can be controlled, but that condition alone is likely not sufficient; the outcome also depends on the values of other parameters to a large degree. See the plot of s_n versus d_c given in Fig. 5.

7. Discussion of results of numerical analysis

In addition to the computational results shown above, in the many simulations run using this model, we observe a variety of behaviors for C for different choices of parameter values. The observed outcomes correlate with some of the disease characteristics typically found in patients with CML, while undergoing treatment or not responding to treatment. These outcomes include: high CML cell counts (Fig. 7); remission and relapse or rebound of CML (Fig. 8); and control of CML (Fig. 6).

Table 2
Statistical results

Parameter	$P > \chi $	Parameter	$P > \chi $	Parameter	$P > \chi $
s_n	0.22	k_n	0.67	C_{max}	0.65
d_n	0.40	η	0.09	r_c	<0.001
d_e	0.61	α_n	0.39	γ_e	0.58
d_c	<0.001	α_e	0.06	γ_c	0.12

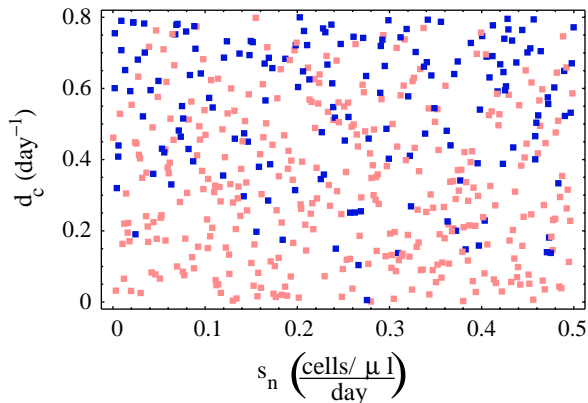


Fig. 5. Plot of LHS sampled values for s_n vs. d_c . The light gray squares are values which caused C to go above the given threshold value of 22 500 cells/ μ l, and the dark gray squares are values for which C remained below the threshold. All other parameters are varied simultaneously.

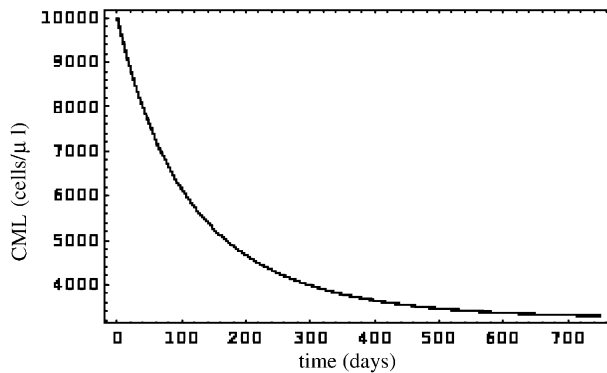


Fig. 6. An example graph of CML versus time, where CML decreases over a 2-year period.

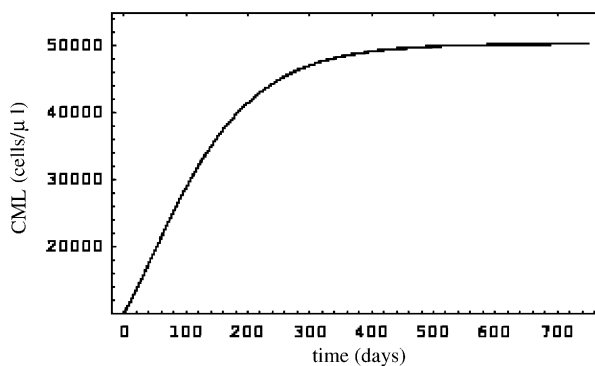


Fig. 7. An example graph of CML versus time, where CML increases over a 2-year period.

We mention this evidence as support for the biological relevance of our model. We show some examples of these graphs below. Table 3 gives the approximate parameter values for the graphs.

The results of the numerical analysis in the previous sections give strong evidence themselves for the validity

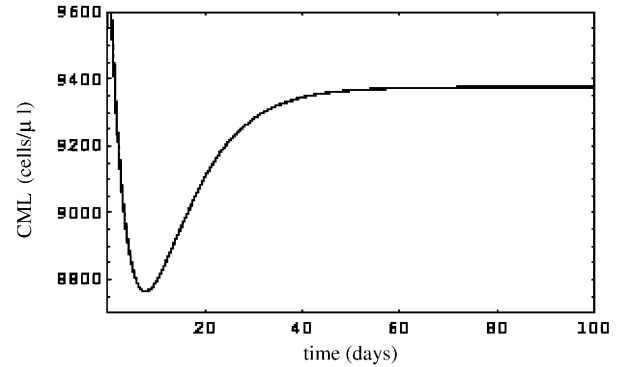


Fig. 8. An example graph of CML versus time, where CML cell counts initially decrease, but later rebound.

of the model presented here. Another measure of the model's accuracy is given by a comparison of known treatments and their effectiveness, to the predictions of the model. The current main treatments for CML are radiation, chemotherapy, bone marrow transplant, and the drug Gleevec (imatinib). (For a review on CML treatments, see Schiffer et al., 2003.)

Radiation and chemotherapy increase the death rates of the CML cells, but they are non-specific treatments, and also increase the death rates of the naive and effector T cells. For the majority of patients, these treatments alone are not effective in eradicating CML (Schiffer et al., 2003). In light of this, in some cases bone marrow transplant is used in conjunction with chemotherapy, either with or without radiation therapy. This combination treatment resets the CML count to zero or very low levels, and increases the death rates of all three cell populations. The combination treatment of bone marrow transplant, radiation, and chemotherapy is the only treatment that can lead to complete permanent cure, however it does not always lead to a cure (Schiffer et al., 2003). These treatment outcomes agree with our numerical results that changing d_c can be significant. Because these treatments also increase d_n and d_e , we do not expect that the treatments would necessarily be sufficient to control CML.

Gleevec was approved in 2001 as a treatment for CML, and completely controls C in 95% of all CML patients for long periods of treatment (Kantarjian et al., 2002). However, CML mutations that cause Gleevec to lose its effectiveness can occur after one or more years, and patients return to disease progression (Kreuzer et al., 2003). Gleevec works by blocking one of the pathways that CML uses to reproduce at an excessive rate (Kreuzer et al., 2003). In our model, this rate is represented by r_c . The LHS results predict that adequately controlling r_c should control CML, which is exactly what is observed in patients treated with Gleevec, until CML mutates around the drug and Gleevec no longer is able to affect the CML growth rate.

Table 3
Parameter values used in graphs

Graph	s_n	d_n	d_e	d_c	k_n	η	α_n	α_e	C_{max}	r_c	γ_e	γ_c
Fig. 6	0.37	0.23	0.30	0.024	0.062	720	0.14	0.98	230 000	0.0057	0.057	0.0034
Fig. 7	0.29	0.35	0.40	0.012	0.066	140	0.39	0.65	160 000	0.011	0.079	0.058
Fig. 8	0.071	0.050	0.12	0.68	0.063	43	0.56	0.53	190 000	0.23	0.0077	0.047

8. Conclusions

The results above suggest several clinically relevant conclusions for treatment of CML. The most striking observation is that the growth rate of CML, r_c , and the natural death rate, d_c , are the most important parameters in the control of CML in this model. The importance of r_c suggests that Gleevec and other drugs which attempt to block excessive growth rates of CML are likely to be successful treatments for CML, for the duration of their effectiveness. Our results also give some indication that increasing d_c may not be sufficient in all cases in controlling CML, which matches observations of current treatments, and which we hope will be investigated more closely by other researchers. Another reason why controlling CML using d_c is likely to be insufficient, is the difficulty of producing a drug which increases d_c without increasing d_n or d_e , the natural death rates of the T cells.

Thus, researchers investigating treatments for CML patients may wish to focus on treatments that can affect the parameter r_c . Our model suggests that treatments that attempt to change the other parameters are less likely to achieve reduction or eradication of CML. Because of the high risk involved in bone marrow transplantation, the resistance that CML develops to Gleevec, and the lack of response of some patients to any of the treatments mentioned here, there is still a pressing need for other treatments or treatment protocols for patients with CML.

Another outcome of our analysis is validation of the model itself. This new model can now be used and refined to include other aspects of the immune-CML interaction. One example would be the incorporation of the effects of immune system-boosting treatments in patients with CML. Another use we hope to see is the development of a more detailed model, with experiments designed to obtain more precision in parameter values.

Having this model allows us to solve the optimal drug dosing problem for drugs such as Gleevec, which we hope to complete in the near future. Given an estimate of the mutation rates of CML, the most effective dosing levels over time can be predicted for Gleevec, or for other drugs or combinations of drugs. Optimal dosing can substantially increase the period in which a patient remains healthy, which is especially important for treatments to which diseases eventually develop resistance or which are expensive. Gleevec currently costs

approximately \$2500 per month, which means that a constrained optimal dosing solution, for limited resource scenarios, would be valuable as well. Further work predicting optimal combinations of chemotherapy, radiation, and drugs such as Gleevec would also be of great benefit in the treatment of CML.

Acknowledgements

Thanks to Peter Lee, MD and Michelle Green, Ph.D., of Stanford University, Denise Kirschner, Ph.D., of the University of Michigan, Alan Perelson, Ph.D., of Los Alamos National Laboratory, and Johnny Lam, statistician, of Palo Alto, CA, for helpful discussions during the development and analysis of this model.

References

- Afenya, E., Calderón, C., 2000. Diverse ideas on the growth kinetics of disseminated cancer cells. *Bull. Math. Biol.* 62, 527–542.
- Athens, J.W., Raab, S.O., Haab, O.P., Boggs, D.R., Ashenbrucker, H., Cartwright, G.E., Wintrobe, M.M., 1965. Leukokinetic studies. X. Blood granulocyte kinetics in chronic myelocytic leukemia. *Leukemia* 44 (5), 765–777.
- Bassukas, I.D., Maurer-Schultze, B., 1988. The recursion formula of the Gompertz function: a simple method for the estimation and comparison of tumor growth curves. *Growth Develop. Aging* 52, 113–122.
- Blower, S., Dowlatabadi, H., 1994. Sensitivity and uncertainty analysis of complex models of disease transmission: an HIV model, as an example. *Int. Stat. Rev.* 2, 229–243.
- Borrelli, R., Coleman, C., 1998. *Differential Equations: A Modeling Perspective*. Wiley, New York.
- Cronkite, E.P., Vincent, P.C., 1969. Granulocytopoiesis. *Ser. Haematol.* 2, 3–43.
- Duvall, C.P., Perry, S., 1968. The use of 51-chromium in the study of leukocyte kinetics in chronic myelocytic leukemia. *J. Lab. Clin. Med.* 71, 614–628.
- Edelstein-Keshet, L., 1988. *Mathematical Models in Biology*. Random House, New York.
- Essunger, P., Perelson, A.S., 1994. Modeling HIV infection of CD4⁺ T-cell subpopulations. *J. Theor. Biol.* 170, 367–391.
- Faderl, S., Talpaz, M., Estrov, Z., O'Brien, S., Kurzrock, R., Kantarjian, H., 1999. The biology of chronic myeloid leukemia. *New Engl. J. Med.* 341 (3), 164–172.
- Fokas, A.S., Keller, J.B., Clarkson, B.D., 1991. A mathematical model of granulocytopoiesis and chronic myelogenous leukemia. *Cancer Res.* 51 (8), 2084–2091.
- Freedman, H., 1980. *Deterministic Mathematical Models in Population Ecology*. Marcel Dekker, New York.

- Gordon, M.Y., Dazzi, F., Marley, S.B., Lewis, J.L., Nguyen, D., Grand, F.H., Davidson, R.J., Goldman, J.M., 1999. Cell biology of CML cells. *Leukemia*. Apr;13 Suppl 1:S65–71.
- Haase, A.T., Henry, K., Zupancic, M., Sedgewick, G., Faust, R.A., Melroe, H., Cavert, W., Gebhard, K., Staskus, K., Zhang, Z.Q., Dailey, P.J., Balfour Jr., H.H., Erice, A., Perelson, A.S., 1996. Quantitative image analysis of HIV-1 infection in lymphoid tissue. *Science* 274, 985–989.
- Janeway, C.A., Travers, P., Walport, M., Shlomchik, M., 2001. *Immunobiology: The Immune System in Health and Disease*. Garland Publishing, New York.
- Kantarjian, H., Sawyers, C., Hochhaus, A., Guilhot, F., Schiffer, C., Gambacorti-Passerini, C., Niederwieser, D., Resta, D., Capdeville, R., Zoellner, U., Talpaz, M., Druker, B., Goldman, J., O'Brien, S.G., Russell, N., Fischer, T., Ottmann, O., Cony-Makhoul, P., Facon, T., Stone, R., Miller, C., Tallman, M., Brown, R., Schuster, M., Loughran, T., Gratwohl, A., Mandelli, F., Saglio, G., Lazzarino, M., Russo, D., Baccarani, M., Morra, E., International STI571 CML Study Group, 2002. Hematologic and cytogenetic responses in patients with accelerated phase chronic myeloid leukemia: results of a phase 2 study. *New Engl. J. Med.* 346, 645–652.
- Kirschner, D., Panetta, J.C., 1998. Modeling immunotherapy of the tumor-immune interaction. *J. Math. Biol.* 37, 235–252.
- Kirschner, D., Webb, G.F., 1996. A model for treatment strategy in the chemotherapy of AIDS. *Bull. Math. Biol.* 58 (2), 367–390.
- Kreuzer, K.A., Le Coutre, P., Landt, O., Na, I.K., Schwartz, M., Schultheis, K., Hochhaus, A., Dorken, B., 2003. Preexistence and evolution of imatinib mesylate-resistant clones in chronic myelogenous leukemia detected by PNA-based PCR clamping technique. *Ann. Hematol.* 82 (5), 284–289.
- Kuznetsov, V., Makalkin, I., Taylor, M., Perelson, A., 1994. Nonlinear dynamics of immunogenic tumors: parameter estimation and global bifurcation analysis. *Bull. Math. Biol.* 56 (2), 295–321.
- Lee, G.R., Foerster, J., Lukens, J., Paraskevas, F., Greer, J.P., Rodgers, G., Wintrobe, M.M. (Eds.), 1999. *Wintrobe's Clinical Hematology*. Lippincott, Williams & Wilkins, Philadelphia, PA.
- Lee, P., 2001. Pers. comm.
- Mohri, H., Perelson, A., Tung, K., Ribeiro, R., Ramratnam, B., Markowitz, M., Kost, R., Hurley, A., Weinberger, L., Cesar, D., Hellerstein, M., Ho, D., 2001. Increased turnover of T lymphocytes in HIV-1 infection and its reduction by antiretroviral therapy. *J. Exp. Med.* 194 (9) (November 5), 1277–1287.
- Nowak, M.A., Bonhoeffer, S., Hill, A.M., Boehme, R., Thomas, H.C., McDade, H., 1996. Viral dynamics in hepatitis B virus infection. *Proc. Natl. Acad. Sci. USA* 93, 4398–4402.
- Nowak, M.A., May, R.M., 2000. *Virus Dynamics: Mathematical Principles of Immunology and Virology*. Oxford University Press, Oxford, UK.
- Othmer, H.G., Adler, F.R., Lewis, M.A., Dallon, J.C., 1997. *Case Studies in Mathematical Modeling Ecology, Physiology, and Cell Biology*. Englewood Cliffs, Prentice-Hall, NJ.
- Perelson, A.S., Nelson, P.W., 1999. Mathematical analysis of HIV-I dynamics *in vivo*. *SIAM Rev.* 41 (1), 3–44.
- Perelson, A.S., Essunger, P., Cao, Y., Vesanen, M., Hurley, A., Saksela, K., Markowitz, M., Ho, D.D., 1997. Decay characteristics of HIV-1-infected compartments during combination therapy. *Nature* 387, 188–191.
- Perelson, A.S., Kirschner, D., De Boer, R., 1993. Dynamics of HIV infection of CD4+ T cells. *Math. Biosci.* 114, 81–125.
- Rubinow, S.I., 1969. A simple model of steady state differentiating cell system. *J. Cell Biol.* 43, 32–39.
- Rubinow, S.I., Lebowitz, J.L., 1975. A mathematical model of neutrophil production and control in normal man. *J. Math. Biol.* 1, 187–225.
- Sawyers, C.L., 1999. Chronic myeloid leukemia. *New Engl. J. Med.* 340 (17), 1330–1340.
- Schiffer, C.A., Hehlmann, R., Larson, R., 2003. Perspectives on the treatment of chronic phase and advanced phase CML and Philadelphia chromosome positive ALL. *Leukemia* 17, 691–699.
- Steel, G.G., 1977. *Growth of Kinetic Tumors*. Oxford University Press, Oxford, UK.
- Stryckmans, P.A., Debusscher, L., Collard, E., 1977. Cell kinetics in chronic granucloctyci leukaemia (CGL). *Clin. Haematol.* 6 (1), 21–40.
- Wigginton, J.E., Kirschner, D., 2001. A model to predict cell-mediated immune regulatory mechanisms during human infection with *Mycobacterium tuberculosis*. *J. Immunol.* 166 (3), 1951–1967.
- Xu, X.L., Ling, Y.B., 1988. A study on the expectational model for tumor growth. *Int. J. Biomed. Comput.* 22 (2), 135–141.