Bulletin of Mathematical Biology (2008) 70: 728–744
DOI 10.1007/s11538-007-9276-z

ORIGINAL ARTICLE

Modeling Imatinib-Treated Chronic Myelogenous Leukemia: Reducing the Complexity of Agent-Based Models

Peter S. Kim[a], Peter P. Lee[b], Doron Levy[c,*]

[a] *Department of Mathematics, Stanford University, Stanford, CA 94305-2125, USA*
[b] *Division of Hematology, Department of Medicine, Stanford University, Stanford, CA 94305, USA*
[c] *Department of Mathematics and Center for Scientific Computations and Mathematical Modeling (CSCAMM), University of Maryland, College Park, MD 20742, USA*

Received: 5 April 2007 / Accepted: 4 September 2007 / Published online: 1 December 2007

Abstract We develop a model for describing the dynamics of imatinib-treated chronic myelogenous leukemia. Our model is based on replacing the recent agent-based model of Roeder et al. (Nat. Med. 12(10):1181–1184, 2006) by a system of deterministic difference equations. These difference equations describe the time-evolution of clusters of individual agents that are grouped by discretizing the state space. Hence, unlike standard agent-base models, the complexity of our model is independent of the number of agents, which allows to conduct simulation studies with a realistic number of cells. This approach also allows to directly evaluate the expected steady states of the system. The results of our numerical simulations show that our model replicates the averaged behavior of the original Roeder model with a significantly reduced computational cost. Our general approach can be used to simplify other similar agent-based models. In particular, due to the reduced computational complexity of our technique, one can use it to conduct sensitivity studies of the parameters in large agent-based systems.

Keywords Chronic myelogenous leukemia · Gleevec · Imatinib · Agent-based models · Difference equations · Steady states

1. Introduction

Chronic Myelogenous Leukemia (CML) is a blood cancer with a common acquired genetic defect resulting in the overproduction of malformed white blood cells. It constitutes nearly 20% of all leukemias, affecting roughly 1 in 100,000 people. The cause of CML is an acquired genetic abnormality in hematopoietic stem cells due to a reciprocal translocation between chromosomes 9 and 22. This translocation, creating the Philadelphia (Ph)

*Corresponding author.
E-mail address: dlevy@math.umd.edu (Doron Levy).

chromosome, has associated oncogenic properties and can be detected in more than 90% of all patients with CML (Thijsen et al., 1999). The BCR-ABL fusion gene, drives increased and aberrant tyrosine kinase activity due to the chromosomal rearrangement. It is this abnormal activity that leads to dysfunctional regulation of cell growth and survival, and consequently to cancer (Thijsen et al., 1999). Treatment and control of CML underwent a dramatic change with the introduction of the new tyrosine kinase inhibitor imatinib, which has proven to be an effective treatment available to nearly all CML patients, especially in chronic phase (Druker and Lydon, 2000). While it is widely agreed that imatinib does not represent a true cure for CML, it does provide, however, an effective control measure available to be undertaken between or in the place of more aggressive treatments such as an allogeneic bone marrow or stem-cell transplantation (Campbell et al., 2001).

In addition to experimental approaches for studying CML, in the past several years, the use of mathematical models has grown rapidly. The first mathematical models of CML date back to 1969 and are due to Vincent et al. (1969), and Rubinow and Lebowitz (1976a, 1976b, 1977). A pioneering work originating from the hematopoietic system is due to Fokas et al. (1991). A simple delayed differential equations model for the immune response in CML is due to Neiman (2002). A more recent work is of Moore and Li (2004). Their ODE-based CML model followed the dynamics of cancer cells, naïve T cells, and killer T cells. Rare versions of CML have been mathematically studied by Bernard et al. (2003). Komarova and Wodarz (2005) have used methods of stochastic networks to study drug resistance with applications to imatinib. A recent mathematical model for CML with imatinib is due to Michor et al. (2005). By matching their ODE model to patients' data, Michor et al. concluded that hematopoietic stem cells are mostly immune to imatinib.

Roeder et al. (2006) present a stochastic agent-based model (ABM) for the interaction between imatinib and CML. It could be considered as an extension of the Michor model (Michor et al., 2005). Both models, (Michor et al., 2005; Roeder et al., 2006), account for the progression of cells through the myeloid lineage. Michor et al. consider four stages of cell differentiation: stem cells, progenitor cells, differentiated cells, and terminally differentiated cells. Similarly, Roeder et al. consider the differentiation of cells through three stages: stem cells, proliferating precursor cells and nonproliferating precursor and mature cells. As an additional assumption, Roeder et al. divide the stem cells into two compartments, proliferating and nonproliferating stem cells. Individual stem cells circulate continually between the two compartments and are affected by imatinib only while proliferating.

In this work we develop a model for describing the dynamics of imatinib-treated chronic myelogenous leukemia. Our model is constructed by replacing the recent agent-based model of Roeder et al. (2006) by a system of deterministic difference equations. These difference equations describe the time-evolution of clusters of individual agents that are grouped together by discretizing the state space. When compared with the original ABM, the resulting system has a significantly lower number of equations that are still simple enough so that the overall computational complexity is significantly reduced. More importantly, the difference equations enjoy much better scalability properties: the population sizes do not affect running times. This is in contrast to ABMs where the number of equations is given by the number of agents. Hence, in our approach it is possible to conduct simulations with a realistic number of cells. Another advantage of our approach is that steady states of the system can be quickly evaluated without having to run the ABM

to its steady state. Due to the reduced computational complexity of our approach, it is also possible to easily conduct sensitivity studies (such as Latin Hypercube Sampling) in order to determine the sensitivity of the system to the choice of parameters. While our technique is demonstrated for this particular leukemia modeling problem, we would like to emphasize that it is not limited to this particular problem as it can be adapted to a wider-variety of similar ABMs.

The structure of this paper is as follows. In Section 2, we briefly review the model of Roeder et al. (2006). In Section 3, we reformulate the Roeder model as a system of difference equations. In Section 4, we compare the steady state behavior, the dynamics of CML genesis and the dynamics of imatinib treatment between the ABM and difference equation model. Concluding remarks are provided in Section 5. These remarks include a discussion on approaches to increasing the numerical accuracy of the difference equation model, an extension to a general method for obtaining deterministic versions of ABMs, and remarks about the utility of deterministic approximations of stochastic ABMs.

2. An overview of the Roeder model

In the Roeder model (2006), hematopoietic stem cells (HSCs) exist in two growth compartments: noncycling, denoted by A, and proliferating, denoted by Ω. At every time step (of 1 hour), each stem cell has a probability of changing its compartment.

At the beginning of every time step, the algorithm stochastically determines whether a cell transfers from A to Ω with probability ω or from Ω to A with probability α. Each stem cell has an affinity, denoted by $a(t)$, and the affinity ranges between $a_{\min}$ and $a_{\max}$ which are estimated to be 0.002 and 1.0, respectively (Roeder et al., 2006). A cell with a high affinity has a high chance of remaining in or transferring to the A environment. Likewise, a cell with a low affinity is more likely to remain in or transfer to the Ω environment, where it starts proliferating.

The probabilities ω and α are given by

$$\begin{aligned} \omega\big(\Omega(t), a(t)\big) &= \frac{a_{\min}}{a(t)} f_{\omega}\big(\Omega(t)\big), \\ \alpha\big(A(t), a(t)\big) &= \frac{a(t)}{a_{\max}} f_{\alpha}\big(A(t)\big). \end{aligned} \tag{1}$$

Here $A(t)$ and $\Omega(t)$ denote the total number of cells in each compartment. In addition, the functions f_{α} and f_{ω} are sigmoidal functions whose definition is given in Eq. (B.1) in Appendix B.

Proliferating cells in the Ω compartment progress through various stages of the cell cycle: G_1, S, G_2, and M. The G_1 phase is the longest period of growth during which the cell generates new organelles. The S phase is the period when DNA synthesis and replication occurs. The G_2 phase is the short period of growth when the cell prepares for mitosis, and the M phase, or mitosis, is when the cell replicates its DNA and divides into two daughter cells. Only Ω cells in the G_1 phase of the cell cycle can transfer to A. In the Roder model, Ω cells spend about two-thirds of their time in the G_1 phase (Roeder et al., 2006).

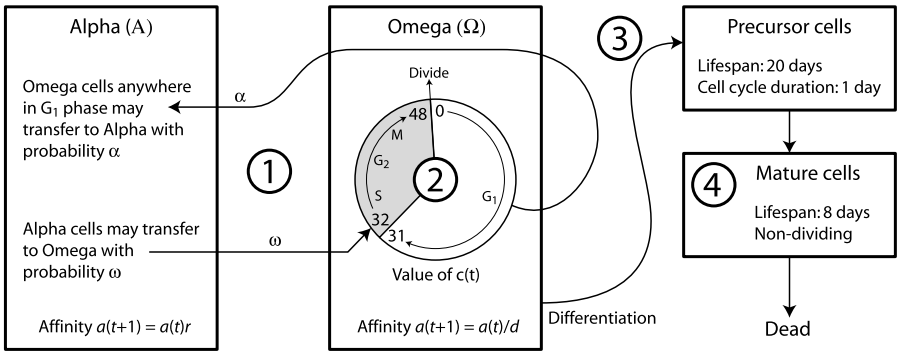


Fig. 1 A state diagram for the Roeder model (2006). (1) At every time step, stem cells may transfer between the A (nonproliferating) and Ω (proliferating) compartments. While in A, a cell's affinity increases by a factor of r up to the maximum affinity, $a_{\max}$. While in Ω, a cell's affinity decreases by a factor of d until it reaches the minimum affinity, $a_{\min}$. (2) Ω cells progress through the G_1, S, G_2, and M phases of the cell cycle. The counter $c(t)$ increases cyclically from 0 to 48. When a cell first enters Ω, $c(t) = 32$ to mark the beginning of S phase. Only cells in the G_1 phase can transfer back to A. (3) When a cell's affinity $a(t)$ drops below $a_{\min}$, it differentiates into a precursor cell. Precursor cells proliferate for 20 days, dividing once per day. (4) At the end of 20 days, precursors differentiate into mature cells and live for 8 additional days without dividing.

For each cell that remains in the A compartment, its affinity increases by a factor of r (estimated in Roeder et al., 2006 as 1.1). Similarly, cells that remain in Ω, decrease their affinity by a factor of d (estimated in Roeder et al., 2006 as 1.05). Once a cell reaches maximum affinity, $a_{\max}$, its affinity stops increasing. Stem cells whose affinity reaches $a_{\min}$, differentiate into a proliferating precursor and then into a nonproliferating mature cell.

In addition, each cell in Ω has a time counter, $c(t)$, that indicates its position in the cell cycle. At each time step, this counter increases by 1. After the counter reaches its maximal value of 48, it recycles back to 0 at the next time step, resulting in a 49-hour cell cycle. Cells entering Ω start with a counter that is set at $c(t) = 32$ corresponding to beginning of the S phase. For the first seventeen hours, the cell progresses through the S, G_2, and M phases and divides into two cells once $c(t) = 48$. Then, for the next 32 hours, $(c(t) = 0, \ldots, 31)$, the cell remains in the G_1 phase. If at the end of this period the cell has not transfered to A, it reenters the S, G_2 and M phases and the cycle repeats. See Fig. 1.

3. A difference equations model

The Roeder model has the advantage of simulating stem cell dynamics to a very high resolution (i.e., each cell is considered individually); however, it has the disadvantage of being computationally demanding. To make the computations feasible, Roeder et al. down-scaled absolute cell numbers to $1/10$ of the realistic values in patients and verified that the scaling did not affect the qualitative behavior of the system (Roeder et al., 2006).

Our goal is to replace the Roeder model by a smaller system of difference equations. In the process, we would like to decrease the total number of variables by discretizing the state space. Time is already discretized in the original model since cell numbers are

updated at fixed time steps of one hour, so the only quantities that have to be discretized are the cell affinities.

We note that $\log r = \log(1.1) = 0.0953$ and $\log d = \log(1.05) = 0.0488$, which means that the ratio $(\log r)/(\log d)$ is approximately 2. If we let $\rho = 0.0488$ and replace r by $e^{2\rho} = 1.10252$ and d by $e^{\rho} = 1.05001$, we obtain the ratio $(\log r)/(\log d) = 2$.

The value of the minimum affinity is $a_{\min} = 0.002$, giving $\log a_{\min} = -6.2146$, which is between -128ρ and -127ρ. Thus, with our modified values of r and d, any cell starting with an affinity $a(t) = e^{-k\rho}$ can only attain affinities of the form $a(t) = e^{-k\rho}$ where $0 \le k \le 127$. Stem cells with affinities other than $e^{-k\rho}$ only exist transiently, because these cells either lose affinity and differentiate or remain in A long enough to attain the maximum affinity of 1. Hence, for the purposes of studying long-term behavior, it suffices to assume that the only possible affinities are of the form $e^{-k\rho}$ for $0 \le k \le 127$.

3.1. System of difference equations

With the new values of r and d there are only finitely many affinity values, so we can fully categorize the stem cell populations at time t by the following variables:

$$A_k(t) = \text{Number of cells in } A \text{ with } \log a(t) = -k\rho,$$

$$\Omega_{k,c}(t) = \text{Number of cells in } \Omega \text{ with } \log a(t) = -k\rho \text{ and } c(t) = c,$$

where $k \in \{0, \dots, 127\}$ and $c \in \{0, \dots, 48\}$. Then, the Roeder model can be reformulated as the following system of difference equations:

$$A_k(t+1) = \begin{cases} (A_0(t) - B_0(t)) + (A_1(t) - B_1(t)) + (A_2(t) - B_2(t)), & k = 0, \\ (A_{k+2}(t) - B_{k+2}(t)) + \sum_{c=0}^{31} \Psi_{k,c}(t), & k = 1, \dots, 125, \\ \sum_{c=0}^{31} \Psi_{k,c}(t), & k = 126, 127, \end{cases} \tag{2}$$

$$\Omega_{k,c}(t+1) = \begin{cases} B_0(t), & k = 0,\ c = 32, \\ 2\Omega_{k-1,48}(t), & k > 0,\ c = 0, \\ \Omega_{k-1,c-1}(t) - \Psi_{k-1,c-1}(t), & k > 0,\ c = 1, \dots, 31, \\ (\Omega_{k-1,31}(t) - \Psi_{k-1,31}(t)) + B_k(t), & k > 0,\ c = 32, \\ \Omega_{k-1,c-1}(t), & k > 0,\ c = 33, \dots, 48, \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

In these equations, the random variables B_k and $\Psi_{k,n}$ denote the number of cells transferring from states A_k and $\Omega_{k,n}$, respectively, and they are given by the following binomial distributions:

$$B_k(t) \sim \text{Bin}\big(A_k(t), \omega\big(\Omega(t), e^{-k\rho}\big)\big),$$

$$\Psi_{k,c}(t) \sim \text{Bin}\big(\Omega_{k,c}(t), \alpha\big(A(t), e^{-k\rho}\big)\big), \quad c = 0, \dots, 31,$$

where $A(t) = \sum_k A_k(t)$ and $\Omega(t) = \sum_{k,c} \Omega_{k,c}(t)$ are the total number of cells in A and Ω, respectively, and the transition probabilities, α and ω, are given by (1).

The three terms in the first line of (2) (corresponding to $k = 0$) denote the number of cells in states A_0, A_1, and A_2 that do not transfer to Ω. At the end of the time step, these

cells end up in state A_0. The first term in the second line of (2) denotes the number of cells in state A_{k+2} that do not transfer to Ω. At the end of the time step, these cells shift to state A_k. The summation corresponds to the number of cells in Ω that transfer to A_k. The third line of (2) is the same as the second line, except that there are no cells with $\log a(t)$ lower than -127ρ that can shift into A_{126} and A_{127}.

The first line in (3) (corresponding to $k = 0, c = 32$) is the number of cells that transfer from state A_0. These cells always start with $c(t+1) = 32$. Since no cells in Ω can have $\log a(t) = 0$ except for those that have just transferred from A, we have $\Omega_{0,c}(t+1) = 0$ for all $k = 0, c \neq 32$, which is the last line of (3).

The second line of (3) represents the number of cells that have completed division and are recycling back to the beginning of the G_1 phase. The factor of 2 accounts for the increase in population due to cell division. The third line of (3) corresponds to cells that have been in the G_1 phase during the previous and current time steps. The expression denote the number of cells that do not transfer to A. These cells shift to the next state, $\Omega_{k,c}$.

The following line ($k > 0, c = 32$) pertains to cells at the beginning of S phase. As above, the first term represents the number of cells that do not transfer to A. These cells shift to the $\Omega_{k,32}$ state. The second term represents the number of cells that transfer from state A_k. The next to last line in (3) applies to cells in the S, G_2 or M phases. These cells do not transfer to A, so they all shift to the next state, $\Omega_{k,c}$.

3.2. Difference equations for differentiated cells

Once a cell's affinity drops below the minimum value ($\log a(t) < -127\rho$), the cell differentiates into a proliferating precursor and later a nonproliferating mature cell. The cell remains in the proliferating precursor state for $\lambda_p = 20$ days (480 hours) and divides every $\tilde{\tau}_c = 24$ hours. At the end of 480 hours, the precursor cell differentiates into a mature cell and lives for $\lambda_m = 8$ additional days (192 hours) without dividing. These durations are given in Roeder et al. (2006) and restated in Table B.1.

The difference equations for the differentiated cells are given by

$$P_j(t+1) = \begin{cases} \sum_{c=0}^{48} \Omega_{124,c}(t) - \sum_{c=0}^{31} \Psi_{124,c}(t), & j = 0, \\ 2P_{j-1}(t), & j = 24, 48, 72, \ldots, 456, \\ P_{j-1}(t), & \text{otherwise}, \end{cases} \tag{4}$$

$$M_j(t+1) = \begin{cases} 2P_{479}(t), & j = 0, \\ M_{j-1}(t), & \text{otherwise}, \end{cases} \tag{5}$$

where $P_j(t)$ denotes the number of cells that have been precursors for j hours, $j = 0, \ldots, 479$, and $M_j(t)$ is the number of cells that have been in the mature state for j hours, $j = 0, \ldots, 191$.

The first line in (4) represents the number of cells with $\log a(t) = -127\rho$ that have not transferred from Ω to A at time t. All these cells differentiate into precursors at time $t+1$. The second line applies to cells that have completed a cycle of division. Cell divisions occur every 24 hours, and the factor of 2 accounts for the increase in population due to division. The last line accounts for cells that increase in age by one hour between times t and $t+1$. The equations in (5) are analogous to those in (4), and we assume that precursor cells perform one final division before differentiating into a mature cell. Once a cell has matured, it no longer divides.

3.3. Modeling leukemia cells and the effects of imatinib

As in the Roeder model, we label leukemia cells as Ph^+ and nonleukemia cells as Ph^-. These labels indicate whether a cell possesses the Philadelphia chromosome, which is present in the majority of CML cases. The Roeder model considers three populations: Ph^- cells, Ph^+ cells, and imatinib-affected Ph^+ cells. In some cases, the model includes a fourth population, imatinib-resistant Ph^+ cells, but in this paper, we only consider the first three. In our model, we formulate a separate set of equations for each population. The difference equations for the three populations are similar to (2) to (5) with a few modifications.

We denote the Ph^- populations by A_k^-, $\Omega_{k,c}^-$, P_j^-, and M_j^-, and their equations are exactly the same as the original equations (2) to (5).

We denote the unaffected Ph^+ populations by A_k^+, $\Omega_{k,c}^+$, P_j^+, and M_j^+. These cells are governed by the transition functions, $f_{\alpha/\omega}$, given by the parameters corresponding to Ph^+ cells in Table B.1. Otherwise, the difference equations (2), (4), and (5) remain the same for A_k^+, P_j^+, and M_j^+. On the other hand, proliferating stem cells, $\Omega_{k,c}^+$, may become imatinib affected with probability r_{inh} at every time step. In our formulation, we assume that these transitions take place at the beginning of a time step, before any other transition. In other words, we first remove the cells that become imatinib-affected and then evaluate all other state transitions and divisions according to (3). More explicitly, let $\Omega_{k,c}^{+/R}(t)$ denote the cells that remain unaffected by imatinib throughout the current time step and $\Omega_{k,c}^{+/I}(t)$ denote the cells that become affected at time t. Then, we have

$$\Omega_{k,c}^{+/R}(t) = \Omega_{k,c}^{+}(t) - \Omega_{k,c}^{+/I}(t),$$

where $\Omega_{k,c}^{+/I}(t) \sim \text{Bin}(\Omega_{k,c}^{+}(t), r_{\text{inh}})$. Finally, we substitute $\Omega_{k,c}^{+/R}(t)$ into the right-hand side of (3) to obtain $\Omega_{k,c}^{+}(t+1)$.

It is unclear in Roeder et al. (2006) whether it is assumed that imatinib-inhibition takes place before or after other transitions within a time interval. However, these two options are almost identical, since the end of one time interval is the beginning of the next. Hence, the only difference between the two options is that the former gains one extra step of imatinib-inhibition at the beginning of treatment. Since the inhibition probability $r_{\text{inh}} = 0.050$ is small, the extra step hardly makes a difference.

Finally, we denote the imatinib-affected Ph^+ populations by $A_k^{+/A}$, $\Omega_{k,c}^{+/A}$, $P_j^{+/A}$, and $M_j^{+/A}$. These cells are governed by the transition functions, $f_{\alpha/\omega}$, given by the parameters corresponding to imatinib-affected Ph^+ cells in Table B.1. Furthermore, at each time step, every proliferating Ph^+ cell (whether imatinib-affected or not), $\Omega^{+/A}$ and Ω^+, may undergo apoptosis with probability r_{deg}. We assume that these cells die and are removed at the beginning of every time step. Other than these adjustments, the difference equations for $A_k^{+/A}$, $\Omega_{k,c}^{+/A}$, $P_j^{+/A}$, and $M_j^{+/A}$ follow (2), (4), (3), and (5). A summary of the algorithm is presented in Appendix A.

3.4. Computational complexity

One of the main advantages of our difference equation model when compared with the Roeder model is that fewer variables have to be updated at every time step.

Specifically, we note that for $k = 0, \ldots, 16$, the variable $\Omega_{k,c}$ is nonzero only when $32 \leq c \leq 32 + k$. Similarly, for $k = 17, \ldots, 48$, the variable $\Omega_{k,c}$ is nonzero only when $32 \leq c$ or $0 \leq c \leq k - 18$. Hence, at every time step of the difference equation model, we only have to update 128 variables A_k and $128 \times 49 - 48 \times 49/2 = 5096$ variables $\Omega_{k,c}$. This makes a total of 5,224 variables. Hence, when we consider the three populations, Ph^-, Ph^+, and Ph^+ cells affected by imatinib, there are about 15,600 variables to update.

By counting the number of variables updated at each time step, our model runs about six times faster than the Roeder model, when assuming a population of about 10^5 stem cells. A much larger improvement will be obtained if the number of cells is rescaled back to the more realistic order of 10^6 stem cells. The number of difference equations does not change with the size of the population, while the number of variables in the Roeder model has to increase proportionally.

An additional gain can be obtained by converting the stochastic model into a deterministic model. This can be done by replacing all binomial random variables $\text{Bin}(N, p)$ with the constant expected value Np and allowing the population sizes to be continuous. This modification eliminates all random number generations, leading to an immense increase in the computational speed. In our case, the deterministic model runs about 80 times faster than the original Roeder model (with 10^5 stem cells). We will show that, on average, the deterministic version does not deviate greatly from the original stochastic, agent-based version.

Whereas the difference equation model can be readily converted into a deterministic form, it is unclear how to rewrite the stochastic, agent-based model as a deterministic process. The main obstacle is that each agent (i.e., cell) is quantized. In particular, we cannot make a fraction of a stem cell transfer to the alternative compartment, while the remainder stays in the current compartment. Indeed, dividing agents into fractional agents is impractical, since this would require doubling the number of agents at every time step.

4. Simulations

4.1. Comparison of steady-state concentrations

Since affinities are quantized in the deterministic model, we can plot the distribution of cells as a bar graph. For comparison, we group the cells of the ABM into affinity wells, where the k^{th} well corresponds to cells with affinities between $e^{-(k+1/2)\rho}$ and $e^{-(k-1/2)\rho}$ for $k = 0, \ldots, 127$. Fig. 2 shows the bar graphs of Ph^- cells for the deterministic and ABM model at steady state.

In the deterministic simulation, the total number of Ph^- precursor cells and Ph^- mature cells at steady state is 8.5308×10^9 and 6.8246×10^{10}, respectively. In the ABM simulation, the steady state population of Ph^- precursor cells is $8.5233 \pm 0.1244 \times 10^9$ and the steady state population of Ph^- mature cells is $6.8232 \pm 0.0555 \times 10^{10}$, where populations are written in the form of mean $\pm$ standard deviation. Hence, the steady state precursor population of the ABM differs from that of the deterministic model by 0.09%, which is more than 16 times less than the standard deviation. Also, the steady state mature population of the ABM differs from that of the deterministic model by 0.02%, which is more than 39 times less than the standard deviation. Thus, in the steady state case for Ph^- cells, the deterministic model approximates the ABM very closely.

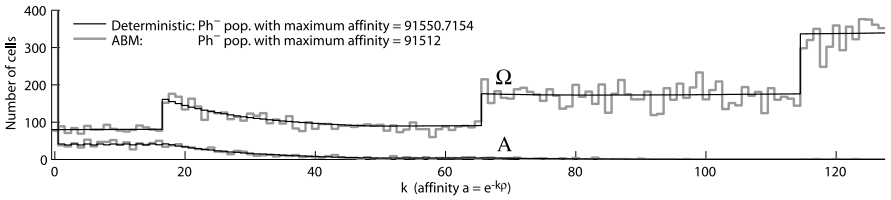


Fig. 2 Steady state profiles for the Ph^- stem cell populations. Cells are grouped into wells based on their affinity values. The upper graph corresponds to Ω cells, and the bottom graph corresponds to A cells.

4.2. Comparison of CML genesis

To simulate CML genesis, we follow the method of Roeder et al. Specifically, we introduce one proliferating Ph^+ cell (i.e., Ω cell) into the Ph^- steady state of the ABM (Roeder et al., 2006). The exact state variables for the initial Ph^+ cell are not included in Roeder et al. (2006), so we choose to introduce one Ω cell with maximum affinity ($a = 1$) and time counter set to the beginning of S-phase ($c = 32$). We add a leukemic cell to the deterministic model by setting $\Omega^+_{0,32}(T_L) = 1$ where T_L is the time of leukemia introduction. Note that since this Ph^+ cell begins at the start of S-phase, it will automatically divide into two leukemia cells after 17 hours.

The law of large numbers does not hold at CML genesis, and so the stochasticity of the ABM greatly affects the rate of initial cancer expansion. To examine the variability in CML progression, we ran 100 simulations of the ABM out to 15 years. The data is reported in Table C.1 of Appendix C. Of the 100 runs, sixty-four resulted in extinction of the cancer stem cells.

In the 36 runs that did not result in extinction, the cancer populations expanded to a BCR-ABL1 level of over 99% within 15 years of CML inception. To calculate the BCR-ABL1 transcript ratio, we applied the same formula used in Roeder et al. (2006), i.e.

$$\text{BCR-ABL1 ratio} = \frac{\#\text{ of mature } Ph^+ \text{ cells}}{(\#\text{ of mature } Ph^+ \text{ cells}) + 2\cdot(\#\text{ of mature } Ph^- \text{ cells})}.$$

We defined the barrier crossing time to be the duration between cancer inception and the moment the BCR-ABL1 ratio first reaches or surpasses 99%. Of the thirty-six simulations in which cancer survived, the average barrier crossing time was $111,300 \pm 6,000$ hours (mean $\pm$ std. dev.). For the deterministic model, the barrier crossing time was 112,911 hours. Figure 3 shows the time evolutions of the mature Ph^- and Ph^+ populations for the deterministic model and for the fastest and slowest CML expansions among the ABM simulations.

The fastest and slowest CML expansion in Fig. 3 have barrier crossing times of 99,286 hours and 127,161 hours, respectively. Of the thirty-six ABM simulations that attained the 99% BCR-ABL1 barrier, twenty-four crossed the barrier earlier than the deterministic solution, while twelve crossed later. As expected, the CML expansion rate of the deterministic solution is slightly slower than the average CML expansion rate of the ABM, because the high extinction probability in the ABM translates to a slower initial growth rate in the deterministic model.

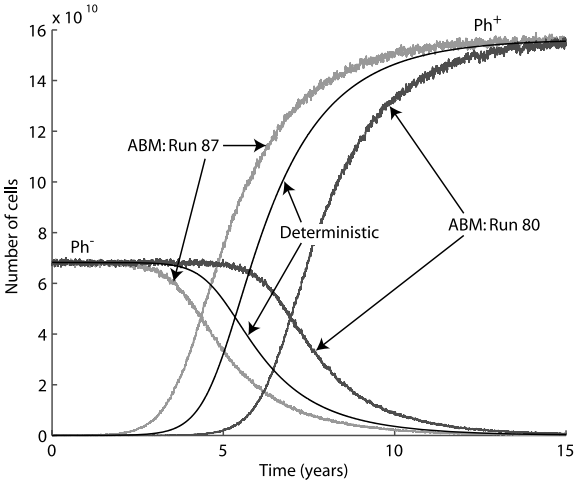


Fig. 3 Time evolution of the Ph^- and Ph^+ populations during CML genesis. The leftmost graph corresponds to run number 87 of the ABM simulations in reported in Table C.1. It crossed the 99% BCR-ABL1 barrier at 99,286 hours. The rightmost graph corresponds to run number 80 in Table C.1. It crosses the 99% BCR-ABL1 barrier at 127,161 hours. The solid black line in the center corresponds to the deterministic solution.

4.3. Comparison of imatinib treatment

To simulate imatinib treatment, we use the end data from Run 4 of the ABM simulations as the initial condition (see Table C.1.) Run 4 crosses the 99% BCR-ABL1 barrier at 112,935 hours, near the crossing time of the deterministic model. Using the method of Section 4.1, we group the cells from Run 4 into affinity wells and translate the ABM data into population variables for the deterministic model. It is impractical to fully specify the end data of Run 4, and therefore, we do not include this information in the paper. In any case, the starting condition hardly makes a difference for the results as long as both models start with the same initial condition.

Setting the imatinib-related parameters to $r_{\text{inh}} = 0.050$ and $r_{\text{deg}} = 0.033$, we simulate the time evolution of the Ph^-, Ph^+, and imatinib-affected Ph^+ populations over 1,000 days. As in Roeder et al. (2006), we conduct 20 ABM simulations and average the results to reduce stochastic variability from low cancer populations.

Figure 4 shows the deterministic solution and the average of 20 ABM simulations of imatinib treatment. The deterministic solution follows the ABM solution very closely over the entire 400 day interval, and if there is any negligible tendency of the deterministic solution to diverge from the ABM solution, it shows up near the end of the 400 days when the cancer population has dropped to such a low level that stochastic effects begin to be noticeable. Both our deterministic and ABM solutions show biphasic declines as noted in Roeder et al. (2006); Michor et al. (2005).

In Fig. 5, we show the absolute and relative errors between the deterministic solution and the average of 20 ABM simulations. We only show errors for the mature Ph^- and mature imatinib-affected Ph^+ populations, because the unaffected Ph^+ population quickly

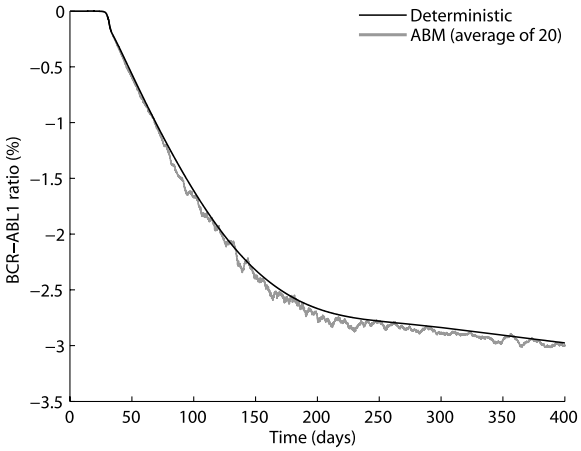


Fig. 4 BCR-ABL1 progression under imatinib treatment. The figure displays the deterministic solution and the average of 20 ABM simulations. Over time, the deterministic solution begins to slightly overestimate the BCR-ABL1 ratio of the ABM simulations.

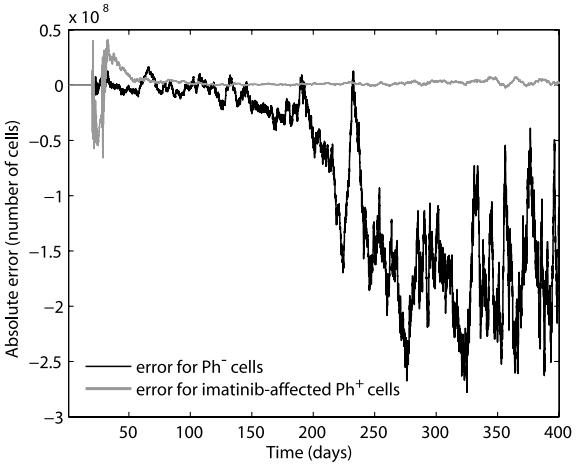
drops to zero in both models. From the figures, we notice that the deterministic model overestimates the mature imatinib-affected Ph^+ population as time progresses. This divergence occurs because both Ph^+ populations become very small under prolonged imatinib treatment. As the Ph^+ populations drop, the stochastic behavior of the ABM becomes more pronounced, leading to a slight divergence between the ABM and the deterministic model.

5. Conclusion

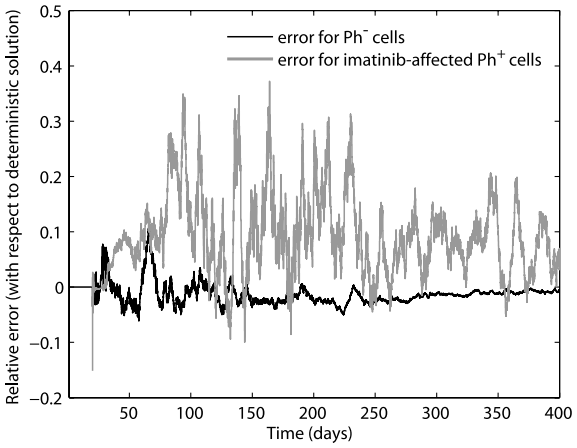
In this paper, we present a paradigm for reformulating stochastic ABMs as deterministic models. This methodology is demonstrated on a mathematical model of chronic myelogenous leukemia. The main advantage of our approach is that the deterministic model can closely reproduce the results of the agent-based model, with a significantly reduced computational cost.

The primary reason for devising ABMs is to capture the individual variability within a population. Most deterministic models lack this variability, because they treat large collections of agents as homogeneous populations that interact homogeneously. However, our method does not eliminate the intra-population diversity of the original ABM. Instead, it groups similar agents into common wells and treats each well as a collective agent. By creating wells that are small enough we can almost entirely recapture the individual variability in the original ABM. Even if we formulate a deterministic version with as many wells as the ABM has agents, the deterministic version has the advantage that it requires no random number generations.

In our particular example, we ran the ABM and deterministic model in Matlab 7.0. Of the 36 ABM simulations of CML genesis that attained the 99% BCR-ABL1 barrier, the average running time was 6 hours and 22 minutes. On the other hand, the average running



(a)



(b)

Fig. 5 Absolute and relative errors between the deterministic solution and the average of 20 ABM simulations. (a) Shown is the absolute error given by $M_{\text{det}} - M_{\text{ABM}}$, where M_{det} denotes the population for the deterministic model and M_{ABM} denotes the average population obtained from 20 ABM simulations. (b) Shown is the relative error given by $(M_{\text{det}} - M_{\text{ABM}})/M_{\text{det}}$. As time progresses, the deterministic solution overestimates the imatinib-affected Ph^+ population.

time for the deterministic version of CML genesis was 4 minutes and 32 seconds. Therefore, a single run of the deterministic version of CML genesis is over eighty times faster than a single run of the ABM. In addition, with ABMs, one usually conducts multiple simulations to obtain the average behavior. These computations were performed with the population sizes reported in Roeder et al. (2006). Furthermore, increasing the number of stem cells by a factor of 10 to reach realistic values would not change the running time of our model. However, the running time of the Roeder model will substantially increase.

The disadvantage of deterministic models is that they eliminate the effects of stochasticity, so they cannot simulate agent extinction or provide confidence intervals for statistical variables, such as the average time to attain the 99% BCR-ABL1 barrier. Stochastic effects, and hence probabilistic and agent-based approaches, become more important when the population size is small. Nonetheless, fast and accurate versions of ABMs are highly useful for parameter fitting and estimation or for large-scale statistical sampling, such as with Latin hypercube sampling (LHS).

Our model provides a fast way for approximating solutions of the Roeder model. In our derivation, we approximated $\log r / \log d$ by 2. Replacing this ratio with a closer estimate to the one used in Roeder et al. (2006) can make both models closer to each other. Alternatively, the resemblance between the two models can increase by adjusting the value of $\rho \approx \log d$.

The methodology used in this paper can be naturally extended to other problems. In particular, in many ABMs, agents are characterized by a set of state variables. The space of state variables can then be discretized, and the populations can be grouped into clusters according to state. Additionally, population variables can be assumed to be continuous. In this manner, one can derive deterministic versions of many ABMs. Since ABMs operate at fixed time steps, constant rates of change naturally give rise to discretized state spaces. Furthermore, most ABMs have traits that change at uniform rates (such as age). However, even if some rates vary, sufficiently fine discretizations can still approximate the continuous range of state values.

Deterministic versions of ABMs generally work best if the number of agents is large enough to approximate their populations with continuous variables. As a further extension, we can take time steps to the continuum limit and convert ABMs into corresponding PDE models. In a forthcoming work, we will demonstrate this idea. A related work that provides a PDE description of a similar model is Roeder (2003).

Acknowledgements

This work was supported by a Research Scholar Award from the American Cancer Society to P.P.L. The work of D.L. was supported in part by the NSF under Career Grant DMS-0133511. The work of P.S.K. was supported in part by the NSF Graduate Research Fellowship Program.

Appendix A: Algorithm for the ABM

At every time step, the ABM performs the following set of actions.

A Preliminary calculations:

1. Calculate the total populations of A and Ω cells.
2. During imatinib treatment:
   - Remove the proliferative Ph^+ cells (Ω^+ and $\Omega^{+/A}$) that undergo apoptosis.
   - Determine which unaffected proliferative Ph^+ cells (Ω^+) become imatinib-affected.

B At this stage in the algorithm, all cells fall into one of three categories: A stem cells, Ω stem cells, of differentiated cells. Depending on the category of the cell at this stage, we perform the following actions:

1. Actions performed on each A stem cell:
   - Determine whether the cell transfers to Ω. If a cell transfers, skip the remaining actions for A cells. Note that the transition function depends on whether the cell is Ph^-, Ph^+, or imatinib-affected. Calculate transition probabilities based on the total population of Ω calculated in Step A1.
   - Increase the cell's affinity by a factor of r.
2. Actions performed on each Ω stem cell:
   - Determine whether the cell transfers to A. If a cell transfers, skip the remaining actions for Ω cells. Calculate transition probabilities based on the total population of A calculated in Step A1.
   - If the cell's affinity is less than or equal to $a_{\min}$, the cell becomes a differentiated cell of age 0. If the cell differentiates, skip the remaining actions for Ω cells.
   - If a cell's affinity is greater than $a_{\min}$, decrease the cell's affinity by a factor of d.
   - Increase the counter c by 1.
   - If the counter c is greater than or equal to 49, set c to 0 and create a new cell with identical attributes and state values as the current cell.
3. Actions performed on each differentiated cell:
   - Increase the cell's age by one.
   - If the cell's age is a multiple of 24 between 24 and 480, inclusively, create a new differentiated cell with the same age as the current cell.
   - If a cell's age reaches 672, that cell dies.

Note that differentiated cells of age less than 480 are considered to be proliferating precursors, whereas differentiated cells of age greater than or equal to 480 are considered to be non-proliferating mature cells.

Appendix B: Parameter estimates

The sigmoidal transition functions are given in Roeder et al. (2006) by

$$f_{\alpha/\omega}(N_{A/\Omega}) = \frac{1}{\nu_1 + \nu_2 \exp(\nu_3 \frac{N_{A/\Omega}}{\bar{N}_{A/\Omega}})} + \nu_4, \tag{B.1}$$

where

$$\nu_1 = \left(h_1 h_3 - h_2^2\right)/(h_1 + h_3 - 2h_2),$$
$$\nu_2 = h_1 - \nu_1,$$
$$\nu_3 = \ln(h_3 - \nu_1/\nu_2),$$
$$\nu_4 = f_{\alpha/\omega}(\infty),$$

and

$$h_1 = 1/\big[f_{\alpha/\omega}(0) - f_{\alpha/\omega}(\infty)\big],$$
$$h_2 = 1/\big[f_{\alpha/\omega}(\tilde{N}_{A/\Omega}/2) - f_{\alpha/\omega}(\infty)\big],$$
$$h_3 = 1/\big[f_{\alpha/\omega}(\tilde{N}_{A/\Omega}) - f_{\alpha/\omega}(\infty)\big].$$

The transition characteristics $f_{\alpha/\omega}(.)$ are given in Table B.1.

Table B.1 Parameters from Roeder et al. (2006)

Param	Description	Ph^-	Ph^+/imatinib-affected
$a_{\min}$	Min value of affinity a	0.002	0.002
$a_{\max}$	Max value of affinity a	1.0	1.0
d	Differentiation coefficient	1.05	1.05
r	Regeneration coefficient	1.1	1.1
τ_c	Cell cycle duration	48 hours	48 hours
τ_S	Duration of S phase	8 hours	8 hours
$\tau_{G_2/M}$	Duration of G_2 and M phases	8 hours	8 hours
λ_p	Lifespan of proliferating precursor cells	20 days	20 days
λ_m	Lifespan of mature cells	8 days	8 days
$\tilde{\tau}_c$	Cell cycle of proliferating precursors	24 hours	24 hours
$f_\alpha(0)$	Transition characteristic for f_α	0.5	1.0
$f_\alpha(\tilde{N}_A/2)$	Transition characteristic for f_α	0.45	0.9
$f_\alpha(\tilde{N}_A)$	Transition characteristic for f_α	0.05	0.058
$f_\alpha(\infty)$	Transition characteristic for f_α	0.0	0.0
$\tilde{N}_A$	Scaling factor	10^5	10^5
$f_\omega(0)$	Transition characteristic for f_ω	0.5	1.0/0.0500
$f_\omega(\tilde{N}_A/2)$	Transition characteristic for f_ω	0.3	0.99/0.0499
$f_\omega(\tilde{N}_A)$	Transition characteristic for f_ω	0.1	0.98/0.0498
$f_\omega(\infty)$	Transition characteristic for f_ω	0.0	0.96/0.0496
$\tilde{N}_A$	Scaling factor	10^5	10^5

Table B.2 Imatinib-related parameters from Roeder et al. (2006). The inhibition intensity, r_{inh}, refers to the probability that a proliferative Ph^+ cell (i.e., an Ω cell) becomes imatinib-affected in a given time interval. The degradation intensity, r_{deg}, refers to the probability that an imatinib-affected, proliferative Ph^+ cell dies in a given interval

Param	Description	Estimate
r_{inh}	Inhibition intensity	0.050 (no resistance) 0.001 (partial resistance) 0.000 (no resistance)
r_{deg}	Degradation intensity	0.033 or 0.028 (no resistance) 0.0021 (partial resistance) 0.000 (complete resistance)

Appendix C: Data for 100 ABM simulations

Table C.1 Extinction and barrier crossing times for 100 ABM simulations. Barrier crossing times are listed between asterisks (*)

Run #	Extinction/ *Barrier* time (hrs)	Run #	Extinction/ *Barrier* time (hrs)	Run #	Extinction/ *Barrier* time (hrs)	Run #	Extinction/ *Barrier* time (hrs)
1	129	26	129	51	1068	76	152
2	137	27	*110593*	52	960	77	*126311*
3	129	28	1184	53	2118	78	5105
4	*112935*	29	1205	54	*109045*	79	906
5	1330	30	*116056*	55	1375	80	*127161*
6	6552	31	*110861*	56	*103524*	81	11602
7	2958	32	*112148*	57	137	82	*109026*
8	5111	33	2131	58	134	83	*109276*
9	690	34	166	59	*119755*	84	2481
10	*111734*	35	129	60	907	85	*117133*
11	741	36	7251	61	*108138*	86	412
12	423	37	*108604*	62	736	87	*99286*
13	146	38	1248	63	*104632*	88	13591
14	332	39	1925	64	129	89	*109257*
15	4189	40	2039	65	*110593*	90	3307
16	1369	41	*106971*	66	1184	91	134
17	4001	42	4657	67	1205	92	*120749*
18	*109724*	43	2121	68	*116056*	93	1358
19	719	44	*105988*	69	824	94	1770
20	4571	45	605	70	*111606*	95	605
21	*109706*	46	*104553*	71	*112950*	96	*116000*
22	570	47	*114991*	72	3830	97	320
23	*108138*	48	166	73	*115972*	98	*109026*
24	736	49	129	74	*105358*	99	1847
25	*104632*	50	935	75	160	100	536

References

Bernard, S., Belair, J., Mackey, M.C., 2003. Oscillations in cyclical neutropenia: new evidence based on mathematical modeling. J. Theor. Biol. 223(3), 283–298.

Campbell, J.D., Cook, G., Holyoake, T.L., 2001. Evolution of bone marrow transplantation—the original immunotherapy. Trends Immunol. 22(2), 88–92.

Druker, B.J., Lydon, N.B., 2000. Lessons learned from the development of an ABL tyrosine kinase inhibitor for chronic myelogenous leukemia. J. Clin. Invest. 105(1), 3–7.

Fokas, A.S., Keller, J.B., Clarkson, B.D., 1991. Mathematical model of granulocytopoiesis and chronic myelogenous leukemia. Cancer Res. 51(8), 2084–2091.

Komarova, N.L., Wodarz, D., 2005. Drug resistance in cancer: principles of emergence and prevention. Proc. Natl. Acad. Sci. USA 102(27), 9714–9719.

Michor, F., Hughes, T.P., Iwasa, Y., Branford, S., Shah, N.P., Sawyers, C.L., Nowak, M.A., 2005. Dynamics of chronic myeloid leukemia. Nature 435(7046), 1267–1270.

Moore, H., Li, N.K., 2004. A mathematical model for chronic myelogenous leukemia (CML) and T cell interaction. J. Theor. Biol. 225(4), 513–523.

Neiman, B., 2002. A mathematical model of chronic myelogenous leukemia. Master's thesis, University College, Oxford University, UK.

Roeder, I., 2003. Dynamical modeling of hematopoietic stem cell organization—design and validation of the new concept of within-tissue plasticity. PhD thesis, University of Leipzig, Germany.

Roeder, I., Horn, M., Glauche, I., Hochhaus, A., Mueller, M.C., Loeffler, M., 2006. Dynamic modeling of imatinib-treated chronic myeloid leukemia: functional insights and clinical implications. Nat. Med. 12(10), 1181–1184.

Rubinow, S.I., Lebowitz, J.L., 1976a. A mathematical model of the acute myeloblastic leukemic state in man. Biophys. J. 16(8), 897–910.

Rubinow, S.I., Lebowitz, J.L., 1976b. A mathematical model of the chemotherapeutic treatment of acute myeloblastic leukemia. Biophys. J. 16(11), 1257–1271.

Rubinow, S.I., Lebowitz, J.L., 1977. A mathematical model of the acute myeloblastic leukemic state in man. Biosystems 8(4), 265–266.

Thijsen, S.F.T., Schuurhuis, G.J., van Oostveen, J.W., Ossenkoppele, G.J., 1999. Chronic mlyeloid leukemia from basics to bedside. Leukemia 13(11), 1646–1674.

Vincent, P.C., Cronkite, E.P., Greenberg, M.L., Kirsten, C., Schiffer, L.M., Stryckmans, P.A., 1969. Leukocyte kinetics in chronic myeloid leukemia. I. DNA synthesis time in blood and marrow myelocytes. Blood 33(6), 843–850.