Optimal Control of an Age-Structured Model of HIV Infection

Hee-Dae Kwon, [*]   Jeehyun Lee, [†]   Sung-Dae Yang [‡]

Abstract

We investigate the optimal treatment strategies with an age-structured model of HIV infection. The age-structured model allows for variations in the virion production rate and the death rate of infected T cells as a function of age, which is the length of time since infection. We derive the optimal therapy protocols by formulating and analyzing an optimal control problem and establish the existence of solutions to the optimal control problem. The optimal treatment strategy is obtained by solving the corresponding optimality system numerically. We demonstrate by numerical simulations that the dynamic treatment strategy delays the time to reach the peak viral load and reduces the viral load. Moreover, we propose that optimal therapy protocols should be changed according to different viral production rates and death rates of infected T cells.

Keywords: HIV Dynamics, Age-Structured Model, Optimal Control, Gradient Method.

1 Introduction

Human immunodeficiency virus (HIV) causes acquired immune deficiency syndrome (AIDS). The virus is characterized by a severe impairment of the immune system and related opportunistic infections. The main target cell of HIV is the CD4+ T helper cell. Recent years have seen the arrival of new drugs that substantially decrease morbidity and mortality in HIV-infected patients. Despite this progress, there is still no treatment protocol that can results in clearance the HIV from patients. In addition, many complications can arise from long-term drug use. For example, drug-resistant strains of HIV can appear, resulting in the resurgence of viral loads after their long-term suppression from treatment [17, 21]. There may also be a number of harmful side effects from such drug use. Moreover, high drug costs and complicated drug regimens make effective Highly Active Anti-Retroviral Therapy (HAART) use burdensome for some patients and impossible for others.

Consequently, a number of researchers have searched for optimal treatment strategies that can decrease virus mutations, pharmaceutical side effects, and complex and expensive medication burdens. The optimal control problems of HIV infection have been examined by using different types of models and objective functionals [1, 2, 9, 14]. These authors suggested the continuous optimal treatment schedules that can be found by solving the corresponding optimality systems.

[*]Department of Mathematics, Inha University, Yonghyundong, Namgu, Incheon 402-751, Republic of Korea (`hdkwon@inha.ac.kr`).

[†]Department of Computational Science and Engineering, Yonsei University, Shinchondong, Seodaemungu, Seoul 120-749, Republic of Korea (`ezhyun@yonsei.ac.kr`).

[‡]National Institute for Mathematical Sciences, 628 Daeduk-Boulevard Yuseong-gu, Daejeon 305-340, Republic of Korea (`sdyang@nims.re.kr`).

B.M. Adams *et al.* considered two different kind of treatment as control functions. One prevents HIV from infecting cells by blocking the integration of the HIV viral code into the host cell genome and the other prevents infected cells from replication of infectious virus particles [1, 2]. L.M. Wein, *et al.* used a control theoretic approach for multi-drug therapies with models allowing mutations [27]. An approximating method was employed because of the high dimensionality of the control problem. The feedback control problems have been explored [3, 5, 7, 23]. In [7], several methods of the stable control of the HIV population were considered by using an external feedback control term that was analogous to the introduction of a therapeutic drug regimen. The optimal feedback control problems and the state estimator problems based on the state dependent Riccati equation (SDRE) approach for HIV infection were considered in [5].

More recently, substantial progress has been made in the on-off type of treatment. This treatment is also known as structured treatment interruption (STI). STI has received considerable attention because it may reduce the risk of HIV mutating to strains resistant to current medication regimens. The STI approach may also reduce the possible long-term toxicity of drugs [1, 2, 6, 15, 28]. A concise summary of clinical STI studies, including protocols and results, is presented in [4]. Some researchers have used a fixed length, prescribed interruption schedule, whereas others have used viral loads and T-cell measurements from patients to determine the interruption period [15, 22]. There is currently no general agreement on which treatment strategies or interruption schemes are optimal. One way to consider the optimal STI is to use a mathematical model for HIV infection in conjunction with control theory. The authors in [1, 2] introduced a method called the direct search approach; this method uses ideas from dynamic programming, to obtain an optimal on-off type of treatment.

To date, many mathematical models have been developed to describe the interaction of CD4+ T cells and HIV in the immune system [8, 16, 19, 24]. Some models of HIV infection have used optimal control theory, generally focusing on the system of ordinary differential equations. However, to our knowledge, optimal control theory based on age-structured models has not been considered in the identification of an optimal methodology for administering HIV treatment. The proposed age-structured model in [18] allows for variations in the death rate of infected CD4+ T cells and the production rate of viral particles. We use this model as constraint equations in the optimal control problem.

The rest of this paper proceeds as follows. Section 2 describes the age-structured HIV model suggested by P.W. Nelson *et al.* [18]. Section 3 present the formulation of the optimal control problem and the corresponding optimality system. We provide a proof of the existence of an optimal control function. We then derive an optimality system that characterizes the optimal control. Section 4 presents the numerical results of the continuous optimal therapy by solving the optimality system. We give concluding remarks in Section 5.

2 Age-structured Model

We now introduce an age-structured model of HIV infection. The model has three state variables: $T(t)$, uninfected CD4+ T cells; $T^*(a, t)$, infected CD4+ T cells structured by the age, a, of their infection; and $V(t)$, the virus particles. The system of two ordinary differential equations and one first order hyperbolic equation describing the HIV dynamics is given by

$$\frac{dT}{dt} = s - dT(t) - (1 - \epsilon(t))kV(t)T(t),$$

$$\frac{\partial T^*}{\partial t} + \frac{\partial T^*}{\partial a}\frac{da}{dt} = -\delta(a)T^*(a,t), \tag{2.1}$$

$$\frac{dV}{dt} = \int_0^\infty P(a)T^*(a,t)da - cV(t).$$

In this model, we assume that uninfected T cells are produced at a constant rate, s and die at a rate, d, per cell. The term kVT represents the infection process wherein infected cells, T^*, are produced by encounters between uninfected target cells, T, and virus particles, V, with an infection rate constant k. The death rate, $\delta(a)$, and the virion production rate, $P(a)$, of T^* are assumed to be the functions of the age of cellular infection, a and virions, V, are assumed to be cleared at a constant rate, c. We also assume $\frac{da}{dt} = 1$; that is, that the time unit for the age of infection is the same as that for clock time. As a first order hyperbolic equation is contained in the model, the boundary and initial conditions should be introduced.

The infected CD4+ T cells of age zero are created by infection; that is,

$$T^*(0,t) = (1 - \epsilon(t))kV(t)T(t).$$

We may also impose specific initial conditions for $T(0) = T_0$, $T^*(a,0) = T_0^*(a)$ and $V(0) = V_0$. The control term $\epsilon(t)$ represents the effectiveness of the reverse transcriptase inhibitors(RTI) that block new infection. Thus the infection rate, k, is reduced to $(1-\epsilon(t))k$, where $0 \leq \epsilon_{min} \leq \epsilon(t) \leq \epsilon_{max} < 1$. Here ϵ_{min} and ϵ_{max} represent minimal and maximal drug efficacy, respectively.

Remark 2.1. *With the above boundary and initial conditions and a smooth enough control function, we note that there exists a unique solution to the system (2.1) which remains bounded and non-negative for $t > 0$ (see [18, 26]).*

The mathematical model (2.1) contains several constant parameters and function parameters that must be assigned for numerical simulations. The descriptions and numerical values for the parameters are summarized in Table 1, which are principally extracted from the paper authored by P.W. Nelson *et al.* [18].

It is reasonable to expect $a \leq a_{max}$ so that the integral in (2.1) is not necessarily an infinite integral. The viral production kernel, $P(a)$, has a maximum production rate, P_{max}, because cellular resources ultimately limit how rapidly virions can be produced. The total number, N, of viral particles produced over the lifespan of an infected cell is called the burst size. We have

$$N = \int_0^\infty P(a)\sigma(a)da,$$

where

$$\sigma(a) = e^{-\int_0^a \delta(s)ds} \tag{2.2}$$

is the probability that an infected cell survives to age a.

In the absence of drug treatment, the model (2.1) has two steady state. One steady state is the trivial or noninfected steady state

$$(\bar{T}, \bar{T}^*(a), \bar{V}) = (\frac{s}{d}, 0, 0). \tag{2.3}$$

3

Table 1: The values of the parameters in the HIV model.

parameter	value	description
s	$0.13\,(cells/\mu l \cdot day)$	production(source) rate of CD4+T cells
d	$0.013\,(1/day)$	death rate of CD4+T cells
k	$0.46 \times 10^{-6}\,(ml/virion \cdot day)$	infection rate of wild-type virus
$\delta(a)$	$\cdot$	infected cell death rate
$P(a)$	$\cdot$	virion production kernel
c	$3\,(1/day)$	virus natural death rate

The other is a non-trivial or infected steady state

$$\bar{T} = \frac{c}{kN}$$

$$\bar{T}^*(a) = (s - \frac{dc}{kN})\sigma(a) \tag{2.4}$$

$$\bar{V} = \frac{s}{c}N - \frac{d}{k}.$$

For a more detailed description of the model and the steady states we refer the reader to [18].

3   An Optimal Control Problem

In this section, we formulate an optimal control problem together with the age-structured model (2.1) to derive the optimal treatment strategies. We minimize not only the virus population but also the systemic costs of drug treatments. The costs of these treatments come from both the actual treatment cost and the severity of unintended side effects of drugs. The objective functional to achieve this goal is defined as

$$J(\epsilon) = \int_0^{t_f} [RV(t) + Q\epsilon(t)^2]dt, \tag{3.1}$$

where Q and R are the weight constants to balance the quantities of virus and the control function, respectively. Our control function, $\epsilon(t)$, represents the drug (RTI) effectiveness satisfying $0 \leq \epsilon_{min} \leq \epsilon(t) \leq \epsilon_{max} < 1$. The control class is chosen to be the measurable functions defined on $[0, t_f]$, with the condition $0 \leq \epsilon_{min} \leq \epsilon(t) \leq \epsilon_{max} < 1$,

Problem 3.1. *We seek an optimal control ϵ^* such that*

$$J(\epsilon^*) = \min\{J(\epsilon) \,|\, \epsilon \text{ is Lebesque-integrable on } [0, t_f] \text{ with values in } U = [\epsilon_{min}, \epsilon_{max}]\,\}$$

subject to a system of a partial differential equation and ordinary differential equations (2.1) with boundary condition $T^*(0, t) = (1 - \epsilon(t))kV(t)T(t)$ and initial conditions $T(0) = T_0$, $T^*(a, 0) = T_0^*(a)$ and $V(0) = V_0$.

The basic framework of an optimal control problem is to prove the existence of an optimal control and then characterizes the optimal control by using the optimality system.

3.1 Existence of an optimal control

We now prove that there exists an optimal control that minimizes the objective functional (3.1) subject to the age-structured model. The existence of an optimal control can be obtained by using a result from Fleming and Rishel [10].

Theorem 3.2. *There exists an optimal control ϵ^* to Problem 3.1.*

Proof. To use the existence result in [10], we recast the system (2.1). First, note that the solution formula for $T^*(a, t)$ can be obtained by using characteristic lines as

$$T^*(a, t) = \begin{cases} B(t - a)\sigma(a) & \text{if } t \geq a, \\ T_0^*(a - t)\frac{\sigma(a)}{\sigma(a-t)} & \text{if } t < a. \end{cases} \tag{3.2}$$

where $T_0^*(a)$ is the initial value, $B(t) = (1 - \epsilon(t))kV(t)T(t)$ is the boundary value of $T^*(a, t)$ and $\sigma(a)$ is given by (2.2) [26].

Substituting (3.2) into (2.1), the system (2.1) can be rewritten as

$$\begin{aligned} \frac{dT}{dt} &= s - dT(t) - (1 - \epsilon(t))kV(t)T(t) \\ \frac{dV}{dt} &= \int_0^t P(a)\sigma(a)\Big(1 - \epsilon(t - a)\Big)kV(t - a)T(t - a)da \\ &\quad + \int_t^\infty P(a)\frac{\sigma(a)}{\sigma(a - t)}T_0^*(a - t)da - cV(t) \end{aligned} \tag{3.3}$$

Since we assume here that $P(a) \leq P_{max} \leq \infty$, we have

$$\lim_{t \to \infty} \int_t^\infty P(a)\frac{\sigma(a)}{\sigma(a - t)}T_0^*(a - t)da = 0.$$

We refer to the conditions in Theorem III. 4.1 and Corollary III. 4.1 in Fleming and Rishel [10]. We now list the requirements from the theorem as follows and verify the nontrivial ones. Let $f(t, \vec{x}, \epsilon)$ be the right hand side of (3.3) with $\vec{x} = [T, V]^T$.

1. f is of class C^1 and there exists a constant C such that

$$|f(t, 0, 0)| \leq C, \ |f_{\vec{x}}(t, \vec{x}, \epsilon)| \leq C(1 + |\epsilon|), \ |f_\epsilon(t, \vec{x}, \epsilon)| \leq C.$$

2. The admissible set $\mathcal{F}$ of all (T_0, V_0, ϵ) such that ϵ is Lebesque-integrable on the interval $[0, t_f]$ with values in U and the solution of (3.3) satisfies the initial conditions is non-empty.

3. For $0 \leq \alpha \leq 1$,

$$f\Big(t, \vec{x}, [(1 - \alpha)\epsilon + \alpha\epsilon]\Big) = (1 - \alpha)f(t, \vec{x}, \epsilon) + \alpha f(t, \vec{x}, \epsilon).$$

4. The control set U is closed, convex and compact.

5. The integrand of the objective functional (3.1) is convex on U and bounded below by $c_1|\epsilon|^\kappa - c_2$ for some $c_1 > 0$ and $\kappa > 1$.

It is clear that $f(t, \vec{x}, \epsilon)$ is of class C^1. Moreover, we have

$$|f(t, 0, 0)| = \left\| \begin{bmatrix} s \\ \int_t^\infty P(a) \dfrac{\sigma(a)}{\sigma(a-t)} T_0^*(a-t) da \end{bmatrix} \right\|$$

$$
\begin{aligned}
|f_{\vec{x}}(t, \vec{x}, \epsilon)| \quad \leq \quad & \left\| \begin{bmatrix} -d - kV & -kT \\ \int_0^t P(a)\sigma(a)kV(t-a)da & \int_0^t P(a)\sigma(a)kT(t-a)da - c \end{bmatrix} \right\| \\
+ \quad & \left\| \begin{bmatrix} kV & kT \\ -\int_0^t P(a)\sigma(a)kV(t-a)da & -\int_0^t P(a)\sigma(a)kT(t-a)da \end{bmatrix} \right\| |\epsilon|
\end{aligned}
$$

$$|f_\epsilon(t, \vec{x}, \epsilon)| = \left\| \begin{bmatrix} kVT \\ -\int_0^t P(a)\sigma(a)kV(t-a)T(t-a)da \end{bmatrix} \right\|$$

Since P, σ, V and T are bounded, there exists a constant C such that

$$|f(t, 0, 0)| \leq C, \ |f_{\vec{x}}(t, \vec{x}, \epsilon)| \leq C(1 + |\epsilon|), \ |f_\epsilon(t, \vec{x}, \epsilon)| \leq C.$$

Because of Remark 2.1, there exists a unique solution for (2.1) for a constant control, which implies Condition 2. The requirement of f being a linear function of ϵ can be replaced by Condition 3 without affecting any argument in the theorem. Conditions 4 and 5 are obvious from the definition. $\qquad \square$

3.2 The optimality system

To compute the optimal control function to Problem 3.1, we derive an optimality system based on the *discretize-then-differentiate approach*. We first discretize the age classes of T^* into an array of age classes between 0 and a_{max} to convert the age-structured model (2.1) into a series of coupled ordinary differential equations. We then differentiate the series of coupled ordinary differential equations to obtain an adjoint(costate) system. See [11] for more information and the advantages of the approach. The integral in the age-structured model can be approximated with a summation. Let $\{a_j\}_{j=0}^n$ be a partition of $[0, a_{max}]$ with $a_0 = 0$, $a_{j-1} \leq a_j$, $\triangle a_j = a_j - a_{j-1} (j = 1, 2, \cdots n)$ and $a_n = a_{max}$. For each $j = 1, 2, \cdots, n$, the discrete equations of the original constraint equations are given by

$$\frac{dT}{dt} = s - dT(t) - (1 - \epsilon(t))kV(t)T(t),$$

$$\frac{dT^*}{dt}(a_j, t) = -\frac{T^*(a_j, t) - T^*(a_{j-1}, t)}{\Delta a_j} - \delta(a_j)T^*(a_j, t),$$

(3.4)

$$\frac{dV}{dt} = \sum_{j=1}^{n} P(a_j)T^*(a_j, t)\Delta a_j - cV(t)$$

$$T^*(a_0, t) = k(1 - \epsilon(t))V(t)T(t) \text{ and } T^*(a_j, 0) = T_0^*(a_j),$$

provided $\frac{da}{dt} = 1$. Using the notations $T_j^* = T^*(a_j, t)$ and $X = [T, T_1^*, T_2^*, \cdots, T_n^*, V]^T$, we have the discrete constraint equations as follows:

$$\dot{X} = \begin{bmatrix} T \\ T_1^* \\ T_2^* \\ \cdots \\ T_n^* \\ V \end{bmatrix}' = \begin{bmatrix} s - dT - (1 - \epsilon)kVT \\ -\dfrac{T_1^* - k(1 - \epsilon)VT}{\Delta a_1} - \delta(a_1)T_1^* \\ -\dfrac{T_2^* - T_1^*}{\Delta a_2} - \delta(a_2)T_2^* \\ \cdots \\ -\dfrac{T_n^* - T_{n-1}^*}{\Delta a_n} - \delta(a_n)T_n^* \\ \sum\limits_{j=1}^{n} P(a_j)T_j^*\Delta a_j - cV \end{bmatrix}$$

(3.5)

with the initial condition $X(0) = [T(0), T_0^*(a_1), T_0^*(a_2), \cdots, T_0^*(a_n), V(0)]^T$.

We now characterize the optimal control function by using Pontryagin's Maximum Principle [10, 13, 20]. For this, we first define our Hamiltonian which is the integrand of the objective functional, coupled with the right hand sides of the discrete constraint equations (3.5) through the adjoint variables $Y = [\xi, \lambda_1, \lambda_2, \cdots, \lambda_n, \eta]^T$. Since the control is bounded, we form the Lagrangian ($\mathcal{L}$) that consists of the Hamiltonian and a penalty multiplier as follows:

$$\mathcal{L} = \left[RV(t) + Q\epsilon^2(t)\right] + \xi\left[s - dT - (1 - \epsilon)kVT\right] + \lambda_1\left[-\frac{T_1^*}{\Delta a_1} + \frac{(1 - \epsilon)kVT}{\Delta a_1} - \delta(a_1)T_1^*\right]$$

$$+ \sum_{j=2}^{n} \lambda_j\left[-\frac{T_j^* - T_{j-1}^*}{\Delta a_j} - \delta(a_j)T_j^*\right] + \eta\left[\sum_{j=1}^{n} P(a_j)T_j^*\Delta a_j - cV\right] - w_1(\epsilon - \epsilon_{min}) - w_2(\epsilon_{max} - \epsilon).$$

where $w_i(t) \geq 0$, $(i = 1, 2)$ is the penalty multiplier satisfying

$$w_1(t)(\epsilon - \epsilon_{min}) = w_2(t)(\epsilon_{max} - \epsilon) = 0 \quad \text{at} \quad \epsilon = \epsilon^*.$$

Here ϵ^* is the optimal control.

Theorem 3.3. *Given an optimal control ϵ^* and solutions to the corresponding discrete constraint equations (3.5) that minimize the objective functional (3.1), there exist adjoint variable $Y = [\xi, \lambda_1, \lambda_2, \cdots, \lambda_n, \eta]^T$ satisfying*

$$
\dot{Y} = \begin{bmatrix} \xi \\ \lambda_1 \\ \lambda_2 \\ \cdots \\ \lambda_n \\ \eta \end{bmatrix}' = - \begin{bmatrix} -d\xi - (1-\epsilon)kV\xi + \dfrac{(1-\epsilon)k}{\Delta a_1}V\lambda_1 \\ -\dfrac{\lambda_1}{\Delta a_1} - \delta(a_1)\lambda_1 + \dfrac{\lambda_2}{\Delta a_2} + P(a_1)\Delta a_1\eta \\ -\dfrac{\lambda_2}{\Delta a_2} - \delta(a_2)\lambda_2 + \dfrac{\lambda_3}{\Delta a_3} + P(a_2)\Delta a_2\eta \\ \vdots \\ -\dfrac{\lambda_n}{\Delta a_n} - \delta(a_n)\lambda_n + P(a_n)\Delta a_n\eta \\ R - (1-\epsilon)kT\xi + \dfrac{(1-\epsilon)k}{\Delta a_1}T\lambda_1 - c\eta \end{bmatrix}
\tag{3.6}
$$

with the terminal condition $Y(t_f) = [0, 0, \cdots, 0]^T$. Moreover, the optimal control function ϵ^* is given by

$$
\epsilon^* = \max\left(\epsilon_{min}, \min\left(\epsilon_{max}, \frac{1}{2Q}\left[\frac{k}{\Delta a_1}VT\lambda_1 - kVT\xi\right]\right)\right).
\tag{3.7}
$$

Proof. The results follow from an application of a version of Pontryagin's Maximum Principal [13]. Using the Lagrangian expression (3.2), we can obtain the adjoint differential equations (3.6) by

$$
\dot{Y} = -\frac{\partial \mathcal{L}}{\partial X}
$$

where $Y(t_f) = [0, 0, \cdots, 0]^T$ is the transversality conditions. To obtain the expression of the optimal control ϵ^*, we explore the necessary optimality condition $\frac{\partial \mathcal{L}}{\partial \epsilon} = 0$, that is,

$$
\frac{\partial \mathcal{L}}{\partial \epsilon} = 2Q\epsilon + kVT\xi - \frac{k}{\Delta a_1}VT\lambda_1 - w_1 + w_2 = 0.
$$

Solving for the optimal control we obtain

$$
\epsilon^* = \frac{1}{2Q}\left[\frac{k}{\Delta a_1}VT\lambda_1 - kVT\xi + w_1 - w_2\right].
$$

To determine an explicit expression for the optimal control without w_1 and w_2, we shall consider all possible values for the control, including the boundary values.

1. In the set $\{t \,|\, \epsilon_{min} < \epsilon^*(t) < \epsilon_{max}\}$, we have $w_1(t) = w_2(t) = 0$ and hence

$$
\epsilon^* = \frac{1}{2Q}\left[\frac{k}{\Delta a_1}VT\lambda_1 - kVT\xi\right].
$$

2. In the set $\{t \,|\, \epsilon^*(t) = \epsilon_{max}\}$, we have $w_1(t) = 0$. Thus the optimal control is characterized as:

$$\epsilon_{max} = \epsilon^* = \frac{1}{2Q}\left[\frac{k}{\Delta a_1}VT\lambda_1 - kVT\xi - w_2\right].$$

Since $w_2 \geq 0$,

$$\frac{1}{2Q}\left[\frac{k}{\Delta a_1}VT\lambda_1 - kVT\xi\right] \geq \epsilon_{max}$$

3. In the set $\{t \,|\, \epsilon^*(t) = \epsilon_{min}\}$, we have $w_2(t) = 0$. Thus we have

$$\epsilon_{min} = \epsilon^* = \frac{1}{2Q}\left[\frac{k}{\Delta a_1}VT\lambda_1 - kVT\xi + w_1\right].$$

This implies that

$$\frac{1}{2Q}\left[\frac{k}{\Delta a_1}VT\lambda_1 - kVT\xi\right] \leq \epsilon_{min} \quad \text{since} \quad w_1 \geq 0.$$

Combining these three cases together, we obtain the desired representation of ϵ^* (3.7). □

The optimality system consists of the state system (3.5) with initial conditions, the adjoint system (3.6) with terminal conditions and the optimality condition (3.7).

4 Numerical simulations

The optimality system was solved numerically to compare optimal treatment with various parameters. In the optimality system, the initial conditions were specified for the state equations (3.5), whereas the terminal conditions were specified for the adjoint equations (3.6). We used a gradient type iterative method to solve this two-point boundary value problem.

1. choose an initial guess of control.

2. solve the state system forward in time by using the initial guess of control.

3. solve the adjoint system backward in time.

4. update the control in each iteration by using the optimality condition.

5. continue the iterations until convergence is achieved.

For more information on such an iterative method, see the reference [11]. The optimal treatment strategy for 200 days was found by solving the optimality system through the use of this iterative scheme. In the numerical runs, we compared optimal treatment strategies by varying the viral production function, the death rate of infected cells and the initial values.

4.1 Functional form of the viral production kernel and the infected-cell death rate

A class of cells has a certain age that corresponds to age-dependent viral production and infected-cell death rates. To represent these rates, we need explicit functional forms of the viral production kernel, $P(a)$ and the death rate of the infected cells, $\delta(a)$. However, such forms are unknown; thus, we have to depend on experimental data to induce the dynamics of viral production and infected-cell death rates. According to prior researches [12, 18], we may consider one functional form of the possible kernels that capture the features of the biology, where a maximum production rate, P_{max} is provided because cellular resources ultimately limit how rapidly virions can be produced. A delayed or nondelayed exponential function can be defined as follows:

$$P(a) = \begin{cases} P_{max}(1 - \exp^{-\beta(a-d_1)}) & \text{if } a \geq d_1 \\ 0 & \text{otherwise} \end{cases} \tag{4.8}$$

where β controls how rapidly the saturation level, P_{max}, is reached. The term d_1 represent the delay in viral production; that is, it takes time d_1 after initial infection for the first viral particles to be produced (see Figure 1).
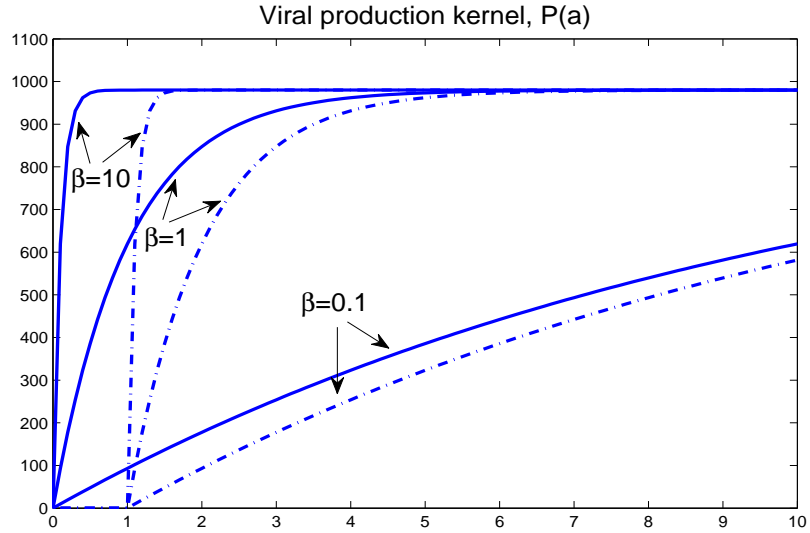

Figure 1: Virion production kernel, $P(a)$ where $P_{max} = 980$ virions/day in (4.8). Solid lines are no delay cases ($d_1=0$) and solid-dash lines are delay cases ($d_1=1$).

For the same reason given in the case of the viral production kernel, we represent the death rate of infected cells as a function of age. We assume that infected T cell death rates increase with the age of the cells. In addition, a minimal infection time is needed before epitopes are expressed, the cells become susceptible to cell-mediated killing, or for enough viral products are made such that the cells die from the infection itself. In [18], the death rate of infected cells is defined by

$$\delta(a) = \begin{cases} \delta_0 & a < d_2, \\ \delta_0 + \delta_m(1 - \exp^{-\gamma(a-d_2)}) & a \geq d_2, \end{cases} \tag{4.9}$$

where $\delta_0 + \delta_m$ is the maximal death rate, γ controls the time to saturation and d_2 is the delay between infection and the onset of cell-mediated killings. The term δ_0 represents the background death rate and $\delta(a)$ explains the exponential loss of infected cells with the age of the cells in Figure 2. In the numerical simulations we used $\delta_0 = 0.05$ and $\delta_m = 1$.
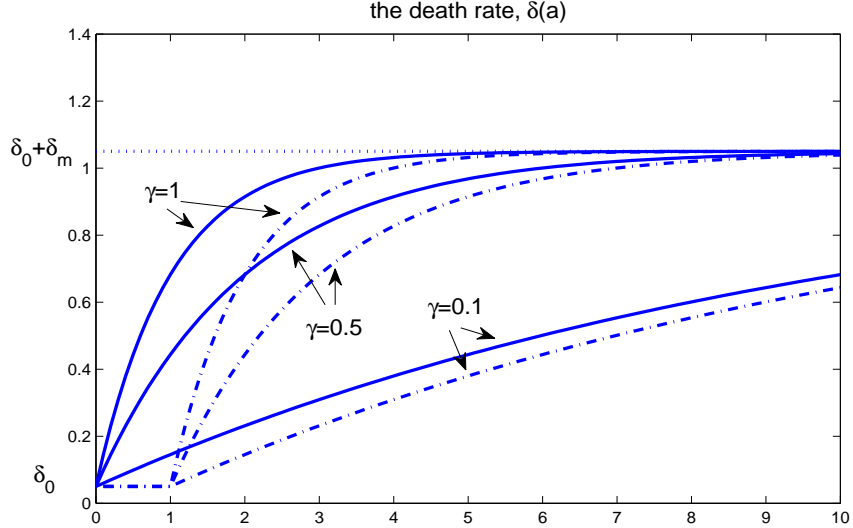

Figure 2: Death rate of infected T cells with $\delta(a)$ given by (4.9) where $\delta_0 = 0.05$ day^{-1} and $\delta_m = 1$. Solid lines are no delay cases (d_2=0) and solid-dash lines are delay cases (d_2=1).

4.2 Optimal control of primary infection

We simulated the optimal control problem by employing the dynamics of viral production rates and death rates depending on the age of infected cells. In particular, we discretized the age classes of T^* into an array of equally sized age classes between 0 and a_{max}. Here a_{max} does not necessarily correspond to the maximum age of infection of a cell; instead, it may corresponds to the age at which the production rate closely approaches the asymptotic value P_{max}. Thus a_{max} is essentially constant with age. See [18] for more information.

Drug efficacies were bounded by $\epsilon_{min} = 0$ and $\epsilon_{max} = 0.8$. Since the magnitudes of the virus population and drug treatment functions in the objective functional (3.1) were on different scales, we balanced them by choosing weight values $R = 0.1$ and $Q = 10^3$. In addition, we chose $a_{max} = 10$ and $n = 100$ such that $\triangle a_j = 0.1\,(j = 1, 2, \cdots, n)$.

We first stimulated early infection by perturbing the noninfected unstable steady state with the introduction of a small amount of virus particles and very low levels of infected T cells. We took the initial conditions $T(0) = 10/\mu l$, $V(0) = 0.02/ml$, and $T_j^*(0) = 10^{-3}$ for $j = 1, \cdots, n$ [18]. The optimal control function and the optimal solutions with $\beta = 1$, $d_1 = 0$, $P_{max} = 980$, $\gamma = 0.5$, $\delta_0 = 0.05$, $\delta_m = 1$, and $d_2 = 0$ in $P(a)$ and $\delta(a)$ are shown in Figure 3. We also present the solutions with no treatment (i.e., $\epsilon = 0$) for comparison purposes. In the first graph, we see the recommendations for the treatment schedule of reverse transcriptase inhibitors(RTI) during a given period of time. The second graph depicts the dynamics of uninfected T cells with and without

treatment. The total population of uninfected T cells corresponding to the optimal control was much higher than the one without treatment over the period. In addition, the use of the optimal therapy protocol increased the number of uninfected T cells as time progressed. By contrast, in the no treatment case, a certain level of uninfected T cells was maintained. As expected, a peak time delay and a reduction in viral load were observed.

4.3 Varying the viral production kernel and the infected-cell death rate

We investigated the effects of the dynamics of viral production rates and death rates as a function of the age of infected cells by performing numerical simulations with different parameter sets.

$$\text{Data 1}: \quad \text{Varying } P(a) \text{ by using } \beta = 10, 1, 0.1 \text{ without a delay } (d_1 = 0)$$

$$\text{Data 2}: \quad \text{Varying } P(a) \text{ by using } \beta = 10, 1, 0.1 \text{ with a delay } (d_1 = 1)$$

$$\text{Data 3}: \quad \text{Varying } \delta(a) \text{ by using } \gamma = 1, 0.5, 0.1 \text{ without a delay } (d_2 = 0)$$

$$\text{Data 4}: \quad \text{Varying } \delta(a) \text{ by using } \gamma = 1, 0.5, 0.1 \text{ with a delay } (d_2 = 1)$$

In Data 1 and 2, fixing the death rate by using $\delta_0 = 0.05$, $\delta_m = 1$, $\gamma = 1$, and $d_2 = 0$, we varied the rate of virion production with and without a delay. The treatment strategies suggested by both runs are illustrated in Figure 4. The graphs in the first column are the results of simulations using various β and no delay. As the age-related virion production rate decreased, there was a decrease in the size of viral loads. As shown in the graphs, the amount of drug treatment also decreased. The second column shows the optimal control when there was a delay between initial infection and viral production. The shape and peak of the optimal control changed depending on the size of viral loads as the delay in viral production was introduced. Note that the amount of drug treatment was reduced as β decreased, consistent with the argument used in the case of no delay.

To understand the impact of viral death rates, a fixed rate of virion production with $P_{max} = 980$, $d_1 = 0$, and $\beta = 0.1$ and various γ were used in Data 3 and 4. As detailed below and illustrated in Figure 5, the recommendations suggested by these runs were consistent with those of previous simulations varying β. The decrease in γ implied decrease in the death rates of productively infected T cells. It raised both infected cells and viral loads and thus required additional treatment efforts. The changes resulting from delays can be explained by using the same argument used for the numerical results of Data 2.

4.4 Sensitivity of optimal control to initial values

In general, the model sensitivity to initial conditions is one of the important issues in primary infection. There is no way to measure the exact amount of virus and the density of target cells at the beginning of infection. In addition, the exact infection time is unknown [18, 25]. Therefore, it is worth exploring how treatment strategies change in response to different initial conditions, which affect the time to reach the peak viral load. We considered three different initial conditions $T(0) = 1/\mu l$, $T(0) = 10/\mu l$, and $T(0) = 100/\mu l$ to observe the changes in optimal control functions. The optimal solutions and the untreated solutions were compared to depict the relation between the dynamics of the treatment schedule and the time to reach the peak viral load.

Larger sets of initial data on infected T cells led to higher and earlier peak viral loads. This explains the shift of drug treatment toward an earlier stage as shown in Figure 6. Noteworthy is

that drug treatment should be started just before the first peak viral load and that higher doses at late stages are beneficial for the effective control of viral loads in HIV patients.

5 Conclusions

This paper examined optimal control theory with an age-structured model to design treatment strategies for the HIV infection. The age-structured model allows for variations in the virion production rate and the death rate of infected T cells as a function of age. We included in the model a control function that represents the efficacy of reverse transcriptase inhibitors(RTIs). We proved the existence of optimal solutions and derived an optimality system based on the *discretize-then-differentiate* approach. The optimal therapy protocol was obtained by solving a large optimality system of equations with opposite orientations. The optimal treatment strategy maintained the virus load at a very low level and delayed the time to reach the peak viral load. The numerical results indicated that therapy protocols should be changed according to different viral production kernels and infected cell death rate. Moreover, we observed that the initiation of drug administration should depend on the initial values of infected CD4 T cells.

Acknowledgements

References

[1] B. M. Adams, H. T. Banks, M. Davidian, H. D. Kwon, H. T. Tran, S. N. Wynne, and E. S. Rosenberg, HIV dynamics: modeling, data analysis, and optimal treatment protocols, *J. Comput. Appl. Math* , 184 (2005) pp. 10-49.

[2] B. M. Adams, H. T. Banks, H. D. Kwon, and H. T. Tran, Dynamic multidrug therapies for HIV: Optimal and STI control approaches, *Math. Biosci. Engrg.*, 1 (2004) pp. 223-241.

[3] J. Alvarez-Ramirez, M. Meraz, and J. X. Velasco-Hernandez, Feedback control of the chemotherapy of HIV, *Int. J. Bifur. Chaos*, 10 (2000) pp. 2207-2219.

[4] S. H. Bajaria, G. Webb, and D. E. Kirschner, Predicting differential responses to structured treatment interruptions during HAART, *Bull, Math. Biol.*, 66 (2004) pp. 1093-1118

[5] H. T. Banks, Hee-Dae Kwon, J. A. Toivanen, H. T. Tran, A state-dependent Riccati equation-based estimator approach for HIV feedback control, *Optim. Contr. Appl. Meth.*, (2006), to appear

[6] S. Bonhoeffer, M. Rembiszewski, G.M. Ortiz and D.F. Nixon, Risks and benefits of structured antiretroviral drug therapy interruptions in HIV-1 infection, *AIDS*, 14 (2000) pp.2313-2322.

[7] M. E. Brandt and G. Chen, Feedback control of a biodynamical model of HIV-1, *IEEE Trans. on Biom. Engrg.*, 48 (2001) pp.754-759

[8] R. J. De Boer and C. A. Boucher , Anti-CD4 therapy for SIDS suggested by mathematical models, *Proc. Biol. Sci.*, 263 (1996) pp.899-905

[9] K. R. Fister, S. Lenhart, and J. S. McNally, Optimizing chemotherapy in an HIV model, *Electronic J. of Differential Equation*, 32 (1998) pp. 1-12

[10] W.H. Fleming and R.W. Rishel, *Deterministic and stochastic optimal control*, Springer-Verlag, New York, 1975.

[11] M.D. Gunzburger, *Perspectives in flow control and optimization*, SIAM, Philadelphia, 2003.

[12] D. Ho and Y. Huang, The HIV-1 vaccine race, *Cell*, 110 (2002) pp. 135-138

[13] M.I. Kamien and N.L. Schwartz, *Dynamic optimization*, North-Holland, Amsterdam, 1991.

[14] D. Kirschner, S. Lenhart, and S. Serbin, Optimal control of the chemotherapy of HIV, *J. Math. Bio.*, 35 (1997) pp. 775-792

[15] J. Lisziewicz and F. Lori, Structured treatment interruptions in HIV/AIDS therapy, *Microbes and Infection*, 4 (2002) pp.207-214.

[16] A. R. McLean and S. D. W. Frost, Zidovudine and HIV: mathematical models of within-host population dynamics, *Reviews in Medical Virology*, 5 (1995) pp.141-147.

[17] H. Moore and W. Gu, A mathematical model for treatment-resistant mutations of HIV, *Math. Biosci. Engrg.*, 2 (2005) pp.363-380.

[18] P.W. Nelson, M.A. Gilchrist, D. Coombs, J.M. Hyman and A.S. Perelson, An age-structured model of HIV infection that allows for variations in the production rate of viral particles and the death rate of productively infected cells, *Math. Biosci. Engrg.*, 1 (2004) pp.267-288.

[19] A. S. Perelson and P. W. Nelson, Mathematical anslysis of HIV-1 dynamics in vivo, *SIAM Rev.*, 41 (1999) pp.3-44.

[20] L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkrelidze and E.F. Mishchenko, *The Mathematical Theory of Optimal Processes*, Gordon and Breach, 1962.

[21] D.D. Richman, D. Havlir, J. Corbeil, et al., Nevirapine resistance mutations of human immunodeficiency virus type 1 selected during therapy, *J. Virol.*, 68 (1994), pp.1660-1666.

[22] L. Ruiz et al., Structured treatment interruption in chronically HIV-1 infected patients after long-term viral suppression, *AIDS*, 14 (2000), pp.397-403.

[23] H. Shim, S. J. Han, C. C. Chung, S. Nam, and J. H. Seo, Optimal scheduling of drug treatment for HIV infection: Continuous dose control and receding horixon control, *Int. J. control Autom. Systems*, 1 (2003) pp. 401-407.

[24] T. Shiri, W. Garira, and S. D. Musekwa, A two-strain HIV-1 mathematical model to assess the effects of chemotherapy on disease parameters, *Math. Biosci. Engrg.*, 2 (2005) pp. 811-832.

[25] M. Stafford, L. Corey, Y. Cao, E Daar, D. Ho, and A. Perelson, Modeling plasma virus condentration during primary infection, *J. Theor. Biol.*, 203 (2000) pp. 285-301.

[26] G. Webb, *Theory of Nonlinear Age-Dependent Population Dynamics*, Marcel Dekker, New York, 1985.

[27] L.M. Wein, S.A. Zenios, and M.A. Nowak, Dynamic multidrug therapies for HIV: A control theoretic approach, *J. theor. Biol.*, 185 (1997) pp.15-29.

[28] D. Wodarz and M.A. Nowak, Specific therapy regimes could lead to long-term immunological control of HIV, *Proceedings of the National Academy of Sciences*, 96 (1999) pp.14464-14469.
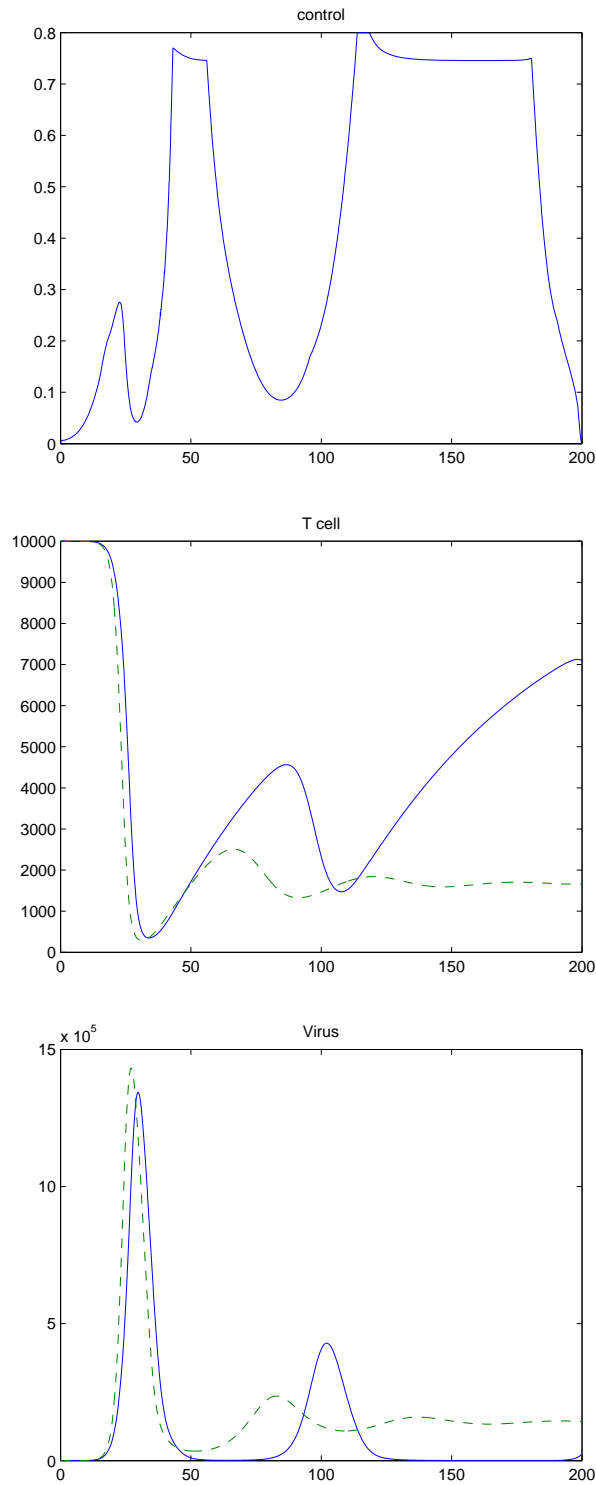
Figure 3: Plots show the optimal control function, optimal solutions (solid line), and untreated solutions (dashed line). We chose $\beta = 1$, $d_1 = 0$, $P_{max} = 980$, $\gamma = 0.5$, $\delta_0 = 0.05$, $\delta_m = 1$, and $d_2 = 0$ in $P(a)$ and $\delta(a)$.
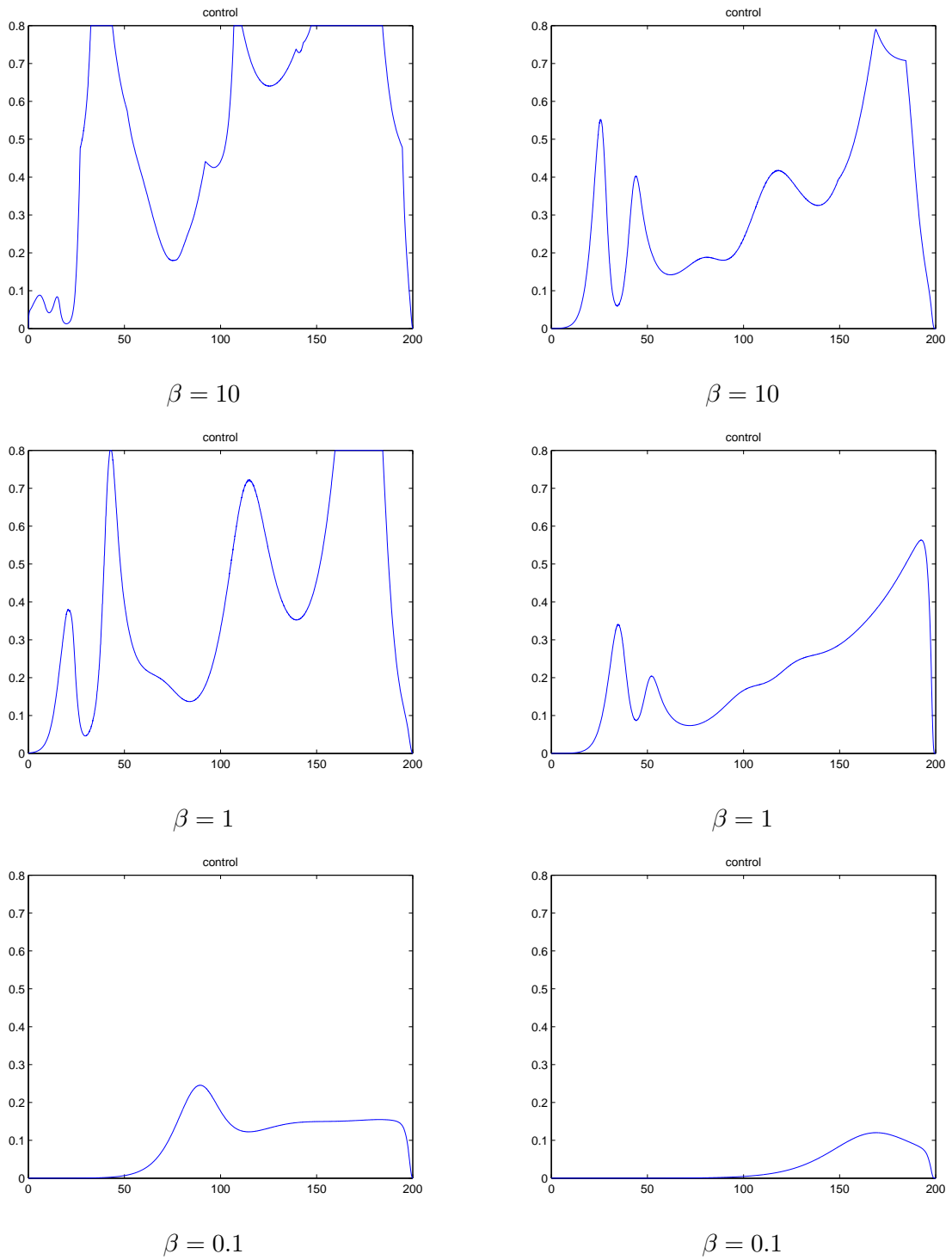
Figure 4: Optimal control functions with various virion production rate $P(a)$. The graphs in the left column are the results of the no-delay case ($d_1 = 0$); those in the right column, the delay case ($d_1 = 1$) using $\delta_0 = 0.05$, $\delta_m = 1$, $\gamma = 1$, and $d_2 = 0$.
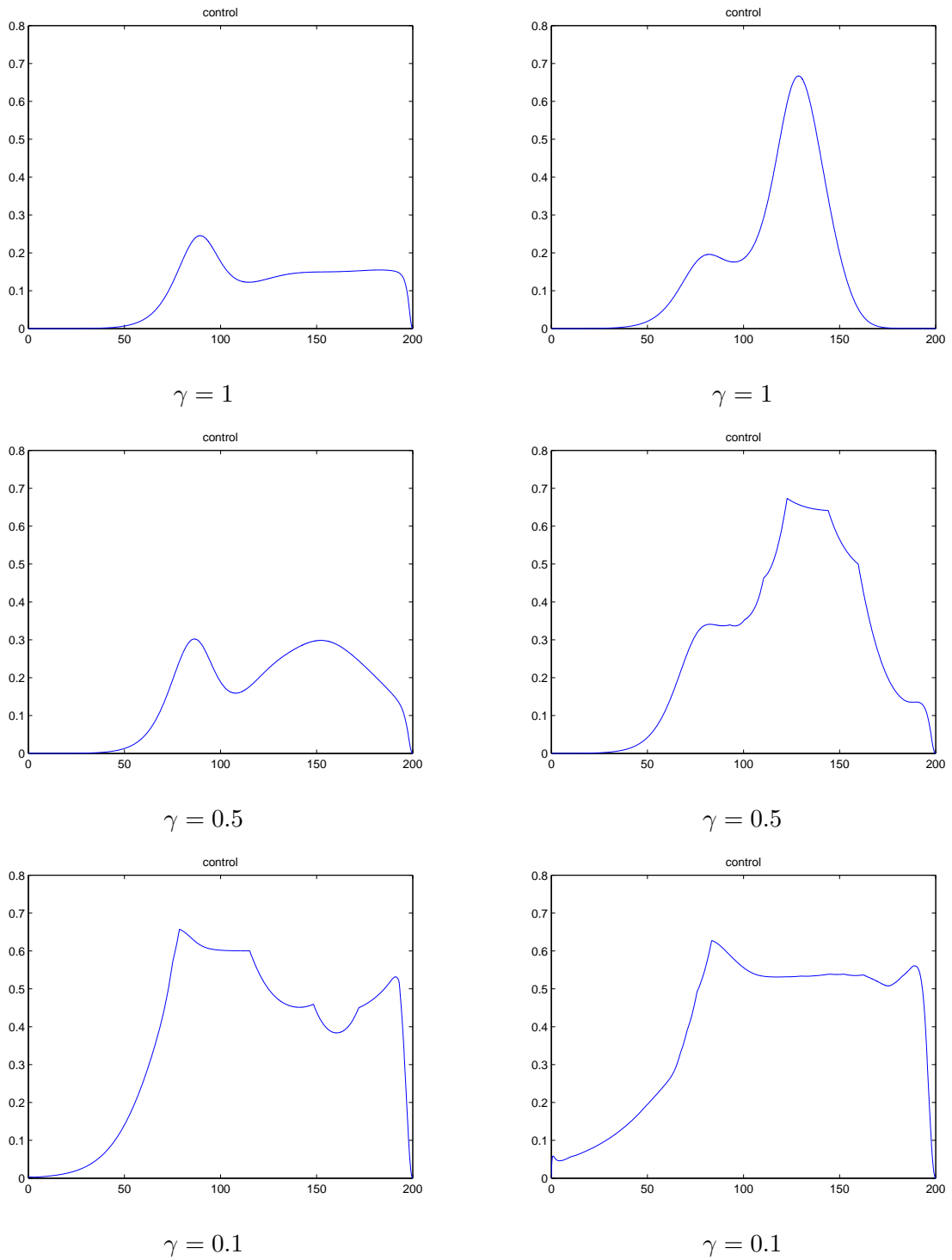
Figure 5: Optimal control functions with various death rate of infected cells $\delta(a)$. The graphs in the left column are the results of the no-delay case ($d_2 = 0$); those in the right column, the delay case ($d_2 = 1$) using $P_{max} = 980$, $d_1 = 0$, and $\beta = 0.1$.
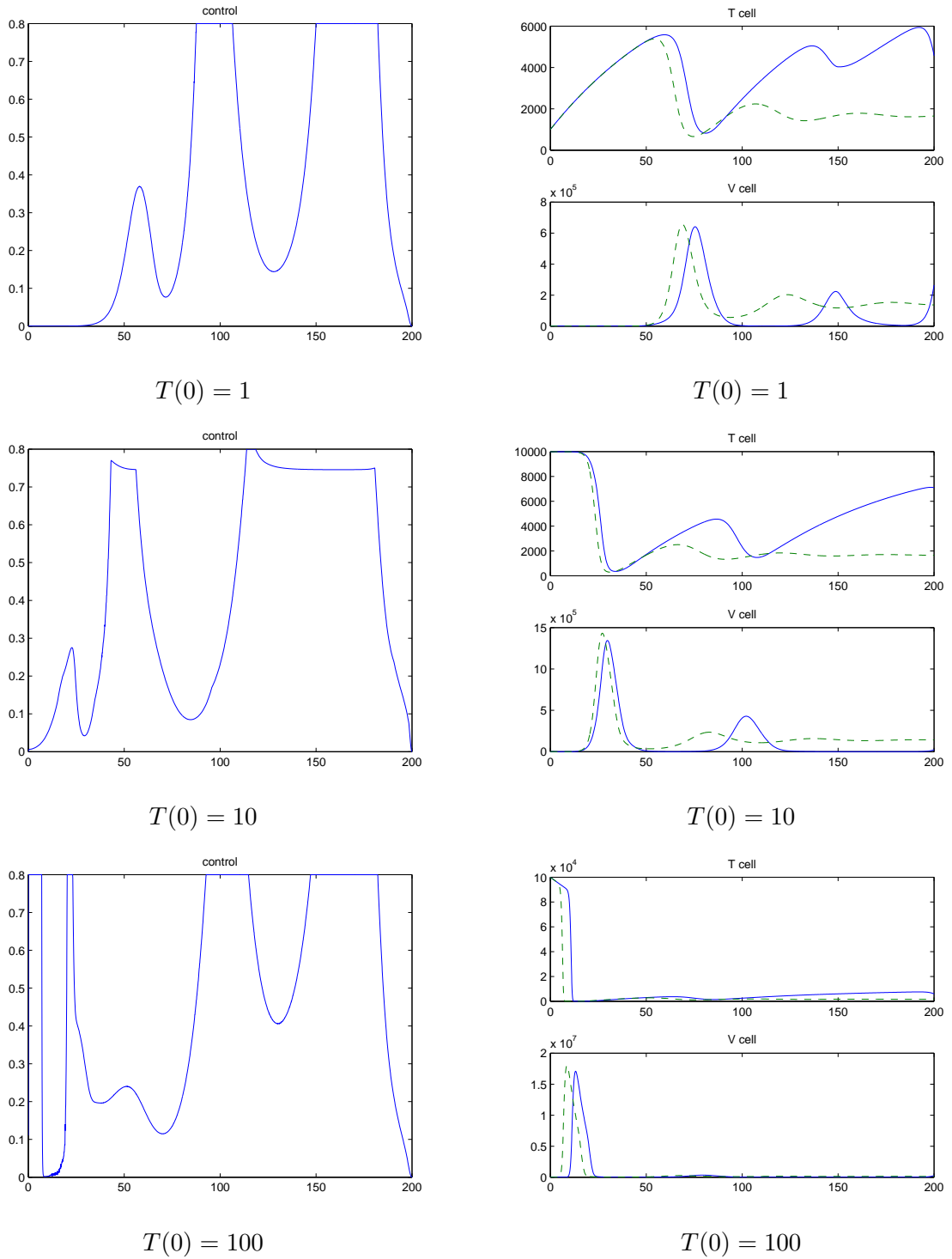
Figure 6: Optimal controls and corresponding solutions with various initial data, $T(0)(1/\mu l)$. The graphs in the left column are the results of the optimal control function, ϵ; those in the right column, uninfected CD4$^+$ T cells $T(t)$ and virus particles $V(t)$. The optimal solutions (solid line) were compared with the untreated solutions (dashed line). For this figure, we used $\beta = 1$, $d_1 = 0$, $P_{max} = 980$, $\gamma = 0.5$, $\delta_0 = 0.05$, $\delta_m = 1$, and $d_2 = 0$ in $P(a)$ and $\delta(a)$.

19