12-2014

Mathematical Modeling of Immune Responses to Hepatitis C Virus Infection

Ivan Ramirez
East Tennessee State University

Mathematical Modeling of Immune Responses to Hepatitis C Virus Infection

---

A thesis

presented to

the faculty of the Department of Mathematics and Statistics

East Tennessee State University

In partial fulfillment

of the requirements for the degree

Master of Science in Mathematical Sciences

---

by

Ivan Ramirez

December 2014

Ariel Cintron-Arias, Ph.D., Chair

Robert Gardner, Ph.D.

Jeff Knisley, Ph.D.

Keywords: hepatitis C virus, sensitivity analysis, optimal control, inverse problem

ABSTRACT

Mathematical Modeling of Immune Responses to Hepatitis C Virus Infection

by

Ivan Ramirez

An existing mathematical model of ordinary differential equations was studied to better understand the interactions between hepatitis C virus (HCV) and the immune system cells in the human body. Three possible qualitative scenarios were explored: dominant CTL response, dominant antibody response, and coexistence. Additionally, a sensitivity analysis was carried out to rank model parameters for each of these scenarios. Therapy was addressed as an optimal control problem. Numerical solutions of optimal controls were computed using a forward-backward sweep scheme for each scenario. Model parameters were estimated using ordinary least squares fitting from longitudinal data (serum HCV RNA measurements) given in reported literature.

DEDICATION

This thesis is dedicated to my parents. I am eternally grateful for all their love, support and encouragement. I also dedicate this thesis to my grandmother. She has played such a significant role throughout my life. Abuelita, you have been and always will be an inspiration for all of us.

ACKNOWLEDGMENTS

Firstly, I would like to express my gratitude to my advisor, Dr. Ariel Cintrón-Arias, for his constant guidance and support. This work would have not been possible without him. I will also like to thank Dr. Lenhart, Dr. Forde and Dr. Ciupe for their valuable input to this project. I thank my committee, Dr. Gardner and Dr. Knisley, for the suggestions they gave during the revision process, as well as for all the help they offered me through out my time as a student in the math department at ETSU. Finally, I thank the Department of Mathematics and Statistics of ETSU for a very rewarding experience as a master student.

TABLE OF CONTENTS

ABSTRACT 2

DEDICATION 4

ACKNOWLEDGMENTS 5

LIST OF TABLES 9

LIST OF FIGURES 12

1 BASIC BACKGROUND 13

1.1 Hepatitis C Virus 13

1.2 Immune Responses 14

2 MATHEMATICAL MODEL 20

2.1 Qualitative Scenarios 21

2.1.1 Dominant CTL Response 24

2.1.2 Dominant Antibody Response 26

2.1.3 Coexistence 28

3 SENSITIVITY ANALYSIS PART I 30

3.1 Introduction 30

3.2 Relative Sensitivities: Dominant CTL Response 40

3.3 Relative Sensitivities: Dominant Antibody Response 44

3.4 Relative Sensitivities: Coexistence 47

4 TREATMENT AS AN OPTIMAL CONTROL PROBLEM 50

4.1 Introduction 50

4.2 Necessary Conditions 53

4.3 Forward-Backward Sweep Method 54

4.4 Optimal Treatment Strategy for the Dominant CTL Response Scenario . . . 55
4.5 Optimal Treatment Strategy for the Dominant Antibody Response Scenario . . . 66
4.6 Optimal Treatment Strategy for the Coexistence Scenario . . . 74
5 INVERSE PROBLEM: PARAMETER ESTIMATION . . . 82
5.1 Ordinary Least Squares . . . 82
5.2 The Bootstrap Method . . . 84
5.2.1 Parameter Estimation for Patient 1 . . . 87
5.2.2 Parameter Estimation for Patient 2 . . . 91
5.2.3 Parameter Estimation for Patient 3 . . . 94
5.2.4 Parameter Estimation for Patient 4 . . . 97
6 SENSITIVITY ANALYSIS PART II . . . 100
6.1 Sensitivity Analysis for Patient 1 . . . 100
6.2 Sensitivity Analysis for Patient 2 . . . 103
6.3 Sensitivity Analysis for Patient 3 . . . 106
6.4 Sensitivity Analysis for Patient 4 . . . 109
7 CONCLUDING REMARKS . . . 112
BIBLIOGRAPHY . . . 115
VITA . . . 119

LIST OF TABLES

1 Ranking of parameters that are most influential for dominant CTL response in the transient phase 41

2 Ranking of parameters that are most influential for dominant CTL response in the long-term phase 43

3 Ranking of parameters that are most influential for dominant antibody response in the transient phase 45

4 Ranking of parameters that are most influential for dominant antibody response in the long-term phase 46

5 Ranking of parameters that are most influential for coexistence in the transient phase 48

6 Ranking of parameters that are most influential for coexistence in the long-term phase 49

7 Total production of virus in the dominant CTL response scenario . . 63

8 Total production of infected cells in the dominant CTL response scenario 63

9 Total production of virus in the dominant antibody response scenario 72

10 Total production of infected cells in the dominant antibody response scenario 72

11 Total production of virus in the coexistence scenario 80

12 Total production of infected cells in the coexistence scenario 81

13 Bootstrap estimates and standard errors for patient 1 87

14 Bootstrap estimates and standard errors for patient 2 92

15 Bootstrap estimates and standard errors for patient 3 94

16 Bootstrap estimates and standard errors for patient 4 97
17 Ranking of parameters that are most influential for the viral load in patient 1 in the transient phase 101
18 Ranking of parameters that are most influential for the viral load in patient 1 in the long-term phase 102
19 Ranking of parameters that are most influential for the viral load in patient 2 in the transient phase 104
20 Ranking of parameters that are most influential for the viral load in patient 2 in the long-term phase 104
21 Ranking of parameters that are most influential for the viral load in patient 3 in the transient phase 107
22 Ranking of parameters that are most influential for the viral load in patient 3 in the long-term phase 108
23 Ranking of parameters that are most influential for the viral load in patient 4 in the transient phase 110
24 Ranking of parameters that are most influential for the viral load in patient 4 in the long-term phase 111

LIST OF FIGURES

1 Schematic illustration of the two types of immune responses 17

2 Dynamics of the immune responses 18

3 Disease free equilibrium 22

4 No immune response endemic equilibrium 23

5 Dominant CTL response equilibrium 25

6 Dominant antibody response equilibrium 27

7 Coexistence equilibrium 29

8 Relative sensitivities for example 1 33

9 Numerical solutions for example 1 33

10 Relative sensitivities for example 2 35

11 Numerical solutions for example 2 fixing parameters 35

12 Numerical solutions for example 3 37

13 Numerical solutions for example 3 fixing parameters 38

14 Relative sensitivities for example 3 38

15 Relative sensitivities for the dominant CTL response scenario in the transient phase 40

16 Relative sensitivities for the dominant CTL response scenario in the long-term phase 42

17 Relative sensitivities for the dominant antibody response scenario in the transient phase 44

18 Relative sensitivities for the dominant antibody response scenario in the long-term phase 46

19 Relative sensitivities for the coexistence scenario in the transient phase 47
20 Relative sensitivities for the coexistence scenario in the long-term phase 48
21 Numerical solutions of the optimal control problem for the dominant CTL response scenario and best treatment strategies. 58
22 Numerical solutions of the optimal control problem for the dominant CTL response scenario and best treatment strategies after cessation of therapy. 61
23 Viral load with treatment and with no treatment for the dominant CTL response scenario with a non logarithmic scale 65
24 Numerical solutions of the optimal control problem for the dominant antibody response scenario and best treatment strategies. 69
25 Numerical solutions of the optimal control problem for the dominant CTL response scenario and best treatment strategies after cessation of therapy. 71
26 Viral load with treatment and with no treatment for the dominant antibody response scenario with a non logarithmic scale. 73
27 Numerical solutions of the optimal control problem for the coexistence scenario and best treatment strategies. 77
28 Numerical solutions of the optimal control problem for the coexistence scenario and best treatment strategies after cessation of therapy. . . . 79
29 Viral load with treatment and with no treatment for the coexistence scenario with a non logarithmic scale. 81
30 Bootstrap cloud for patient 1 88

31 Best fit solutions for patient 1 90
32 Bootstrap cloud for patient 2 92
33 Best fit solutions for patient 2 93
34 Bootstrap cloud for patient 3 95
35 Best fit solutions for patient 3 96
36 Bootstrap cloud for patient 4 98
37 Best fit solutions for patient 4 99
38 Relative sensitivities for patient 1 in the transient phase 101
39 Relative sensitivities for patient 1 in the long-term phase 102
40 Relative sensitivities for patient 2 in the transient phase 103
41 Relative sensitivities for patient 2 in the long-term phase 105
42 Relative sensitivities for patient 3 in the transient phase 106
43 Relative sensitivities for patient 3 in the long-term phase 107
44 Relative sensitivities for patient 4 in the transient phase 109
45 Relative sensitivities for patient 4 in the long-term phase 110

1 BASIC BACKGROUND

1.1 Hepatitis C Virus

Hepatitis C, which is an infectious liver disease caused by the hepatitis C virus (HCV), is commonly transmitted through direct contact with the blood of an infected person (blood transfusion, injection drug use, etc). About 150 million people worldwide are chronically infected with hepatitis C virus, and more than 350,000 people die every year from hepatitis C-related liver diseases [28]. After exposure to HCV, a strong host immune response is launched. However, in a majority of patients the response fails to eradicate the virus, leading to chronic infection [14]. As a matter of fact, HCV is a major cause of chronic liver cirrhosis and liver cancer [28]. Antiviral medicines such as combinations of pegylated-interferon$-\alpha$ and ribavirin (PEG-INF/RBV) have been used as medication for several years. Nevertheless, this treatment is far from being ideal, it is associated with significant side effects, is expensive, and only has a cure rate of 50% or less in patients infected with HVC genotype 1 [6, 12, 14]. More recently treatments are focused on direct-acting antiviral agents (DAAs) that target specific steps of the HCV life cycle. More details about these new types of treatments can be found in [24, 11, 23]. There is currently no vaccine for hepatitis C virus. Nevertheless, research in this area is outgoing.

In this thesis we want to understand the interactions between the hepatitis C virus and the immune system. For this reason it is necessary to understand the characteristics of the immune responses against foreign pathogens.

1.2 Immune Responses

The immune system is an organization of cells, organs, and molecules that protect the body from infection. The purpose of the immune system is to defend the body against foreign pathogens such as certain bacteria, viruses, and fungi. There are two fundamental categories of immune responses against invading pathogens [8, 26]: (*i*) innate or nonspecific responses, and (*ii*) specific or adaptive responses. Innate immune responses form a first barrier to protect the body against infection. Immune cells such as macrophages and natural killer cells are examples of cells that fight pathogens in a nonspecific way. In other words, these type of responses act as reactions of the body against an invader, no matter what type of pathogen they are fighting. Skin, coughing, sneezing, and fever are also examples of the nonspecific immune responses. The adaptive responses are due to cells and molecules that are able to recognize the physical structure of a pathogen [26]. They can perceive proteins from which the pathogen is built. Once these cells are activated they start to divide and expand in number. Probably the most important component of the specific immune system are the lymphocytes. These type of cells are white blood cells that arise continuously in the bone marrow (flexible tissue in the interior of the bones). One microlitre of human blood contains about 2500 lymphocytes and in total there are about 10^{12} lymphocytes in an adult human [20]. Lymphocytes can be grouped into two major branches: B-lymphocytes, referred to as B cells, and T-lymphocytes, referred to as T cells [8, 20, 21, 26].

B cells carry antibody molecules on their surface membrane that serve as receptors that can specifically recognize the pathogen [8, 20, 26]. During B cell development,

the immune system creates billions of different antibodies with a limited number of genes by rearranging DNA segments. Mutation can also increase genetic variation in antibodies. Basically, for any pathogen that enters the body, there is a specific antibody molecule that can recognize it [20].

When HCV enters the human body, most B-cells do not have the specific receptors to recognize this foreign intruder. However, some B-cells will be able to bind some viral proteins. There is a specific antibody molecule that can bind to the pathogen in a lock-and-key system manner. Once a B-cell is bound to the virus particle, a process known as endocytosis starts. In this process, the virus is broken into pieces and part of it is presented in association with the so-called major histocompability complex type II (MHC II), which is a molecule on the surface of the B-cell [8, 20, 26].

T-helper cells or CD4 cells are another type of T-cells that play an important role in the activation of B-cells. If CD4 cells can bind the presented part of the virus on the surface of the B-cell, a release of molecules will occur and the activation of the B-cell will be complete [20].

Once B-cells are activated they start to divide into memory and effector cells. This production of new B-cells are replications with the same specific receptor that can recognize HCV. The memory B-cells will take no action but stay in the system in case this pathogen enters the host in the future [20, 26]. The effector B-cells, also called plasma cells, are responsible in the release of new antibodies. These antibodies then bind to the virus particles and tag them as foreign pathogens for elimination by a macrophage [20].

T cells can be grouped into two types of cells: T helper cells or CD4 and cytotoxic T lymphocytes (CTL) or CD8. The acronyms CD4 and CD8 refer to some proteins on the surface of these cells [20, 26]. As we mentioned above, T helper cells play an important role in the activation of the B-cells and therefore in the release of antibodies. The CTL have the ability to take lytic action against infected cells by killing them. These cells can recognize and eliminate infected cells. Once a cell has been infected it becomes a factory of production of new virus. The viral proteins inside the infected cell are presented on the surface in combination with the major histocompability complex type I (MHC I), which is presented essentially in every cell of the human body. If a specific CTL with the correct T-cell receptor can recognize this presented portions on the surface of the infected cells then it will bind to the cell. The CD8 cell will become activated and start to produce chemicals that kill the target cell [20]. After activation, the CTL will proliferate into effector and memory cells, much like in the B-cell activation process. The effector cells will have the same functions as the parent cell: to destroy infected cells while the memory cells will stay in the host in case of a future infection. A schematic illustration of the types of immune responses can be found in Figure 1. A summarized explanation of the dynamics of the immune system is shown in Figure 2.

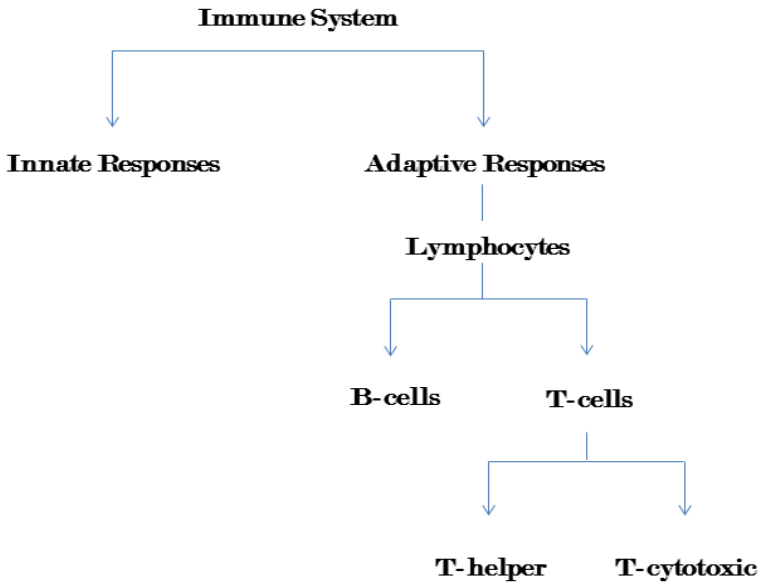

Figure 1: Schematic illustration of the two types of immune responses: the innate and the adaptive responses. The most important compartment of the adaptive responses are the lymphocytes that can be grouped into B-cells and T-Cells.

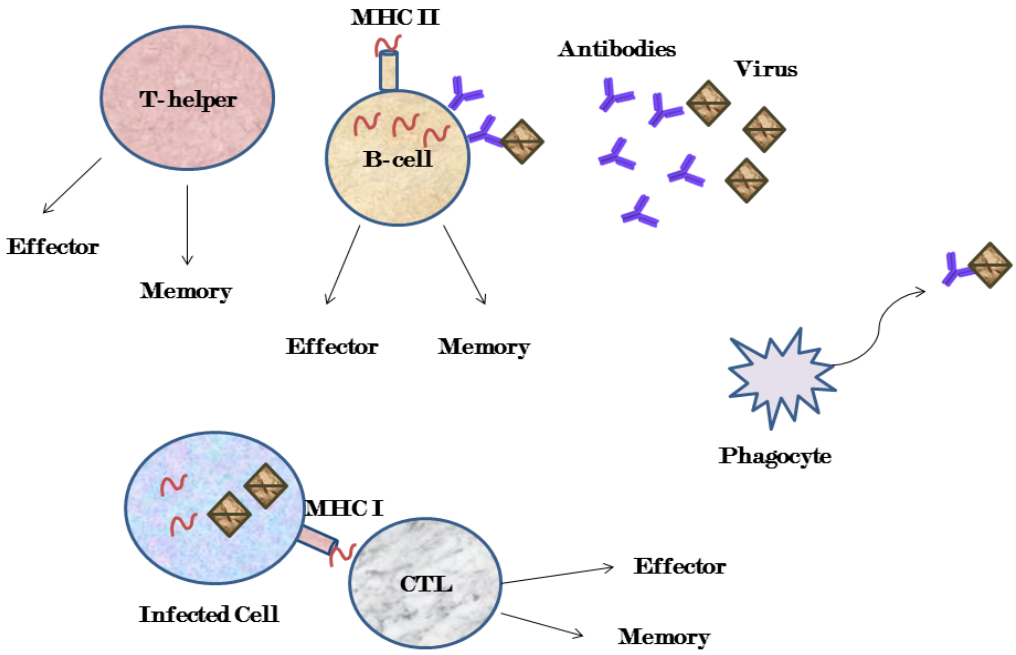

Figure 2: Schematic illustration of the adaptive and innate responses. The adaptive responses can be grouped into B-cells and T-cells. The B-cells secrete antibodies that neutralize free virus particles. The cytotoxic T lymphocytes (CTLs or CD8) eliminate infected cells. The CD4 T helper cells ensure that CTLs and B-cells develop efficiently. Phagocytes are cells in charge of eliminating virus particles that have been bound by antibodies.

The interactions between immune responses and pathogens can be considered a predator-prey system. The virus is the prey and the immune cells like the B-cells and T-cells are the predators. In the absence of the pathogen, the immune cells die. The predator species that is most efficient at capturing the prey can reduce the food resource to levels which are too low for other predator species to survive. This can result in competitive exclusion where only one predator species remains [27]. If antibodies are more efficient at fighting a virus, they can reduce virus load to levels that are too low to stimulate the CTLs. This outcome is going to be called a dominant antibody response. Likewise, if CTLs are more efficient in killing the pathogen, they can reduce the virus load to levels that are too low to stimulate the antibodies. This outcome is called a dominant CTL response. In addition to these exclusion outcomes, there is also a coexistence outcome. Even if the CTLs are very efficient in reducing the number of infected cells, the number of viruses could be sufficiently large to stimulate the antibody response. More details about immunity can be found in [20, 26].

2 MATHEMATICAL MODEL

In literature, several models have been used to describe HCV dynamics [19, 6, 25]. The following model, proposed by Wodarz [25], describes the competition dynamics between CTL and antibodies as responses to a viral infection. In particular, this model can be used to study the interaction between HCV and immune responses in a host. We denote the number of uninfected hepatocytes (cells of the liver), at time t, by $T(t)$, infected hepatocytes by $I(t)$, the viral load by $V(t)$, the number of CTLs by $Z(t)$, and the number of antibodies by $W(t)$. The system of differential equations that describes the change of these populations over time is:

$$\frac{dT}{dt} = s - dT - \beta VT \tag{1}$$

$$\frac{dI}{dt} = \beta VT - aI - pIZ \tag{2}$$

$$\frac{dV}{dt} = kI - \mu V - qVW \tag{3}$$

$$\frac{dZ}{dt} = cIZ - bZ \tag{4}$$

$$\frac{dW}{dt} = gVW - hW \tag{5}$$

This model assumes uninfected host cells (healthy hepatocytes) are produced at a rate s, undergo natural decay at a rate dT and become infected by the interaction with virus at a rate βTV. Infected cells die at a rate aI and are killed by the CTL response at a rate pIZ. Free virus is produced by infected cells at a rate kI, decays at a rate μV and is neutralized by antibodies at a rate qVW. CTL expand in response to viral antigen derived from infected cells at a rate cIZ and decay in the absence of antigenic stimulation at a rate bZ. Antibodies develop in response to free virus at a rate gVW and decay at a rate hW.

2.1 Qualitative Scenarios

Definition 2.1 *The basic reproductive number of the virus R_0 is defined as the average number of newly infected cells produced by a single infected cell at the beginning of the infection.*

It is a well known result that $R_0 = 1$ is a dynamic threshold [20, 26]. In other words, R_0 can be used as a metric to determine whether or not an infectious disease will spread through a population. In fact, if $R_0 > 1$ the infection becomes established and if $R_0 < 1$ the virus will not spread [26, 20].

The system given by equations (1)-(5) has the basic reproductive number:

$$R_0 = \frac{s\beta k}{da\mu}$$

A stability analysis of the model, as well as properties of the solutions such as positivity, boundedness, non periodicity, and limiting behavior can be found in [29]. The system of equations (1)-(5) supports five equilibrium points or steady states. The stability of these points is determined by conditions on the parameters of the system [29]. These equilibria are considered as qualitative scenarios that can be observed in a typical host.

The first equilibrium point, also called disease free equilibrium, is globally asymptotically stable if $R_0 < 1$ [29]. This scenario represents the absence of virus and immune responses. The equilibrium point is given by the expression:

$$\left(\overline{T}^{(1)}, \overline{I}^{(1)}, \overline{V}^{(1)}, \overline{Z}^{(1)}, \overline{W}^{(1)}\right) = \left(\frac{s}{d}, 0, 0, 0, 0\right)$$

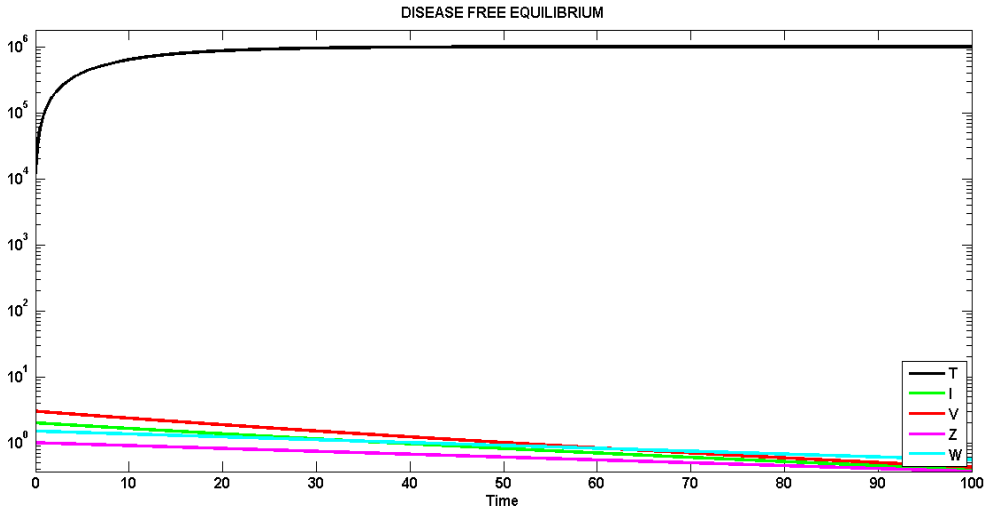

Figure 3: A disease free equilibrium point is observed if $R_0 \leq 1$. Parameter values were chosen as follows: $s = 1.0 \times 10^5, d = 1.0 \times 10^{-1}, \beta = 0, a = 1.0 \times 10^{-2}, p = 1.0 \times 10^{-2}, k = 0, \mu = 1.0 \times 10^{-2}, q = 1.0 \times 10^{-2}, g = 0, c = 0, h = 1.0 \times 10^{-2}, b = 1.0 \times 10^{-2}$. Initial conditions: $T(0) = 1.0 \times 10^4, I(0) = 2.0 \times 10^0, V(0) = 3.0 \times 10^0, Z(0) = 1.0 \times 10^0, W(0) = 1.5 \times 10^0$. According to the values of the parameters, $R_0 = 0$.

In Figure 3, parameters have been chosen properly to give an illustration of this first scenario. The proliferation of uninfected cells reaches its equilibrium value at $\frac{s}{d}$ while the other four populations converge to zero. In this numerical illustration $T(t)$ converges to $\frac{s}{d} = 10^6$.

Under the assumption that $R_0 > 1$ and that there is not an immune response, the system converges to the equilibrium point given by:

$$\left(\overline{T}^{(2)}, \overline{I}^{(2)}, \overline{V}^{(2)}, \overline{Z}^{(2)}, \overline{W}^{(2)}\right) = \left(\frac{s}{d}\frac{1}{R_0}, \frac{\mu}{k}\frac{d}{\beta}\left(R_0 - 1\right), \frac{d}{\beta}\left(R_0 - 1\right), 0, 0\right)$$

This equilibrium is globally asymptotically stable [29, 26] if

$$\frac{gkb}{\mu c} < h \qquad \text{and} \qquad \frac{c\beta hs}{a(dg + \beta h)} < b$$

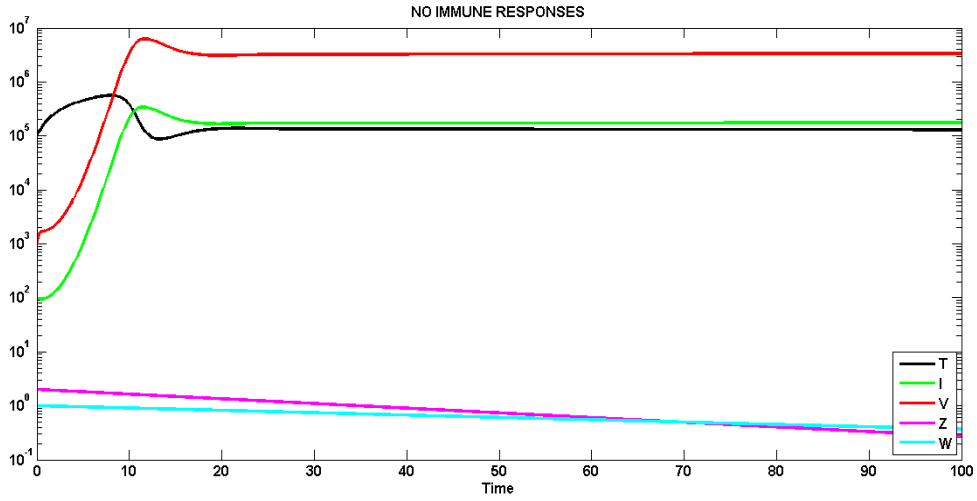

Figure 4: An endemic equilibrium where there is not an immune response is observed. Parameter values were chosen as follows: $s = 1.0 \times 10^5; d = 1.0 \times 10^{-1}, \beta = 2.0 \times 10^{-7}, a = 5.0 \times 10^{-1}, p = 6.4 \times 10^{-4}, k = 1.0 \times 10^2, \mu = 5.0 \times 10^0, q = 5 \times 10^{-1}, g = 0, c = 0, h = 1 \times 10^{-2}, b = 2 \times 10^{-2}$. Initial conditions: $T(0) = 1.0 \times 10^5, I(0) = 1.0 \times 10^2, V(0) = 1.0 \times 10^3, Z(0) = 2.0 \times 10^0, W(0) = 1.0 \times 10^0$. According to the values of the parameters, $R_0 > 1$.

In Figure 4, parameters have been chosen properly to give an illustration of this scenario. The proliferation of uninfected cells, the number of viruses, and infected cells reach their equilibrium values while the immune responses converge to zero.

The first two scenarios are mentioned for completeness. However, the focus in this thesis is on the three scenarios in which $R_0 > 1$ and there are immune responses. These three scenarios are the following: dominant CTL response, dominant antibody response, and coexistence (both CTL and antibody response are activated).

2.1.1 Dominant CTL Response

The third equilibrium observed is called dominant CTL response. A strong proliferation of CTLs combined with a weak antibody production leads to an eventual extinction of this response.

This scenario is described by the following equilibrium:

$$\left(\overline{T}^{(3)}, \overline{I}^{(3)}, \overline{V}^{(3)}, \overline{Z}^{(3)}, \overline{W}^{(3)}\right) = \left(\frac{s\mu c}{d\mu c + \beta kb}, \frac{b}{c}, \frac{k\overline{I}^{(3)}}{\mu}, \frac{\beta\overline{T}^{(3)}\overline{V}^{(3)} - a\overline{I}^{(3)}}{p\overline{I}^{(3)}}, 0\right)$$

Notice that here we have that $\overline{T}^{(3)} = \dfrac{s\mu c}{d\mu c + \beta kb}$, $\overline{I}^{(3)} = \dfrac{b}{c}$, and $\overline{V}^{(3)} = \dfrac{kb}{\mu c}$.
This equilibrium is globally asymptotically stable [29, 26] if

$$\frac{gkb}{\mu c} < h \qquad \text{and} \qquad \frac{c\beta hs}{a(dg + \beta h)} > b$$

In Figure 5, values of parameters were chosen to illustrate this scenario where the infection has become established. Large values of the viral load and infected cells were chosen for the initial conditions to indicate the case of a chronic infection. See caption of Figure 5 for details on the values of parameters and initial conditions. The

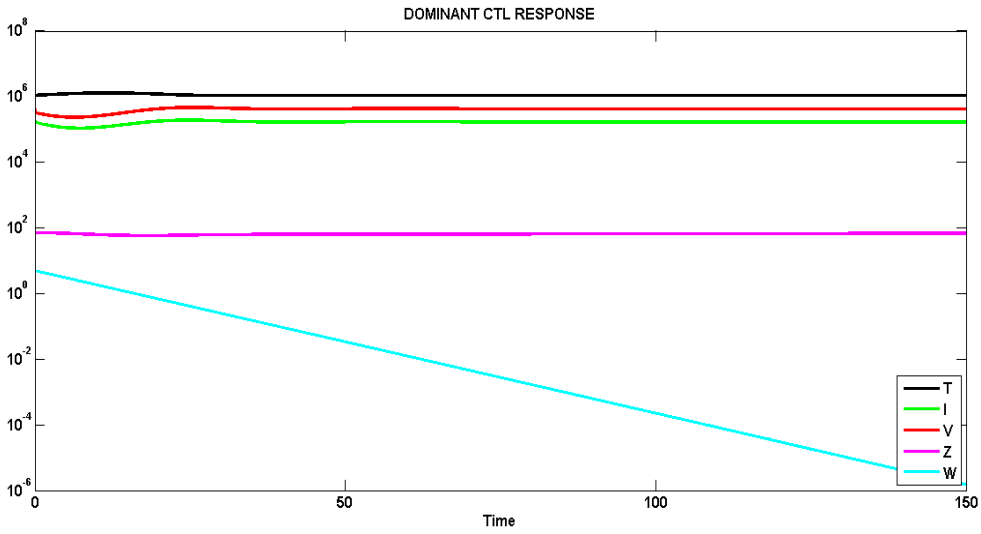

Figure 5: Numerical solutions for equations (1)-(5) for the dominant CTL response scenario with $R_0 > 1$ where the third equilibrium is stable. An endemic equilibrium is observed. Parameter values were chosen as follows: $s = 2.0 \times 10^5, d = 1.0 \times 10^{-1}, \beta = 2.0 \times 10^{-7}, a = 5.0 \times 10^{-1}, p = 6.4 \times 10^{-4}, k = 2.0 \times 10^1, \mu = 8.0 \times 10^0, q = 5.0 \times 10^{-1}, g = 1.0 \times 10^{-11}, c = 3.0 \times 10^{-7}, h = 1.0 \times 10^{-1}, b = 5.0 \times 10^{-2}$. Initial conditions: $T(0) = 1.0 \times 10^6, I(0) = 1.6 \times 10^5, V(0) = 4.1 \times 10^5, Z(0) = 7.1 \times 10^1, W(0) = 5.0 \times 10^0$.

virus persists in the presence of a dominant CTL response leading system going to an endemic equilibrium.

Because parameter c is the proliferation rate of CTLs, for larger values of this parameter, damped oscillations will be observed at the beginning of the simulation. Infection will not become totally extinct, but a considerable reduction of the viral load and infected cells will be observed.

2.1.2 Dominant Antibody Response

The fourth equilibrium that is observed is called dominant antibody response. The proliferation rate of antibody is assumed to be much stronger than the natural production rate of CTLs. For this reason, the antibody response develops and the CTL response is unsuccessful.

This scenario is described by the following equilibrium:

$$\left(\overline{T}^{(4)}, \overline{I}^{(4)}, \overline{V}^{(4)}, \overline{Z}^{(4)}, \overline{W}^{(4)}\right) = \left(\frac{sg}{dg+\beta h}, \frac{\beta hs}{a(dg+\beta h)}, \frac{h}{g}, 0, \frac{k\overline{I}^{(4)} - \mu\overline{V}^{(4)}}{q\overline{V}^{(4)}}\right)$$

Notice that here we have that $\overline{T}^{(4)} = \dfrac{sg}{dg+\beta h}$, $\overline{I}^{(4)} = \dfrac{\beta hs}{a(dg+\beta h)}$, and $\overline{V}^{(4)} = \dfrac{h}{g}$. This equilibrium is globally asymptotically stable [29, 26] if

$$\frac{gkb}{\mu c} > h \qquad \text{and} \qquad \frac{c\beta hs}{a(dg+\beta h)} < b$$

In Figure 6, parameters were chosen to give an illustration of this scenario. Large values of the viral load and infected cells were chosen for the initial conditions to indicate the case of a chronic infection. The virus persists in the presence of a dominant antibody response leading to an endemic equilibrium. Damped oscillations are observed at the beginning of the simulation indicating that the antibody response is unsuccessful in reducing the viral load.

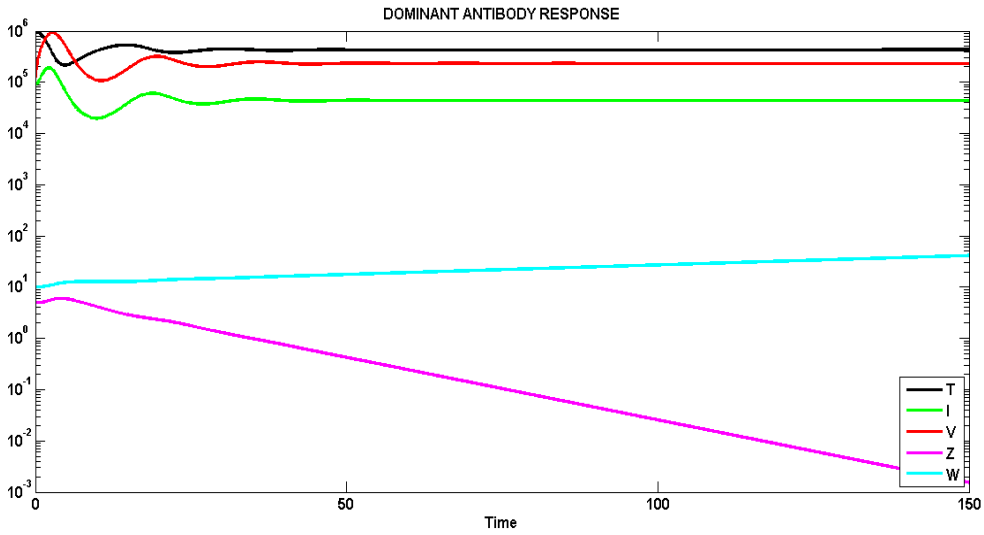

Figure 6: Numerical solutions for equations (1)-(5) for the dominant antibody response scenario with $R_0 > 1$ where the fourth equilibrium is stable. An endemic equilibrium is observed. Parameter values were chosen as follows: $s = 1.0 \times 10^5, d = 5.0 \times 10^{-2}, \beta = 8.0 \times 10^{-7}, a = 1.8, p = 6.4 \times 10^{-4}, k = 8.0 \times 10^0, \mu = 1.5 \times 10^0, q = 5.0 \times 10^{-4}, g = 8.0 \times 10^{-8}, c = 6.0 \times 10^{-6}, h = 1.0 \times 10^{-2}, b = 1.0 \times 10^{-1}$. Initial conditions: $T(0) = 1.0 \times 10^6, I(0) = 1.0 \times 10^5, V(0) = 1.0 \times 10^5, Z(0) = 5.0 \times 10^0, W(0) = 1.0 \times 10^1$.

2.1.3 Coexistence

The fifth and last equilibrium that is observed is called coexistence because both CTL and antibody responses are equally established. These two immune responses compete against each other with the same objective: eradication of the infection. This scenario is described by the following equilibrium:

$$\left(\overline{T}^{(5)}, \overline{I}^{(5)}, \overline{V}^{(5)}, \overline{Z}^{(5)}, \overline{W}^{(5)}\right) = \left(\frac{sg}{dg+\beta h}, \frac{b}{c}, \frac{h}{g}, \frac{\beta \overline{T}^{(5)} \overline{V}^{(5)} - a\overline{I}^{(5)}}{p\overline{I}^{(5)}}, \frac{k\overline{I}^{(5)} - \mu \overline{V}^{(5)}}{q\overline{V}^{(5)}}\right)$$

Notice that here we have that $\overline{T}^{(5)} = \dfrac{sg}{dg+\beta h}$, $\overline{I}^{(5)} = \dfrac{b}{c}$, and $\overline{V}^{(5)} = \dfrac{h}{g}$. This equilibrium is globally asymptotically stable [29, 26] if

$$\frac{gkb}{\mu c} > h \qquad \text{and} \qquad \frac{c\beta hs}{a(dg+\beta h)} > b$$

In Figure 7, parameters are chosen so as illustrate this scenario were the infection has become established. Large values of the viral load and infected cells were chosen for the initial conditions to indicate the case of a chronic infection. Even though the viral load was brought to lower levels in comparison to the two previous scenarios, infection still persists in the presence of antibody and CTL responses. In other words, $V(t)$ and $I(t)$ do not tend to zero as t grows.

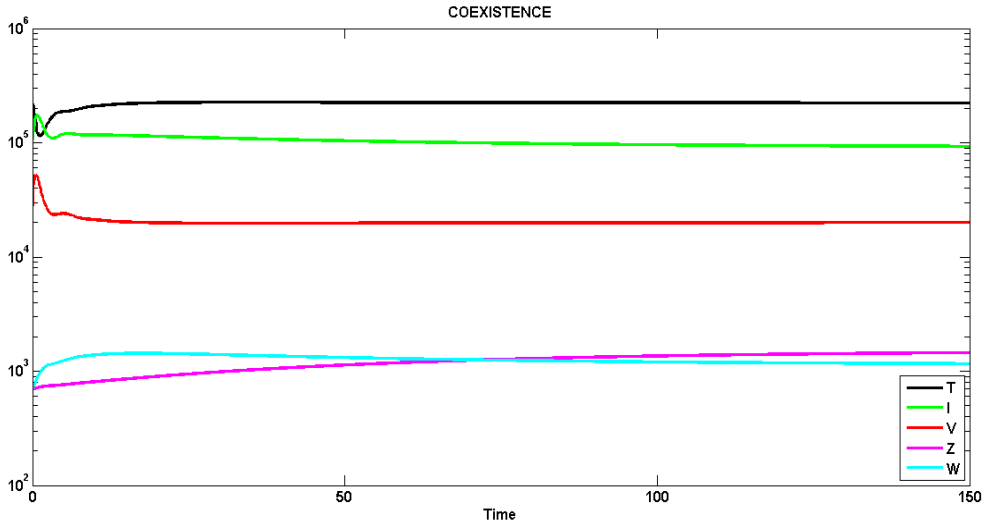

Figure 7: Numerical solutions for equations (1)-(5) for the coexistence scenario with $R_0 > 1$ where fifth equilibrium is stable. An endemic equilibrium is observed. Parameter values were chosen as follows: $s = 2.0 \times 10^5, d = 1.0 \times 10^{-1}, \beta = 4.0 \times 10^{-5}, a = 9.9 \times 10^{-1}, p = 6.4 \times 10^{-4}, k = 5.0 \times 10^2, \mu = 2.9 \times 10^0, q = 2.0 \times 10^0, g = 1.0 \times 10^{-5}, c = 4.4 \times 10^{-7}, h = 2.0 \times 10^{-1}, b = 4.0 \times 10^{-2}$. With initial conditions: $T(0) = 2.2 \times 10^5, I(0) = 1.0 \times 10^5, V(0) = 1.9 \times 10^4, Z(0) = 7.0 \times 10^2, W(0) = 7.0 \times 10^2$.

3 SENSITIVITY ANALYSIS PART I

3.1 Introduction

Consider the system of differential equations given by

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{f}(t, \boldsymbol{x}(t, \boldsymbol{\theta}), \boldsymbol{\theta}) \tag{6}$$

where $\boldsymbol{x}(t, \boldsymbol{\theta}) \in \mathbb{R}^n$ denotes the state variable vector at time t and $\boldsymbol{\theta} \in \mathbb{R}^p$ is the parameter vector. It is of our interest to determine which of these parameters are more important in the behavior of a particular state variable.

Definition 3.1 *We define* $\phi(t) = \dfrac{\partial \boldsymbol{x}(t)}{\partial \boldsymbol{\theta}}$ *to be the sensitivity matrix of* $\boldsymbol{x}(t)$*. This matrix measures the sensitivity of the solution of (6) with respect to changes in the parameters* θ_k *for* $k = 1, \cdots, p$*.*

By the chain rule for several variables and by assuming smoothness we have that

$$\frac{d}{dt}\phi(t) = \frac{d}{dt}\left(\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{\theta}}\right) = \frac{\partial}{\partial \boldsymbol{\theta}}\left(\frac{d\boldsymbol{x}}{dt}\right) = \frac{\partial}{\partial \boldsymbol{\theta}}\left(\boldsymbol{f}(t, \boldsymbol{x}(t), \boldsymbol{\theta})\right) = \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{\theta}} + \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}} \cdot \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{\theta}}$$

To compute the sensitivities of (6) we solve what is called the system of sensitivity equations given by:

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{f}(t, \boldsymbol{x}(t, \boldsymbol{\theta}), \boldsymbol{\theta}) \tag{7}$$

$$\frac{d}{dt}\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{\theta}} = \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{\theta}} + \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}} \cdot \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{\theta}} \tag{8}$$

Traditional sensitivity analysis is the quantification of the effect that changes in parameters have on model solutions. Assume $x(t)$ is a state variable of the model given in (6). We can now define the traditional sensitivity function for $x(t)$ [1].

Definition 3.2 *The traditional sensitivity function for the solution model $x(t)$ with respect to θ_k is defined by $\dfrac{\partial x}{\partial \theta_k}(t)$ where $k = 1, \cdots, p$.*

Because parameters can have different units and state variables can be of different orders of magnitude, we define the relative sensitivities of the state variable $x(t)$ with respect to the parameter θ_k as the product

$$\frac{\theta_k}{x(t)} \frac{\partial x(t)}{\partial \theta_k} \tag{9}$$

Once the system of sensitivity equations (7)-(8) is solved, we can determine how a specific state variable changes with respect to a given parameter θ_k. For instance, if $\dfrac{\partial x}{\partial \theta_k} = 0$ then the parameter θ_k has no influence in the behavior of the variable $x(t)$. If $\dfrac{\partial x}{\partial \theta_k} > 0$ then the parameter θ_k has a positive influence in the behavior of $x(t)$. That is, for lager values of θ_k, $x(t)$ will increase. Similarly, if $\dfrac{\partial x}{\partial \theta_k} < 0$ then the parameter θ_k has a negative impact in the behavior of $x(t)$. That is, for larger values of θ_k, $x(t)$ will decrease.

In order to illustrate this concept we will find the sensitivities for some basic examples. In our first two examples there is no need to compute the sensitivity equations since the solutions of the models can be found analytically by standard methods.

Example 1 Consider the exponential growth model given by the initial value problem

$$\frac{dx}{dt} = rx \tag{10}$$

$$x(0) = x_0 \tag{11}$$

We know this model has solution $x(t) = x_0 e^{rt}$.

Since this model has only one parameter, the traditional sensitivity and the relative sensitivity are given respectively by:

$$\frac{\partial x}{\partial r} = x_0 t e^{rt} \quad \text{and} \quad \frac{r}{x(t)}\frac{\partial x}{\partial r} = rt$$

Notice that

$$\frac{r}{x(t)}\frac{\partial x}{\partial r} > 0 \text{ if } r > 0 \text{ and } \frac{r}{x(t)}\frac{\partial x}{\partial r} < 0 \text{ if } r < 0$$

Figure 8 shows the relative sensitivities of the model. If $r > 0$, then $\frac{r}{x(t)}\frac{\partial x}{\partial r}$ is a monotonically increasing and positive function. In other words, the state function $x(t)$ increases as r grows. Similarly, if $r < 0$ then $\frac{r}{x(t)}\frac{\partial x}{\partial r}$ is a monotonically decreasing and negative function. Which implies that the solution $x(t)$ decreases as r grows. In Figure 9 we have a numerical solutions for the model given by equations (10)-(11) taking different values of r. In this figure, we can see that for positive values of r, as this parameter grows $x(t)$ will grow as well. For negative values of r, as this parameter grows, the state variable $x(t)$ decays faster.

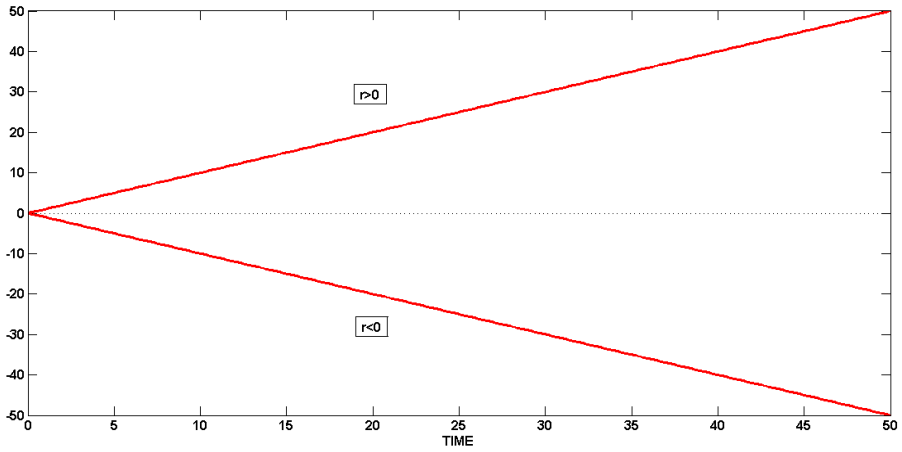

Figure 8: Relative sensitivities for the model given in Example 1

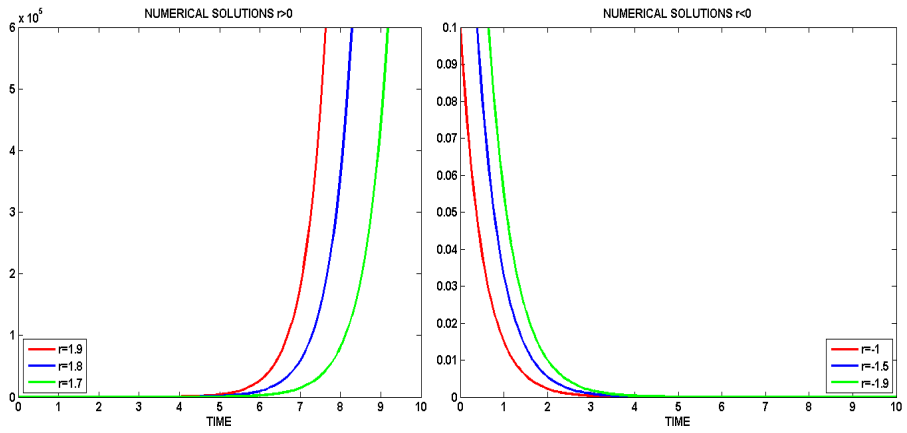

Figure 9: Numerical solutions of Example 1 with positive and negative values of r

Example 2 Consider the initial value problem:

$$\begin{aligned} \frac{dx}{dt} &= a - bx \\ x(0) &= x_0 \end{aligned}$$

The vector of parameters is $\theta = (a, b)$.

For this example, the solutions of the sensitivities can be computed analytically. In fact:

$$x(t) = \frac{a}{b} + \left(x_0 - \frac{a}{b}\right) e^{-bt}$$

Then the traditional sensitivities are

$$\begin{aligned} \frac{\partial x}{\partial a} &= \frac{1}{b}\left(1 - e^{-bt}\right) \\ \frac{\partial x}{\partial b} &= -\frac{a}{b^2}\left(1 - e^{-bt}\right) + \left(\frac{a}{b} - x_0\right) te^{-bt} \end{aligned}$$

The relative sensitivities are: $\dfrac{a}{x(t)}\dfrac{\partial x}{\partial a}$ and $\dfrac{b}{x(t)}\dfrac{\partial x}{\partial b}$.

Notice that

$$\lim_{t\to+\infty} \frac{a}{x}\frac{\partial x}{\partial a} = \frac{1}{b^2} > 0, \quad \text{and} \quad \lim_{t\to+\infty} \frac{b}{x}\frac{\partial x}{\partial b} = -\frac{a^2}{b^4} < 0$$

This is, the relative sensitivity of $x(t)$ with respect to a is always positive, which implies that the parameter a has a positive influence in the behavior of $x(t)$. In other words, for larger values of a the function $x(t)$ will increase. Likewise, the relative sensitivity of $x(t)$ with respect to b is always negative, which implies that for larger values of b the function $x(t)$ will decrease. The numerical solutions of the sensitivities are given in Figure 10 and an illustration of the behavior of $x(t)$ with respect to the parameters a and b is given in Figure 11.

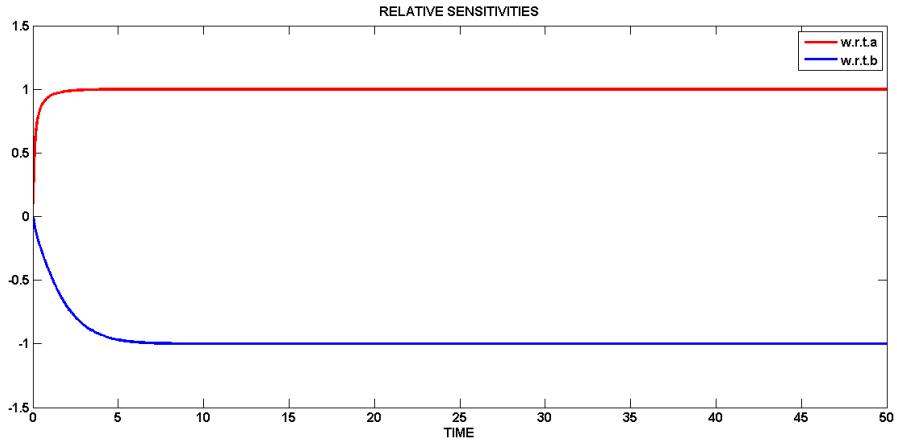

Figure 10: Relative Sensitivities for the initial value problem in Example 2. Parameters and initial condition were chosen as follows: $a = 1$, $b = 1$, and $x(0) = 0.1$.

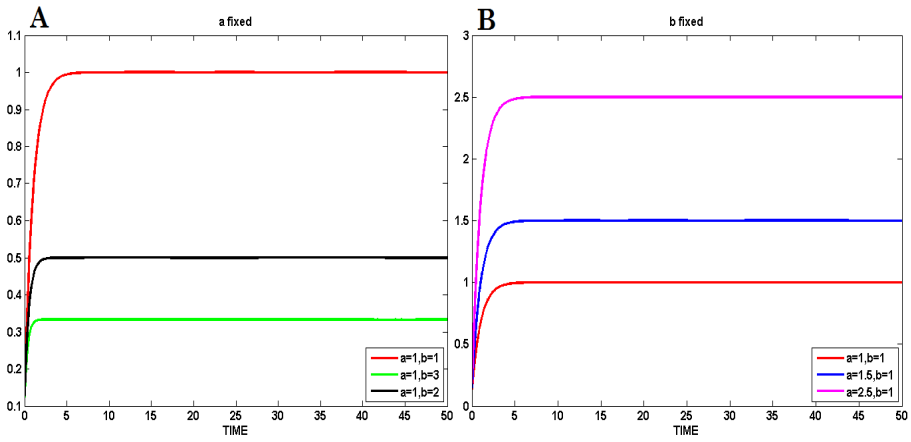

Figure 11: (A) Numerical solution for $x(t)$ fixing a and letting b vary. (B) Numerical solution for $x(t)$ fixing b and letting a vary. If we fix the value of a and we increase parameter b we can see how function $x(t)$ will reach smaller values. On the other hand, by fixing parameter b and letting a increase, $x(t)$ will reach larger values.

Example 3 Consider the SIR model:

$$\frac{ds}{dt} = -\beta si \tag{12}$$

$$\frac{di}{dt} = \beta si - \gamma i \tag{13}$$

$$1 = s(0) + i(0) \tag{14}$$

Unlike Examples 1 and 2, the sensitivities for this model have to be calculated numerically because an analytic solution does not exist. The sensitivities are found by solving the following system:

$$\begin{aligned} \frac{d\boldsymbol{x}}{dt} &= \boldsymbol{f}(t, \boldsymbol{x}(t, \boldsymbol{\theta}), \boldsymbol{\theta}) \\ \frac{d}{dt}\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{\theta}} &= \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{\theta}} + \frac{\partial \boldsymbol{f}}{\partial t} \cdot \frac{\partial \boldsymbol{x}}{\partial \boldsymbol{\theta}} \end{aligned}$$

where $\boldsymbol{\theta} = (\beta, \gamma)$ is the vector of parameters.

Figure 12 exhibits the numerical solution of the model given by equations (12)-(14). The population represented by $s(t)$ decays from its initial condition to attain its non-zero equilibrium point. On the other hand, the population represented by $i(t)$ shows a bell shaped curve having a maximum point nearly at $t = 6.5$ and eventually going extinct.

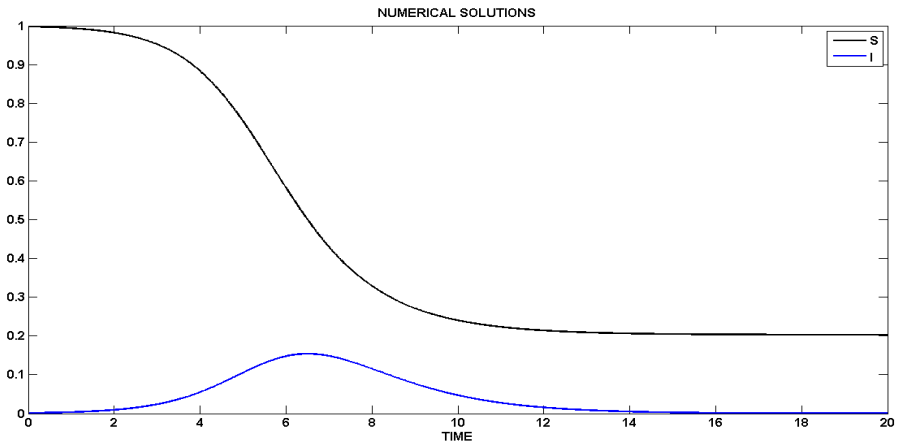

Figure 12: Numerical Solutions for the SIR model. Parameters and initial conditions were chosen as follows: $\beta = 2$, $\gamma = 1$, $s(0) = 0.998$, and $i(0) = 0.0013$.

According to Figure 13, the larger the values of γ the smaller the values $i(t)$ attains. Similarly, the larger the values of β, the larger the values of $i(t)$. This is totally intuitive since β is the rate of production of $i(t)$ and γ is the clearance rate. Nevertheless, Figure 14 tells us that this behavior is not always like that. Even though at the beginning of the simulation the sensitivities of β are positive and the sensitivities of γ are negative, the role of these parameters changes. For values of $t > 7.2$, the sensitivity of β becomes negative and the sensitivity of γ becomes positive for values of $t > 9$. This means that for $t \in (7.2, \infty)$, if we increase β then $i(t)$ will reach smaller values. Additionally, for $t \in (9, \infty)$, if we increase γ then $i(t)$ will attain larger values.

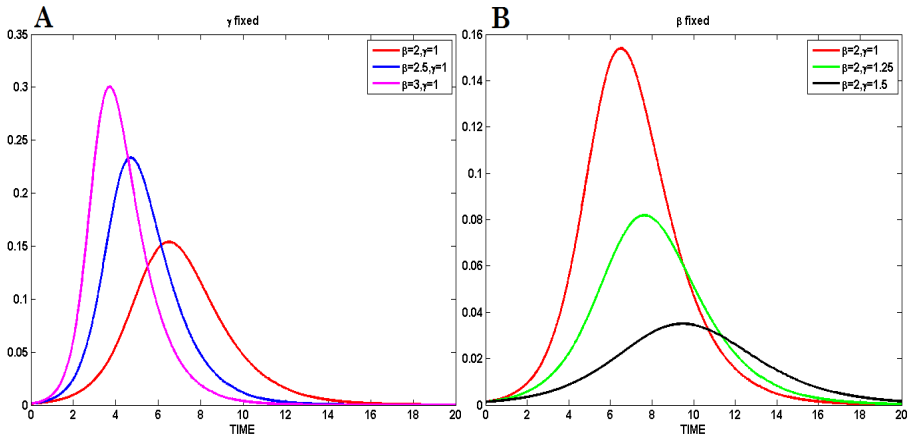

Figure 13: Numerical Solutions of $i(t)$ fixing β and γ show the effect of increasing one parameter at a time. Initial conditions were chosen as follows: $s(0) = 0.998$, and $i(0) = 0.0013$. The parameter values are displayed in the legend of the figure.

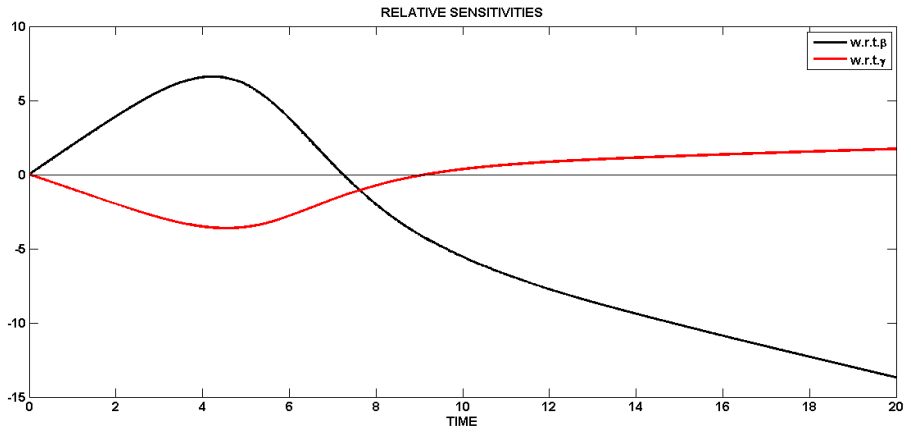

Figure 14: Relative sensitivities for $i(t)$ versus time t show how the role of the parameters β and γ eventually switches. Parameters and initial conditions were chosen as follows: $\beta = 2$, $\gamma = 1$, $s(0) = 0.998$, and $i(0) = 0.0013$.

We are interested in calculating the relative sensitivities to the model given in equations (1)-(5). This model can be written as

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{f}(t, \boldsymbol{x}(t), \boldsymbol{\theta}) \tag{15}$$

where $\boldsymbol{x}(t) = [T(t), I(t), V(t), Z(t), W(t)]$, $\boldsymbol{\theta} = (s, d, \beta, a, p, k, \mu, q, c, b, g, h)$ is the vector of the twelve parameters of the system, and $\boldsymbol{f}(t, \boldsymbol{x}(t), \boldsymbol{\theta})$ is the right hand side of the system of equations.

In particular, we are interested in distinguishing what parameters play an important role in the increasing or decreasing behavior of the viral load $V(t)$. In order to find the sensitivity matrix $\phi(t) = \dfrac{\partial \boldsymbol{x}}{\partial \boldsymbol{\theta}}$, which is a 5×12 matrix, numerical solutions of corresponding sensitivity equations were computed. The third row of this matrix is given by $\dfrac{\partial V(t)}{\partial \boldsymbol{\theta}}$, the change in $V(t)$ with respect the parameters $\boldsymbol{\theta}$. Each entry will tell us information of how $V(t)$ changes with respect to each parameter. We use the same values of parameters and initial conditions that we used in Section 2 to find the numerical solutions of the system of sensitivity equations (7)-(8) for each scenario: dominant CTL response, dominant antibody response, and coexistence.

3.2 Relative Sensitivities: Dominant CTL Response

Using the values of parameters and initial conditions used in Section 2.1.1, numerical simulations of the relative sensitivities were computed for a transient and long-term phases for the dominant CTL response scenario.

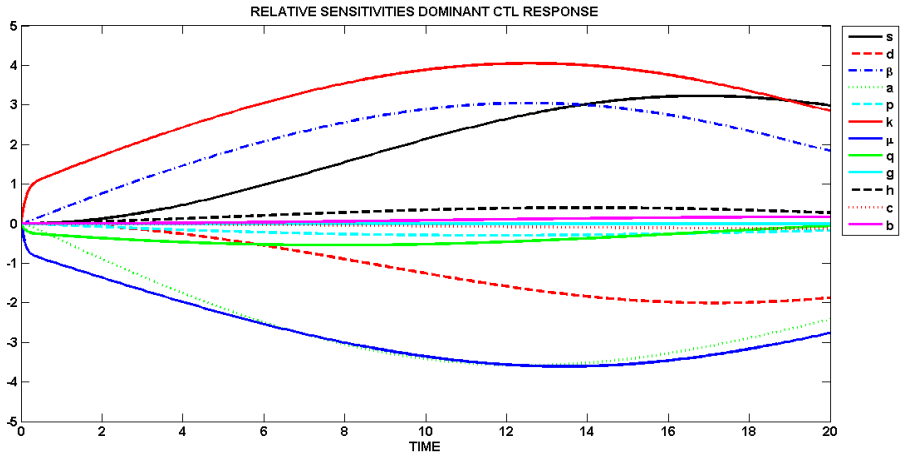

Figure 15: Numerical solutions of the relative sensitivities $\frac{\theta_k}{V}\frac{\partial V}{\partial \theta_k}$ for the dominant CTL response scenario in a transient phase from 0 to 20

Figure 15 exhibits the transient phase where the parameters with more impact in the behavior of $V(t)$ are k and μ. This is because the relative sensitivities, $\dfrac{k}{V(t)}\dfrac{\partial V(t)}{\partial k}$ and $\dfrac{\mu}{V(t)}\dfrac{\partial V(t)}{\partial \mu}$, have the most positive and negative values, respectively. This is not surprising since k is the rate of proliferation and μ is the rate of natural decay of the viral load.

As seen in Figure 15, other parameters such as, β, a, s, and d also play an important role in the behavior of $V(t)$. In fact, if parameters β and s increase during this time window of $[0, 20]$ we would expect that the viral load will reach larger values. On the other hand, parameters a and d have a negative influence in the behavior of $V(t)$. This is, if the rates of natural decay of uninfected and infected cells increase over this period of time, then the number of viruses decreases. Once again this makes total sense because the virus depends on these cells to get reproduced. A summary of the parameters with more influence in the behavior of the viral load in the transient phase can be found in Table 1.

Table 1: Ranking of the most influential parameters for dominant CTL response in the transient phase. Increasing parameters k, β, and s will increase $V(t)$ during this phase. Similarly, increasing the values of parameters μ, a, and d will produce a reduction in the number of viruses.

	Positive	**Negative**
Parameter	k	μ
	β	a
	s	d

In the long-term phase shown in Figure 16, the parameters k and μ still are the most influential parameters because their sensitivities are again the most positive and negative functions. However, now these parameters are also joined by c and b, the proliferation and natural decay rates of the CTL, respectively. It was not surprising that these two parameters eventually would be responsible for the behavior of $V(t)$ since we are in the dominant CTL response scenario. It is also important to notice that parameters β, a, s, and d that were important in the transient phase eventually will become irrelevant. A summary of the most important parameters in the behavior of $V(t)$ in the long-term phase is given in Table 2.

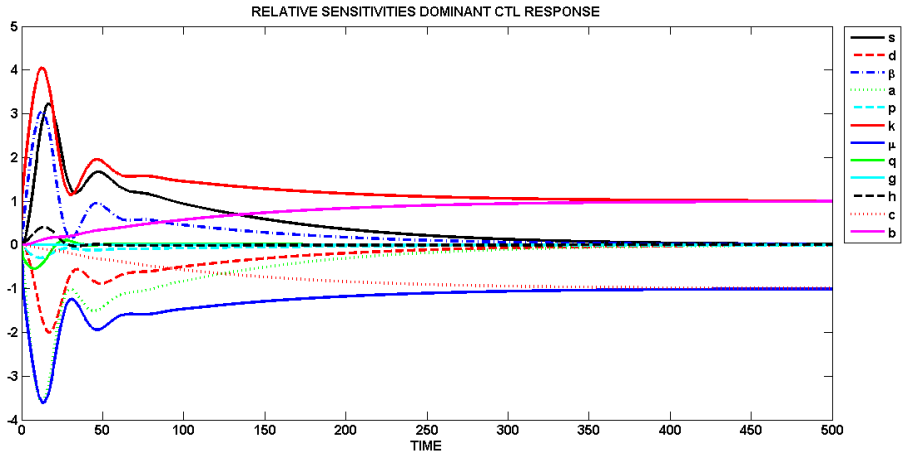

Figure 16: Numerical solutions of the relative sensitivities $\frac{\theta_k}{V}\frac{\partial V}{\partial \theta_k}$ for the dominant CTL response scenario in a long-term phase from 0 to 500

Table 2: Ranking of the most influential parameters for the dominant CTL response in the long-term phase. Increasing parameters k and b will increase $V(t)$ during this phase. Similarly, increasing the values of parameters μ and c will produce a reduction in the number of viruses.

	Positive	**Negative**
Parameter	k	μ
	b	c

3.3 Relative Sensitivities: Dominant Antibody Response

Numerical simulations of the sensitivities for the dominant antibody response scenario were computed using the same values of the parameters used in Section 2.1.2. These numerical simulations were calculated in transient and long-term phases.

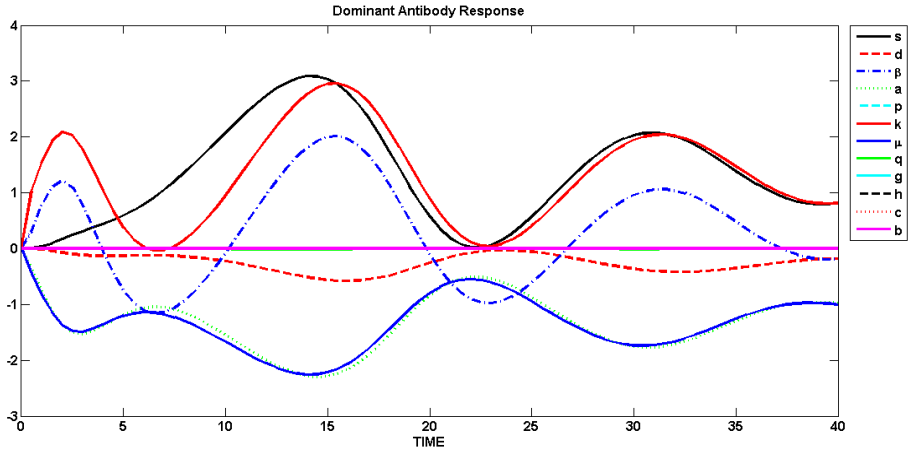

Figure 17: Numerical solutions of the relative sensitivities $\frac{\theta_k}{V}\frac{\partial V}{\partial \theta_k}$ for the dominant antibody response scenario in a transient phase from 0 to 40

For the transient phase with $t \in [0, 40]$, which is shown in Figure 17, the parameters that are most influential in the behavior of the virus load are μ, k, s, a, and β because their sensitivities reach the most positive and negative values among all the other sensitivities. However, because of the presence of an oscillatory behavior on these sensitivities there are periods of time where their influence is stronger. A summary of the most influential parameters in the behavior of the viral load is shown in Table 3.

Table 3: Ranking of parameters that are most influential for dominant antibody response in the transient phase. Increasing parameters k, s and β will increase $V(t)$ during this phase. Similarly, increasing the values of parameters μ and a will produce a reduction in the number of viruses.

	Positive	**Negative**
Parameter	k	μ
	s	a
	β	

A most interesting behavior is found in the long-term phase shown in Fig 18. The oscillations that were observed in the transient phase continued until $t = 70$. After this point, the sensitivities of s and k are clearly the most positive functions and the sensitivities of μ and a are the most negative functions. However, at $t = 450$ something even more interesting and counterintuitive happened: the role of these four parameters switch and two new parameters that so far were almost irrelevant become the most influential. These new parameters are g and h: the rates of proliferation and natural decay of the antibody response, respectively. Finally, for values of $t >$ 1400 only these two parameters are in charge in the behavior of $V(t)$. The other parameters s, k, μ, and a eventually become irrelevant. A summary of the most influential parameters in the long-term phase is shown in Table 4.

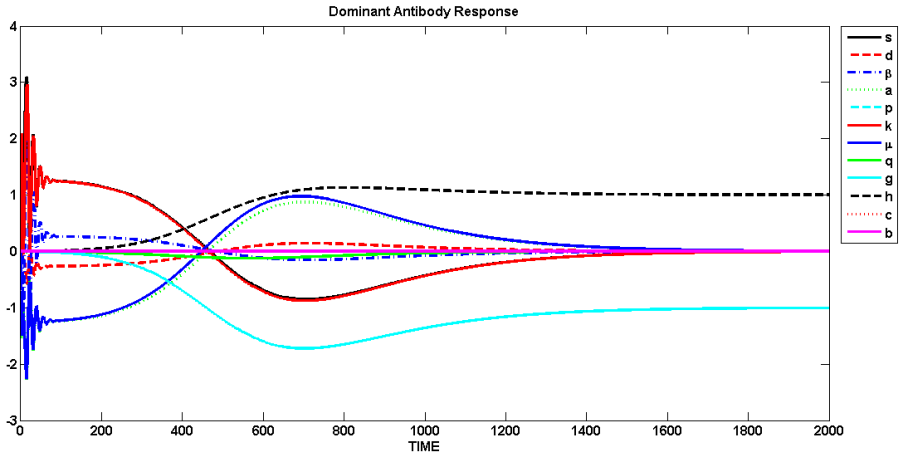

Figure 18: Relative sensitivities $\frac{\theta_k}{V}\frac{\partial V}{\partial \theta_k}$ for the dominant antibody response scenario in a long-term phase from 0 to 2000

Table 4: Ranking of parameters for the relative sensitivities in the dominant antibody response scenario in a long-term phase. Increasing parameter h will increase $V(t)$ during this phase. Similarly, increasing the value of parameter g will produce a reduction in the number of viruses.

	Positive	**Negative**
Parameter	h	g

3.4 Relative Sensitivities: Coexistence

Numerical solutions of the sensitivities for the coexistence scenario were calculated using the same values of the parameters taken in Section 2.1.3. These solutions were computed for transient and long-term phases.

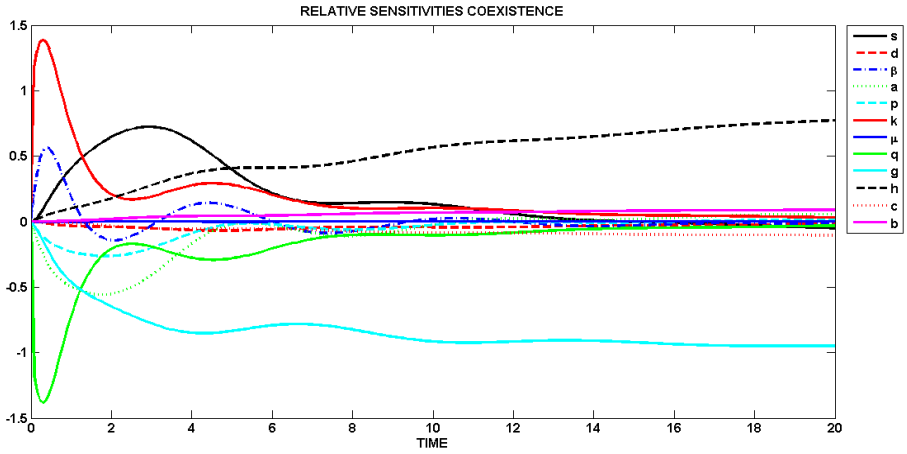

Figure 19: Numerical solutions of the relative sensitivities $\frac{\theta_k}{V}\frac{\partial V}{\partial \theta_k}$ for the coexistence scenario in a transient phase from 0 to 20

It is reasonable to think that, in this scenario, $V(t)$ will be mostly affected by the values of parameters k, μ, c, b, h, and g since these parameters are proliferation and natural decay rates of virus, CTL, and antibodies, respectively. However, as shown in Figure 19, at the beginning of the simulation, from $t = 0$ to $t = 2$, $V(t)$ looks to be affected by the values of parameters k and q. After this point, for a short period of time with $t \in (1.8, 5)$ the sensitivities of s and g are the most positive and the most negative, respectively. However, a unexpected outcome was obtained, in the transient phase for $t > 5$ and in the long-term phase, shown in Figure 20, the viral load $V(t)$ is

affected only by the parameters g and h. In other words, in the coexistence scenario the behavior of the viral load is affected only by the levels of production and reduction of antibodies. A summary of the most influential parameters in the transient phase and in the long-term phase is shown in Tables 5 and 6, respectively.

Table 5: Ranking of the most influential parameters of $V(t)$ for the coexistence scenario in a transient phase. Increasing parameters h, s, and k will increase $V(t)$ during this phase. Similarly, increasing the value of parameters g and q will produce a reduction in the number of viruses.

	Positive	**Negative**
Parameter	h	g
	s	q
	k	

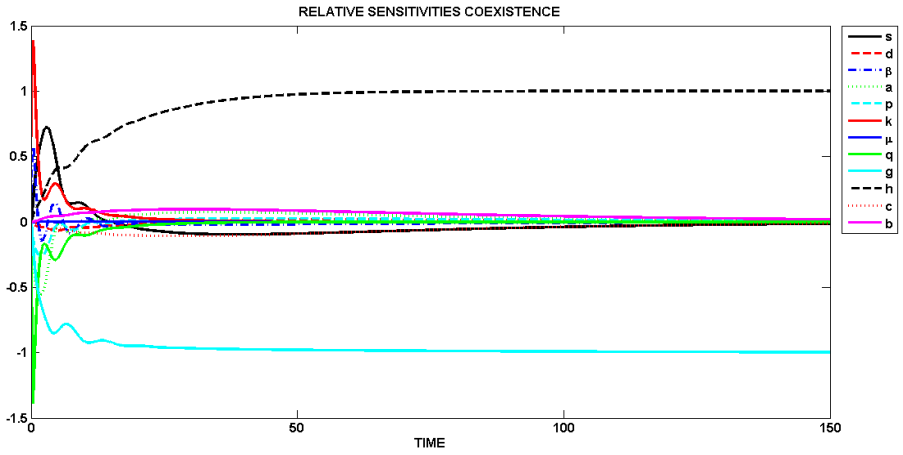

Figure 20: Relative sensitivities $\frac{\theta_k}{V}\frac{\partial V}{\partial \theta_k}$ for the coexistence scenario in a long-term phase from 0 to 150

Table 6: Ranking of the most influential parameters of $V(t)$ for the coexistence scenario in a long-term phase. Increasing parameter h will increase $V(t)$ during this phase. Similarly, increasing the value of parameter g will produce a reduction in the number of viruses.

	Positive	**Negative**
Parameter	h	g

4 TREATMENT AS AN OPTIMAL CONTROL PROBLEM

4.1 Introduction

For several years antivirals such as interferon (IFN), pegylated-interferon (PEG-INF) and ribavirin (RVB) have been used to reduce the levels of HCV RNA. In fact, in approximately 50% of HCV treated patients with combinations of PEG-INF/RVB, the virus is not eradicated [6, 14]. More recently treatments are focused on direct-acting antiviral agents (DAAs) that target specific steps of the HCV life cycle [24]. Nevertheless, the use of these treatments is always restricted by their side effects.

In this section, we want to determine a best treatment strategy of a combination of two antiviral medicines that reduces the infection and that minimizes the "cost" effect to the body. We want to determine optimal treatment schedules for each of the three scenarios discussed in the previous chapters.

The idea is to extend the system given by equations $(1) - (5)$, with two piecewise continuous functions $u_1(t)$ and $u_2(t)$ that describe a treatment strategy of a combination of two drugs over a certain fixed time horizon $[0, t_f]$. A best treatment strategy has not only to consider the time where the medicines have to be administered but also the side effects of the treatment to the human body. To approach the problem of determining a best treatment strategy, we solve an optimal control problem. In this thesis we provide definitions and concepts to our particular control problems. The interested reader can find more general details in optimal control theory applied to biological models in [18].

Consider the extension of the equations (1)-(5) given by the system of differential equations

$$\begin{aligned}
\frac{dT}{dt} &= s - dT - (1-u_1)\beta VT && (16)\\
\frac{dI}{dt} &= (1-u_1)\beta VT - aI - pIZ && (17)\\
\frac{dV}{dt} &= (1-u_2)kI - \mu V - qVW && (18)\\
\frac{dZ}{dt} &= cIZ - bZ && (19)\\
\frac{dW}{dt} &= gVW - hW && (20)
\end{aligned}$$

where $u_i(t) = 1$ represents maximal use of therapy and $u_i(t) = 0$ corresponds to minimal use of therapy, for $i = 1, 2$.

We rewrite the system of equations $(16) - (20)$ as

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{f}(t, \boldsymbol{x}(t), u_1(t), u_2(t)) \tag{21}$$

where

$$\boldsymbol{x}(t) = [T(t), I(t), V(t), Z(t), W(t)]$$

We define the set of admissible controls by

$$\mathscr{U} = \{(u_1, u_2) \mid u_1, u_2 \,\text{measurable}, 0 \le u_1 \le 1, 0 \le u_2 \le 1, t \in [0, t_f]\}\,.$$

In this context, an optimal controls problem consists on finding piecewise continuous functions $u_1(t)$ and $u_2(t)$, called control functions in $\mathscr{U}$, and the associated states variables $\boldsymbol{x}(t)$ that minimize a given objective functional of the form

$$J(u_1(t), u_2(t)) = \int_0^{t_f} G(t, \boldsymbol{x}(t), u_1(t), u_2(t))dt$$

In other words, we are interested in solving the optimal control problem

$$\begin{aligned} \min_{u_1,u_2 \in \mathscr{U}} \quad & \int_0^{t_f} G(t, \boldsymbol{x}(t), u_1(t), u_2(t))dt \\ \text{subject to} \quad & \frac{d\boldsymbol{x}}{dt} = \boldsymbol{f}(t, \boldsymbol{x}(t), u_1(t), u_2(t)) \\ & \boldsymbol{x}(t_0) = \boldsymbol{x}_0 \end{aligned}$$

The main technique for such optimal control problems consists of solving a set of necessary conditions that an optimal control pair and corresponding state variables must satisfy. The necessary conditions we derive were developed by Pontryagin *et al.* [22].

Before giving the necessary conditions we need to give the following definition:

Definition 4.1 *The Hamiltonian H is defined by*

$$H(t, \boldsymbol{x}(t), u_1(t), u_2(t)) = G(t, \boldsymbol{x}(t), u_1(t), u_2(t)) + \boldsymbol{\lambda}(t) \cdot \boldsymbol{f}(t, \boldsymbol{x}(t), u_1(t), u_2(t))$$

where $\boldsymbol{\lambda}(t) = [\lambda_1(t), \cdots, \lambda_5(t)]$ is a piecewise differentiable vector-valued function, and each λ_i is called the adjoint variable corresponding to x_i.

We propose three different control problems to determine an optimal treatment for each scenario: dominant CTL response, dominant antibody response and coexistence.

In other words, we define three different objective functionals for each scenario. To find the solution to each problem we need to solve a set of necessary conditions, that in general, are given in the next subsection but that will be studied in detail for each scenario.

4.2 Necessary Conditions

The following theorem is an extension of Pontryagin's Maximum Principle (Pontryagin *et al.* [22]) given in Lenhart and Workman [18].

Theorem 4.2 *Given the optimal control pair* $(u_1^*(t), u_2^*(t))$ *and solutions* $\boldsymbol{x}^*(t)$ *of the state system given by equations (16)-(20), there exists* $\boldsymbol{\lambda}(t)$*, a piecewise differentiable vector-valued function, satisfying*

$$H(t, \boldsymbol{x}^*, u_1(t), u_2(t), \boldsymbol{\lambda}(t)) \geq H(t, \boldsymbol{x}^*, u_1^*(t), u_2^*(t), \boldsymbol{\lambda}(t))$$

for all $u_1(t), u_2(t) \in \mathscr{U}$ *at each time* t*, where*

$$\frac{d\lambda_j(t)}{dt} = -\frac{\partial H}{x_j} \quad (\textit{Adjoint Equations})$$

and $\lambda_j(t_f) = 0 \quad$ *for* $j = 1, \cdots, 5 \quad$ *usually called the transversality conditions. The optimality conditions are given by*

$$0 = \frac{\partial H}{\partial u_k} \text{ at } u_k^* \quad \textit{for} \quad k = 1, 2 \quad \textit{(Optimality Condition)}$$

To solve the set of necessary conditions given in Theorem 4.2 we applied a forward-backward sweep method. A much more detailed version of this method can be found in [18].

4.3 Forward-Backward Sweep Method

The optimality system given in Theorem 4.2 is solved numerically using a forward-backward sweep method [18]. The following is an outline of the algorithm:

1. Make an initial guess for $\boldsymbol{u}$ over the interval.

2. Using the initial condition $\boldsymbol{x} = \boldsymbol{x}(t_0)$, and the values for $\boldsymbol{u}$, solve $\boldsymbol{x}$ forward in time according to its differential equation in the optimality system.

3. Using the transversality condition and the values for $\boldsymbol{u}$ and $\boldsymbol{x}$, solve $\boldsymbol{\lambda}$ backward in time according to its differential equation in the optimality system.

4. Update $\boldsymbol{u}$ by entering the new $\boldsymbol{x}$ and $\boldsymbol{\lambda}$ values into the characterization of the optimal control.

5. Check for convergence. If values of the variables in this iteration and the last iteration are negligibly close, output the current values as solutions. If values are not close, return to step 2.

To the best of our knowledge, Chakrabarty and Joshi [4] were the first ones in approaching the problem of finding an optimal combination treatment strategy using optimal control theory. However, in their model they did not consider immune responses. Similar problems for optimizing chemotherapy in HIV models can be found in [15, 4, 17]. Our aim is to determine an optimal treatment strategy for the three scenarios seen in Section 2: dominant CTL response, dominant antibody response, and coexistence.

4.4 Optimal Treatment Strategy for the Dominant CTL Response Scenario

Under the assumption that the CTL response is much stronger than the antibody response, we seek to minimize the number of infected cells $I(t)$ and the "cost" based on the effect of the treatment to the human body. For this reason, the objective functional to be minimized is

$$J(u_1, u_2) = \int_0^{t_f} \left(A_1 I(t) + \frac{1}{2} \left(A_2 u_1^2(t) + A_3 u_2^2(t) \right) \right) dt$$

subject to the system

$$\begin{aligned}
\frac{dT}{dt} &= s - dT - (1-u_1)\beta VT \\
\frac{dI}{dt} &= (1-u_1)\beta VT - aI - pIZ \\
\frac{dV}{dt} &= (1-u_2)kI - \mu V - qVW \\
\frac{dZ}{dt} &= cIZ - bZ \\
\frac{dW}{dt} &= gVW - hW
\end{aligned}$$

The positive parameters A_1, A_2, and A_3 balance the size terms. With higher cost parameters A_2 and A_3, the system has controls where maximum treatment is continued for a shorter period of time. In other words, with large values of these parameters the cost effect of treatment is virtually not important. The severity of therapy in the human body is described by the terms u_1^2 and u_2^2. The reason behind considering a finite time window is that the administration of treatment is usually restricted to a limited time period.

Necessary conditions are derived by using the extension of Potryagin's Maximum Principle [22] given in Theorem 4.2.

Theorem 4.3 *Given the optimal control pair* $(u_1^*(t), u_2^*(t))$ *and solutions* T^*, I^*, V^*, Z^*, W^* *of the state system, there exists adjoint variables* $\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5$ *satisfying:*

$$\begin{aligned}
\lambda_1' &= \lambda_1\left(d + (1-u_1)\beta V\right) - \lambda_2(1-u_1)\beta V \\
\lambda_2' &= -A_1 + \lambda_2(a + pZ) - \lambda_3 k(1-u_2) - \lambda_4 cZ \\
\lambda_3' &= \lambda_1(1-u_1)\beta T - \lambda_2(1-u_1)\beta T + \lambda_3(\mu + qW) - \lambda_5 gW \\
\lambda_4' &= \lambda_2 pI + \lambda_4(b - cI) \\
\lambda_5' &= \lambda_3 qV + \lambda_5(h - gV)
\end{aligned}$$

with $\lambda_1(t_f) = \lambda_2(t_f) = \lambda_3(t_f) = \lambda_4(t_f) = \lambda_5(t_f) = 0$*, which are the transversality conditions. Furthermore, the controls are characterized by*

$$\begin{aligned}
u_1^* &= \mathit{min}\left\{1, \mathit{max}\left\{0, \frac{(\lambda_2 - \lambda_1)\beta V^* T^*}{A_2}\right\}\right\} \\
u_2^* &= \mathit{min}\left\{1, \mathit{max}\left\{0, \frac{\lambda_3 k I^*}{A_3}\right\}\right\}
\end{aligned}$$

Proof:

The Hamiltonian is defined as follows:

$$\begin{aligned}
H(t, \boldsymbol{x}(t), u_1(t), u_2(t), \boldsymbol{\lambda}(t)) &= A_1 I + \frac{1}{2}\left(A_2 u_1^2(t) + A_3 u_2^2(t)\right) \\
&\quad + \lambda_1(s - dT - (1-u_1(t))\beta V T) \\
&\quad + \lambda_2((1-u_1(t))\beta V T - aI - pIZ) \\
&\quad + \lambda_3((1-u_2(t))kI - \mu V - VW) \\
&\quad + \lambda_4(cIZ - bZ) \\
&\quad + \lambda_5(gVW - hW)
\end{aligned}$$

We obtain the adjoint equations:

$$\begin{aligned}
\lambda_1' &= -\frac{\partial H}{\partial T} = \lambda_1\left(d + (1-u_1)\beta V\right) - \lambda_2(1-u_1)\beta V \\
\lambda_2' &= -\frac{\partial H}{\partial I} = -A_1 + \lambda_2(a + pZ) - \lambda_3 k(1-u_2) - \lambda_4 cZ \\
\lambda_3' &= -\frac{\partial H}{\partial V} = \lambda_1(1-u_1)\beta T - \lambda_2(1-u_1)\beta T + \lambda_3(\mu + qW) - \lambda_5 gW \\
\lambda_4' &= -\frac{\partial H}{\partial Z} = \lambda_2 pI + \lambda_4(b - cI) \\
\lambda_5' &= -\frac{\partial H}{\partial W} = \lambda_3 qV + \lambda_5(h - gV)
\end{aligned}$$

with $\lambda_1(t_f) = \lambda_2(t_f) = \lambda_3(t_f) = \lambda_4(t_f) = \lambda_5(t_f) = 0$. In order to obtain the optimality condition, we maximize the Hamiltonian with respect to u_1 and u_2 at the optimal u_1^* and u_2^*, respectively.

This is, $\dfrac{\partial H}{\partial u_i} = 0$ at u_i^* for $i = 1, 2$.

Therefore,

$$u_1^* = \frac{(\lambda_2 - \lambda_1)\beta V^* T^*}{A_2} \text{ and } u_2^* = \frac{\lambda 3 k I^*}{A_3}$$

By standard routines of optimal control problems with bounded controls [4, 9, 15, 17, 18] the characterizations of the optimal controls are given by:

$$\begin{aligned}
u_1^* &= \mathsf{min}\left\{1, \mathsf{max}\left\{0, \frac{(\lambda_2 - \lambda_1)\beta V^* T^*}{A_2}\right\}\right\} \\
u_2^* &= \mathsf{min}\left\{1, \mathsf{max}\left\{0, \frac{\lambda_3 k I^*}{A_3}\right\}\right\}
\end{aligned}$$

□

The forward-backward-sweep method was used to find the numerical solutions of the optimal control problem in the time window $[0, t_f]$ with $t_f = 50$. These solutions are given in Figure 21.

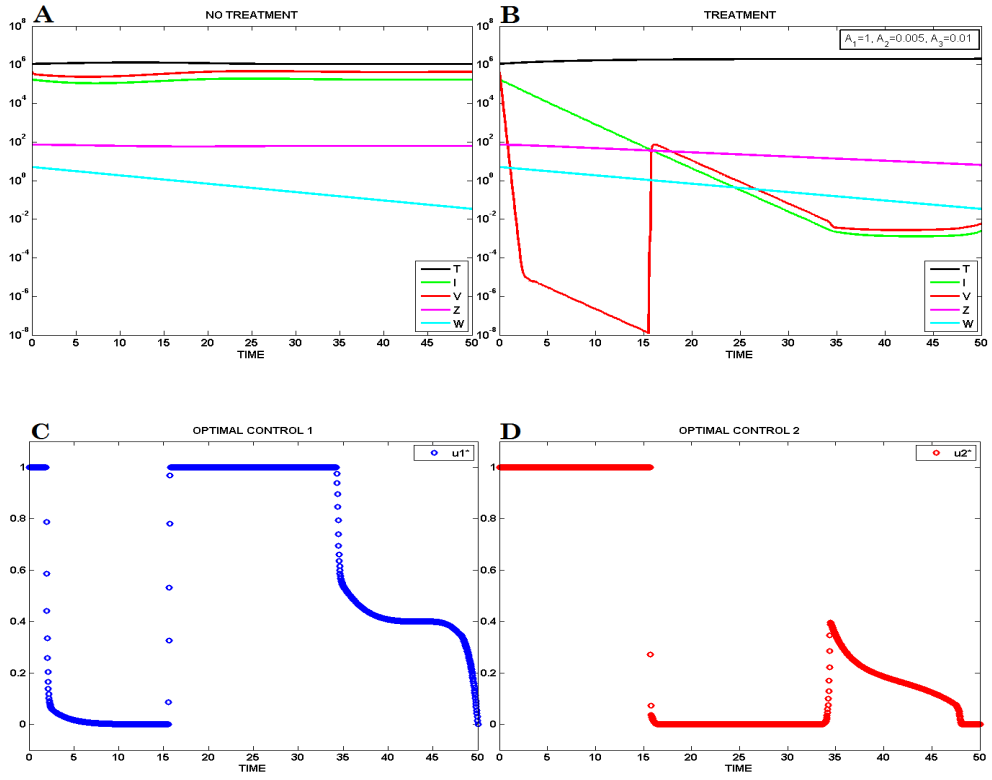

Figure 21: Numerical solutions and best treatment strategies. (A) shows the numerical solutions of the system with no treatment. (B) shows the effect of one session of treatment.(C) and (D) show the best treatment strategy. Parameters were chosen as follows: $s = 2.0\times10^{5}, d = 1.0\times10^{-1}, \beta = 2.0\times10^{-7}, a = 5.0\times10^{-1}, p = 6.4\times10^{-4}, k = 2.0 \times 10^{1}, \mu = 8.0 \times 10^{0}, q = 5.0 \times 10^{-1}, g = 1.0 \times 10^{-11}, c = 3.0 \times 10^{-7}, h = 1.0 \times 10^{1-}, b = 5.0 \times 10^{-2}, A_1 = 1.0 \times 10^{0}, A_2 = 5.0 \times 10^{-3}, A_3 = 1.0 \times 10^{-2}, t_f = 5.0 \times 10^{1}$. Initial conditions: $T(0) = 1.0 \times 10^{6}, I(0) = 1.6 \times 10^{5}, V(0) = 4.1 \times 10^{5}, Z(0) = 7.1 \times 10^{1}, W(0) = 5.0 \times 10^{0}$.

Figure 21 (C) and (D) exhibits the controls u_1^* and u_2^*, that are the optimal treatment routines suggested by the model. According to the numerical solutions, the drug that suppresses the production of infected cells, i.e. u_1, must be administered in full scale from $t = 0$ to $t = 2$. Then, u_1^* is tapered off from $t = 2$ to $t = 16$. Finally, u_1^* has to be administered in full scale again, this time from $t = 16$ to $t = 35$. On the other hand, the model suggests that the drug that interferes in the production of virus, u_2^*, must be administered in full scale from $t = 0$ to $t = 15$. Then, the dose is tapered off from $t = 15$ to $t = 35$ and administered again in a much lower scale for $35 < t < 50$.

Figure 21 (A) shows the numerical solutions of the model given by equations (16)-(20) assuming no treatment has been administered. In other words, $u_1(t) = 0$ and $u_2(t) = 0$. The value of the parameters and initial conditions are the same that were used in Section 2.1.1.

In Figure 21 (B) we have the numerical solutions of the optimal control problem for the state variables. This solution shows how the viral load and number of infected cells are affected by the effect of treatment. With the treatment schedule suggested by the model, we see a sharp decrease in the log of the viral load describing a biphasic decline followed by a rebound at the moment treatment u_2 is tapered off. However, this rebound is controlled by the drug u_1 that keeps the viral load at low levels.

To understand what happens after cessation of therapy, we solved the optimal control problem in two parts: first, we let u_1 and u_2 to be non zero functions from $t = 0$ to 50 to compute the forward-backward method in this time window. Second, we let $u_1 = 0$ and $u_2 = 0$ from $t = 50$ to $t = 150$ and calculated the numerical solutions. These numerical solutions are shown in Figure 22.

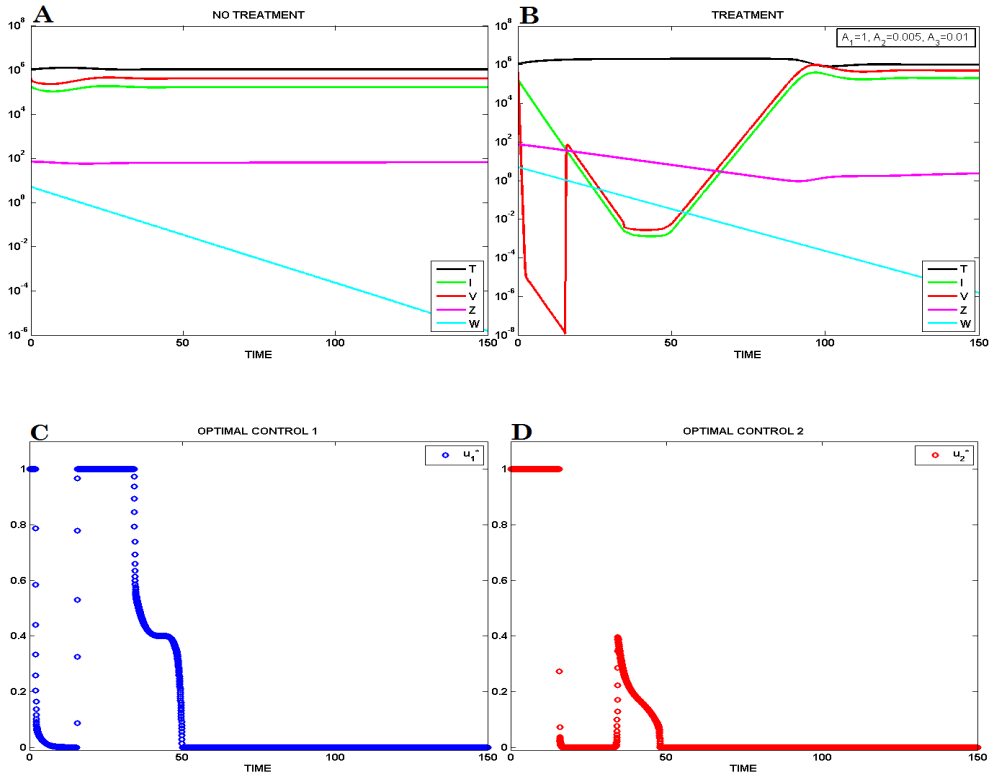

Figure 22: Numerical solutions after cessation of therapy. A continuation of the first treatment schedule was computed. We let u_2 and u_2 to be activated in the time window $[0, 50]$. Then we let $u_1 = 0$ and $u_2 = 0$ for $t \in [50, 150]$. (A) numerical solutions of the system with no treatment; (B) levels of virus rebound after cessation of treatment for $t > 50$; (C) and (D) are the continuation of the best treatment schedules.

It is reasonable to ask, what is the benefit of therapy if after treatment the viral load rebounds to its pretreatment levels? To provide a plausible answer to this question, we calculate the number of new viruses produced and the number of new infected cells produced during treatment and with no treatment. The rate of production of new viruses is given by the term kI, where k has units of $\frac{1}{time}$, then kI has units of $\frac{cells}{time}$.
Therefore, if the patient is not under treatment,

$$\int_0^{t_f} kI\, dt$$

is an approximation of the total production of infected cells over the interval $[0, t_f]$. Similarly, during therapy

$$\int_0^{t_f} (1 - u_2)kI\, dt$$

is the total production of virus over the interval $[0, t_f]$.

Table 7 shows the comparison of the total production of virus during treatment and with no treatment. During treatment the production of new virus decreases in five orders of magnitude. However, after cessation of treatment and running the simulation until $t = 150$, the number of viruses is of the same order of magnitude.

Table 7: The total production of virus was computed to compare the dominant CTL response scenario under treatment and with no treatment

	No Treatment	Treatment
Total production of new virus	$\int_0^{50} kI\,dt = 1.58 \times 10^8$	$\int_0^{50} (1-u_2)kI\,dt = 1.43 \times 10^3$
	$\int_0^{150} kI\,dt = 4.96 \times 10^8$	$\int_0^{150} (1-u_2)kI\,dt = 2.66 \times 10^8$

Table 8: The total production of infected cells was computed to compare the dominant CTL response scenario under treatment and with no treatment

	No Treatment	Treatment
Total production of infected cells	$\int_0^{50} \beta VTdt = 4.27 \times 10^6$	$\int_0^{50} (1-u_1)\beta VTdt = 1.21 \times 10^{-2}$
	$\int_0^{150} \beta VTdt = 1.31 \times 10^7$	$\int_0^{150} (1-u_1)\beta VTdt = 6.87 \times 10^6$

The production of new virus is assumed to be the number of interactions between free virus particles and uninfected cell. Therefore, the total production of new infected cells before treatment is given by the integral

$$\int_0^{t_f} \beta VT \, dt.$$

Similarly, the total production of new infected cells during treatment is calculated by

$$\int_0^{t_f} (1 - u_1)\beta VT \, dt.$$

Table 8 shows the comparison of these two quantities. During treatment for $t \in [0, 50]$ the production of new virus decreases in eight orders of magnitude. However, after cessation of therapy and computing the numerical solutions for $t \in [0, 150]$ the production of new infected cells is reduced only in one order of magnitude.

According to Tables 7 and 8, the benefit of therapy is precisely during the period of time the drug is being administered. If the goal of therapy is to reduce the levels of the viral load to undetectable numbers in a certain fixed time horizon, the optimal schedule suggested here is the desired treatment routine.

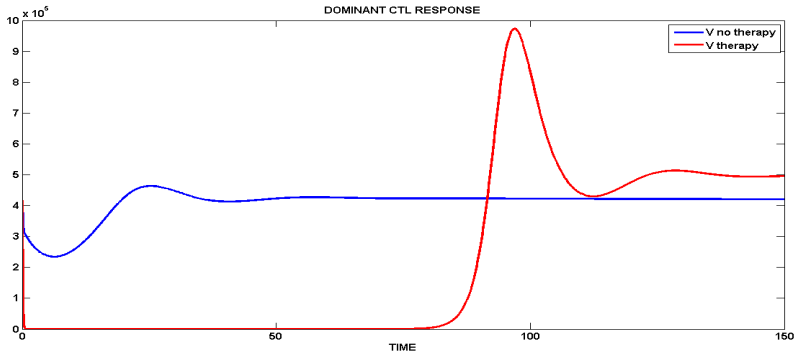

Figure 23: Comparison of the viral load under treatment and without treatment in a non logarithmic scale for the dominant CTL response scenario. The viral load with a non logarithmic scale during treatment (red curve) and with no treatment (blue curve) show that on the short term, the viral load decays significantly. However, on the long-term the steady state level is higher than without treatment. Worst than leaving untreated.

Figure 23 shows the numerical solution of $V(t)$ in a non logarithmic scale during treatment (red curve) and with no treatment (blue curve). During the administration of drugs, the viral load is maintained in considerably lower levels. This reduction in the number of viruses continues even after cessation of therapy. However, eventually $V(t)$ rebounds to levels that are higher than without treatment. This is the reason why in Table 7 the number of viruses is of the same order of magnitude during treatment than with no treatment in the long-term.

4.5 Optimal Treatment Strategy for the Dominant Antibody Response Scenario

As stated in Section 2, in this scenario we assume a vigorous antibody response combined with a weak CTL response. Because the antibodies play the role of neutralizing the production of new virus, we seek to minimize the function $V(t)$ and the "cost" based on the effect of the therapy to the body. The objective functional to be minimized is

$$J(u_1, u_2) = \int_0^{t_f} \left(A_1 V(t) + \frac{1}{2} \left(A_2 u_1^2(t) + A_3 u_2^2(t) \right) \right) dt$$

subject to the system

$$\begin{aligned} \frac{dT}{dt} &= s - dT - (1-u_1)\beta VT \\ \frac{dI}{dt} &= (1-u_1)\beta VT - aI - pIZ \\ \frac{dV}{dt} &= (1-u_2)kI - \mu V - qVW \\ \frac{dZ}{dt} &= cIZ - bZ \\ \frac{dW}{dt} &= gVW - hW \end{aligned}$$

The parameters A_1, A_2, and A_3 balance the size terms. With higher cost parameters A_2 and A_3, the system has controls where maximum treatment is continued for a shorter period of time. In other words, with large values of these parameters the cost effect of treatment is virtually not important. The severity of therapy in the human body is described by the terms u_1^2 and u_2^2. The reason for considering a finite time window is that the administration of treatment is usually restricted to a limited time period. Necessary conditions are derived by using the extension of Pontryagin's Maximum Principle [22] given in Theorem 4.2.

Theorem 4.4 *Given the optimal control pair* $(u_1^*(t), u_2^*(t))$ *and solutions* T^*, I^*, V^*, Z^*, W^* *of the state system, there exists adjoint variables* $\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5$ *satisfying:*

$$\begin{aligned}
\lambda_1' &= \lambda_1\left(d + (1-u_1)\beta V\right) - \lambda_2(1-u_1)\beta V \\
\lambda_2' &= \lambda_2(a + pZ) - \lambda_3 k(1-u_2) - \lambda_4 cZ \\
\lambda_3' &= -A_1 + \lambda_1(1-u_1)\beta T - \lambda_2(1-u_1)\beta T + \lambda_3(\mu + qW) - \lambda_5 gW \\
\lambda_4' &= \lambda_2 pI + \lambda_4(b - cI) \\
\lambda_5' &= \lambda_3 qV + \lambda_5(h - gV)
\end{aligned}$$

with $\lambda_1(t_f) = \lambda_2(t_f) = \lambda_3(t_f) = \lambda_4(t_f) = \lambda_5(t_f) = 0$*, which are the transversality conditions. Furthermore, the characterization of the control is given by*

$$\begin{aligned}
u_1^* &= \mathit{min}\left\{1, \mathit{max}\left\{0, \frac{(\lambda_2 - \lambda_1)\beta V^* T^*}{A_2}\right\}\right\} \\
u_2^* &= \mathit{min}\left\{1, \mathit{max}\left\{0, \frac{\lambda_3 k I^*}{A_3}\right\}\right\}
\end{aligned}$$

Proof:

The Hamiltonian is defined as follows:

$$\begin{aligned}
H(t, \boldsymbol{x}(t), u_1(t), u_2(t), \boldsymbol{\lambda}(t)) &= A_1 V + \frac{1}{2}\left(A_2 u_1^2(t) + A_3 u_2^2(t)\right) \\
&\quad + \lambda_1(s - dT - (1-u_1(t))\beta V T) \\
&\quad + \lambda_2((1-u_1(t))\beta V T - aI - pIZ) \\
&\quad + \lambda_3((1-u_2(t))kI - \mu V - VW) \\
&\quad + \lambda_4(cIZ - bZ) \\
&\quad + \lambda_5(gVW - hW)
\end{aligned}$$

We obtain the adjoint equations:

$$\begin{aligned}
\lambda_1' &= -\frac{\partial H}{\partial T} = \lambda_1\left(d + (1-u_1)\beta V\right) - \lambda_2(1-u_1)\beta V \\
\lambda_2' &= -\frac{\partial H}{\partial I} = \lambda_2(a+pZ) - \lambda_3 k(1-u_2) - \lambda_4 cZ \\
\lambda_3' &= -\frac{\partial H}{\partial V} = -A_1 + \lambda_1(1-u_1)\beta T - \lambda_2(1-u_1)\beta T + \lambda_3(\mu + qW) - \lambda_5 gW \\
\lambda_4' &= -\frac{\partial H}{\partial Z} = \lambda_2 pI + \lambda_4(b - cI) \\
\lambda_5' &= -\frac{\partial H}{\partial W} = \lambda_3 qV + \lambda_5(h - gV)
\end{aligned}$$

with $\lambda_1(t_f) = \lambda_2(t_f) = \lambda_3(t_f) = \lambda_4(t_f) = \lambda_5(t_f) = 0$. In order to obtain the optimality condition, we maximize the Hamiltonian with respect to u_1 and u_2 at the optimal u_1^* and u_2^*, respectively.

This is, $\dfrac{\partial H}{\partial u_i} = 0$ at u_i^* for $i = 1, 2$.

Therefore,

$$u_1^* = \frac{(\lambda_2 - \lambda_1)\beta V^* T^*}{A_2} \text{ and } u_2^* = \frac{\lambda 3 k I^*}{A_3}$$

By standard routines of optimal control problems with bounded controls [4, 9, 15, 17, 18] the characterizations of the optimal controls are given by:

$$\begin{aligned}
u_1^* &= \min\left\{1, \max\left\{0, \frac{(\lambda_2 - \lambda_1)\beta V^* T^*}{A_2}\right\}\right\} \\
u_2^* &= \min\left\{1, \max\left\{0, \frac{\lambda_3 k I^*}{A_3}\right\}\right\}
\end{aligned}$$

□

The forward-backward-sweep method was used to find the numerical solutions of the optimal control problem in the time window $[0, t_f]$ with $t_f = 40$. These numerical solutions are shown in Figure 24.

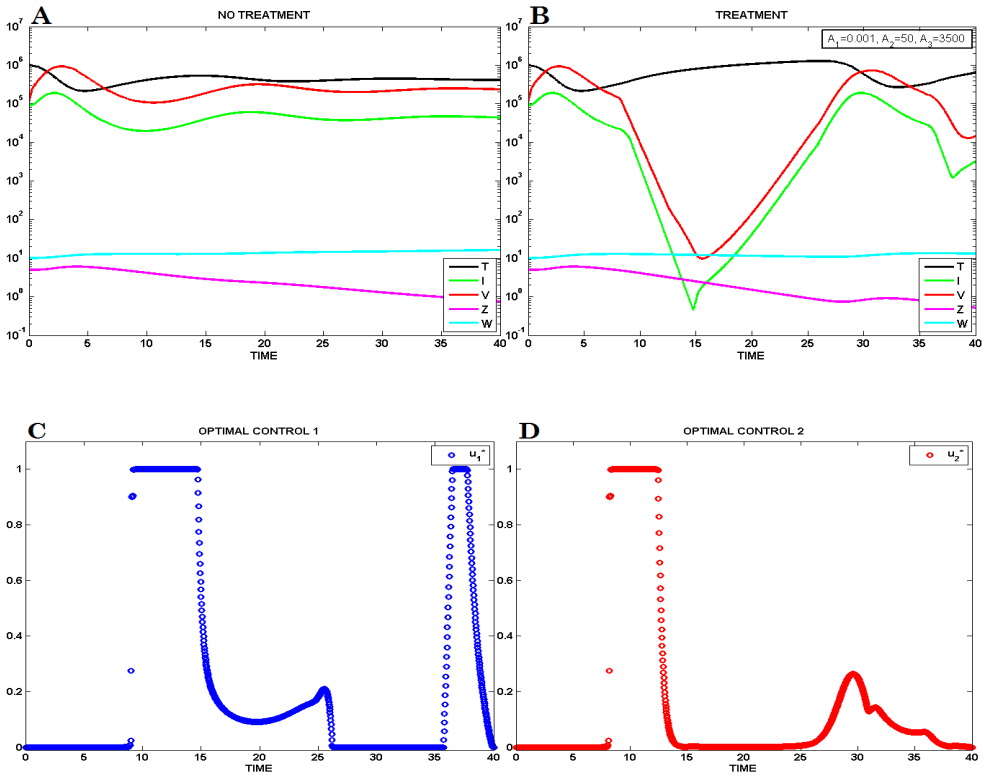

Figure 24: Numerical solutions and best treatment strategies. (A) shows the numerical solutions of the system with no treatment. (B) shows the numerical solutions of the system with treatment. (C) and (D) describe the best treatment strategies. Parameter were chosen as follows: $s = 1 \times 10^5, d = 5 \times 10^{-2}, \beta = 8 \times 10^{-7}, a = 1.8, p = 6.4 \times 10^{-4}, k = 8, \mu = 1.5, q = 5 \times 10^{-4}, g = 8 \times 10^{-8}, c = 6 \times 10^{-6}, h = 0.01, b = 0.1, A_1 = 001, A_2 = 50, A_3 = 3500, t_f = 40.$

In Figure 24 (C) and (D), the controls u_1^* and u_2^* represent the drug administration schedule. The model suggests that u_1^* must be administered in full scale from $t = 9$ to $t = 15$, then it must be reduced to much lower levels and turned off from $t = 26$ to $t = 36$. Finally, it has to be administered again for a short period of time from $t = 36$ to $t = 38$ and eventually tapered off again. On the other hand, u_2^* must start being fully administered from $t = 8$ to $t = 14$, it is turned off until $t = 26$ and finally for values from $t = 26$ to $t = 40$, this drug must be administered in levels no larger than 40% of the total dosage.

We computed a continuation of the numerical solutions letting the controls being activated from $t = 0$ to $t = 40$ and then letting them be zero from $t = 40$ to $t = 150$. In Figure 25 (B), we can see how after cessation of treatment the viral load and infected cells rebound to their pretreatment levels.

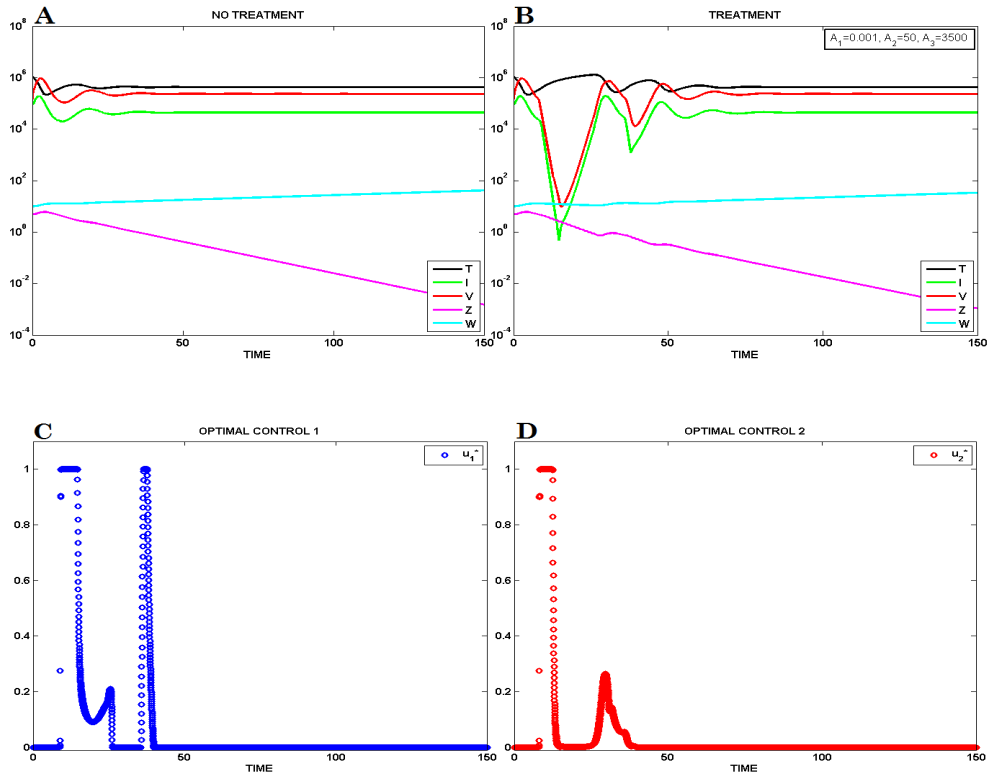

Figure 25: Numerical solutions after cessation of therapy. A continuation of the first treatment schedule was computed. We let u_1 and u_2 to be activated in the time window $[0, 40]$. Then we let $u_1 = 0$ and $u_2 = 0$ for $t \in [40, 150]$. (A) numerical solutions of the system with no treatment; (B) levels of virus and infected cells rebound after cessation of therapy at $t = 40$; (C) and (D) are the best treatment schedules.

It is clear that the effect of the therapy in this scenario is not as promising as in the CTL dominance scenario. In fact, in Table 9 we can notice that with this treatment schedule the production of new virus does not decrease in orders of magnitude. However, a reduction of one order of magnitude in the number of newly infected cells can be observed in Table 10.

Table 9: The total production of virus was computed to compare the dominant antibody response scenario under treatment and with no treatment

	No Treatment	Treatment
Total production	$\int_0^{40} kI\,dt = 1.67 \times 10^7$	$\int_0^{40} (1-u_2)kI\,dt = 1.24 \times 10^7$
of new virus	$\int_0^{150} kI\,dt = 5.52 \times 10^7$	$\int_0^{150} (1-u_2)kI\,dt = 5.12 \times 10^7$

Table 10: The total production of infected cells was computed to compare the dominant antibody response scenario under treatment and with no treatment

	No Treatment	Treatment
Total production	$\int_0^{40} \beta VT dt = 3.72 \times 10^6$	$\int_0^{40} (1-u_1)\beta VT dt = 3.00 \times 10^6$
of infected cells	$\int_0^{150} \beta VT dt = 1.23 \times 10^7$	$\int_0^{150} (1-u_1)\beta VT dt = 1.17 \times 10^7$

Figure 26 shows the comparison of the viral load with treatment and with no treatment for the dominant antibody response scenario. The rebound of $V(t)$ during treatment explains why the production of new viruses, shown in Table 9, does not decrease drastically as in the dominant CTL scenario. In the long-term it is neither worse nor better, because steady state levels are the same.

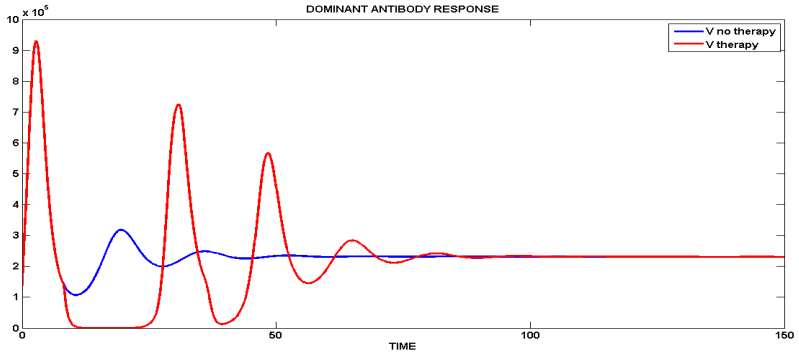

Figure 26: Comparison of the viral load under treatment and without treatment in a non logarithmic scale for the dominant antibody response scenario. The viral load with a non logarithmic scale during treatment (red curve) and with no treatment (blue curve) show that on the short term, the viral load oscillates about the equilibrium point. In the long-term, the values of the viral load during treatment and with no treatment stay at the equilibrium point.

4.6 Optimal Treatment Strategy for the Coexistence Scenario

Under the assumption that both CTL and antibody responses develop, we want to minimize the number of viruses $V(t)$, the number of infected cells $I(t)$, and the "cost" based on the effect of the therapy to the body. For this reason, the objective functional to be minimized is

$$J(u_1, u_2) = \int_0^{t_f} \left(A_1 I(t) + A_2 V(t) + \frac{1}{2} \left(A_3 u_1^2(t) + A_4 u_2^2(t) \right) \right) dt$$

subject to the system

$$\begin{aligned}
\frac{dT}{dt} &= s - dT - (1-u_1)\beta VT \\
\frac{dI}{dt} &= (1-u_1)\beta VT - aI - pIZ \\
\frac{dV}{dt} &= (1-u_2)kI - \mu V - qVW \\
\frac{dZ}{dt} &= cIZ - bZ \\
\frac{dW}{dt} &= gVW - hW
\end{aligned}$$

The parameters A_1, A_2, and A_3 balance the size terms. With higher cost parameters A_2 and A_3, the system has controls where maximum treatment is continued for a shorter period of time. In other words, with large values of these parameters the cost effect of treatment is virtually not important. The severity of therapy in the human body is described by the terms u_1^2 and u_2^2. The reason for considering a finite time window is that the administration of treatment is usually restricted to a limited time period. Necessary conditions are derived by using the extension of Pontryagin's Maximum Principle [22] given in Theorem 4.2.

Theorem 4.5 *Given the optimal control pair* $(u_1^*(t), u_2^*(t))$ *and solutions* T^*, I^*, V^*, Z^*, W^* *of the state system, there exists adjoint variables* $\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5$ *satisfying:*

$$\begin{aligned}
\lambda_1' &= \lambda_1\left(d + (1-u_1)\beta V\right) - \lambda_2(1-u_1)\beta V \\
\lambda_2' &= -A_1 + \lambda_2(a + pZ) - \lambda_3 k(1-u_2) - \lambda_4 cZ \\
\lambda_3' &= -A_2 + \lambda_1(1-u_1)\beta T - \lambda_2(1-u_1)\beta T + \lambda_3(\mu + qW) - \lambda_5 gW \\
\lambda_4' &= \lambda_2 pI + \lambda_4(b - cI) \\
\lambda_5' &= \lambda_3 qV + \lambda_5(h - gV)
\end{aligned}$$

with $\lambda_1(t_f) = \lambda_2(t_f) = \lambda_3(t_f) = \lambda_4(t_f) = \lambda_5(t_f) = 0$*, which are the transversality conditions. Furthermore, the optimal controls are characterized by*

$$\begin{aligned}
u_1^* &= \mathit{min}\left\{1, \mathit{max}\left\{0, \frac{(\lambda_2 - \lambda_1)\beta V^* T^*}{A_3}\right\}\right\} \\
u_2^* &= \mathit{min}\left\{1, \mathit{max}\left\{0, \frac{\lambda_3 k I^*}{A_4}\right\}\right\}
\end{aligned}$$

Proof:

The Hamiltonian is defined as follows:

$$\begin{aligned}
H(t, \boldsymbol{x}(t), u_1(t), u_2(t), \boldsymbol{\lambda}(t)) &= A_1 I + A_2 V + \frac{1}{2}\left(A_3 u_1^2(t) + A_4 u_2^2(t)\right) \\
&\quad + \lambda_1(s - dT - (1-u_1(t))\beta V T) \\
&\quad + \lambda_2((1-u_1(t))\beta V T - aI - pIZ) \\
&\quad + \lambda_3((1-u_2(t))kI - \mu V - VW) \\
&\quad + \lambda_4(cIZ - bZ) \\
&\quad + \lambda_5(gVW - hW)
\end{aligned}$$

We obtain the adjoint equations:

$$\begin{aligned}
\lambda_1' &= -\frac{\partial H}{\partial T} = \lambda_1\left(d + (1-u_1)\beta V\right) - \lambda_2(1-u_1)\beta V \\
\lambda_2' &= -\frac{\partial H}{\partial I} = -A_1 + \lambda_2(a + pZ) - \lambda_3 k(1-u_2) - \lambda_4 cZ \\
\lambda_3' &= -\frac{\partial H}{\partial V} = -A_2 + \lambda_1(1-u_1)\beta T - \lambda_2(1-u_1)\beta T + \lambda_3(\mu + qW) - \lambda_5 gW \\
\lambda_4' &= -\frac{\partial H}{\partial Z} = \lambda_2 pI + \lambda_4(b - cI) \\
\lambda_5' &= -\frac{\partial H}{\partial W} = \lambda_3 qV + \lambda_5(h - gV)
\end{aligned}$$

with $\lambda_1(t_f) = \lambda_2(t_f) = \lambda_3(t_f) = \lambda_4(t_f) = \lambda_5(t_f) = 0$. In order to obtain the optimality condition, we maximize the Hamiltonian with respect to u_1 and u_2 at the optimal u_1^* and u_2^*, respectively.
This is, $\dfrac{\partial H}{\partial u_i} = 0$ at u_i^* for $i = 1, 2$.
Therefore,

$$u_1^* = \frac{(\lambda_2 - \lambda_1)\beta V^* T^*}{A_3} \text{ and } u_2^* = \frac{\lambda 3 k I^*}{A_4}$$

By standard routines of optimal control problems with bounded controls [4, 9, 15, 17, 18] the characterizations of the optimal controls are given by:

$$\begin{aligned}
u_1^* &= \min\left\{1, \max\left\{0, \frac{(\lambda_2 - \lambda_1)\beta V^* T^*}{A_2}\right\}\right\} \\
u_2^* &= \min\left\{1, \max\left\{0, \frac{\lambda_3 k I^*}{A_3}\right\}\right\}
\end{aligned}$$

□

The forward-backward-sweep method was used to find the numerical solutions of the optimal control problem in the time window $[0, t_f]$ with $t_f = 15$. These numerical solutions are shown in Figure 27.

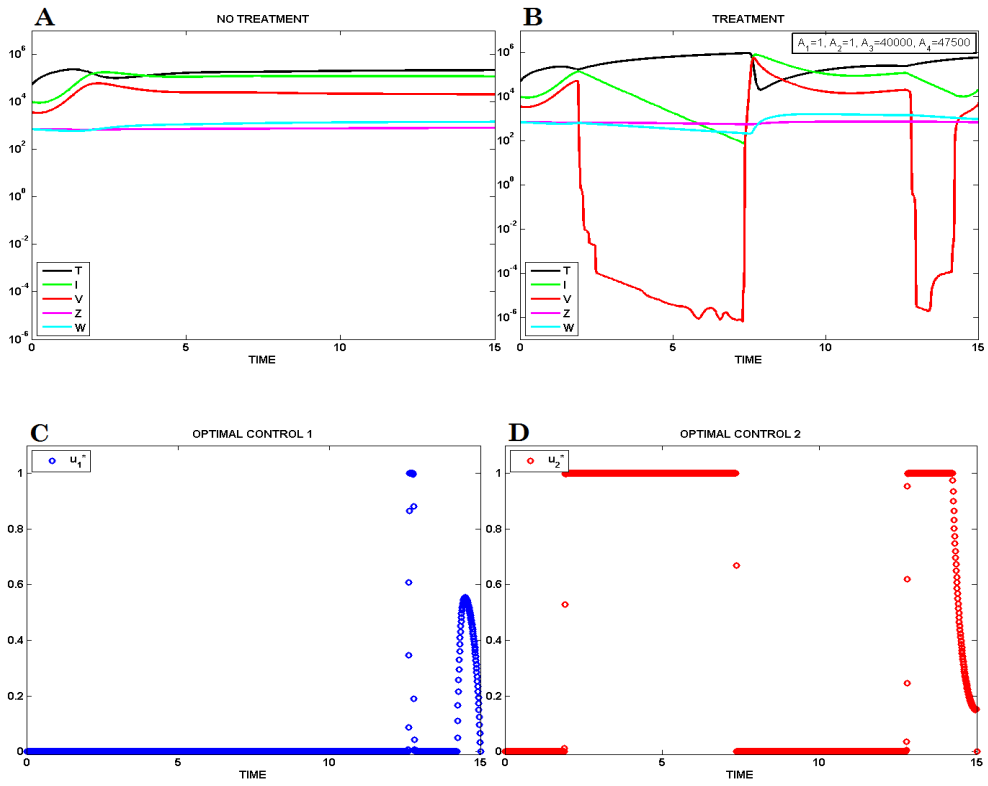

Figure 27: Numerical solutions and best treatment strategies. (A) shows the numerical solutions of the system with no treatment. (B) shows the numerical solutions of the system with treatment. (C) and (D) describe the best treatment strategies. Parameters were chosen as follows: $s = 2.0 \times 10^5, d = 1.0 \times 10^{-1}, \beta = 4.0 \times 10^{-5}, a = 9.9 \times 10^{-1}, p = 6.4 \times 10^{-4}, k = 5.0 \times 10^2, \mu = 2.9 \times 10^0, q = 2.0 \times 10^0, g = 1.0 \times 10^{-5}, c = 4.4 \times 10^{-7}, h = 2.0 \times 10^{-1}, b = 4.0 \times 10^{-2}, A_1 = 1.0 \times 10^0, A_2 = 1.0 \times 10^0, A_3 = 4.0 \times 10^4, A_4 = 4.7 \times 10^4, t_f = 1.5 \times 10^1$. The large values taken for the parameters A_3 and A_4 indicate the cost of the therapy on the human body. In other words, the side effects of the therapy are assumed to be substantial.

Figure 27 (A) shows the numerical solutions of the model with no treatment. The parameter values and initial conditions used in this calculations were the same used in Section 2.1.3. Figure 27 (C) and (D) represent the drug administration schedule. The model suggests that drug represented by the control function u_2^* must be administered in full scale from $t = 2$ to $t = 7$. After this point, this drug is tapered off and fully administered again from $t = 12.7$ to $t = 14.3$. This regime will drop the number of viruses by about 10 orders of magnitude. On the other hand, the treatment represented by the control u_1^* is hardly used and it has to be administered just for a short period of time at $t = 14$. For this reason, the level of infected cells does not decrease as much as the viral load. In Figure 27 (B), we have displayed the numerical solutions of the optimal control problem during a time horizon from $t = 0$ to $t = 15$. These solutions show the outcomes of applying the suggested treatment schedule. After cessation of treatment, the viral load and the number of infected cells return to their pretreatment levels. This is illustrated in Figure 28 (B), which shows the continuation of the numerical solutions of the optimal control problem after cessation of therapy.

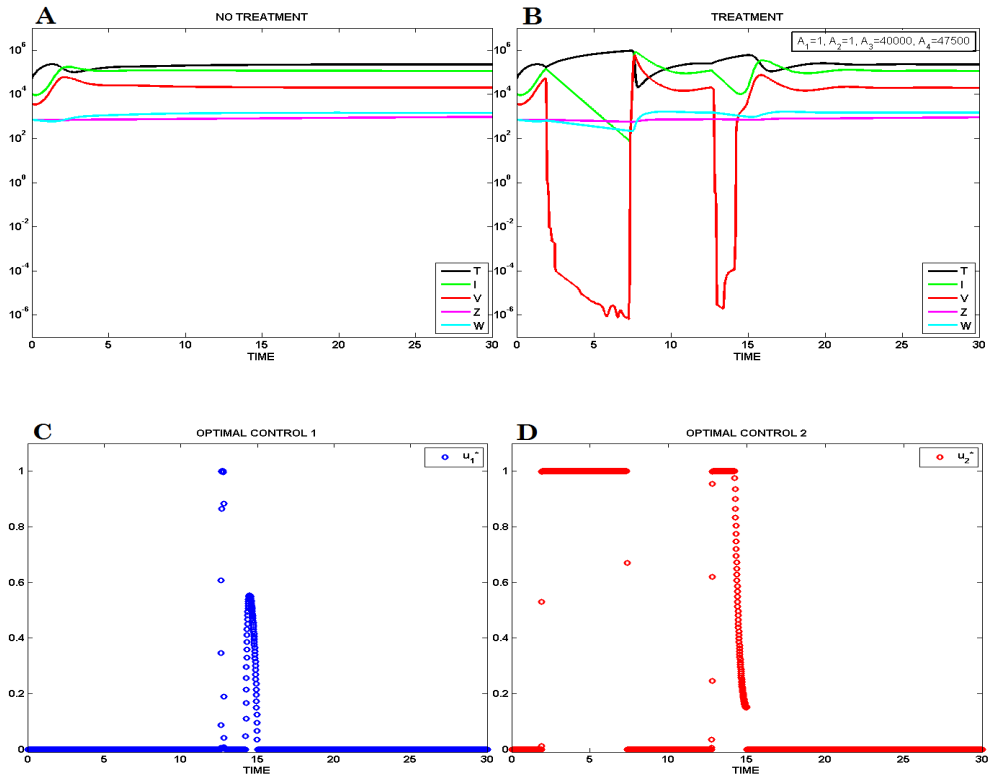

Figure 28: Numerical solutions after cessation of therapy. A continuation of the first treatment schedule was computed. We let u_1 and u_2 to be activated in the time window $[0, 15]$. Then we let $u_1 = 0$ and $u_2 = 0$ for $t \in [15, 30]$. (A) shows the numerical solutions of the system with no treatment. (B) shows how the levels of virus and infected cells rebound after cessation of therapy at $t = 15$. (C) and (D) describe the best treatment strategies.

Tables 11 and 12 show that there is not a significant reduction in the number of newly formed viruses and infected cells during treatment. In fact, these numbers do not change in orders of magnitude regardless of the administration of treatment. In order to understand such a contradicting result, we refer the reader to Figure 29, that shows the viral load in a non logarithmic scale during treatment and with no treatment. From this figure, it is clear that even though the number of viruses was reduced during therapy, when the treatment is tapered off, $V(t)$ rebounds to its pretreatment levels through damp oscillations. In some of these oscillations, the viral load reaches numbers much larger than in the non treatment scenario. Again, the benefit of therapy is just observed during short periods of time.

Table 11: The total production of virus in a non logarithmic scale was computed to compare the coexistence scenario under treatment and with no treatment

	No Treatment	Treatment
Total production	$\int_0^{15} kI\,dt = 8.40 \times 10^8$	$\int_0^{15} (1-u_2)kI\,dt = 5.86 \times 10^8$
of new virus	$\int_0^{30} kI\,dt = 1.70 \times 10^9$	$\int_0^{30} (1-u_2)kI\,dt = 1.55 \times 10^9$

Table 12: The total production of infected cells was computed to compare the coexistence scenario under treatment and with no treatment

	No Treatment	Treatment
Total production of infected cells	$\int_0^{15} \beta VTdt = 2.56 \times 10^6$ $\int_0^{30} \beta VTdt = 5.22 \times 10^6$	$\int_0^{15} (1-u_1)\beta VTdt = 1.90 \times 10^6$ $\int_0^{30} (1-u_1)\beta VTdt = 4.92 \times 10^6$

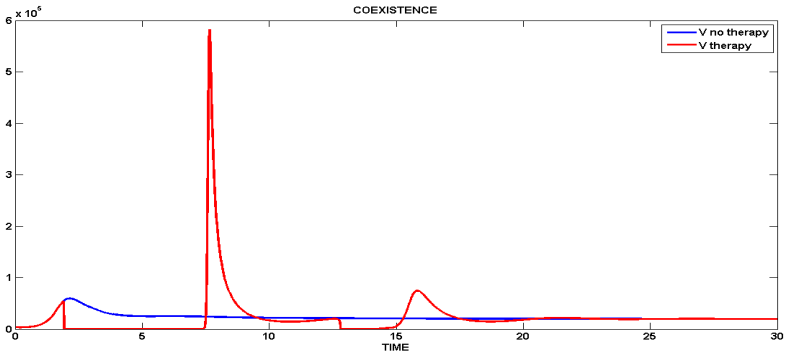

Figure 29: Comparison of the viral load under treatment and without treatment in a non logarithmic scale for the coexistence scenario. The numerical solutions of the viral load with a non logarithmic scale during treatment (red curve) and with no treatment (blue curve) show that at $t = 8$, $V(t)$ rebounds to large values. However, this rebound is just for a short period of time. Eventually, $V(t)$ decreases and in the the long-term, the values of the viral load during treatment and with no treatment stay at the equilibrium point.

5 INVERSE PROBLEM: PARAMETER ESTIMATION

Experimental data sets exhibiting HCV RNA levels from chronic HCV patients treated with several types of drugs can be found in existing literature [6, 12, 19]. With these data sets, one can estimate a set of parameters that gives the best fit of the data with a model. This problem is called the inverse problem. We address this problem using an ordinary least squares (OLS) method to find a first estimation of the parameters and then with a bootstrap method to find a new set of parameter estimates and their corresponding standard errors.

5.1 Ordinary Least Squares

The formulation of the standard ordinary least squares usually involves two models: a mathematical model and a statistical model [1]. Assume we have a problem modeled by a system of differential equations of the form

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{f}(t, \boldsymbol{x}(t, \boldsymbol{\theta}), \boldsymbol{\theta})$$

where $x(t, \boldsymbol{\theta}) \in \mathbb{R}^n$ denotes the state variable vector at time t and $\boldsymbol{\theta} \in \mathbb{R}^p$ denotes the parameter vector. The output of the mathematical model is denoted by $z(t_i, \boldsymbol{\theta_0})$. In this case, $z(t_i, \boldsymbol{\theta_0})$ is a functional of the state variable $\boldsymbol{x}(t, \boldsymbol{\theta})$, that is, $z(t_i, \boldsymbol{\theta_0}) \in \mathscr{F}(x(t, \boldsymbol{\theta}))$ and $\boldsymbol{\theta_0}$ is the true parameter vector. The statistical model for the observation process is given by the random variables

$$Y_i = z(t_i, \boldsymbol{\theta_0}) + \mathscr{E}_i \qquad \text{for } i = 1, 2, \cdots, n.$$

For example, in our model the statistical model for the observation process is

$$Y_i = \ln\left(V(t_i, \boldsymbol{\theta_0})\right) + \mathscr{E}_i$$

The errors $\mathscr{E}_i$ are assumed to be random variables that satisfy [1]:

1. $\mathscr{E}_i$ for for $i = 1, 2, \cdots, n$ are independent and identically distributed random variables. This is, $\text{cov}(\mathscr{E}_i, \mathscr{E}_j) = 0$ whenever $i \neq j$.

2. $E[\mathscr{E}_i] = 0$ for every i.

3. $\text{var}(\mathscr{E}_i) = \sigma_0^2 < \infty$, for every i.

We use the ordinary least squares method to minimize $[Y_i - \ln(V(t_i, \boldsymbol{\theta}))]^2$ over the set of parameter vectors $\boldsymbol{\theta}$ constrained by a pre-specified feasible region denoted by Θ. The minimizer is a random variable, called the estimator $\boldsymbol{\theta}_{OLS}$ given by

$$\boldsymbol{\theta}_{OLS} = \mathsf{argmin}_{\boldsymbol{\theta} \in \Theta} \sum_{i=1}^{n} [Y_i - \ln(V(t_i, \boldsymbol{\theta}))]^2$$

The theoretical quantities $\boldsymbol{\theta_0}$ and σ_0^2 are in general unknown. In practice, one has only the data associated with a single realization y_i of the observation process Y_i for $i = 1, \cdots, n$ and has no option but to compute an estimate for $\hat{\boldsymbol{\theta}}_{OLS}$. Using the genetic algorithm in MATLAB we carry out the following minimization process:

$$\hat{\boldsymbol{\theta}}_{OLS} = \mathsf{argmin}_{\boldsymbol{\theta} \in \Theta} \sum_{i=1}^{n} [y_i - \ln(V(t_i, \boldsymbol{\theta}))]^2$$

The reader can find more details and examples of using the OLS method for parameter estimation in [1].

5.2 The Bootstrap Method

The bootstrap method not only gives us parameter point-estimates but also an interval of possible values or confidence intervals. The bootstrap method uses an empirical distribution for the statistic based just on the data at hand. This empirical distribution can be generated through resampling. Resampling takes random samples with replacement(of the same size) from the original sample.

The following is the outline of the algorithm [2] :

1. Let $\hat{\boldsymbol{\theta}}^{(0)}$ to be the vector of parameters estimates from the entire dataset $\{y_1, y_2, \cdots, y_n\}$ using OLS.

2. Using this estimate we define the standard residuals:

$$\overline{r}_j = \sqrt{\frac{n}{n-p}} \left(y_j - \ln \left(V(t_i, \hat{\boldsymbol{\theta}}^{(0)}) \right) \right)$$

   for $j = 1, \cdots, n$. Then $\{\overline{r}_1, \cdots, \overline{r}_n\}$ are realizations of independent identically distribute ($i.i.d$) random variables $\overline{R}_j$ from an empirical distribution $\mathscr{F}_n$.

3. Set $m = 0$ and create a bootstrap sample of size n using random sampling with replacement from $\{\overline{r}_1, \cdots, \overline{r}_n\}$ to form a bootstrap sample $\left\{ \overline{r}_1^{(m)}, \cdots, \overline{r}_n^{(m)} \right\}$.

4. Create bootstrap sample points (for $j = 1, \cdots, n$)

$$y_j^{(m)} = \ln \left(V(t_i, \hat{\boldsymbol{\theta}}^{(0)}) \right) + r_j^{(m)}$$

5. Obtain a new estimate $\hat{\boldsymbol{\theta}}^{(m+1)}$ from the bootstrap sample $\left\{ y_j^{(m)} \right\}$.

6. Set $m = m + 1$ and repeat steps 3-5.

We carried out the iterative process $m = 10,000$ times. Once this process is finished, we compute mean, covariance matrix, and standard errors using the following formulas:

$$\hat{\boldsymbol{\theta}}_{\mathsf{boot}} = \frac{1}{M}\sum_{m=1}^{M}\hat{\boldsymbol{\theta}}^{(m)} \tag{22}$$

$$\mathsf{Cov}\left(\hat{\boldsymbol{\theta}}_{\mathsf{boot}}\right) = \frac{1}{M-1}\sum_{m=1}^{M}\left(\hat{\boldsymbol{\theta}}^{(m)}-\hat{\boldsymbol{\theta}}_{\mathsf{boot}}\right)\left(\hat{\boldsymbol{\theta}}^{(m)}-\hat{\boldsymbol{\theta}}_{\mathsf{boot}}\right)^{T} \tag{23}$$

$$\mathsf{SE}_k\left(\hat{\boldsymbol{\theta}}_{\mathsf{boot}}\right) = \sqrt{\mathsf{Cov}\left(\hat{\boldsymbol{\theta}}_{\mathsf{boot}}\right)_{kk}} \tag{24}$$

where $\left(\hat{\boldsymbol{\theta}}_{\mathsf{boot}}\right) \in \mathbb{R}^p$ is assumed to be a column vector.

To evaluate the model, we fit data sets taken from existing literature [6, 7, 12, 19]. First of all, we find point estimates for the twelve parameters of the system using OLS. Secondly, we run the bootstrap method to find new point estimates along with their standard errors. However, the problem of finding the point estimates and intervals of confidence for these parameters having only data sets of the viral load is a complex problem due to the lack of information on the other variables that are involved. For this reason, for some parameters we were able to find only point estimates; their standard errors were not confident enough.

Different HCV RNA profiles of patients under treatment have been observed, namely: rebound to baseline values, biphasic decay, triphasic decay, and flat second phases after cessation of therapy. The model given by equations (1)-(5) could be fit to all of these profile observations with exception of the triphasic decay. The consideration of a logistic term in the equations for the proliferation of hepatocytes may yield to a more realistic model that can also be fit to a triphasic decay data [6]. However, the inclusion of such a term would imply an increase in the number of parameters. More details about the bootstrapping method can be found at [2].

5.2.1 Parameter Estimation for Patient 1

The data set taken from Neumann *et al.* [19] was used to find point estimates of the parameters. A MATLAB code was run to solve an OLS optimization problem using the genetic algorithm. Initial conditions for the uninfected and infected cells were also computed in the optimization process. With these first point estimates, a bootstrap method was computed with 10,000 iterations to get a new set of parameter estimates with their corresponding standard errors. Point estimates and their standard errors are summarized in Table 13.

Table 13: Bootstrap estimates and standard errors for patient 1

Parameter	**Bootstrap Estimate**	**Standard Error (SE)**
s	8.7386×10^{3}	t.l.n.r.
d	1.0700×10^{-2}	1.5000×10^{-2}
β	1.9677×10^{-5}	2.0170×10^{-5}
a	1.6010×10^{-1}	1.5090×10^{-1}
p	6.5859×10^{-4}	7.7706×10^{-4}
k	5.4780×10^{-1}	2.0880×10^{-1}
μ	2.8595×10^{0}	4.6500×10^{-1}
q	9.5861×10^{-4}	t.l.n.r.
c	4.5438×10^{-7}	t.l.n.r.
b	1.8060×10^{-1}	t.l.n.r.
g	9.6712×10^{-7}	t.l.n.r.
h	6.2800×10^{-1}	9.7510×10^{-1}

Table 13 shows the parameter estimates and standard errors found in the bootstrapping. For some parameters, large standard errors were obtained. This is not surprising due to the large variability that was observed for some data points in the cloud provided in Figure 30. This large variability is due to the fact that the observations in the data points vary in several orders of magnitude. This explains the fact that for the data points that were taken after day 5 we observe larger variability. These issues reveal the limitations of the OLS method. The use of generalized ordinary least squares (GOLS) [1] might be a possible solution to this weakness in our method.

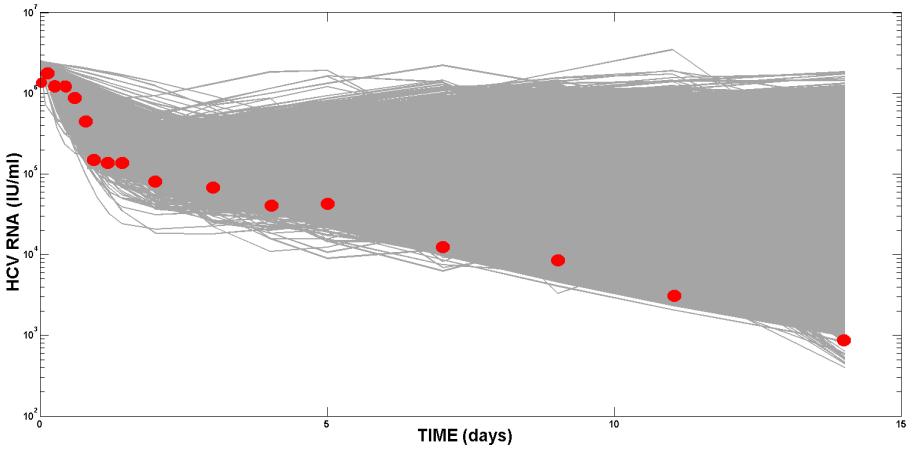

Figure 30: Bootstrap cloud for patient 1. Running the bootstrap algorithm for 10,000 iterations a cloud of best fit solutions is plotted. Parameter estimates and standard errors are calculated after the iterative process.

For those parameters with standard errors that are larger than the estimation, only their point estimates are shown on Table 13. Instead we use the abbreviation "t.l.n.r.", which stands for "too large not reported". In Figure 30, the best fits found with the bootstrap and the data points were plotted. The thickness of the cloud indicates large variability of the data points. This variability is larger for those measurements that were taken after day 5.

It is important to notice that since the data set was taken from a patient that has been under treatment for 14 days, the parameter estimates found are affected by the treatment. For this reason, even though in Figure 31 the numerical solutions show that patient 1 has no immune response, with the information we have we are not able to claim that this is the case.

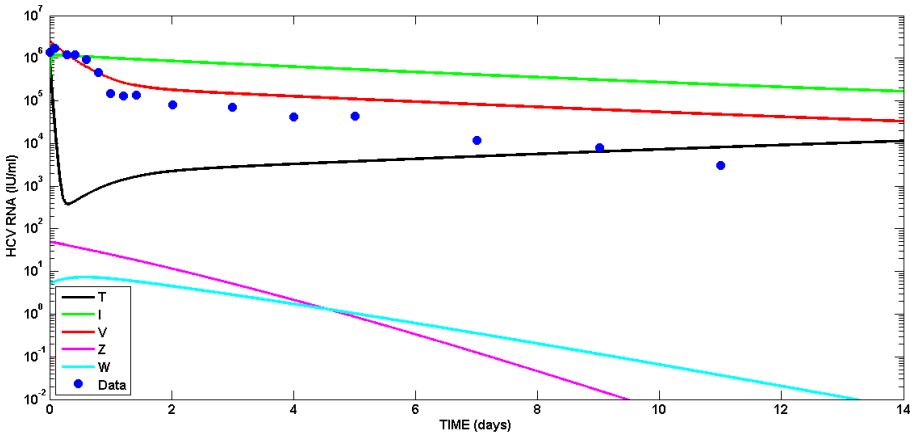

Figure 31: Comparison of viral load data with model predictions. We fit HCV RNA levels from a chronic HCV infected patient treated with interferon $\alpha - 2b$ from Neumann *et al.*[19]. Numerical solutions of system given by equations (1)-(5) were calculated using the set of parameters given in Table 30. Initial conditions were chosen as follows: $T(0) = 1.0 \times 10^6, I(0) = 2.0 \times 10^5, V(0) = 2.5 \times 10^6, Z(0) = 5.0 \times 10^1, W(0) = 5.0 \times 10^0$.

5.2.2 Parameter Estimation for Patient 2

The model given by equations (1)-(5) is not exclusive for HCV infections. Other type of viruses can also be understood through this model. As a matter of fact, data from a hepatitis B infected patient during drug therapy was taken from Dahari *et al.* [7] to fit the model and to estimate parameters and their corresponding standard errors. As stated at the beginning of this section, a first iteration of the OLS method with the genetic algorithm was run in MATLAB to find the point estimates for the twelve parameters of the system. Initial conditions for the uninfected and infected cells were also computed in the optimization process. With this set of parameter estimates, called $\hat{\theta}^{(0)}$ in the bootstrap algorithm, we run 10,000 iterations to find new parameter estimates with their corresponding standard errors. For this data set, we obtained decent standard errors for ten out of the twelve parameters. For parameters d and p that have standard errors that are larger than their estimation, only their point estimates are shown on Table 14. Instead we use the abbreviation "t.l.n.r.", which stands for "too large not reported". In other words, we do not report standard errors with corresponding relative errors bigger than one.

Figure 32 shows the best fits solutions found with the bootstrap and the data points. Just like for patient 1, the cloud tends to be thicker for those measurements that were taken after day 5.

Table 14: Bootstrap estimates and standard errors for patient 2

Parameter	**Bootstrap Estimate**	**Standard Error (SE)**
s	1.2025×10^{4}	1.5652×10^{4}
d	8.0000×10^{-3}	t.l.n.r.
β	9.3000×10^{-3}	2.9000×10^{-3}
a	7.7400×10^{-2}	3.0700×10^{-2}
p	4.8000×10^{-3}	t.l.n.r.
k	7.8300×10^{-2}	1.4900×10^{-2}
μ	4.9410×10^{-1}	3.0430×10^{-1}
q	1.6659×10^{-4}	9.9770×10^{-5}
c	3.0638×10^{-7}	7.1389×10^{-8}
b	3.5750×10^{-1}	2.3820×10^{-1}
g	6.6439×10^{-7}	8.7652×10^{-7}
h	7.2260×10^{-1}	1.5590×10^{-1}

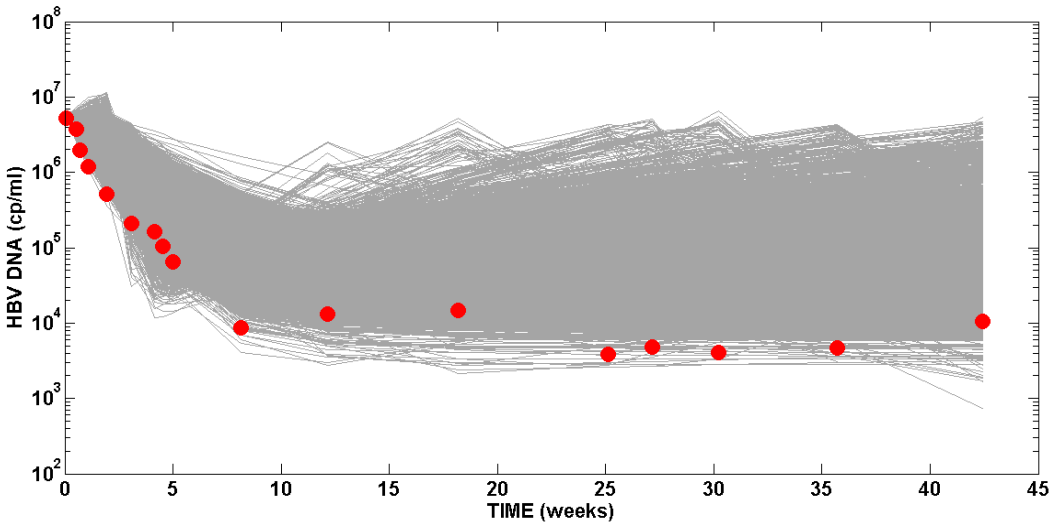

Figure 32: Bootstrap cloud for patient 2. Running the bootstrap algorithm for 10,000 iterations a cloud of best fit solutions is plotted. Parameter estimates and standard errors are calculated after the iterative process.

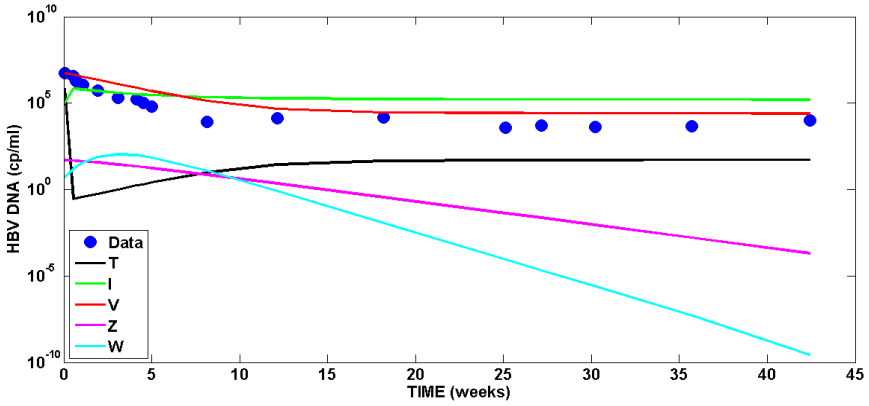

Figure 33: Comparison of viral load data with model predictions. We fit HBV DNA levels from patient that was treated with pegylated interferon-2αa [7]. Numerical solutions for patient 2 were calculated using the set of parameters given in Table 14. Initial conditions were chosen as follows: $T(0) = 7.0 \times 10^5, I(0) = 9.7 \times 10^4, V(0) = 5.5 \times 10^6, Z(0) = 50, W(0) = 5$.

Figure 33 shows the data points and the numerical solutions of the system (1)-(5) using the set of parameters shown in Table 14. Observations made for patient 1 are also valid for patient 2. According to this figure and to the values of the estimates, this looks to be an example of an endemic equilibrium with no immune response (second equilibrium seen in Section 2). However, this could be misleading because the data set is taken from a patient that has been under treatment and, as a consequence, some parameters are affected. A possible approach to solve the problem of detecting a difference between the parameters and the treatment that has been applied is solving the inverse problem simultaneously with an optimal control problem. Nevertheless, this process is computationally intensive.

5.2.3 Parameter Estimation for Patient 3

HCV RNA levels from a patient treated with interferon $\alpha - 2b$ were taken from Neumann *et al.*[19]. A first iteration of the OLS optimization method with the genetic algorithm was run in MATLAB to find the point estimates for the twelve parameters of the system. Initial conditions for the uninfected and infected cells were also computed in the optimization process. With this set of parameter estimates, called $\hat{\theta}^{(0)}$ in the bootstrap algorithm, we run 10,000 iterations of this method to find the new parameter estimates and their corresponding standard errors. Point estimates and their standard errors are summarized in Table 15. For those parameters with standard errors that are larger than the estimation, only their point estimates are shown on Table 15. Instead we use the abbreviation "t.l.n.r.", which stands for "too large not reported".

Table 15: Bootstrap estimates and standard errors for patient 3

Parameter	**Bootstrap Estimate**	**Standard Error (SE)**
s	5.5687×10^{4}	8.0768×10^{4}
d	9.5042×10^{-4}	t.l.n.r.
β	5.0346×10^{-4}	t.l.n.r.
a	6.3480×10^{-1}	t.l.n.r.
p	9.7000×10^{-3}	t.l.n.r.
k	2.3581×10^{1}	2.5610×10^{1}
μ	3.4945×10^{0}	1.2655×10^{0}
q	1.8000×10^{-3}	8.3000×10^{-3}
c	1.0767×10^{-7}	8.2202×10^{-7}
b	1.1361×10^{0}	1.4274×10^{0}
g	6.8124×10^{-7}	t.l.n.r.
h	9.4230×10^{-1}	t.l.n.r.

In Figure 34 the best fits found with the bootstrap and the data points were plotted. The thickness of the cloud indicates large variability of the data points.

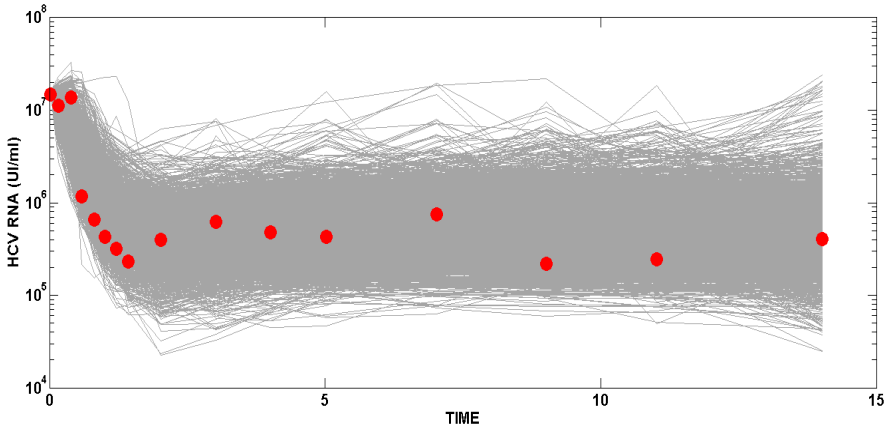

Figure 34: Bootstrap cloud for patient 3. Running the bootstrap algorithm for 10,000 iterations a cloud of best fit solutions is calculated to later compute the standard errors.

Figure 35 shows the numerical solutions of the system (1)-(5) using the set of parameters found in Table 15. It also shows the data set taken from [19]. According to this figure and to the values of the estimates, this looks to be an example of an endemic equilibrium with no immune response (second equilibrium seen in Section 2). However, this could be misleading because the data set is taken from a patient that has been under treatment and as a consequence some parameters are affected. A possible approach to solve the problem of making a difference between the parameters and the treatment that has been applied could be solving the inverse problem along with an optimal control problem. However this process is computationally intensive and therefore out of our possibilities.

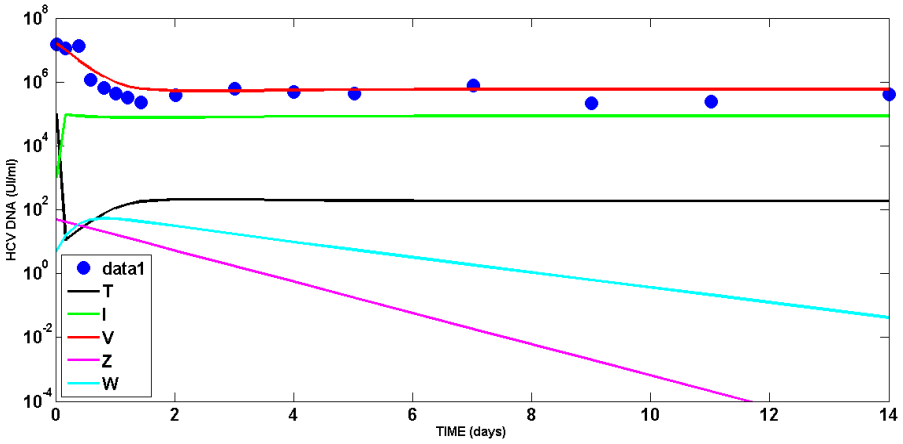

Figure 35: Comparison of viral load data with model predictions. We fit HCV RNA levels from patient that was treated with interferon-2αb [19]. Numerical solutions for patient 3 were calculated using the set of parameters given in Table 15. Initial conditions were chosen as follows: $T(0) = 1.0 \times 10^5, I(0) = 1.0 \times 10^3, V(0) = 1.6 \times 10^7, Z(0) = 50, W(0) = 5$.

5.2.4 Parameter Estimation for Patient 4

Data from a hepatitis B infected patient during drug therapy was taken from Dahari *et al.* [7]. A first iteration of the OLS method with the genetic algorithm was run in MATLAB to find the point estimates for the twelve parameters of the system. Initial conditions for the uninfected and infected cells were also computed in the optimization process. With this set of parameter estimates, called $\hat{\theta}^{(0)}$ in the bootstrap algorithm, we run 10,000 iterations to find new parameter estimates with their corresponding standard errors. For this data set, we obtained decent standard errors for all twelve parameters. Table 16 shows a summary of the parameter estimates and their corresponding standard errors.

Table 16: Bootstrap estimates and standard errors for patient 4

Parameter	**Bootstrap Estimate**	**Standard Error (SE)**
s	9.0505×10^{3}	2.7321×10^{3}
d	9.4000×10^{-3}	1.4000×10^{-3}
β	6.6000×10^{-3}	1.2000×10^{-3}
a	1.7000×10^{-2}	4.0000×10^{-3}
p	1.1000×10^{-3}	1.5422×10^{-4}
k	3.8530×10^{-1}	8.3900×10^{-2}
μ	2.9267×10^{0}	3.0380×10^{-1}
q	1.0000×10^{-3}	1.7529×10^{-4}
c	1.5526×10^{-7}	2.3676×10^{-7}
b	2.6030×10^{-1}	5.2600×10^{-2}
g	5.2049×10^{-7}	6.4187×10^{-8}
h	9.0200×10^{-1}	1.4810×10^{-1}

In Figure 36, the best fits found with the bootstrap and the data points were plotted. From this figure we can see how the data points in the second phase have large variability compare to the first data points.

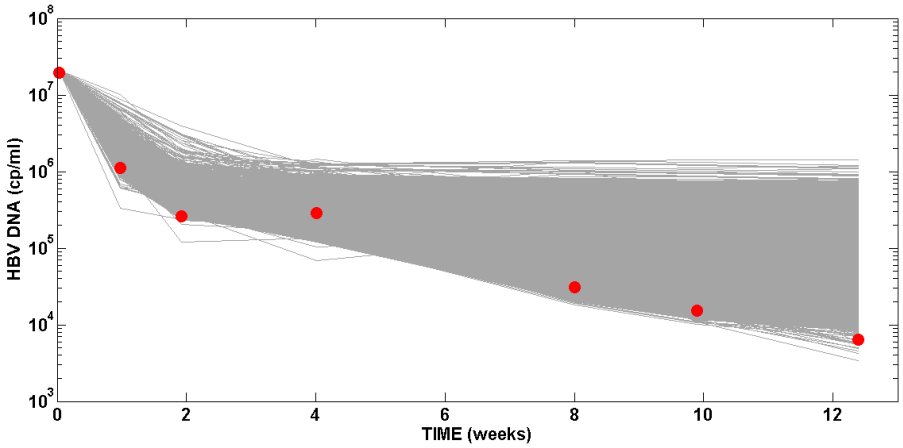

Figure 36: Bootstrap cloud for patient 4. Running the bootstrap algorithm for 10,000 iterations a cloud of best fit solutions is calculated to later compute the standard errors.

Figure 37 shows the numerical solutions of equations (1)-(5) using the parameters obtained in the bootstrapping and shown in Table 16. As explained in the previous sections, due to the fact the data set is taken from a patient under treatment, we do not have enough information to claim that there is not an immune response.

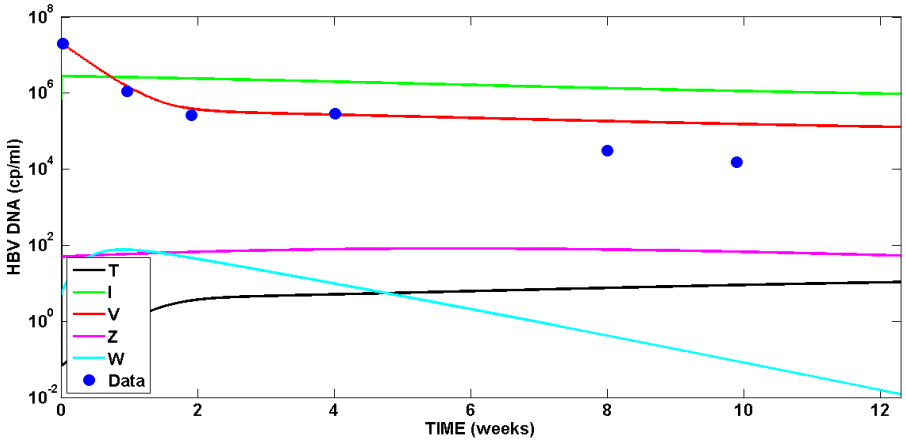

Figure 37: Comparison of viral load data with model predictions. We fit HBV DNA levels from a patient that was treated with adefovir dipivoxil [7]. Numerical solutions for patient 4 were calculated using the set of parameters given in Table 16. Initial conditions were chosen as follows: $T(0) = 2.6 \times 10^6, I(0) = 1.7 \times 10^5, V(0) = 2.1 \times 10^7, Z(0) = 50, W(0) = 5$.

6 SENSITIVITY ANALYSIS PART II

The parameters estimates found in Section 5 were used to perform a sensitivity analysis for each patient during transient and long-term phases. Just like in Section 3, we want to determine what are the most influential parameters in the behavior of the viral load for patients that have been under treatment.

6.1 Sensitivity Analysis for Patient 1

For patient 1, we found that during treatment the parameters that are the most influential are k, μ, and a. These are the natural production and natural decay of viruses and the natural decay of infected cells, respectively. In fact, the larger the values of k the larger the values of the viral load. Similarly, the larger the values of the parameters μ and a, the smaller the values that $V(t)$ will reach. Fig 38 shows the relative sensitivities of $V(t)$ in the transient phase. In Table 17, we have a summary of the most influential parameters in the behavior of the viral load in the transient phase.

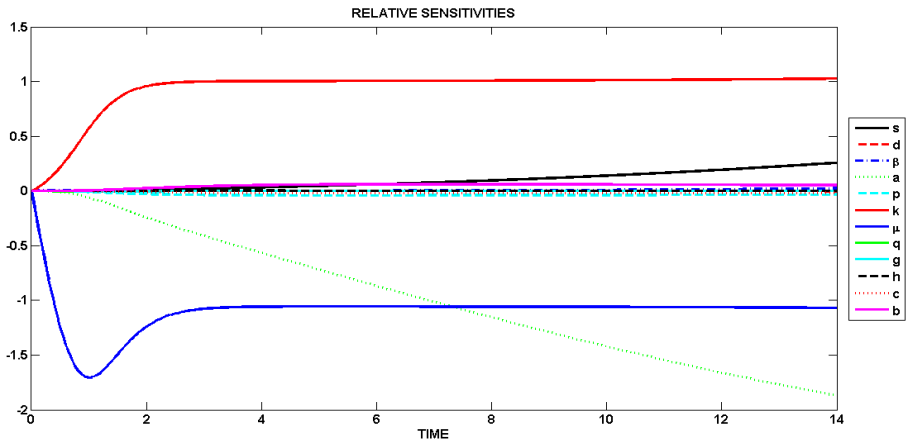

Figure 38: Numerical solutions of relative sensitivities $\frac{\theta_k}{V}\frac{\partial V}{\partial \theta_k}$ for patient 1 in a transient phase from 0 to 14

Table 17: Ranking of the most influential parameters for the viral load for patient 1 in the transient phase. During the period of time the patient is under treatment, increasing parameter k will increase the viral load. Similarly, increasing the values of parameters μ and a will produce a reduction in the number of virus particles.

	Positive	**Negative**
Parameter	k	μ a

The simulation suggests that, in a longer time window (in the long-term phase), the parameter s also becomes important in the behavior of $V(t)$. As a matter of fact, the sensitivity of s is as positive as the sensitivity of parameter k. See Figure 39 and Table 18.

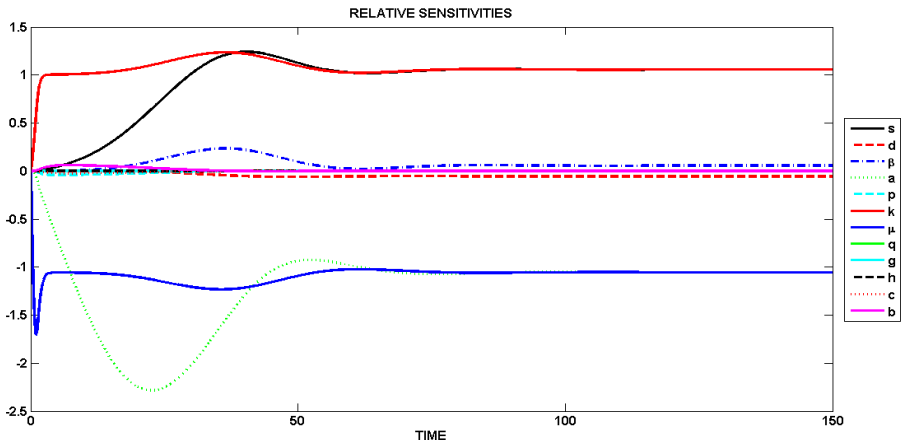

Figure 39: Numerical solutions of relative sensitivities $\frac{\theta_k}{V}\frac{\partial V}{\partial \theta_k}$ for patient 1 in a long-term phase from 0 to 150

Table 18: Ranking of the most influential parameters in the viral load for patient 1 in the long-term. The model suggests that in a long-term phase, increasing parameters k and s will increase the viral load. Similarly, increasing the values of parameters μ and a will produce a reduction in the number of virus particles.

	Positive	**Negative**
Parameter	k	μ
	s	a

6.2 Sensitivity Analysis for Patient 2

According to Figure 40, the parameters that are most influential in the behavior of $V(t)$ for patient 2 are k, s in a positive manner and μ, a negatively. However, it is important to mention that from $t = 10$ up to $t = 25$ the simulation suggests that larger values of b would increase the viral load. This seems to be counterintuitive since b is the rate of natural decay of CTLs. Table 19 shows a summary of the most important parameters in the behavior of $V(t)$ in the transient phase fr patient 2.

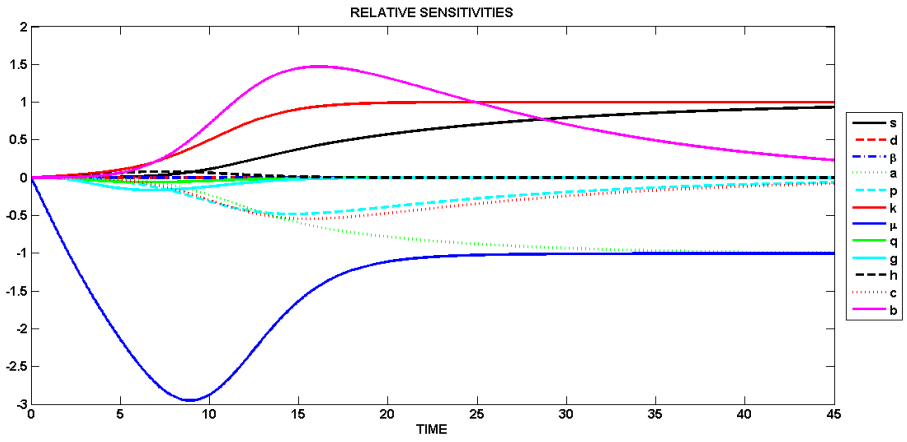

Figure 40: Numerical solutions of relative sensitivities $\frac{\theta_k}{V}\frac{\partial V}{\partial \theta_k}$ for patient 2 in a transient phase from 0 to 45

Table 19: Ranking of the most influential parameters for the viral load for patient 2 in the transient phase. During the period of time the patient is under treatment, increasing parameters b, k, and s will increase the viral load. Similarly, increasing the values of parameters μ and a will produce a reduction in the number of virus particles.

	Positive	**Negative**
Parameter	b	μ
	k	a
	s	

In Figure 41 we notice that, in the long-term the effect of b eventually becomes irrelevant. In this figure is also shown that, parameters k and s are equally influential in a positive manner. Likewise, parameters μ and a are equally influential in a negative manner. Table 20 shows a summary of these results.

Table 20: Ranking of the most influential parameters for the viral load for patient 2 in the long-term. The model suggests that in a long-term phase, increasing parameters k and s will increase the viral load. Similarly, increasing the values of parameters μ and a will produce a reduction in the number of virus particles.

	Positive	**Negative**
	k	μ
Parameter	s	a

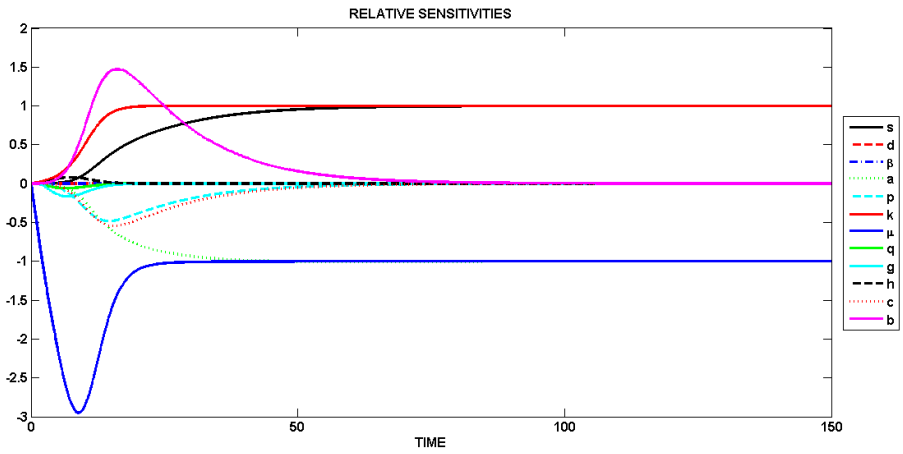

Figure 41: Numerical solutions of relative sensitivities $\frac{\theta_k}{V}\frac{\partial V}{\partial \theta_k}$ for patient 2 in a transient phase from 0 to 150

6.3 Sensitivity Analysis for Patient 3

Sensitivities were computed for patient 3 using the estimates found in Section 5. According to Figure 42, during the period of time the patient was under treatment (transient phase), the parameters found to be the most influential were k, s in a positive manner and μ, a negatively. That is, increasing the proliferation of hepatocytes and virus would make the viral load increase while the clearance rate of virus and infected cells would make $V(t)$ decrease. In Table 21, we have displayed the most influential parameters in this phase.

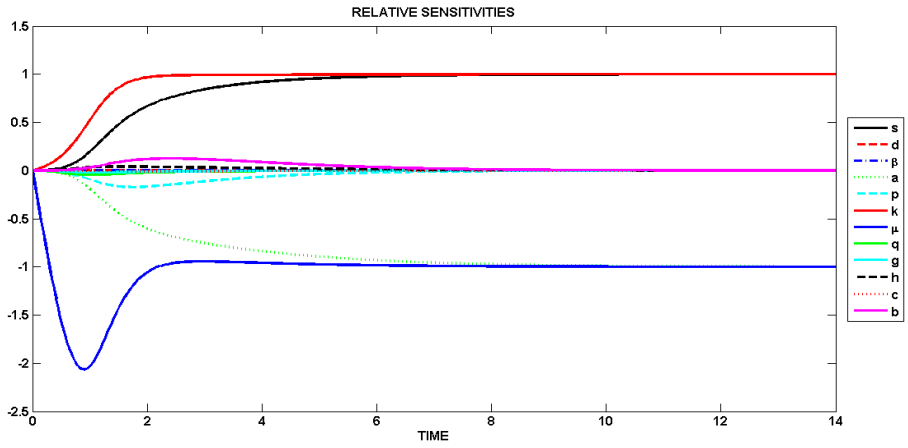

Figure 42: Numerical solutions of relative sensitivities $\frac{\theta_k}{V}\frac{\partial V}{\partial \theta_k}$ for patient 3 in a transient phase from 0 to 14

Table 21: Ranking of the most influential parameters for the viral load of patient 3 in a transient phase. During the period of time the patient is under treatment, increasing parameters k, and s will increase the viral load. Similarly, increasing the values of parameters μ and a will produce a reduction in the number of virus particles.

	Positive	**Negative**
Parameter	k	μ
	s	a

In the long-term phase the behavior of $V(t)$ is still determined by the parameters k, s, μ, and a. This is shown in Figure 43 and summarized in Table 22.

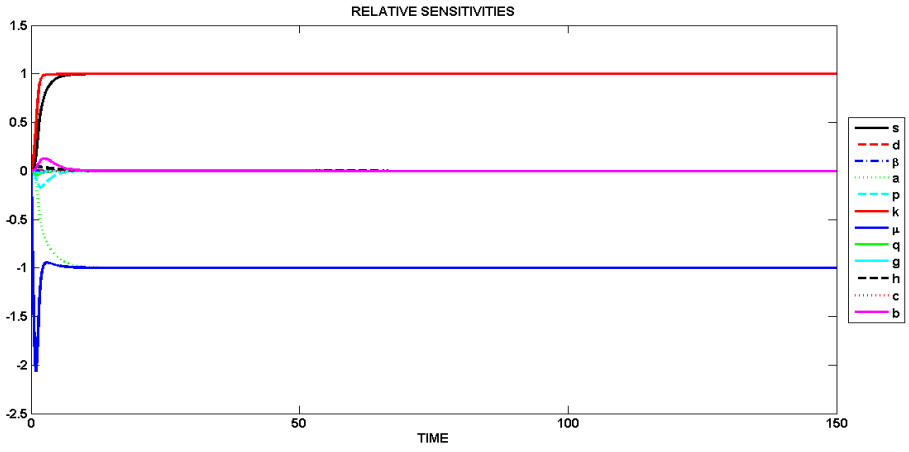

Figure 43: Numerical solutions of relative sensitivities $\frac{\theta_k}{V}\frac{\partial V}{\partial \theta_k}$ for patient 3 in a long-term phase from 0 to 150

Table 22: Ranking of the most influential parameters in the viral load for patient 3 in the long-term. The model suggests that in a long-term phase, increasing parameters k and s will increase the viral load. Similarly, increasing the values of parameters μ and a will produce a reduction in the number of virus particles.

	Positive	**Negative**
Parameter	k	μ
	s	a

6.4 Sensitivity Analysis for Patient 4

A sensitivity analysis was done using the set of parameters estimated in Section 5 for patient 4. The numerical solutions shown in Figure 44 suggest that, at the beginning of the treatment, the most influential parameters in the behavior of the viral load are k and μ. However, for $t > 5$ parameters b and c play an important role during the administration of treatment. These results are summarized in Table 23 that displays the most important parameters for the transient phase.

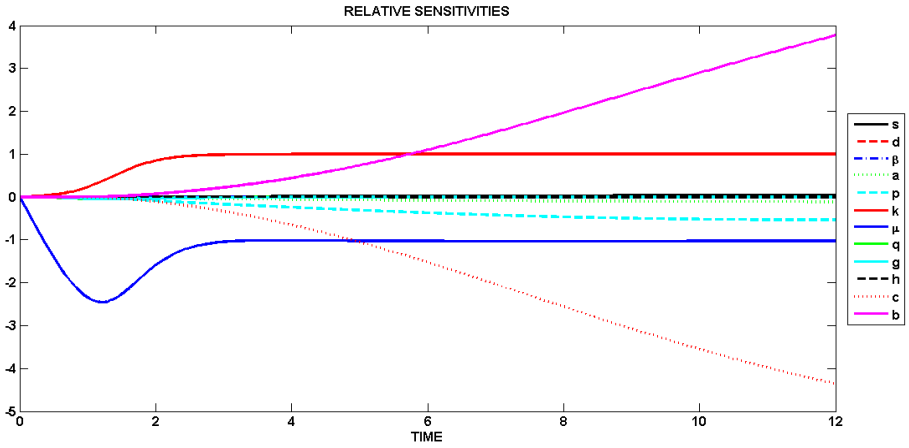

Figure 44: Numerical solutions of relative sensitivities $\frac{\theta_k}{V}\frac{\partial V}{\partial \theta_k}$ for patient 4 in a transient phase from 0 to 12

Table 23: Ranking of the most influential parameters for the viral load of patient 4 in a transient phase. During the period of time the patient is under treatment, increasing parameters k, and b will increase the viral load. Similarly, increasing the values of parameters μ and c will produce a reduction in the number of virus particles.

	Positive	**Negative**
Parameter	k	μ
	b	c

Even though at the transient phase one would expect that parameters b and c dominate among the other parameters, according to Figure 45, for $t > 150$, parameters k, μ, s, and a become the most important in the long-term phase. These results are summarized in Table 24.

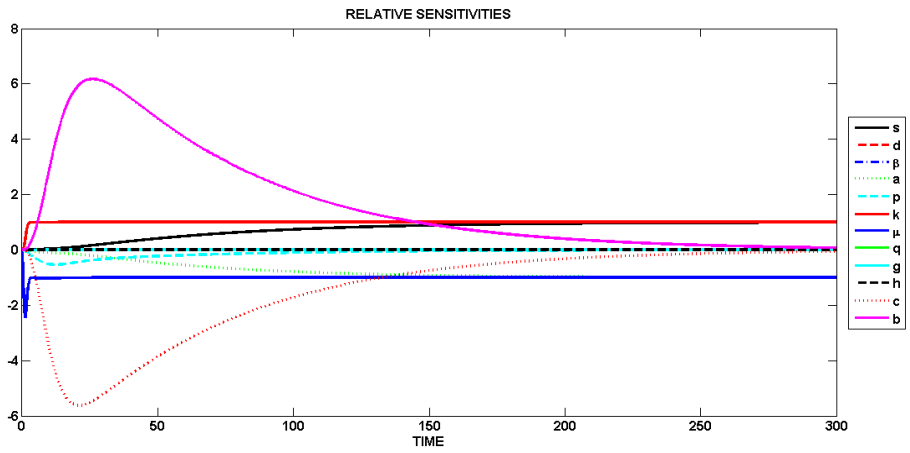

Figure 45: Numerical solutions of relative sensitivities $\frac{\theta_k}{V}\frac{\partial V}{\partial \theta_k}$ for patient 4 in a long-term phase from 0 to 300

Table 24: Ranking of the most influential parameters in the viral load for patient 4 in the long-term. The model suggests that in a long-term phase, increasing parameters b, k and s will increase the viral load. Similarly, increasing the values of parameters c, μ and a will produce a reduction in the number of virus particles.

	Positive	**Negative**
Parameter	b	c
	k	μ
	s	a

7 CONCLUDING REMARKS

In this thesis, we have used an existing model from [25] of ordinary differential equations to describe the interaction between the responses of the human immune system and a viral infection. In particular, in the context of the HCV infection, three possible scenarios were studied, namely: dominant CTL response, dominant antibody response, and coexistence.

In Section 2, by considering artificial values of parameters, a general illustration of these scenarios was given. Each scenario can be characterized by stability conditions given by the parameters. This conditions are studied deeply in [29].

Our first contribution to this study was a sensitivity analysis to determine the most influential parameters in the behavior of the viral load. Transient and long-term phases were explored with numerical simulations. In general in a transient phase, the most influential parameters for dominant CTL response are k and μ. These are the parameters of proliferation and natural decay of the viral load, respectively. For the dominant antibody response scenario, we found that the most influential parameters in the transient phase for the dominant antibody response scenario are also k and μ and that this behavior eventually switches in the long-term phase. Ultimately, their sensitivities converge to zero, and V(t) becomes affected only by parameters g and h, the proliferation and dead rates of the antibodies, respectively. In Section 3.4, a total counterintuitive behavior was observed in the sensitivities for the coexistence scenario: the viral load is affected only by parameters g and h.

Our second task in this thesis was to use optimal control theory to determine optimal treatment strategies for each scenario mentioned above. We were able to suggest treatment schedules that minimize the infection and the side effects of the therapy in a finite time horizon. We conclude that if the goal of therapy is to reduce the production of new virus and to reduce the production of infected cells, then the dominant CTL response scenario seems to be more favorable than the dominant antibody response and coexistence scenarios. However, in all the scenarios the viral load and the number of infected cells rebound to pretreatment levels upon therapy cessation.

Parameter estimates and standard errors were calculated by implementing the OLS optimization method followed by the bootstrap method. Large variability was observed in some of the parameter estimates. This variability can be explained in the bootstrap cloud found for each data set (see Figures 30,32, 34, and 36). These clouds have all the same behavior; large uncertainty for the data points in the second phase.

Because the parameter estimates are coming from data sets of patients that were under treatment, it is not possible to use the conditions given in Section 2 to classify the dynamical system associated with each data set as dominant CTL, dominant antibody, or coexistence. In fact, based on Figures 31,33,35, and 37, it seems these systems settle into an endemic equilibrium without immune response. However, this might not be the case. A possible approach to solve this issue involves the simultaneous solution of an optimal control and optimization problems which is a heavy computationally intensive problem.

A sensitivity analysis was done to understand what is the role of the estimated parameters in the behavior of the state variable $V(t)$ for data from 4 patients taken from existing literature. As a result we can conclude that, in general, k and μ are the most important parameters.

BIBLIOGRAPHY

[1] H.T. Banks and H.T. Tran, "Mathematical and Experimental Modeling of Physical and Biological Process", Taylor and Francis Group, New York,(2009).

[2] H.T. Banks, K. Holm, and F. Kappel, *Comparison of optimal design methods in inverse problems*, Inverse Problems **27** (2011) 075002 (31pp).

[3] F. Brauer and C. Castillo-Chavez, "Mathematical Models in Population Biology and Epidemiology", Springer, New York, (2010).

[4] S.P. Chakrabarty and H.R. Joshi, *Optimally controlled treatment strategy using interferon and ribavirin for hepatitis C*, Journal of Biological Systems. **17** (2009), 97-110.

[5] E.A. Coddington and N. Levinson, "Theory of Ordinary Differential Equations", McGraw-Hill, (1972).

[6] H. Dahari, A. Lo , R.M. Ribeiro, and A.S. Perelson, *Modeling hepatitis C virus dynamics: Liver regeneration and critical drug efficacy*, Journal of Theoretical Biology. **247** (2007), 371-381.

[7] H. Dahari, E. Shudo, R.M. Ribeiro, and A.S. Perelson, *Modeling complex decay profiles of hepatitis B virus during antiviral therapy*, Hepatology. **49** (2009), 32-38.

[8] P.J. Delves and I.M Roitt, *Advances in Immunology*, The New England Journal of Medicine. **343** (2000), 37-49.

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

[10] W.H. Fleming and R.W. Rishel, "Deterministic and Stochastic Optimal Control", Springer Verlag, New York, (1975).

[11] J. Guedj, H. Dahari, L. Rong, N.D. Sansone, R.E. Nettles, S.J. Cotler, T.J. Layden, S.L. Uprichard, and A.S. Perelson, *Modeling shows that the Ns5A inhibitor Daclatasvir has two modes of action and yields a shorter estimate of the hepatitis c virus half-life*, PNAS. **110** (2013), 3991-3996.

[12] E. Herrmann, J.H Lee, G. Marinos, M. Modi, S. Zeuzem, and N.P. Lam, *Effect of ribavirin on hepatitis C viral kinetics in patients treated with pegylated interferon*, Hepatology. **37** (203), 1351-1358.

[13] H. Hethcote, "The mathematics of infectious diseases", SIAM Rev. **42** (2000), 599-653.

[14] E. Jirillo, "Hepatitis C virus Disease: Immunobiology and Clinical Applications", Springer, (2008).

[15] H.R. Joshi, *Optimal control of an HIV immunology model*, Optimal Control Application and Methods. **23** (2002), 199-213.

[16] W. Kermack and A. McKendrick, *Contributions to the Mathematical Theory of Epidemics I*, P. R. Soc. Lon. Ser. A **115** (1927), 700-721.

[17] D. Kirschner, S. Lenhart, and S. Serbin, *Optimal control of the chemotherapy of HIV*, Journal of Mathematical Biology. Springer. **35** (1997), 775-792.

[18] S. Lenhart and J.T. Workman, "Optimal control applied to biological models", Chapman and Hall/CRC, (2007).

[19] A.U. Neumann, N.P. Lam, H. Dahari, D.R Gretch, T.E. Wiley, T.J. Layden, and A.S. Perelson, *Hepatitis C viral dynamics in vivo and the antiviral efficacy of interferon-α therapy*, Science. **282** (1998), 103-107.

[20] M.A. Nowak and R.M. May, "Virus dynamics. Mathematical principles of immunology and virology", Oxford University Press, (2000).

[21] J.K. Percus, "Mathematical Methods in Immunology", Courant Lecture Notes in Mathematics. **23** (2012).

[22] L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkrelidze, and E.F. Mishchenko, "The Mathematical Theory of Optimal Processes", Vol 4, Gordon and Breach Science Publishers, (1986).

[23] L. Rong, H. Dahari, R. M. Ribeiro, and A. S. Perelson, *Rapid Emergence of Protease Inhibitor Resistance in Hepatitis C Virus*, Science Translational Medicine. **2** (2010), 30ra32.

[24] L. Rong, J. Guedj, H. Dahari, D.J. Coffield, M. Levi, P. Smith, and A. S. Perelson, *Analysis of a Hepatitis C Virus Decline during Treatment with the Protease Inhibitor Danoprevir Using a Multiscale Model*, PLOS Computational Biology. **9** (2013), e1002959.

[25] D. Wodarz, *Hepatitis C virus dynamics and pathology: the role of CTL and antibody responses*, Journal of General Virology. **84** (2003), 1743-1750.

[26] D. Wodarz, "Killer cell dynamics. Mathematical and computational approaches to immunology", Interdisciplinary Applied Mathematics. Springer, (2005).

[27] D. Wodarz, *Ecological and Evolutionary Principles in Immunology*, Ecology Letters. **9** (2006), 694-705.

[28] World Health Organization (April 2014) Hepatitis C. Fact sheet no.164.

[29] N. Yousfi, K. Hattaf, and M. Rachik, *Analysis of a HCV Model with CTL and Antibody Responses*, Applied Mathematical Sciences. **3** (2009), 2835-2846.

VITA

IVAN RAMIREZ

Education:	B.S. Mathematics, University of Costa Rica, San Jose, Costa Rica 2010 M.S. Mathematical Sciences, East Tennessee State University Johnson City, Tennessee 2014
Professional Experience:	Instructor, Universidad de Costa Rica, San Jose, Costa Rica, 2009-2012 Graduate Assistant, East Tennessee State University, Johnson City, Tennessee, January 2013- July 2013 Teaching Assistant, East Tennessee State University, Johnson City, Tennessee, August 2013 - August 2014