

ABSTRACT

MURAD, NEHA. Quantitative Modeling and Optimal Control of Immunosuppressant Treatment Dynamics in Renal Transplant Recipients. (Under the direction of H.T. Banks.)

Individuals undergoing kidney transplants are put on a lifetime supply of immunosuppression to prevent an allograft rejection. However suboptimal immunosuppressive therapy can put a renal transplant recipient at a risk of infection. The key to a successful transplant is finding the optimal balance between over-suppression and under-suppression of the immune response, a task which is difficult to achieve in renal transplant patients due to the narrow therapeutic index of immunosuppressants. BK virus is a pathogen known to be a leading cause of kidney failure in renal transplant recipients. This pathogen has no antiviral therapy and can only be controlled by the body's immune response. This sole dependance on the immune system to curb BKV infections makes optimization of the immunosuppression therapy an even more pertinent problem to solve. In this dissertation, we first examine uncertainty in clinical data from a kidney transplant recipient infected with BK virus and investigate mathematical model and statistical model misspecification in the context of least squares methodology. A difference-based method is directly applied to data to determine the correct statistical model that represents the uncertainty in data. We then carry out an inverse problem with the corresponding iterative weighted least squares technique and use the resulting modified residual plots to detect mathematical model discrepancy. This process is implemented using both clinical and simulated data. Our results demonstrate mathematical model misspecification when both simpler and more complex models are assumed compared to data dynamics. Our ultimate goal is to apply control theory to adaptively predict the optimal amount of immunosuppression in renal transplant recipients; however, we first need to formulate a biologically realistic model. The process of quantitatively modeling biological processes is iterative and often leads to new insights with every iteration. We illustrate this iterative process of modeling for renal transplant recipients infected by BK virus. We analyze and improve on the current mathematical model by modifying it to be more biologically realistic and amenable for designing an adaptive treatment strategy. We use the improved model to design our feedback control loop applying Receding Horizon Control (RHC) methodology. Since data is not available for all model states we use Non Linear Kalman Filtering (specifically Extended Kalman Filter (EKF)) to estimate the non-measurable states. Combining RHC and EKF we design an adaptive treatment plan which predicts the optimal immunosuppression therapy for renal transplant recipients.

© Copyright 2018 by Neha Murad

All Rights Reserved

Quantitative Modeling and Optimal Control of Immunosuppressant Treatment Dynamics
in Renal Transplant Recipients

by
Neha Murad

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Biomathematics

Raleigh, North Carolina

2018

APPROVED BY:

Alun Lloyd

Justin Post

Hien Tran

H.T. Banks
Chair of Advisory Committee

DEDICATION

The journey that has led me here, hasn't always been easy. Over the years, I have come to believe that impediments, while they may seem like annoyances and hindrances, are one's biggest assets. If

I have succeeded today it's not in spite of my challenges, it is because of them.

This dissertation is dedicated to all the people who have stood by me through thick and thin.

This journey isn't just mine, it's *ours*.

BIOGRAPHY

Neha Murad was born on December 18th, 1988 in the city of joy, Kolkata, India. She was raised in Kolkata and attended an all girls convent school, Our Lady Queen of the Missions. In 2007, she enrolled in the Mathematics honors program at St. Xavier's College, Kolkata. She graduated First Class (a distinction secured by less than 1% candidates) and received her B.Sc in Mathematics honors degree in 2010. She went on to get an M. Sc. degree in Mathematics from Jadavpur University. It was in Jadavpur University that she was first exposed to the scope of mathematical applications in biology. After graduating with an M.Sc. in Mathematics in 2013, Neha left her hometown Kolkata for the first time to go to University of Iowa to pursue her Ph.D. in Applied Mathematics. While exploring various applications of mathematics at graduate school in Iowa, Neha realized her growing passion for mathematical modeling in biological systems. She successfully passed her Ph.D. Qualifying Examination at Iowa and received an M.Sc. in Mathematics in 2015. She transferred to the Biomathematics Ph.D. program at North Carolina State University to further explore the research areas of ecology, epidemiology and personalized medicine. In the future Neha hopes to be an active contributor in the field of mathematics and computing as it applies to understanding complex biological systems.

ACKNOWLEDGEMENTS

This dissertation would not have been possible without the support and encouragement of the countably infinite set of individuals I have had the privilege of knowing over the last 29 years of my life.

I would first and foremost like to thank my Ph.D. advisor Dr. H.T. Banks for his constant guidance and unwavering support. My journey would have been inconceivable without his supervision and contagious passion for mathematics. I would also like to thank Susie Banks for always being a constant calming source in not only Dr. Banks' life, but all our lives. She reminded us from time to time to take a break from the frenzy that we would so often create. Dr. Banks and Susie, you were a home away from home, your Thanksgiving dinners will be truly missed!

I am very grateful to the rest of my advisory committee for firstly agreeing to be part of my Ph.D. journey and secondly providing all the help I could have truly asked for, in not only finishing this dissertation but also completing graduate school. Thank you Dr. Hien Tran, Dr. Alun Lloyd and Dr. Justin Post for always being willing to provide all the assistance and support I could have asked for and more.

I wholeheartedly acknowledge the constructive contributions of my collaborators and esteemed colleagues, Dr. Rebecca Everett, Dr. Shuhua Hu, Dr. Eric Rosenberg and Dr. Hien Tran. The task of conducting research and presenting it in this dissertation would have been insurmountable without their pragmatic suggestions and all-embracing support.

A big thank you to Lesa Denning for making all the paperwork a breeze and for always having an answer for my myriad questions.

I would also like to thank Mr. Debdutta Das for being the funniest and the most encouraging mathematics teacher any high school student could ask for!

I am thankful to my mother, Dora, for bringing me into this world and helping me be the woman I am today. Her unconditional love for me and infinite belief in me has helped me get through even the toughest of days. Thank you my little sister Nida, for always looking up to me (while looking past my flaws) and inspiring me to strive harder and be a better person every single day. Thank you to my father for encouraging the mathematician in me, I push myself harder each day because of him.

I would also like to thank my newly acquired parents (in-laws), Sankar and Chirosree, for loving me like a daughter and celebrating even my smallest of victories with aplomb.

A massive shout out to all my friends who have always been a major source of comfort and cheer when things would get a bit discouraging or frustrating. Amanda L., Brandon, Michael, Marcella, Nick and Rebecca, thank you for putting up with me, listening to me, being there during my car troubles, having Sammy's wings with me, looking after my plants and so much more. I would also like to thank my friends from the University of Iowa, Anh, Gabo, Nick, Rolando, Cole, Alex, Kathryn, Maria and Alejandro, for being an immense source of love and encouragement during my first years away from my home and country. You all were a big source of warmth in those unimaginably cold Iowa winters. To my forever friends (some of whom are now legally family), Abhishek, Arjun, Eshan, Rachit, Shefa, Sneha, Upa and Varsha, thank you for being the people I can count on for literally anything, for the rest of my life, irregardless of which part of the world you all are in.

Lastly, thank you Avi, my husband, my best friend, my forever decision tree generator, for bearing witness to it all, not just my big moments of success and failure, but even the most mundane inconsequential aspects of my life in graduate school and more. Thank you for loving every detail of my life and making it yours. This would have not been possible without you.

TABLE OF CONTENTS

LIST OF TABLES	viii
LIST OF FIGURES	ix
Chapter 1 Introduction	1
1.1 Thesis Outline	2
1.2 Motivation	2
1.3 Recent Modeling Efforts in Kidney Transplantation	3
1.4 The Immune System and BK Virus in Context of Kidney Transplants	5
Chapter 2 Mathematical and Statistical Model Misspecification in Modeling Immune Response in Renal Transplant Recipients	8
2.1 Introduction	8
2.2 Mathematical Model and Clinical Data	9
2.2.1 Mathematical Model Synopsis	9
2.2.2 Log-scaled Model	11
2.2.3 Clinical Data	13
2.3 Inverse Problem Method and Statistical Model Selection	14
2.4 Difference-based Methods and Modified Residuals	16
2.5 Results	18
2.5.1 Clinical Data	18
2.5.2 Simulated Data	21
2.6 Conclusion	24
Chapter 3 Immunosuppressant Treatment Dynamics in Renal Transplant Recipients: An Iterative Modeling Approach	26
3.1 Introduction	26
3.2 Iteration I: Preliminary Model	28
3.2.1 Data Collection and Biological Model	28
3.2.2 Mathematical Model	28
3.2.3 Statistical Error Model	30
3.2.4 Model Analysis	31
3.2.5 Biological Interpretation of Model and Changes in Understanding	33
3.3 Iteration II: Improved Model	34
3.3.1 Mathematical Model	34
3.3.2 Model Analysis and Biological Interpretation	38
3.4 Iteration III: Further Prospective Improvements	41
3.4.1 Mathematical Model	41
3.4.2 Model Analysis and Inference	42
3.5 Discussion	46
Chapter 4 Optimal Control of Immunosuppressants in Renal Transplant Recipients Susceptible to BKV Infection	47

4.1	Introduction	47
4.2	Mathematical Model	49
4.3	Optimal Control	51
4.3.1	Open Loop Control	52
4.3.2	Feedback Loop Control: Receding Horizon Methodology	53
4.4	Filtering: Continuous-Discrete Extended Kalman Filter	55
4.5	Numerical Results	58
4.5.1	Numerical Results: Fixed Immunosuppressant Dosages	58
4.5.2	Numerical Results: Open Loop Control	61
4.5.3	Numerical Results: Feedback Control with Perfect Information	66
4.5.4	Numerical Results: Extended Kalman Filter	68
4.5.5	Numerical Results: Feedback Control with State Estimation	70
4.6	Conclusion	72
Chapter 5	Future Directions	74
5.1	Dynamics of Immunosuppressive Treatments	74
5.1.1	Induction Therapy	74
5.1.2	Maintenance Therapy	75
5.1.3	Discussion	77
5.2	Role of Genetic Variation in the Optimal Control of Immunosuppressants in Renal Transplant Recipients	78
5.2.1	Genetic Variation as a Cause of Variability in Immune Responses and Transplant Outcomes	78
5.2.2	Renal Pharmacogenomics	80
5.2.3	Discussion	81
Chapter 6	Conclusion	82
BIBLIOGRAPHY	84
APPENDIX	91
Appendix A	Mathematical and Statistical Model Misspecification in Modeling Immune Response in Renal Transplant Recipients	92
A.1	Derivation of Modified Pseudo Errors	92
A.2	Simulated Data from the Original Model (3.1)	94
A.3	Simulated Data from the Simpler Model (2.7)	98

LIST OF TABLES

Table 2.1	Description of state variables.	11
Table 2.2	Description of model parameters and corresponding fixed values.	11
Table 3.1	Summary of cell dynamics under influence of immunosuppression.	32
Table 3.2	Original (Iteration I) and new model (Iteration II) parameters.	37
Table 3.3	Initial conditions both original (Iteration I) and new (Iteration II).	37

LIST OF FIGURES

Figure 1.1	Urothelial cell infected by BKV Copyright ©2011 Michael Bonert	6
Figure 2.1	Model diagram of the BKV virus affecting renal cells [8].	12
Figure 2.2	Patient TOS003 BKV viral plasma loads and plasma creatinine levels.	13
Figure 2.3	Viral load modified pseudo errors vs. time for various γ_1 values.	18
Figure 2.4	Creatinine modified pseudo errors vs. time for various γ_2 values.	19
Figure 2.5	Model (2.1) solution and clinical data with $\gamma=(0.5,0)$	20
Figure 2.6	Modified residuals for V and C with $\gamma = (0.5,0)$	20
Figure 2.7	Model (2.7) solution and simulated data created from model (2.1) with $\gamma=(0.5,0)$	22
Figure 2.8	Modified residuals for V and C with $\gamma = (0.5,0)$	23
Figure 2.9	Model (2.1) solution and simulated data created from model (2.7) with $\gamma=(0,0.7)$	24
Figure 2.10	Modified residuals for V and C with $\gamma = (0,0.7)$	24
Figure 3.1	Schematic of the iterative modeling process [11].	27
Figure 3.2	Plot illustrating the balance between under and over suppression of the immune response.	32
Figure 3.3	BK viral load in blood, creatinine, susceptible and infected cell dynamics for highest and lowest immunosuppressant dosages.	34
Figure 3.4	Model simulations for both Iteration I and II of modeling for different ϵ_I values.	39
Figure 3.5	Figures showing the step functions used in Iteration II model (3.3) and their smooth approximations used in Iteration III model (3.5).	43
Figure 3.6	Model simulations for Iteration II and III of modeling for different ϵ_I values.	44
Figure 4.1	A sample plot showing the concept of RHC [24].	54
Figure 4.2	A schematic diagram showing the RHC algorithm [24].	55
Figure 4.3	Model dynamics when ϵ dosages are fixed. (Red dashed line is the upper bound on healthy biomarker values.)	59
Figure 4.4	Open loop control for weights $(W_V, W_C) = (1, 1)$ with varying ϵ_0	62
Figure 4.5	Open loop control for weights $(W_V, W_C) = (1, 1)$ with varying ϵ_0	63
Figure 4.6	Open loop control with varying W_V with initial immunosuppressant, $\epsilon_0 = 0.45$	64
Figure 4.7	Open loop control with varying W_C with initial immunosuppressant, $\epsilon_0 = 0.45$	65
Figure 4.8	Optimal immunosuppressant dosage for first 340 days of treatment with an initial guess for control estimation $\epsilon_0 = 0.45$ when perfect information is available on the patient every 20 days. The blue dots are the daily recommended optimal dose. (Red dashed line is the upper bound on healthy biomarker values.)	67
Figure 4.9	State estimation when $Q \gg R$: Filter trusts data more.	68
Figure 4.10	State estimation when $R \gg Q$: Filter trusts model more.	69
Figure 4.11	State estimation when $R \approx Q$: Current Settings.	69
Figure 4.12	Daily optimal immunosuppressant dosages for first 340 days of treatment with an initial control guess $\epsilon_0 = 0.45$ when the patient visited the doctor every 20 days. The blue dots are the daily recommended optimal dose. (Red dashed line is the upper bound on healthy biomarker values.)	71

Figure A.1	Simulated viral load modified pseudo errors vs. time for various γ_1 values. . .	95
Figure A.2	Simulated creatinine modified pseudo errors vs. time for various γ_2 values. . .	96
Figure A.3	Inverse problem model (3.1) solution and simulated data with $\gamma = (0.5, 0)$. . .	96
Figure A.4	Modified residuals for V and C with $\gamma = (0.5, 0)$	97
Figure A.5	Modified pseudo errors for the viral load and creatinine with $\gamma = (0, 0.7)$	98

CHAPTER

1

INTRODUCTION

Abbreviations

BKV	BK virus
CKD	Chronic Kidney Disease
ESRD	End Stage Renal Disease
GvHD	Graft-Versus-Host Disease
GFR	Glomerular Filtration Rate
HCMV	Human Cytomegalovirus
HIV	Human Immunodeficiency Virus
PVAN	Polyomavirus type BK-associated Nephropathy
SOT	Solid Organ Transplantation

1.1 Thesis Outline

The rest of this chapter is dedicated to explaining the motivation, previous works and some background biology for modeling immune response in renal transplant recipients. Chapter 2 explains a method to choose the correct statistical model directly from a set of observations and then identify any mathematical model misspecification in context to the problem defined in Chapter 1. Chapter 3 is dedicated to illustrating the iterative process of modeling biological systems by improving the existing model introduced in Chapter 2 and making it biologically more realistic. Using the model developed in Chapter 3, we present in Chapter 4, the algorithm and tools used in the development of an optimal treatment schedule for renal transplant recipients and the consequent results. In Chapter 5, we explore future directions in the field of renal pharmacogenomics and specific drug dynamics of immunosuppressants as it applies to personalized medicine. A summary and conclusion is given in Chapter 6.

1.2 Motivation

According to the Organ Procurement and Transplantation Network (OPTN) as of March 11, 2018, kidney transplants are the highest number of solid organ transplants, comprising of 428,298 transplants between January 1, 1988 to February 28, 2018 [92]. There are currently 114,949 people waiting for lifesaving organ transplants in the U.S. Of these, 95,301 await kidney transplants [78]. Glomerular Filtration Rate (GFR) is often used as an indicator for kidney health and function; it measures the rate at which the kidney clears toxic waste from the blood. A GFR number of 90 or less in adults is used as an indicator for kidney disease [30]. Chronic kidney disease (CKD), also commonly known as chronic kidney failure, is characterized by gradual but progressive loss of kidney function. The fifth stage of CKD, called End Stage Renal Disease (ESRD), occurs when kidney function reduces to less than 15% and leads to permanent kidney failure [30]. Patients with ESRD have two choices of therapy - dialysis or kidney transplantation. Kidney transplantation is often chosen since transplants (grafts) can improve survival and lower healthcare costs compared to dialysis [44].

While most talk about the success rate of kidney transplants, the National Kidney foundation [91] points out that although the official statistics is that at the the end of the first month 97% of total renal transplant recipients have a working transplant, that number decreases to 93% by the end of the first year and becomes 83% by the end of 3 years. At 10 years, only 54% of transplant kidneys are still working. In fact, over 20% of kidney transplants every year are re-transplants. (Note that the transplant statistics are the most recent overall numbers from the Scientific Registry of Transplant Recipients [1]. The results are different for deceased donors and living donors. More details on the

results from a particular transplant program is available at [1].) The authors in [89] studied kidney transplants that progressed to failure after a biopsy for clinical indications and narrowed down the top three causes for renal failure. The most prevalent cause for failure was due to organ rejection, followed by glomerulonephritis caused in patients with infections in the throat or skin. The third most prevalent cause for kidney rejection is polyomavirus-associated nephropathy (PVAN) (7%). PVAN is mainly caused by high-level replication of the human polyomavirus type 1, also called BK virus (BKV), in renal tubular epithelial cells [35]. Currently there are no BKV-specific antiviral therapy, but in some cases, BKV replication may be controlled by reducing the level of immunosuppression [48].

In order to prevent the body from rejecting the transplant, patients are usually put on a lifetime regime of immunosuppressive medications to prevent the body from rejecting the allograft [51]. However these immunosuppressive treatments often leave the recipient susceptible to opportunistic bacterial and viral pathogens and can even reactivate latent viruses preexisting in the recipient and/or donor's organ. Common viruses that impact transplant recipients include human cytomegalovirus (HCMV), Epstein-Barr virus (EBV), human herpes virus (HHV)-6, HHV-7, and human polyomavirus 1 (BK virus) [8]. Thus for a renal transplant to be successful, a crucial but fragile balance needs to be struck between over-suppression and under-suppression of the immune system. While the former can weaken the body's immune response making it susceptible to infections, the latter can cause the immune response to fight the renal graft leading to kidney rejection. Currently, immunosuppressive treatment protocols in the United States are in a state of flux with varying treatment regimens across different organ transplant centers. One possible reason for this inconsistency is that centers are implementing individual or group based treatment protocols [22, 39].

The lack of consistent immunosuppressive treatment strategies across transplant centers and the narrow therapeutic index of immunosuppressants motivates us to mathematically and statistically model dynamics of the immune system in the context of BKV infections. Our goal is to build a robust dynamical systems model which emulates the human immune response to the allograft and BKV infection in the renal tubular epithelial cells of the transplanted kidney. We then aim to use this model to formulate an adaptive personalized treatment strategy for renal transplant recipients using feedback optimal control methods.

1.3 Recent Modeling Efforts in Kidney Transplantation

Several research groups have recently contributed to developing mathematical models to study the mechanisms of the complex biological process involving solid organ transplantation (SOT),

specifically kidney transplantation.

Funk *et al.*, [37] use simple mathematical equations to investigate the relationship between BKV replication and polyomavirus type BK-associated nephropathy (PVAN), one of the most common viral complications to develop in renal transplant recipients. It is a prominent cause of renal transplant dysfunction and graft loss. Since PVAN was first reported in 1995, an occurrence rate from 1% to 10% has been reported [47]. Prevalence of PVAN is mainly attributable to BK virus. The authors in [37] assume the BK virus grows and decays exponentially and then present the corresponding equations to calculate the viral doubling times and half-lives. The generation time and basic reproductive ratio (R_0) equations are also given. The authors then perform a retrospective mathematical analysis on 15 individual patient datasets with the given equations. Their results indicate rapid replication of the BK virus, elucidating the progressive nature of PVAN, contrary to the general perception in clinical practice. The authors also propose the use of R_0 as a measure of the efficiency of anti-viral interventions in future studies.

The authors of [35] extend this work and first perform statistical analysis on datasets from 223 kidney transplant recipients to help understand the relationship between BKV in the plasma and urine. The authors present a dynamical model that considers four cell populations and two virus populations: uninfected kidney tubular epithelial cells, infected tubular epithelial cells, plasma virus, uninfected urothelial cells, infected urothelial cells, and urine virus. Antiviral interventions are represented as a time-dependent function that affects the growth of both viral populations. The basic reproductive ratio for the kidney and urinary compartments are determined and used in parameter sensitivity analysis. The authors simulate the mathematical model to explore the dynamics of BKV for various scenarios and compare simulation results to 7 individual kidney transplant patient datasets. The scenario that best matches the clinical data assumes viral replication starts in the kidney, infects urothelial cells, and then a bidirectional viral flux increases replication in both compartments. The results provide insights into the relationship between the two compartments with respect to BKV replication and suggest further awareness into PVAN progression.

Kepler *et al.*, [55] also consider the effect of the immune response on the viral infection in the presence of immunosuppression therapy. The authors specifically consider human cytomegalovirus (HCMV) infection (one of eight known human herpes viruses) in SOT recipients and present a model that describes the dynamics of the viral load, immune response, actively-infected cells, susceptible cells, and latently-infected cells. The authors show that the model can describe the three types of infection: primary, latent, and reactivation. Model simulations are given for varying amounts of immunosuppression. Due to the limited amount of corresponding data in literature, the authors do not compare their model to patient data. However, the authors state how inverse problems can be performed with individual datasets to estimate parameter values for individual patients, providing

insights into the heterogeneity in disease progression.

Banks *et al.*, [7] modify the model in [55] to include the body's immune response to a donor kidney. Their model considers susceptible cells, infected cells, free HCMV, HCMV-specific CD8+T-cells, allospecific CD8+T-cells that target the donor kidney, and serum creatinine, a biological marker for renal function. Due to a lack of data, model validation is not provided; model simulations with varying antiviral and immunosuppressive drug efficacy are produced only to verify that results match clinical trends. The authors then use simulated data to demonstrate an optimal control problem to design an adaptive anti-viral and immunosuppressant treatment schedule that balances over-suppression and under-suppression of the immune system. Banks *et al.*, [8] further modify this model to consider BKV infection and partially validate the resulting model using clinical data. Due to the large number of parameters and limited data, the authors implement an iterative process to identify the most sensitive model parameters to be estimated.

1.4 The Immune System and BK Virus in Context of Kidney Transplants

BK Virus was first detected in 1971 in a Sudanese kidney transplant recipient with initials "BK" [4, 17, 84]. This virus is found in over 80% of the world's population, but infection doesn't cause illness in healthy immunocompetent people [4], replication of the virus often occurs during states of immune suppression [17]. BKV belongs to the polyomaviridae family which also includes JC virus (JCV) and simian virus (SV40) [84]. BK Virus is the primary etiologic agent in most cases where patients exhibit symptoms of PVAN, although JC virus can also cause PVAN. BK Virus is a more recently recognized viral infection that can affect the renal graft early and late after transplantation. It's detection and treatment are best managed in a transplantation center. It is a ubiquitous virus that remains in a latent state. About 30% to 60% of kidney transplant recipients develop BK viruria after transplantation, and 10% to 20% develop BK viremia. Among those who develop BK viremia, 5% to 10% develop BK nephropathy; of these, approximately 70% lose the allograft and the remainder exhibit some kidney dysfunction. BK infection may be associated with ureteral stenosis and possible obstruction, tubulointerstitial nephritis, and a progressive rise in the serum creatinine level, with ultimate allograft failure. Such infection must be evaluated in any episode of renal dysfunction and prospectively evaluated approximately every 3 to 6 months in the first year after transplantation [50]. Reactivation and/or spike in BKV load in renal transplant recipients is often due to the high dosage of the potent immunosuppressants prescribed to decrease the odds of graft rejection. There is currently no approved antiviral drug therapy available to fight BKV; vigilant screening for early detection and monitoring the immunosuppressant dosage are the only prevention methods for symptomatic BKV nephropathy [47].

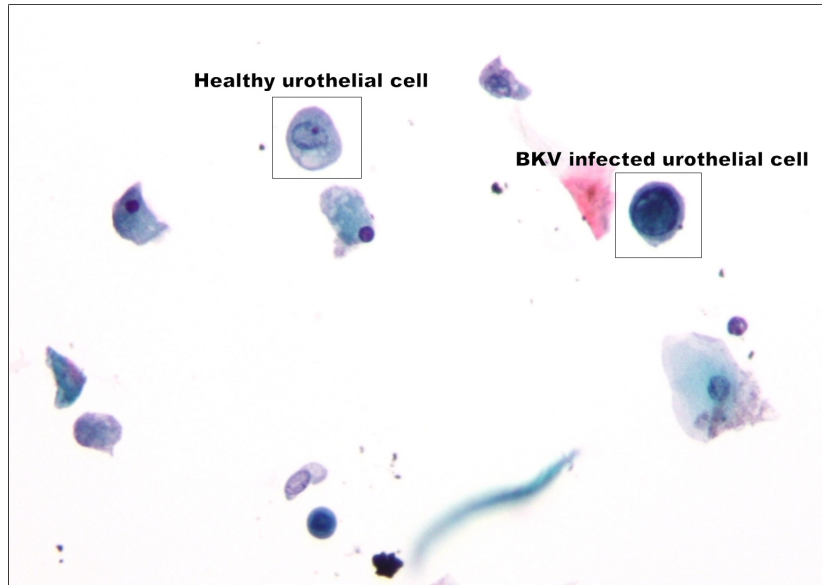


Figure 1.1: Urothelial cell infected by BKV

Copyright ©2011 Michael Bonert

The immune response can be divided into two kinds, innate immune response (natural, same with each encounter) and acquired immune response (adaptive, improves on repeated exposure). Some cells that depict innate immune response are phagocytic cells (neutrophils, monocytes, macrophages), inflammatory mediator cells ((basophils, mast cells, and eosinophils) and natural killer cells. The acquired immune response on the other hand is responsible for the proliferation of antigen-specific B and T cells. B cells create immunoglobulins (antigen specific antibodies responsible for eliminating foreign microorganisms). T cells help B cells to make antibodies and can annihilate intracellular pathogens by activating macrophages and killing virally infected cells. White blood cells (WBCs), (also called leukocytes), are the cells of the immune system that are involved in protecting the body against both infectious disease and foreign invaders. A CD8+ T-cell (also known as cytotoxic T cell or killer T-cell) is a type of white blood cell that kills cancer cells, cells that are infected (particularly with viruses), or cells that are damaged in other ways. A CD4 T-cell (also known as mature T helper cell) is a type of WBC that sends signals to other types of immune cells, including CD8+ T-cells. All compartments of the immune system seem to be involved in keeping pathogens like BKV at bay, virus-specific T-cells being of particular importance [2].

As mentioned earlier in Section 1.2, Glomerular Filtration Rate(GFR) is used to measure kidney health and function [29]. The compound serum creatinine is produced as a byproduct of muscle metabolism (breakdown of a product called phosphocreatine in the muscle) and excreted in the urine. Low levels of creatinine in the blood is an indicator of good renal health and is often used

as a surrogate to assess GFR. Thus the two biomarkers that can be readily measured to ascertain kidney health and infection status from blood plasma samples in a renal transplant patient are the BKV load and creatinine.

CHAPTER

2

MATHEMATICAL AND STATISTICAL MODEL MISSPECIFICATION IN MODELING IMMUNE RESPONSE IN RENAL TRANSPLANT RECIPIENTS

2.1 Introduction

Mathematical and statistical models are useful tools to investigate the mechanisms of any complex biological process. In Chapter 1 we introduced our problem and our motivations behind building a mathematical model for the immune response and formulating an adaptive treatment regimen for renal transplant recipients. In the presence of data, a statistical model is necessary to account for the discrepancy between the actual phenomenon and the observation process. In this chapter, in context of the problem introduced in Chapter 1, we build on the work of Banks *et al.*, [8] and investigate the uncertainty in clinical data from a BKV infected kidney transplant recipient.

The standard protocol for choosing the correct statistical model is to make an informed guess for the statistical model, solve an inverse problem and fit the mathematical model to data. Then using the

predictions from the mathematical model, compute and plot residuals or modified residuals against time and predictions to see if they are random. If residual plots are not randomly distributed around the horizontal axis a new statistical model is chosen and the process repeated [9]. There are two major disadvantages of this method, first that it is time consuming and computationally expensive, as it might take several attempts of performing an inverse problem to identify the correct statistical model associated with the observation process. Second, it determines the correct statistical model under the tacit assumption that one has a correct mathematical model, which when modeling complex biological systems, may not always be the case.

To overcome these drawbacks, in this chapter, we introduce a second order difference-based method to identify the correct statistical model directly from the time series data, without making any assumptions about the accuracy of the mathematical model. After performing an inverse problem with this appropriate statistical model, we demonstrate how modified residual or residual plots can reveal error in the mathematical model if any. The remainder of the chapter is as follows. Section 2.2 contains an overview of the BKV model and clinical data. The inverse problem and difference-based methodologies are given in Sections 2.3 and 2.4, respectively. Section 2.5 includes our results using both clinical and simulated data. Lastly, we present our conclusions in Section 2.6.

2.2 Mathematical Model and Clinical Data

2.2.1 Mathematical Model Synopsis

We consider the following BKV model in [8], which describes the dynamics of the free BK viral load (V), susceptible cells (H_S), BKV infected cells (H_I), BKV-specific CD8+ T cells (E_V), allospecific CD8+ T cells (E_K), and the surrogate for GFR, serum creatinine (C)

$$\dot{H}_S = \lambda_{HS} \left(1 - \frac{H_S}{\kappa_{HS}} \right) H_S - \beta H_S V \quad (2.1a)$$

$$\dot{H}_I = \beta H_S V - \delta_{HI} H_I - \delta_{EH} E_V H_I \quad (2.1b)$$

$$\dot{V} = \rho_V \delta_{HI} H_I - \delta_V V - \beta H_S V \quad (2.1c)$$

$$\dot{E}_V = (1 - \epsilon_I) [\lambda_{EV} + \rho_{EV}(V) E_V] - \delta_{EV} E_V \quad (2.1d)$$

$$\dot{E}_K = (1 - \epsilon_I) [\lambda_{EK} + \rho_{EK}(H_S) E_K] - \delta_{EK} E_K \quad (2.1e)$$

$$\dot{C} = \lambda_C - \delta_C(E_K, H_S) C \quad (2.1f)$$

where

$$\rho_{EV}(V) = \frac{\bar{\rho}_{EV} V}{V + \kappa_V}, \quad (2.1g)$$

$$\rho_{EK}(H_S) = \frac{\bar{\rho}_{EK} H_S}{H_S + \kappa_{KH}}, \quad (2.1h)$$

$$\delta_C(E_K, H_S) = \frac{\delta_{C0} \kappa_{EK}}{E_K + \kappa_{EK}} \cdot \frac{H_S}{H_S + \kappa_{CH}}, \quad (2.1i)$$

and initial conditions,

$$(H_S(0), H_I(0), V(0), E_V(0), E_K(0), C(0)) = (H_{S0}, H_{I0}, V_0, E_{V0}, E_{K0}, C_0). \quad (2.1j)$$

Susceptible cells proliferate logistically at a maximum rate λ_{HS} with carrying capacity κ_{HS} . Susceptible cells and free virions are both lost due to interacting with each other at rate β , resulting in infected cells. Infected cells lyse at rate δ_{HI} due to the cytopathic effect of the virus and produce p_V virions. BKV-specific CD8+ T cells also eliminate infected cells at rate δ_{EH} . The free virus is naturally cleared at rate δ_V . Both the BKV-specific CD8+ T cells and the allospecific CD8+ T cells are inversely related to the immunosuppressant dosage efficiency ϵ_I . The source rates for E_V and E_K are given by λ_{EV} and λ_{EK} respectively. The BKV-specific CD8+ T cells proliferate in the presence of free virions at maximum rate $\bar{\rho}_{EV}$ with half saturation level κ_V . Similarly, the allospecific CD8+ T cells proliferate in the presence of susceptible cells at maximum rate $\bar{\rho}_{EK}$ with half saturation level κ_{KH} . Both E_V and E_K die at constant rates δ_{EV} and δ_{EK} respectively. Creatinine is produced at rate λ_C . The clearance rate for C is dependent upon both E_K and H_S with a maximum clearance rate of δ_{C0} and saturation levels κ_{EK} and κ_{CH} . The efficiency of the immunosuppressant ϵ_I is approximated by the following piecewise constant function

$$\epsilon_I(t) = \begin{cases} \epsilon_1 & t \in [0, 21] \\ \epsilon_2 & t \in (21, 60] \\ \epsilon_3 & t \in (60, 120] \\ \epsilon_4 & t \in (120, 450]. \end{cases} \quad (2.1k)$$

The state variable descriptions and units can be found in Table 2.1. The diagrammatic representation of model (2.1) is given in Figure 2.1. Table 2.2 contains the description of the model parameters and fixed values. Note that these values that we treat as fixed parameters are estimated in [8] using an initial guess obtained from literature or through simulation (for more details, see [8]). A value of est. indicates a free parameter to be estimated in the current chapter.

Table 2.1: Description of state variables.

State	Description	Unit
H_S	Concentration of susceptible host cells	cells/mL
H_I	Concentration of infected host cells	cells/mL
V	Concentration of free BKV	copies/mL
E_V	Concentration of BKV-specific CD8+ T cells	cells/mL
E_K	Concentration of allospecific CD8+ T cells that target kidney	cells/mL
C	Concentration of serum creatinine	mg/dL

Table 2.2: Description of model parameters and corresponding fixed values

Parameter	Description	Unit	Value
λ_{HS}	Proliferation rate for H_S	1/day	0.030
κ_V	Saturation constant	copies/mL	180.676
κ_{HS}	Saturation constant	cells/mL	1025.888
λ_{EK}	Source rate of E_K	cells/(mL·day)	0.002
β	Infection rate of H_S by V	mL/(copies·day)	est.
δ_{EK}	Death rate of E_K	1/day	est.
δ_{HI}	Death rate of H_I by V	1/day	0.085
λ_C	Production rate for C	mg/(dL·day)	0.007
ρ_V	# Virions produced by H_I before death	copies/cells	4292.398
δ_{C0}	Maximum clearance rate for C	1/day	0.014
δ_{EH}	Elimination rate of H_I by E_V	mL/(cells·day)	0.002
κ_{EK}	Saturation constant	cells/mL	0.200
δ_V	Natural clearance rate of V	1/day	0.372
κ_{CH}	Saturation constant	cells/mL	10.000
λ_{EV}	Source rate of E_V	cells/(mL·day)	0.001
$\bar{\rho}_{EK}$	Maximum proliferation rate for E_K	1/day	est.
δ_{EV}	Death rate of E_V	1/day	est.
κ_{KH}	Saturation constant	cells/mL	84.996
$\bar{\rho}_{EV}$	Maximum proliferation rate for E_V	1/day	est.
ϵ_I	Efficacy of immunosuppressive drugs		

2.2.2 Log-scaled Model

Due to a scale difference among model states and model parameters as well as to ensure that state variables do not become negative, we use log transformation to resolve any scaling issues during numerical simulations and implementation of the inverse problem (see [8] for details). We can rewrite model (2.1) as the vector system,

$$\frac{dy}{dt} = \mathbf{h}(\mathbf{y}, \bar{\mathbf{q}}, \mathbf{y}_0)$$

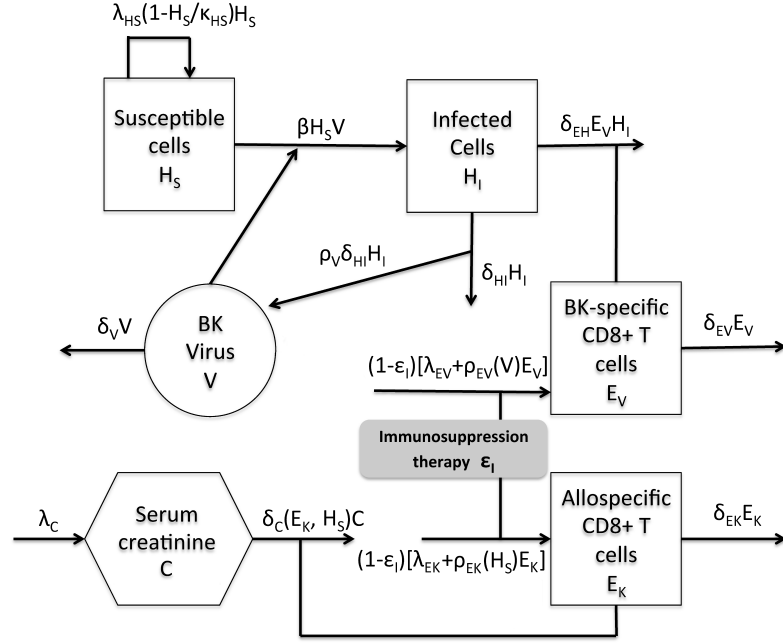


Figure 2.1: Model diagram of the BKV virus affecting renal cells [8].

where

$$\mathbf{y} = [H_S, H_I, V, E_V, E_K, C]^T,$$

$$\bar{\mathbf{q}} = [\lambda_{HS}, \lambda_{EK}, \lambda_{EV}, \lambda_C, \beta, \delta_{EH}, \delta_V, \bar{\rho}_{EV}, \delta_{EV}, \delta_{EK}, \delta_{C0}, \delta_{HI}, \kappa_{CH}, \kappa_{KH}, \kappa_{HS}, \kappa_{EK}, \kappa_V, \rho_V, \bar{\rho}_{EK}, \epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4]^T,$$

and

$$\mathbf{y}(0) = \mathbf{y}_0.$$

We make the afore-mentioned log transformation by defining variables

$$x_i = \log_{10}(y_i), \quad i = 1, 2, 3, 4, 5$$

$$x_6 = y_6,$$

$$x_{0i} = \log_{10}(y_{0i}), \quad i = 1, 2, 3, 4, 5$$

$$x_{06} = y_{06},$$

$$q_j = \log_{10}(\bar{q}_j), \quad j = 1, 2, \dots, 19$$

$$q_j = q_j, \quad j = 20, \dots, 23.$$

Then the log-scaled model becomes

$$\frac{d\mathbf{x}}{dt} = \mathbf{g}(\mathbf{x}, \mathbf{q}, \mathbf{x}_0),$$

where $g_i(\mathbf{x}, \mathbf{q}, \mathbf{x}_0)$ is given by

$$\begin{aligned} g_i(\mathbf{x}, \mathbf{q}, \mathbf{x}_0) &= \frac{dx_i}{dt} = \frac{dx_i}{dy_i} \frac{dy_i}{dt} = \frac{1}{y_i \ln(10)} h_i(\mathbf{y}, \bar{\mathbf{q}}, \mathbf{y}_0), \quad i = 1, 2, 3, 4, 5 \\ &= \frac{1}{10^{x_i} \ln(10)} h_i(10^{(x_1, x_2, \dots, x_5)}, x_6, 10^{(q_1, q_2, \dots, q_{19})}, q_{20}, q_{21}, \dots, q_{23}, 10^{(x_{01}, x_{02}, \dots, x_{05})}, x_{06}), \end{aligned}$$

and

$$\begin{aligned} g_6(\mathbf{x}, \mathbf{q}, \mathbf{x}_0) &= h_6(\mathbf{y}, \bar{\mathbf{q}}, \mathbf{y}_0) \\ &= h_6(10^{(x_1, x_2, \dots, x_5)}, x_6, 10^{(q_1, q_2, \dots, q_{19})}, q_{20}, q_{21}, \dots, q_{23}, 10^{(x_{01}, x_{02}, \dots, x_{05})}, x_{06}). \end{aligned}$$

2.2.3 Clinical Data

We investigate uncertainty in the clinical data [8]. This data set consists of eight BK viral plasma load (DNA copies/mL) measurements and sixteen plasma creatinine level (mg/dL) measurements for patient TOS003 from Massachusetts General Hospital. The patient was diagnosed with BKV infection in the first 3 months of transplantation. With every visit, dosage and combination of immunosuppressants were updated. Figure 2.2 contains the plots of the data.

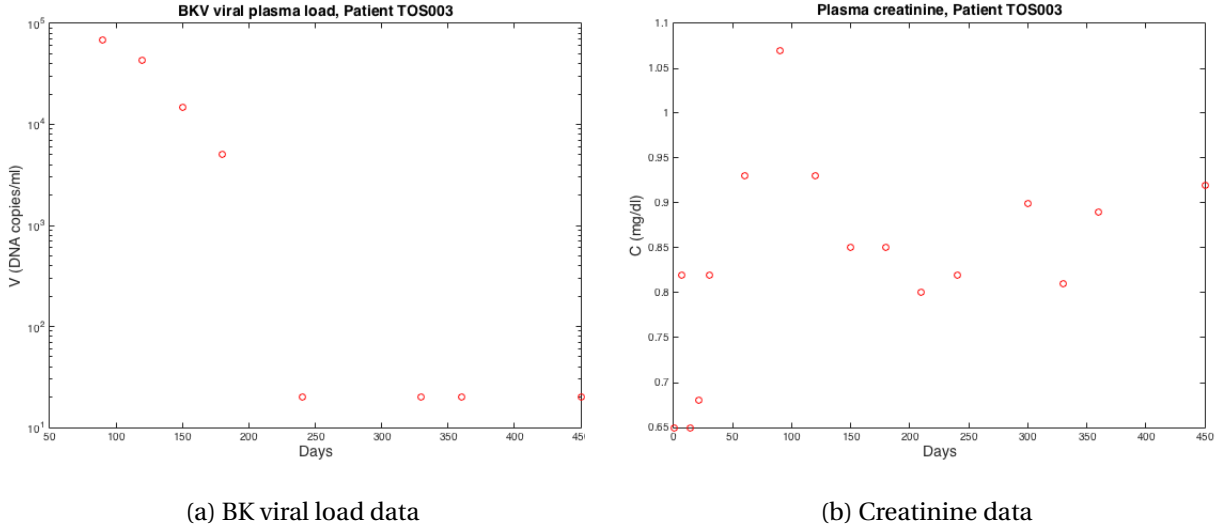


Figure 2.2: Patient TOS003 BKV viral plasma loads and plasma creatinine levels.

2.3 Inverse Problem Method and Statistical Model Selection

We follow standard inverse problem procedures to estimate parameters in our mathematical model [9, 11, 25, 88]. Consider a general N -dimensional dynamical system with parameter vector \mathbf{q} ,

$$\begin{aligned}\frac{d\mathbf{x}}{dt}(t) &= \mathbf{g}(t, \mathbf{x}(t); \mathbf{q}), \\ \mathbf{x}(t_0) &= \mathbf{x}_0,\end{aligned}$$

with an m dimensional observation process

$$f(t; \boldsymbol{\theta}) = \mathcal{C} \mathbf{x}(t; \boldsymbol{\theta}),$$

where $\boldsymbol{\theta} = (\mathbf{q}^\top, \mathbf{x}_0^\top)^\top$ is the vector of parameters along with the initial conditions to be estimated and \mathcal{C} is the $m \times N$ observation matrix.

Our data set consists of observed values for the plasma viral load and creatinine levels. Thus our observation matrix is the following

$$\mathcal{C} = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

as $\mathbf{x} = [\log_{10} H_S, \log_{10} H_I, \log_{10} V, \log_{10} E_V, \log_{10} E_K, C]^\top$ and $\mathbf{f} = [\log_{10} V, C]^\top$.

Let y_i^1 represent the free BK viral load measurements and y_j^2 represent the plasma creatinine load measurements at time points $t_i^1, i = 1, 2, \dots, n_1$ and $t_j^2, j = 1, 2, \dots, n_2$ respectively. Here $n_1 = 8$ and $n_2 = 16$. We note that there is some discrepancy between the actual phenomenon, which is represented through the data, and the above observation process. We account for this uncertainty with the statistical models,

$$\begin{aligned}Y_i^1 &= f_1(t_i^1; \boldsymbol{\theta}_0) + f_1(t_i^1; \boldsymbol{\theta}_0)^{\gamma_1} \mathcal{E}_i^1, & i = 1, 2, \dots, n_1, \\ Y_j^2 &= f_2(t_j^2; \boldsymbol{\theta}_0) + f_2(t_j^2; \boldsymbol{\theta}_0)^{\gamma_2} \mathcal{E}_j^2, & j = 1, 2, \dots, n_2,\end{aligned}$$

where $\gamma \geq 0$ and the $p \times 1$ vector $\boldsymbol{\theta}_0 \in \boldsymbol{\Omega}$ is the “true” or nominal parameter set. Here $f_1(t_i^1; \boldsymbol{\theta}_0) = x_3(t_i^1; \boldsymbol{\theta}_0)$ and $f_2(t_j^2; \boldsymbol{\theta}_0) = x_6(t_j^2; \boldsymbol{\theta}_0)$. The $n_1 \times 1$ and $n_2 \times 1$ random error vectors \mathcal{E}^1 and \mathcal{E}^2 respectively are assumed to be independent and identically distributed (*i.i.d*) with mean zero and $\text{Var}(\mathcal{E}_i^1) = \sigma_{01}^2$

and $\text{Var}(\mathcal{E}_j^2) = \sigma_{02}^2$. The corresponding realizations are,

$$\begin{aligned} y_i^1 &= f_1(t_i^1; \theta_0) + f_1(t_i^1; \theta_0)^{\gamma_1} \epsilon_i^1, & i = 1, 2, \dots, n_1, \\ y_j^2 &= f_2(t_j^2; \theta_0) + f_2(t_j^2; \theta_0)^{\gamma_2} \epsilon_j^2, & j = 1, 2, \dots, n_2. \end{aligned}$$

This multiplicative structure of the observational error in the above statistical model exists because often in biological models the size of the observation error is proportional to the size of the observations. For $\gamma \geq 0$, a generalized least squares method or an iterative weighted least squares method as used below is appropriate to perform the inverse problem. In order to estimate θ_0 , we want to minimize the distance between the collected data and mathematical model, where the observables are weighted according to their variability and, for each observable, the observations over time are weighted unequally.

The iterative weighted least squares estimate $\hat{\theta}_{IWL S}$ is numerically determined by iteratively solving the following system :

$$\begin{aligned} \hat{\theta}_{IWL S} = \underset{\theta \in \Omega}{\text{argmin}} & \left(\sum_{i=1}^{n_1} [y_i^1 - f_1(t_i^1; \theta)]^T \hat{V}_1^{-1}(t_i^1) [y_i^1 - f_1(t_i^1; \theta)] \right. \\ & \left. + \sum_{j=1}^{n_2} [y_j^2 - f_2(t_j^2; \theta)]^T \hat{V}_2^{-1}(t_j^2) [y_j^2 - f_2(t_j^2; \theta)] \right) \end{aligned} \quad (2.2)$$

$$\hat{V}_1(t_i^1) = \frac{\hat{\omega}_i^1}{n_1 - p} \sum_{i=1}^{n_1} \frac{(y_i^1 - f_1(t_i^1; \hat{\theta}_{IWL S}))^2}{\hat{\omega}_i^1} \quad (2.3)$$

$$\hat{V}_2(t_j^2) = \frac{\hat{\omega}_j^2}{n_2 - p} \sum_{j=1}^{n_2} \frac{(y_j^2 - f_2(t_j^2; \hat{\theta}_{IWL S}))^2}{\hat{\omega}_j^2}, \quad (2.4)$$

where $\hat{\omega}_i^1 = f_1(t_i^1; \hat{\theta}_{IWL S})^{2\gamma_1}$ and $\hat{\omega}_j^2 = f_2(t_j^2; \hat{\theta}_{IWL S})^{2\gamma_2}$.

We use the following iterative procedure [9, 11, 25, 88] :

1. Estimate $\hat{\theta}_{IWL S}^{(0)}$ using (2.2) with $\hat{V}_1(t_i^1) = 1$ and $\hat{V}_2(t_j^2) = 1$. Set $k = 0$.
2. Compute weights $\hat{\omega}_i^{1(k)} = f_1(t_i^1; \hat{\theta}_{IWL S}^{(k)})^{2\gamma_1}$ and $\hat{\omega}_j^{2(k)} = f_2(t_j^2; \hat{\theta}_{IWL S}^{(k)})^{2\gamma_2}$.
3. Solve for $\hat{V}_1(t_i^1)^{(k)}$ and $\hat{V}_2(t_j^2)^{(k)}$ using $\hat{\theta}_{IWL S}^{(k)}$, $\hat{\omega}_i^{1(k)}$, and $\hat{\omega}_j^{2(k)}$ in equations (2.3) and (2.4) respectively.
4. Estimate $\hat{\theta}_{IWL S}^{(k+1)}$ using $\hat{V}_1(t_i^1)^{(k)}$ and $\hat{V}_2(t_j^2)^{(k)}$ in equation (2.2).

5. Set $k := k + 1$ and return to step 2. Terminate when two successive estimates for $\hat{\theta}_{IWLS}$ are sufficiently close.

Note that this is not the same as taking the derivative of the argument in the right side of (2.2) and setting it equal to zero. For more details see page 63 of [9] and page 89 of [88].

The estimated variances for each observable σ_{01}^2 and σ_{02}^2 are approximated by the following:

$$\hat{\sigma}_{01}^2 = \frac{1}{n_1 - p} \sum_{i=1}^{n_1} \left(\frac{y_i^1 - f_1(t_i^1; \hat{\theta}_{IWLS})}{f_1(t_i^1; \hat{\theta}_{IWLS})^{\gamma_1}} \right)^2$$

$$\hat{\sigma}_{02}^2 = \frac{1}{n_2 - p} \sum_{j=1}^{n_2} \left(\frac{y_j^2 - f_2(t_j^2; \hat{\theta}_{IWLS})}{f_2(t_j^2; \hat{\theta}_{IWLS})^{\gamma_2}} \right)^2.$$

If we assume $\gamma = (\gamma_1, \gamma_2) = (0, 0)$, then our statistical model is called an absolute error model and an ordinary least squares method is appropriate for parameter estimation. Banks *et al.* [8] consider an absolute error model and additionally assume that the variances for each observable are equal (i.e., $\sigma_{01}^2 = \sigma_{02}^2$). While the statistical model choice in [8] yields good results, we believe it is more biologically realistic to assume the variance in observation errors are not equal and the size of the observation error is proportional to the size of the observed quantity.

2.4 Difference-based Methods and Modified Residuals

We use a second order difference-based method to determine the correct statistical model (γ value) [6, 13]. Another method often implemented consists of performing an inverse problem with some γ value and computing the modified residuals

$$\mathcal{M}_{kl} = \frac{y_l^k - f_k(t_l^k; \hat{\theta})}{f_k(t_l^k; \hat{\theta})^{\gamma_k}}, \quad (2.5)$$

for each observable k at time $t_l, l = 1, \dots, n_k$. The plots of \mathcal{M}_{kl} vs. t_l should be randomly scattered around the horizontal axis. If an undesired non-random shape is present (e.g., a megaphone or inverted megaphone shape), then a different γ value is chosen and the process is repeated until a γ value produces the desired random scatter plot. However, this method does not consider both the mathematical model and statistical model misspecifications; it determines the correct statistical model under the tacit assumption that one has a correct mathematical model. It is also time consuming as it might take several attempts of performing an inverse problem (each of which

may be computationally expensive) and plotting the modified residuals until a good statistical model is chosen.

We follow [6] and first apply the second order difference-based method directly to the data to determine the correct γ value, which is both computationally economical as well as time efficient and independent of any assumed correct mathematical model. We first calculate the following pseudo measurement errors for observable k at time $t_l, l = 1, \dots, n_k$

$$\hat{\epsilon}_l^k = \begin{cases} \frac{1}{\sqrt{2}}(y_{l+1}^k - y_l^k) & \text{for } l = 1 \\ \frac{1}{\sqrt{6}}(y_{l-1}^k - 2y_l^k + y_{l+1}^k) & \text{for } l = 2, \dots, n_k - 1 \\ \frac{1}{\sqrt{2}}(y_l^k - y_{l-1}^k) & \text{for } l = n_k. \end{cases}$$

Next we calculate the modified pseudo errors

$$\eta_l^k = \frac{\hat{\epsilon}_l^k}{|y_l^k - \hat{\epsilon}_l^k|^{\gamma_k}}$$

for observable k at time $t_l, l = 1, \dots, n_k$ for different values of γ_k . (For derivation of modified pseudo errors see Appendix A.1.) We plot these modified pseudo errors vs. time for different γ_k values to find the γ_k value that produces a random scatter plot. Once the correct observational error is accounted for, we perform the inverse problem with this statistical model and compute the modified residuals in (2.5). If the modified residual plots are not randomly distributed around the horizontal axis, then the error must be due to mathematical model misspecification, implying another iteration of the modeling process is needed. Note that just using modified residuals is insufficient in detecting mathematical model misspecification; using modified pseudo errors in combination with modified residuals as described above would assist in identifying any discrepancies in mathematical model formulation.

2.5 Results

2.5.1 Clinical Data

Using second order differencing, we plot the modified pseudo errors for both viral load and creatinine versus time and visually assess the plots to choose an appropriate γ value. Figure 2.3 contains the graphs of the viral load modified pseudo errors vs. time for various γ_1 values. As can be seen due to the limited amount of data, it is difficult to determine the correct γ_1 value through visual assessment. The value $\gamma_1 = 0.5$ provides an approximately symmetric distribution around the horizontal axis with relatively small modified pseudo error values. The creatinine modified pseudo errors for different γ_2 values are given in Figure 2.4. The modified pseudo errors with $\gamma_2 = 0$ appear to be randomly distributed whereas the modified pseudo errors with $\gamma_2 = 1$ reveal a slight non-random (megaphone) shape. Even though we visually assess the plots to pick a suitable γ value to the best of our ability, the sparseness of the data set makes it difficult to make a stronger case for a particular statistical model.

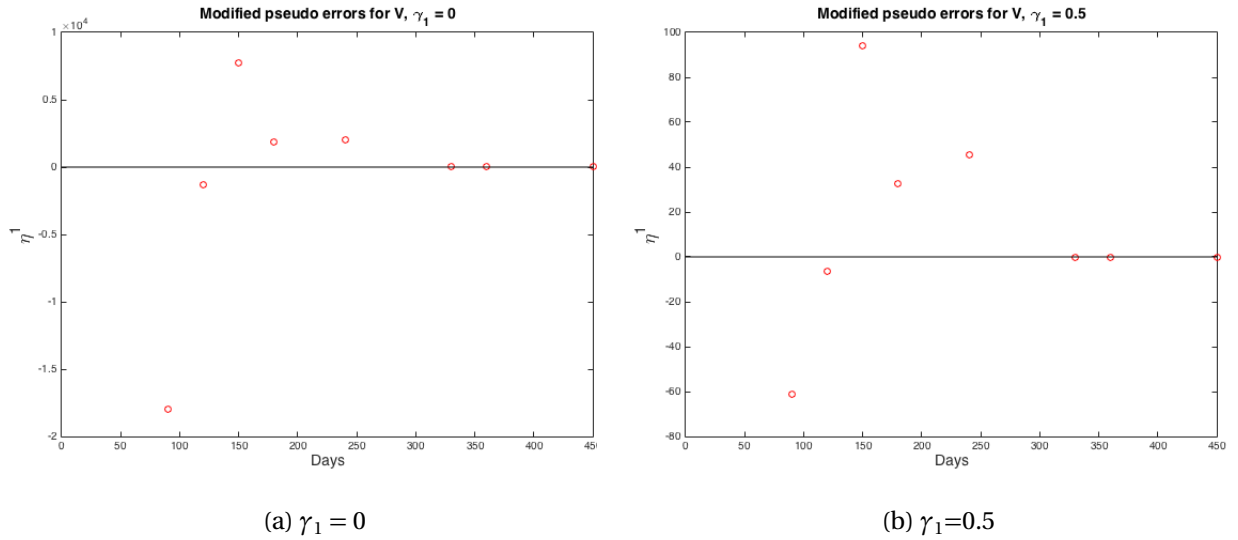
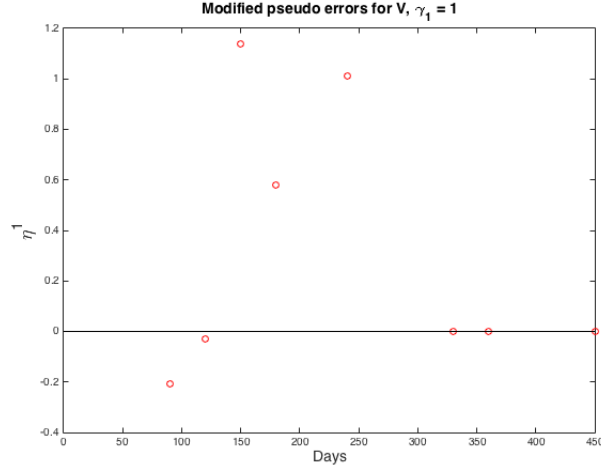
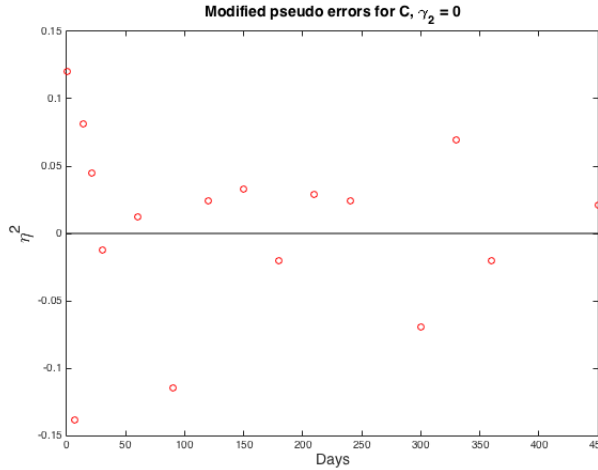


Figure 2.3: Viral load modified pseudo errors vs. time for various γ_1 values.

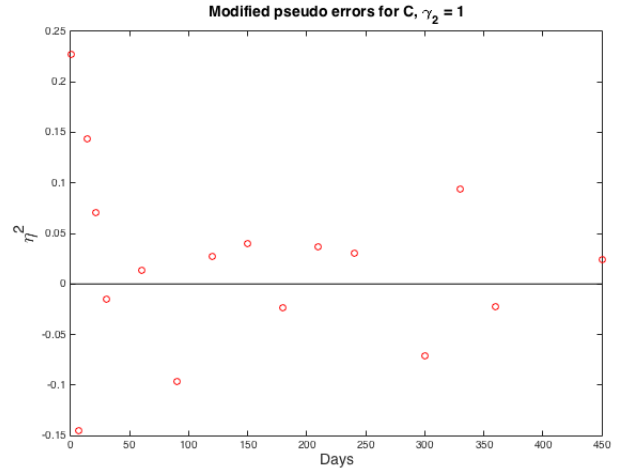


(c) $\gamma_1=1$

Figure 2.3 (cont.): Viral load modified pseudo errors vs. time for various γ_1 values.



(a) $\gamma_2 = 0$



(b) $\gamma_2=1$

Figure 2.4: Creatinine modified pseudo errors vs. time for various γ_2 values.

With our best guess of the correct statistical model, $\gamma = (0.5, 0)$, we next perform an inverse problem for the 5 most sensitive parameters [8] and obtain the modified residuals in order to detect the presence of mathematical model error. We perform the inverse problem using the MATLAB function `fmincon`. The initial guesses for the parameters are those used in [8] and lower and upper bounds are set for each of the 5 parameters for computational efficiency. The estimated parameters, $[\log_{10} \beta, \log_{10} \bar{\rho}_{EV}, \log_{10} \delta_{EV}, \log_{10} \delta_{EK}, \log_{10} \bar{\rho}_{EK}] = [-7.061, -0.632, -1.007, -0.92, -0.704]$. We can see from Figure 2.5 that the model solution fits the data well and the corresponding modified resid-

uals in Figure 2.6 appear to form a random band around the horizontal axis. This suggests that the mathematical model accurately describes the biological process, although again it is difficult to conclude this with conviction due to the limited amount of data.

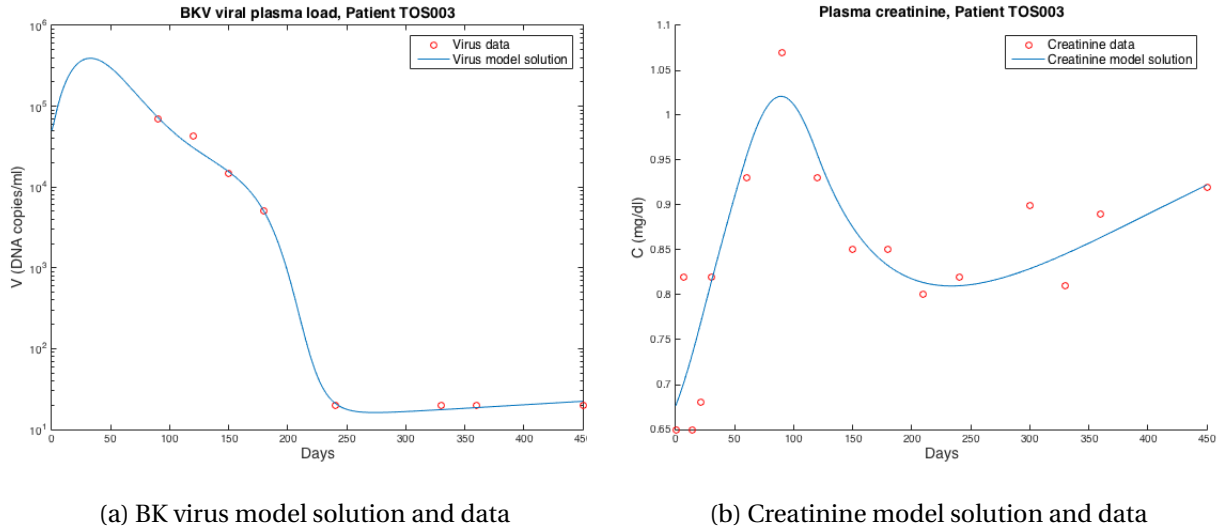


Figure 2.5: Model (2.1) solution and clinical data with $\gamma=(0.5,0)$.

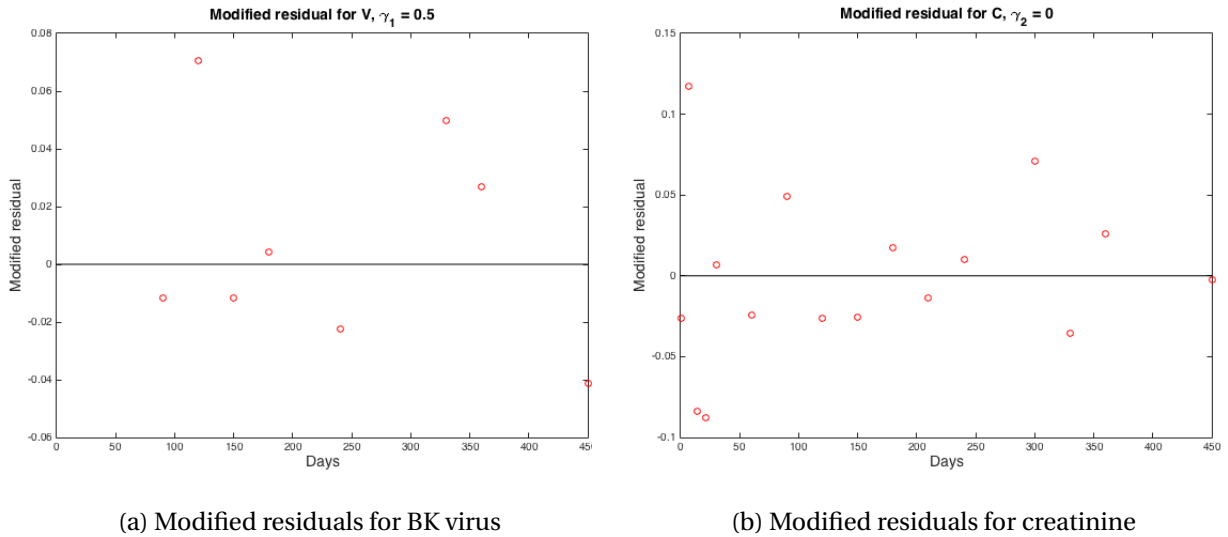


Figure 2.6: Modified residuals for V and C with $\gamma = (0.5,0)$.

2.5.2 Simulated Data

When using the second order difference-based method with sparse clinical data, it is not very easy to pick a statistical model or make a strong case for the presence/absence of mathematical model misspecification. To illustrate the need for a denser data set, we repeat the above process with the following simulated data created by adding noise to the “true” model solution

$$V_i = f_1(t_i^1; \theta_0) + f_1(t_i^1; \theta_0)^{\gamma_1} \epsilon_i^1, \quad (2.6a)$$

$$C_j = f_2(t_j^2; \theta_0) + f_2(t_j^2; \theta_0)^{\gamma_2} \epsilon_j^2, \quad (2.6b)$$

where $\epsilon^1 \sim N(0, 0.3)$, $\epsilon^2 \sim N(0, 0.03)$, $\gamma = (0.5, 0)$, and the estimated parameters $[\log_{10} \beta, \log_{10} \bar{\rho}_{EV}, \log_{10} \delta_{EV}, \log_{10} \delta_{EK}, \log_{10} \bar{\rho}_{EK}] = [-7.067, -0.601, -0.964, -0.995, -0.785]$. We assume that “data” is collected every week for $t^1 = t^2 = [0, 7, 14, \dots, 448]$. As expected, the second order differencing method produces the desired scatter plot for $\gamma = (0.5, 0)$ and undesired megaphone shapes for other γ values (see Appendix A.2 for corresponding figures). The modified residual plots also exhibit no mathematical model misspecification, which is expected since the data was created using the mathematical model (see Appendix A.3 for modified residual plots).

We now demonstrate how the modified residual plots can detect mathematical model error or misspecification by performing an inverse problem with a simpler version of model (2.1), given in (2.7). While the original model (2.1) assumes the susceptible population grows logistically where cell proliferation and the eventual plateauing of the population is dependent on the susceptible population size, the simpler model (2.7) assumes a growth rate of $\lambda_{HS} - \delta_{HS} H_S$, where H_S cells are produced at a constant rate λ_{HS} (cells/(mL·day)) independent of the cell population and die at a rate proportional to H_S with a proportionality constant or death rate of δ_{HS} (1/day). The simpler model is given by the following

$$\dot{H}_S = \lambda_{HS} - \delta_{HS} H_S - \beta H_S V \quad (2.7a)$$

$$\dot{H}_I = \beta H_S V - \delta_{HI} H_I - \delta_{EH} E_V H_I \quad (2.7b)$$

$$\dot{V} = \rho_V \delta_{HI} H_I - \delta_V V - \beta H_S V \quad (2.7c)$$

$$\dot{E}_V = (1 - \epsilon_I)[\lambda_{EV} + \rho_{EV}(V)E_V] - \delta_{EV} E_V \quad (2.7d)$$

$$\dot{E}_K = (1 - \epsilon_I)[\lambda_{EK} + \rho_{EK}(H_S)E_K] - \delta_{EK} E_K \quad (2.7e)$$

$$\dot{C} = \lambda_C - \delta_C(E_K, H_S)C \quad (2.7f)$$

where

$$\rho_{EV}(V) = \frac{\bar{\rho}_{EV} V}{V + \kappa_V}, \quad (2.7g)$$

$$\rho_{EK}(H_S) = \frac{\bar{\rho}_{EK} H_S}{H_S + \kappa_{KH}}, \quad (2.7h)$$

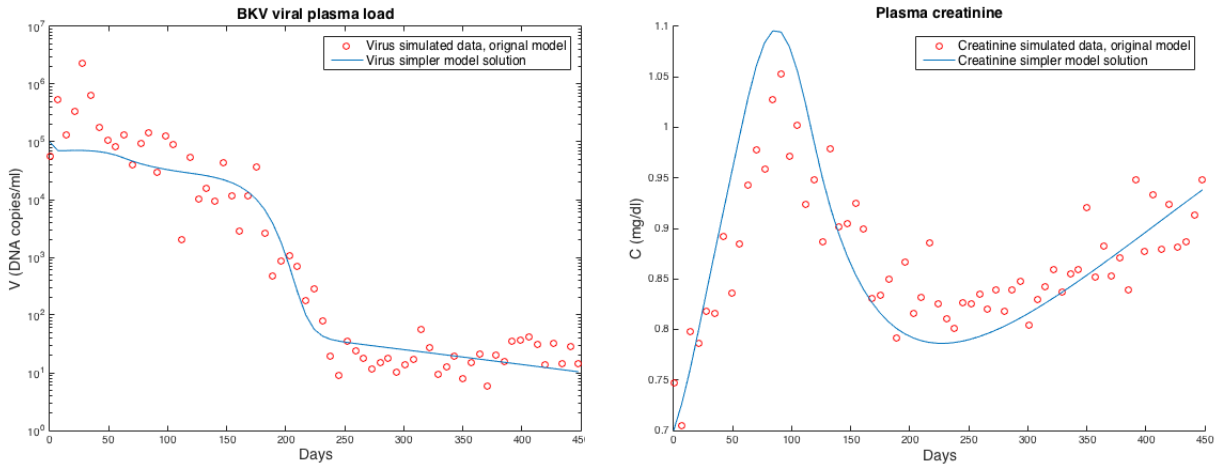
$$\delta_C(E_K, H_S) = \frac{\delta_{C0} \kappa_{EK}}{E_K + \kappa_{EK}} \cdot \frac{H_S}{H_S + \kappa_{CH}}, \quad (2.7i)$$

and initial conditions,

$$(H_S(0), H_I(0), V(0), E_V(0), E_K(0), C(0)) = (H_{S0}, H_{I0}, V_0, E_{V0}, E_{K0}, C_0). \quad (2.7j)$$

The immunosuppressant efficiency is defined by the piecewise constant function (2.1k).

We perform the inverse problem using $\gamma = (0.5, 0)$ to estimate the 6 parameters $\beta, \bar{\rho}_{EV}, \delta_{EV}, \delta_{EK}, \delta_{HS}, \bar{\rho}_{EK}$ and obtain the solutions in Figure 2.7. The estimated parameters, $[\log_{10} \beta, \log_{10} \bar{\rho}_{EV}, \log_{10} \delta_{EV}, \log_{10} \delta_{EK}, \log_{10} \delta_{HS}, \log_{10} \bar{\rho}_{EK}] = [-7.061, -0.632, -0.632, -1.007, -0.92, -0.704]$. Even though the simpler model produces a reasonable fit to the data, the modified residuals produce a strong non-random pattern (Figure 2.8). Since we already eliminated statistical error model discrepancy (through the difference-based method), these non-random modified residuals indicate a mathematical model misspecification. That is, the simpler model (2.7) is unable to accurately capture the dynamics represented in the data.



(a) BK virus model solution and simulated data

(b) Creatinine model solution and simulated data

Figure 2.7: Model (2.7) solution and simulated data created from model (2.1) with $\gamma=(0.5, 0)$.

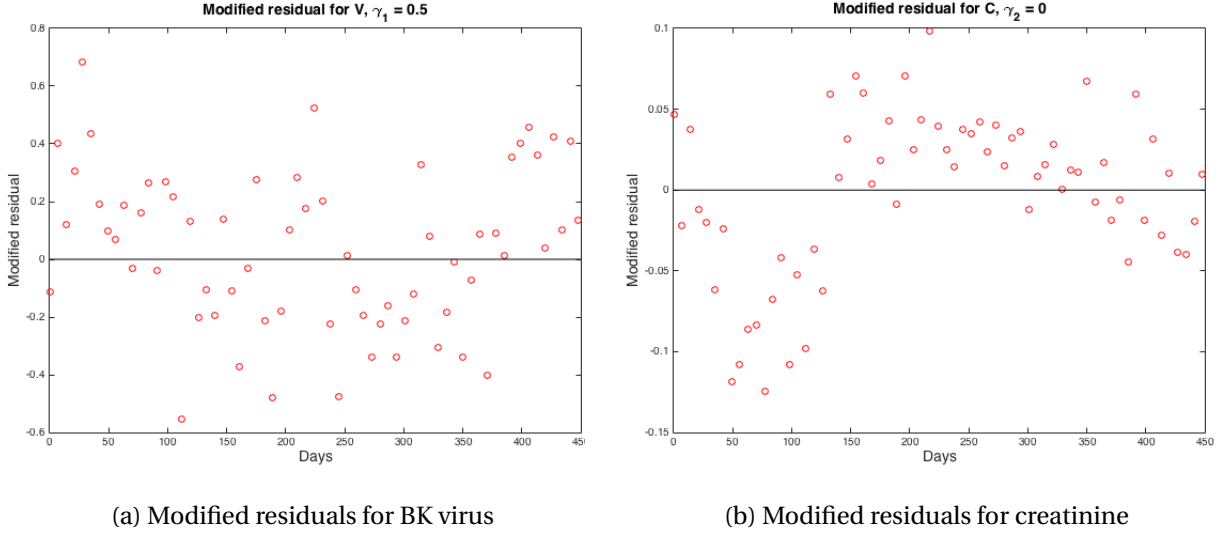
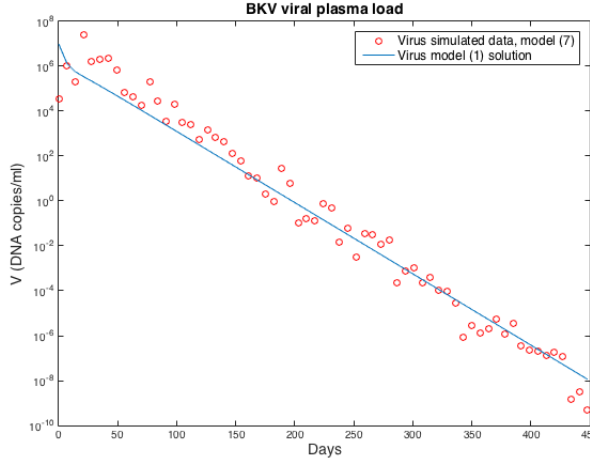


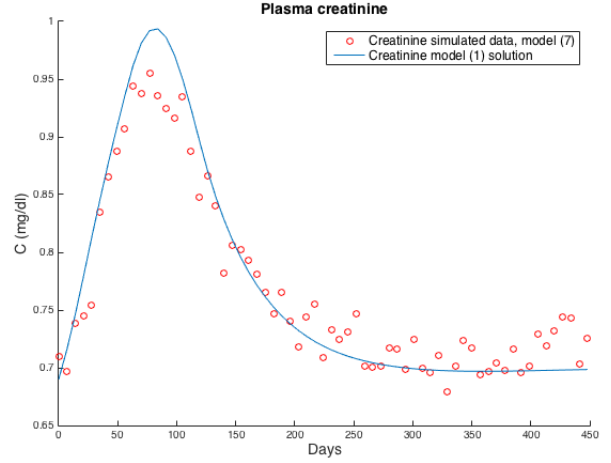
Figure 2.8: Modified residuals for V and C with $\gamma = (0.5, 0)$.

We next investigate mathematical model misspecification when a more complex model than warranted is assumed. To do so, we create a new simulated data set for $\mathbf{t}^1 = \mathbf{t}^2 = [0, 7, 14, \dots, 448]$ using (2.6) where \mathbf{f}_1 and \mathbf{f}_2 now represent $\log_{10} V$ and C in model (2.7), $\gamma = (0, 0.7)$, $\mathcal{E}^1 \sim N(0, 0.5)$, and $\mathcal{E}^2 \sim N(0, 0.02)$. Parameter values from Table 2.2 are used to create the new simulated data set with free parameters $[\log_{10} \beta, \log_{10} \bar{\rho}_{EV}, \log_{10} \delta_{EV}, \log_{10} \delta_{EK}, \log_{10} \bar{\rho}_{EK}] = [-7.067, -0.601, -0.964, -0.995, -0.785]$ and additional parameter $\delta_{HS} = 0.003/\text{day}$ [35].

As expected, the difference-based method with $\gamma = (0, 0.7)$ produces random scatter plots (see Appendix A.3). We perform the inverse problem with this data set and the original model (2.1). The estimate values for the 5 parameters are, $[\log_{10} \beta, \log_{10} \bar{\rho}_{EV}, \log_{10} \delta_{EV}, \log_{10} \delta_{EK}, \log_{10} \bar{\rho}_{EK}] = [-8.020, -0.744, -0.223, -0.863, -0.675]$. The model (2.1) solutions and corresponding modified residuals are plotted in Figure 2.9 and Figure 2.10. Even though the fit between the model and data looks acceptable, the modified residuals display a strong non-random pattern, indicating incorrect mathematical model assumptions.

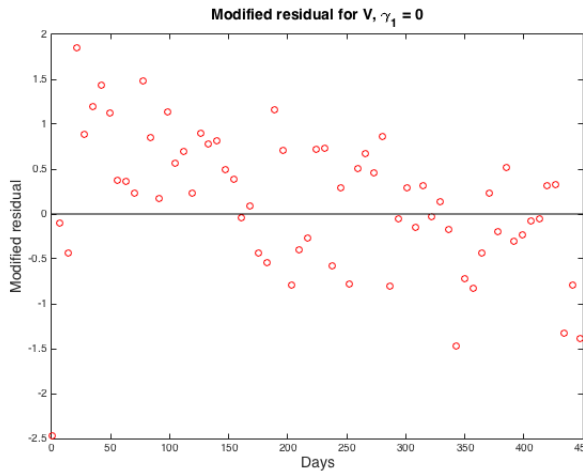


(a) BK virus model solution and simulated data

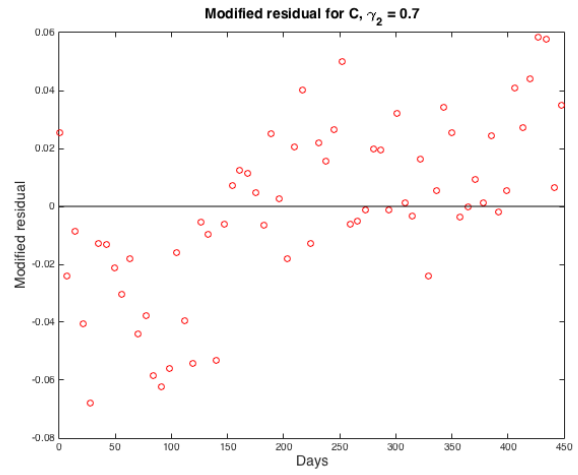


(b) Creatinine model solution and simulated data

Figure 2.9: Model (2.1) solution and simulated data created from model (2.7) with $\gamma=(0,0.7)$.



(a) Modified residuals for BK virus



(b) Modified residuals for creatinine

Figure 2.10: Modified residuals for V and C with $\gamma = (0,0.7)$.

2.6 Conclusion

We investigate mathematical model and statistical model misspecifications in the context of least squares methodology using a BKV model and both clinical and simulated data. Banks *et al.*, [8] use ordinary least squares techniques to perform an inverse problem with clinical data. We build on this work and assume what we believe is a more biologically realistic statistical error model; we

consider different variances for different observables and allow the error to depend on the size of the observables (measurements). We follow [6] and demonstrate how difference-based methods can be applied directly to data to determine the correct statistical model and further, we illustrate the use of modified residuals to detect mathematical model discrepancy. The presence of mathematical model error suggests possibly another iteration of modeling might be needed. However, due to the limited amount of clinical data, no strong conclusion can be reached.

We thereby demonstrate these methods using dense simulated data. We first create data using the BKV model (2.1) with an associated statistical model. The difference-based method correctly identifies the assumed γ value. Using this statistical model, we perform an inverse problem using a simpler model (2.7) where the nonlinearity is removed from the susceptible cell population growth dynamics. While the model (2.7) solution fits the data reasonably well, the modified residual plots depict a strong pattern, identifying the mathematical model discrepancy. We then repeat the process by creating data using the simpler model (2.7) and perform an inverse problem with the original model (2.1). That is, we now assume a more complex dynamical system in comparison to the biological process represented by the data. Again, the modified residuals indicate error in the mathematical model. Therefore this method can reveal mathematical model misspecification when either simpler or more complex models are assumed as compared to the data dynamics.

Previously, modified residual plots were solely used to determine the correct statistical model by iteratively performing multiple inverse problems until the correct statistical model was chosen [9]. Using both the difference-based method as well as modified residual plots is computationally more efficient; the difference-based method can be applied directly to the data and thus multiple inverse problems need not be performed. However, and even more notably, the previous method (using only modified residual plots) determines the statistical model under the possibly uninformed assumption of a correct mathematical model; the use of both the difference-based method and modified residuals accounts for both types of error in the inverse problem without prior model assumptions.

CHAPTER

3

IMMUNOSUPPRESSANT TREATMENT DYNAMICS IN RENAL TRANSPLANT RECIPIENTS: AN ITERATIVE MODELING APPROACH

3.1 Introduction

A patient undergoing a solid organ transplantation is usually put on a lifetime regime of immunosuppressive medications to prevent the body from rejecting the allograft [51]. However, these immunosuppressive treatments often leave the recipient susceptible to opportunistic pathogens including viruses. Achieving the delicate balance between under-suppression and over-suppression of the immune system is key to successful and sustainable transplantation. Mathematical and statistical models can be important and beneficial tools that contribute to the improvement and optimization of treatment protocols. Modeling of any biological process is an iterative one, as seen in Figure 3.1. The biologist first presents a research question about some biological process as well as knowledge about biological relationships and mechanisms, often depicted by a schematic or diagram. The mathematician then represents the biological hypothesis of the relationships through

a mathematical model. Analytic or numerical analysis of the model produces results which are interpreted and compared to the biological system, possibly leading to a change in the understanding of the biological relationships. The process is continuously repeated, sometimes through multiple research efforts, using the resulting biological insights (See [11] for more details on the iterative process of modeling).

In this chapter we first explain and analyze the existing model (2.1) seen in Chapter 2 which captures the biological mechanisms of the immune response in renal transplantation patients with respect to infection caused by the human polyomavirus type 1, named “BK virus” (BKV) [8]. Our results show evidence of the lapses in biological understanding and implementation of the corresponding mathematical model. We then attempt to address and correct for the discrepancies between the mathematical model and the biological system by revising components of the mathematical model, thereby illustrating the iterative process of modeling. Once a mathematical model is developed, it could in turn be used for control theory applications to predict optimal drug regimens, an important problem we will address in the coming chapters.

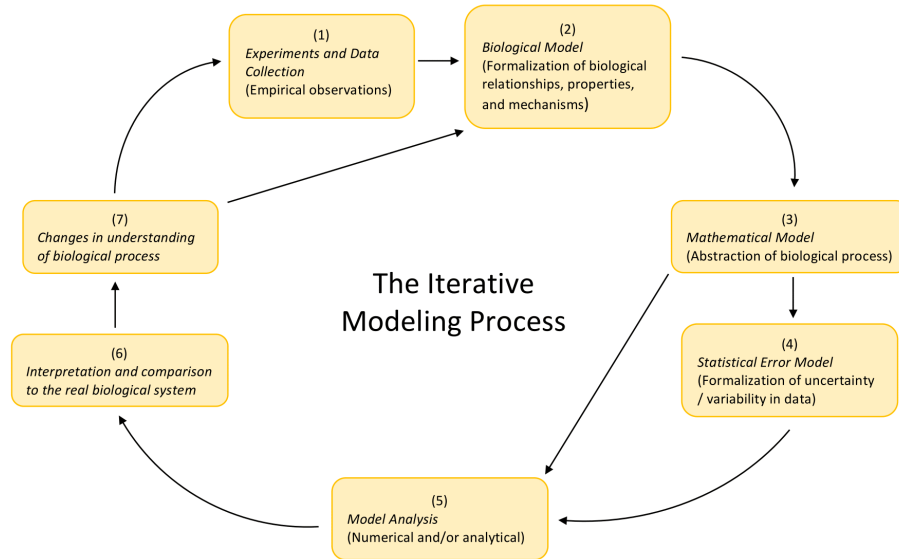


Figure 3.1: Schematic of the iterative modeling process [11].

3.2 Iteration I: Preliminary Model

3.2.1 Data Collection and Biological Model

Data used to fit the model and obtain some of the model parameters in [8] was collected at Massachusetts General Hospital from a renal transplant patient TOS003 diagnosed with BKV in the first 3 months of kidney transplantation. (This data was collected with approval of the MGH human subjects (IRB) committee. Furthermore data was shared with NCSU in a de-identified manner.) Eight BK viral plasma load (DNA copies/mL) measurements and sixteen plasma creatinine level (mg/dL) measurements were collected (See Chapter 3, Figure 2.2 for plot of datasets). Creatinine levels are used as a surrogate for GFR (Glomerular Filtration Rate), to measure kidney function. Due to the sparsity of data collected, we do not estimate parameters for Iteration II and Iteration III of our modeling efforts, instead we use most of the parameters from literature.

The authors in [8] describe a model schematic of the compartmentalized biological model as shown in Figure 2.1 in Chapter 2.

Since no effective anti-viral treatment exists for BKV, the authors in [8] modify the model in [7] to only consider the efficiency of immunosuppressant treatment. Additionally, the authors model the effect of the susceptible cells on creatinine clearance and on the allospecific CD8+T-cell population growth. As opposed to Funk et al. [35], the authors in [8] do not consider the urothelial cells and only consider infection in tubular epithelial cells, in part due to the availability of data. While this makes the model more specific, one can hope to expand the current model scope to include urothelial cells once relevant data becomes available.

3.2.2 Mathematical Model

The mathematical model as seen in Chapter 2, Equation (2.1), describes the concentrations of susceptible cells (H_S), infected cells (H_I), free BKV (V), BKV-specific CD8+T-cells (E_V), allospecific CD8+T-cells (E_K) that target the kidney, and serum creatinine (C), by the following system of ordinary differential equations [8]:

$$\dot{H}_S = \lambda_{HS} \left(1 - \frac{H_S}{\kappa_{HS}} \right) H_S - \beta H_S V \quad (3.1a)$$

$$\dot{H}_I = \beta H_S V - \delta_{HI} H_I - \delta_{EH} E_V H_I \quad (3.1b)$$

$$\dot{V} = \rho_V \delta_{HI} H_I - \delta_V V - \beta H_S V \quad (3.1c)$$

$$\dot{E}_V = (1 - \epsilon_I) [\lambda_{EV} + \rho_{EV}(V) E_V] - \delta_{EV} E_V \quad (3.1d)$$

$$\dot{E}_K = (1 - \epsilon_I) [\lambda_{EK} + \rho_{EK}(H_S) E_K] - \delta_{EK} E_K \quad (3.1e)$$

$$\dot{C} = \lambda_C - \delta_C(E_K, H_S) C, \quad (3.1f)$$

where

$$\rho_{EV}(V) = \frac{\bar{\rho}_{EV} V}{V + \kappa_V}, \quad (3.1g)$$

$$\rho_{EK}(H_S) = \frac{\bar{\rho}_{EK} H_S}{H_S + \kappa_{KH}}, \quad (3.1h)$$

$$\delta_C(E_K, H_S) = \frac{\delta_{C0} \kappa_{EK}}{E_K + \kappa_{EK}} \cdot \frac{H_S}{H_S + \kappa_{CH}}. \quad (3.1i)$$

The initial conditions are given by,

$$(H_S(0), H_I(0), V(0), E_V(0), E_K(0), C(0)) = (H_{S0}, H_{I0}, V_0, E_{V0}, E_{K0}, C_0). \quad (3.1j)$$

$$\epsilon_I(t) = \begin{cases} \epsilon_1 & t \in [0, 21] \\ \epsilon_2 & t \in (21, 60] \\ \epsilon_3 & t \in (60, 120] \\ \epsilon_4 & t \in (120, 450]. \end{cases} \quad (3.2)$$

Detailed explanation of the dynamics of the model and model states can be found in Subsection 2.2.1 of Chapter 2.

3.2.3 Statistical Error Model

The authors in [8] assume the simplest statistical error model, an absolute error model, where the variances of the error for each observable (viral load and creatinine) are equal and constant over time. That is, the authors account for the uncertainty in the dataset by the following statistical error model

$$\begin{aligned} Y_i^1 &= f_1(t_i^1; \theta_0) + \mathcal{E}_i^1, & i = 1, 2, \dots, n_1, \\ Y_j^2 &= f_2(t_j^2; \theta_0) + \mathcal{E}_j^2, & j = 1, 2, \dots, n_2. \end{aligned}$$

The functions $f_1(t_i^1; \theta_0)$ and $f_2(t_j^2; \theta_0)$ represent the model solution for viral load and creatinine at times t_i^1 and t_j^2 respectively, assuming a “true” or nominal parameter set θ_0 . The existence of this “true” parameter set is a standard assumption in statistical models [9].

The $\mathcal{E}_i^1, \mathcal{E}_j^2$ terms represent the measurement error that causes the measurements to differ from the model solution with the “true” parameter set. We assume the $n_1 \times 1$ and $n_2 \times 1$ random vectors \mathcal{E}_i^1 and \mathcal{E}_j^2 respectively, are independent and identically distributed with mean zero and $Var(\mathcal{E}_i^1) = \sigma_{01}^2$ and $Var(\mathcal{E}_j^2) = \sigma_{02}^2$. The authors in [8] assume $\sigma_{01}^2 = \sigma_{02}^2$ (Note the model is log scaled as shown in Chapter 2). The corresponding method for parameter estimation is ordinary least squares (OLS).

In Chapter 2, we build on the works of [8] and consider a more general (and possibly more biologically realistic) statistical error model [13]. We assume the variances of observation errors are not equal and allow for the errors to depend on the size of the observed quantity. That is, we assume the following statistical error model, a relative error model, given by

$$\begin{aligned} Y_i^1 &= f_1(t_i^1; \theta_0) + f_1(t_i^1; \theta_0)^{\gamma_1} \mathcal{E}_i^1, & i = 1, 2, \dots, n_1, \\ Y_j^2 &= f_2(t_j^2; \theta_0) + f_2(t_j^2; \theta_0)^{\gamma_2} \mathcal{E}_j^2, & j = 1, 2, \dots, n_2, \end{aligned}$$

for $\gamma_k \geq 0, k = 1, 2$. The measurement error term now can depend on the size of the model solution of the observables. Note that if both $\gamma_1 = 0$ and $\gamma_2 = 0$, the two statistical error models are equivalent. The corresponding method, for parameter estimation, assuming a relative error model is iterative weighted least squares (IWLS).

Recall in Chapter 2, we use a second order difference-based method to eliminate statistical error model misspecification by selecting the correct statistical error model directly from the data. We show how modified residuals from the inverse problem can then be used to detect discrepancies in

mathematical model formulation. However, we also point out that due to sparsity of data available to us, it is difficult to ascertain which statistical model is suitable for this specific dataset [13].

3.2.4 Model Analysis

The authors of [8] numerically analyze the model by performing an inverse problem with the data to estimate parameters. That is, the authors seek to find a parameter set that minimizes the distance between the collected data and mathematical model. However, the model has a large number of parameters (29 parameters) and thus not all the parameters can be reliably estimated with such a small dataset of 24 observations. Thus, the authors of [8] implement an iterative procedure to determine the most sensitive parameters. An inverse problem is then performed to estimate the 5 most sensitive parameters, and the resulting model solutions provide a reasonable fit to the data (see [8] for details). However, in Chapter 2 we present examples which illustrate that a good fit is not necessarily enough to conclude the accuracy of a model. We conclude that further data collection endeavors are needed to reduce the uncertainty in parameter choices made for the model [13].

Our primary motivation of mathematically modeling the immune response to the allograft and BK virus is to eventually formulate a control problem to adaptively predict patient specific optimal immunosuppressant dosage as treatment progresses. Specifically, the immunosuppressant dosage level should keep viral loads low (less than 10,000 copies/ml [28, 63, 64]) and creatinine levels within a healthy range (0.6 – 1.1 mg/dL [7]). Figure 3.2 depicts the general dynamics of both the BK virions and CD8+ T-cells for varying amounts of immunosuppressants.

As a first step towards a feedback control strategy, we implemented a simple open loop control problem (where the output has no influence or effect on the control input). The results (not shown) provided inadequate validation to the mathematical model's robustness for formulation and design of the control. Hence, we considered a simpler test: to simulate model solutions for the highest and lowest immunosuppressant dosage and observe the model dynamics. Note that, for all simulations, we assume a constant immunosuppression efficiency (i.e., $\epsilon_1 = \epsilon_2 = \epsilon_3 = \epsilon_4$ in (3.2)), whereas the authors in [8] considered different constant values for immunosuppression dosage, $\epsilon_1 = 0.1009$, $\epsilon_2 = 0.3658$, $\epsilon_3 = 0.5999$, and $\epsilon_4 = 0.3649$.

Recall that a drug efficiency of 0% ($\epsilon_I = 0$) implies that the immune system is not compromised and the body fights the virus effectively (i.e., BK viral loads are under control and there is negligible infection); however the immune response treats the kidney transplant as a foreign object and attacks it, killing the susceptible cells and causing the creatinine levels to increase. On the other hand, a drug efficiency of 100% ($\epsilon_I = 1$) causes the immune system to be significantly weakened (Note this is

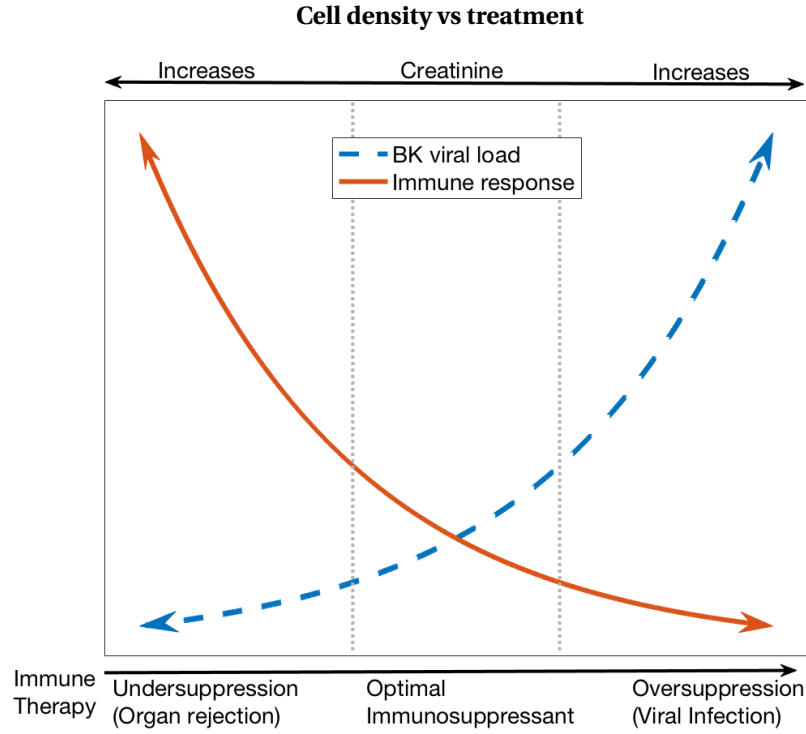


Figure 3.2: Plot illustrating the balance between under and over suppression of the immune response.

an extreme case of immunosuppression therapy and is usually not prescribed for individuals). While it is no longer a threat to the allograft, the immune system is now unable to defend itself against outside infections (in our case BKV infection), causing an increase in viral load. These dynamics are summarized in Table 3.1.

Table 3.1: Summary of cell dynamics under influence of immunosuppression.

ϵ	CD8+ T-cells	BKV	Infected cells	Susceptible cells	Creatinine
Low	↑	↓	↓	↓	↑
High	↓	↑	↑	↓	↑

The model solutions in Figure 3.3 depict that the viral load, the infected cells and the susceptible cells are all fairly impervious to the immune response which is contradictory to the dynamics depicted in Table 3.1. Furthermore, Figure 3.3b illustrates that creatinine is sensitive to the control where in fact it should be increasing for both extremes of immunosuppression, a phenomenon not captured

by the current model (3.1).

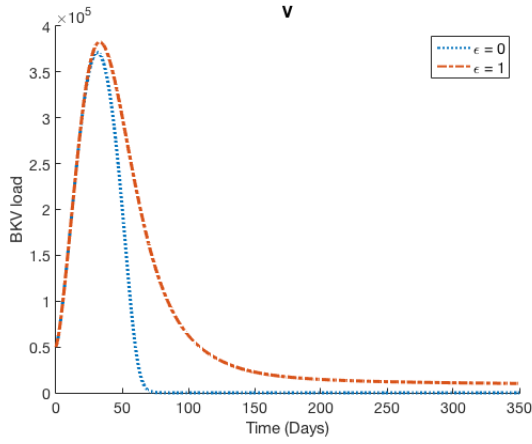
We also notice from the model solutions with $\epsilon_I = 1$ in Figure 3.3, both the BK viral load and infected cell population increase as expected, but then decrease. A large viral load indicates BK virus-associated nephropathy (viral loads $> 185,000$ copies/mL) [28], where the kidney susceptible cells are damaged due to infection. The decrease observed in Figures 3.3a and 3.3d implies that there are no more remaining susceptible cells for the virus to infect. However, we see in Figure 3.3c that the susceptible population remains around carrying capacity, suggesting a discrepancy between the model and biological understanding.

These findings from our analysis prompted us to delve further in understanding the biological interpretation and the parameter values for this model.

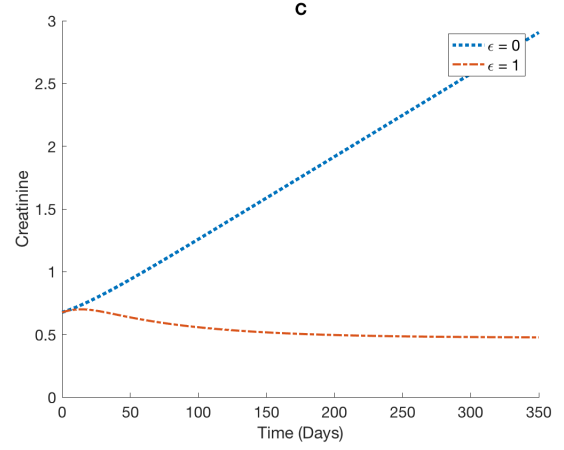
3.2.5 Biological Interpretation of Model and Changes in Understanding

Based on our findings in the previous section, we renewed our effort to understand the biological interpretation of the mathematical model and analysis from [8]. Enumerated below is a list of biological discrepancies in the model.

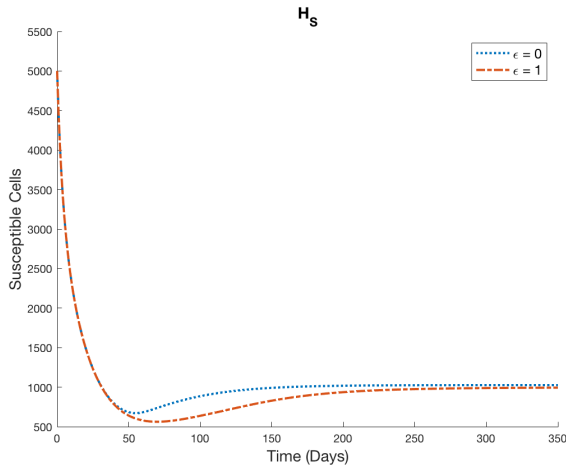
1. The mammalian kidney is a non-regenerative organ. While the kidneys can self-repair certain small sections of the nephron, loss of nephrons due to chronic kidney injury is irreversible, resulting in permanent damage and impaired renal function [19, 34, 49]. In model (3.1) the logistic growth term represents the regeneration and proliferation of susceptible kidney cells. This term is not representative of the true biological phenomenon.
2. The factors contributing to the decay in the susceptible cell population as seen in (3.1a) are due to infection by the BK virus. However, a large population of allo-specific CD8+ T-cells also attack susceptible cells, causing a decrease in both the total H_S population and kidney function (observed via creatinine levels). This biological mechanism is not captured in model (3.1).
3. The estimated initial condition for viral load in [8] was approximately 50,000 copies/mL which is higher than the threshold viral load for initial detection of viremia (10,000 copies/mL [28, 63, 64]). This would imply that the patient had an active viral infection just before and during their kidney transplantation, which again represents an unusual situation.



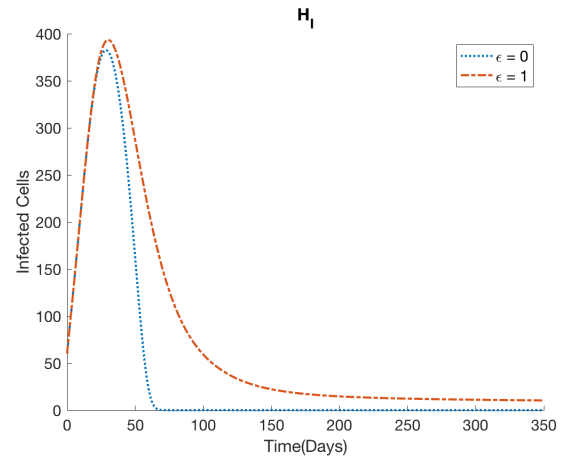
(a) BK viral load in blood.



(b) Creatinine.



(c) Susceptible Cells.



(d) Infected Cells.

Figure 3.3: BK viral load in blood, creatinine, susceptible and infected cell dynamics for highest and lowest immunosuppressant dosages.

3.3 Iteration II: Improved Model

3.3.1 Mathematical Model

Based on our model analysis and our renewed biological understanding, we propose another iteration of modeling which encapsulates our new observations. The updated BKV model is presented

below:

$$\dot{H}_S = -\chi_{(E_K > E_K^*)} \tilde{\beta} H_S E_K - \chi_{(V > V^*)} \beta H_S V \quad (3.3a)$$

$$\dot{H}_I = \chi_{(V > V^*)} \beta H_S V - \delta_{HI} H_I - \chi_{(E_V > E_V^*)} \delta_{EH} E_V H_I \quad (3.3b)$$

$$\dot{V} = \rho_V \delta_{HI} H_I - \delta_V V - \chi_{(V > V^*)} \beta H_S V \quad (3.3c)$$

$$\dot{E}_V = (1 - \epsilon_I) [\lambda_{EV} + \rho_{EV}(V) E_V] - \delta_{EV} E_V \quad (3.3d)$$

$$\dot{E}_K = (1 - \epsilon_I) [\lambda_{EK} + \rho_{EK}(H_S) E_K] - \delta_{EK} E_K \quad (3.3e)$$

$$\dot{C} = \lambda_C - \delta_C(H_S) C \quad (3.3f)$$

where

$$\rho_{EV}(V) = \frac{\bar{\rho}_{EV} V}{V + \kappa_V}, \quad (3.3g)$$

$$\rho_{EK}(H_S) = \frac{\bar{\rho}_{EK} H_S}{H_S + \kappa_{KH}}, \quad (3.3h)$$

$$\delta_C(H_S) = \delta_{C0} \cdot \frac{H_S}{H_S + \kappa_{CH}}. \quad (3.3i)$$

As before initial conditions are given by,

$$(H_S(0), H_I(0), V(0), E_V(0), E_K(0), C(0)) = (H_{S0}, H_{I0}, V_0, E_{V0}, E_{K0}, C_0). \quad (3.3j)$$

Following the justification presented in Section 3.2.5, the logistic term in (1a) has been removed. The additional term $\tilde{\beta} H_S E_K$ represents the loss of healthy susceptible cells when under attack from allo-specific CD8 + T-cells. The parameter $\tilde{\beta}$ represents the death rate of H_S by E_K . The term $\beta H_S V$ continues to represent the infection of susceptible cells by free virions. However, we additionally assume that for trace levels of both allospecific CD8 + T-cells and viral load, the susceptible cells are not destroyed and are constant. That is, at trace population levels there is negligible interaction. We approximate this phenomenon mathematically with the following characteristic or indicator function χ

$$\chi_{(x > x^*)} = \begin{cases} 1, & \text{for } x > x^* \\ 0, & \text{otherwise.} \end{cases} \quad (3.4)$$

Again we see the presence of the characteristic function with the infection term $\beta H_S V$ in equation (3b), indicating that for low levels of viral load there is no infection. The infected cell population

can also decrease due to elimination by the BK-specific CD8+T-cells at rate δ_{EH} only when there is a sufficient number of immune cells present, hence the characteristic function. The dynamics of BK virus remain the same as in (3.1) except for the additional characteristic term in (3c).

The parameter $\epsilon_I \in [0, 1]$ continues to represent the efficiency of immunosuppressive drugs where we assume a drug efficiency of 0% ($\epsilon_I = 0$) indicates that the treatment does not affect the immune system and a drug efficiency of 100% ($\epsilon_I = 1$) is the highest dose that curbs the immune response fully. However, we also know that the dosage, type, and concentration of drugs often change over the course of treatment. With devising a control problem formulation as our eventual goal in improving this model, we replace the stepwise function from the previous iteration (3.2) with a single control parameter, ϵ_I , which we aim to be able to control and predict over time.

In model (3.1) the loss of creatinine was assumed to be a function of both the susceptible cells and the allo-specific CD8+T-cells that target the kidney. Our revised model more explicitly incorporates the effect of allo-specific CD8+T-cells on the healthy susceptible cells in (3a), thus we do not need the E_K dependent decay function in (3i). Note that the model is very sensitive to the threshold concentrations, V^* , E_K^* and E_V^* . We chose a V^* value that was significantly below detection level for BK Virus related viremia (10,000 copies/mL [28, 63, 64]). Thresholds E_K^* and E_V^* were deduced from total populations of CD8+T-cells observed during numerical simulations.

A description of the state variables, parameters and initial conditions are given in Tables 2.1, 3.2 and 3.3 respectively.

Table 3.2: Original (Iteration I) and new model (Iteration II) parameters.

Parameter	Description	Unit	Iteration I	Iteration II	Justification
λ_{HS}	Proliferation rate for H_S	/day	0.03	-	
κ_V	Half saturation constant	copies/mL	180	10^6	See A
κ_{HS}	Saturation constant	cells/mL	1025	-	
β	Attack rate on H_S by E_K	mL/(cells-day)	-	0.0001	
λ_{EK}	Source rate of E_K	cells/(mL-day)	0.002	285	[7] See B
β	Infection rate of H_S by V	mL/(copies-day)	8.22×10^{-8}	8.22×10^{-8}	
δ_{EK}	Death rate of E_K	/day	0.103	0.09	
δ_{HI}	Death rate of H_I by V	/day	0.085	0.085	
λ_C	Production rate for C	mg/(dL-day)	0.01	0.01	
ρ_V	# Virions produced by H_I before death	copies/cells	4292.4	15000	3 – 44,000 [36]
δ_{C0}	Maximum clearance rate for C	/day	0.014	0.2	[7]
δ_{EH}	Elimination rate of H_I by E_V	mL/(cells-day)	0.0018	0.0018	
κ_{EK}	Half saturation constant	cells/mL	0.2	-	
δ_V	Natural clearance rate of V	/day	0.37	0.05	0.04 – 15.12 [8, 35, 36]
κ_{CH}	Half saturation constant	cells/mL	10	10^4	See A
λ_{EV}	Source rate of E_V	cells/(mL-day)	0.001	285	[7] See B
ρ_{EK}	Maximum proliferation rate for E_K	/day	0.164	0.137	
δ_{EV}	Death rate of E_V	/day	0.11	0.17	
κ_{KH}	Half saturation constant	cells/mL	84.996	10^3	See A
ρ_{EV}	Maximum proliferation rate for E_V	/day	0.25	0.36	
V^*	Threshold concentration of BKV	copies/mL	-	1000	
E_K^*	Threshold concentration of Allospecific CD8+ T-cells	cells/mL	-	2500	
E_V^*	Threshold concentration of BKV specific CD8+ T-cells	cells/mL	-	500	

Table 3.3: Initial conditions both original (Iteration I) and new (Iteration II).

State	Iteration I IC	Iteration II IC	Justification
H_{S0}	5×10^3 cells/mL	1025 cells/mL	Assume $H_{S0} = \kappa_{HS}$ from [8]
H_{I0}	60 cells/mL	2×10^{-16} cells/mL	Trace infection before transplant
V_0	5×10^4 copies/mL	1200 copies/mL	Minimal V of 10,000 copies/mL for low BK viremia detection [28, 63, 64]
E_{K0}	0.04 cells/mL	2×10^{-16} cells/mL	Negligible amounts of Allospecific CD8+ T-cells
E_{V0}	0.4 cells/mL	100 cells/mL	Low level of BKV specific CD8+ T-cells
C_0	1.07 mg/dL	0.7 mg/dL	Range 0.6 -1.1 [7]

Below are the detailed justifications for parameter value changes, as shown in Table 3.2.

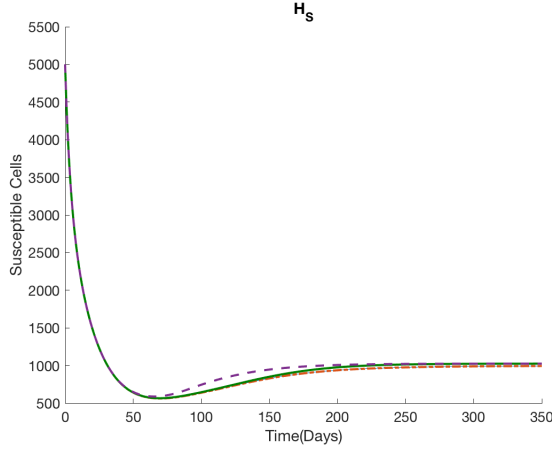
A. The value for parameter κ_V changed from 180 to 10^6 copies/mL. Recall that κ_V represents the half saturation constant in the term $\rho_{EV}(V) = \frac{\bar{\rho}_{EV} V}{V + \kappa_V}$. In Figure 3.3, we notice that viral loads reach approximately 4×10^5 copies/mL. If $\kappa_V \ll 4 \times 10^5$, $\rho_{EV}(V) \approx \bar{\rho}_{EV}$, a constant. Similarly, if $\kappa_V \approx 4 \times 10^5$, $\rho_{EV}(V) \approx \frac{\bar{\rho}_{EV}}{2}$. These choices for κ_V make the BK dependent growth rate of E_V impervious to viral load values. Thus we set $\kappa_V = 10^6$ in order to capture the sensitivity of the growth rate of E_V to V . Similarly we notice that susceptible cell populations are primarily between orders of magnitude 10^2 cells/mL and 10^3 cells/mL, hence we chose $\kappa_{KH} = 10^3$ and $\kappa_{CH} = 10^4$.

B. The authors in [8] assume low source rates λ_{EV} and λ_{EK} for BKV specific and allospecific cells respectively, resulting in the BKV load proliferating at a much higher rate compared to the CD8+T-cells. Thus, the BKV dynamics become impervious to the immune response, rendering it insensitive to the control over time. The authors in [7] consider $\lambda_{EV} = \lambda_{EK} = 0.5$ cells/ μ L-blood-day. Using the conversion 1 cell/ μ L-blood-day = 570 cells/mL-plasma-day, we obtain $\lambda_{EV} = \lambda_{EK} = 285$ cells/mL-plasma-day.

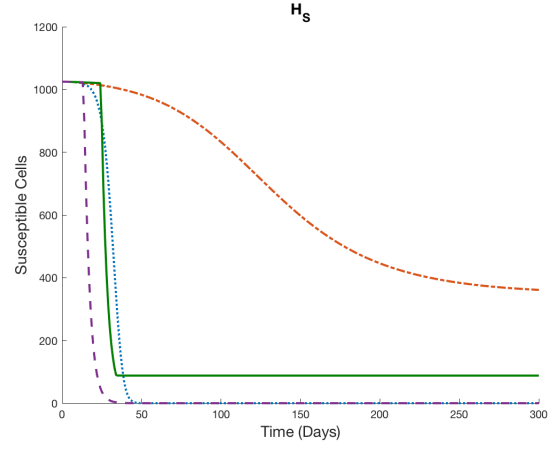
3.3.2 Model Analysis and Biological Interpretation

Figure 3.4 contains model solutions with parameter values and initial conditions in Tables 3.2 and 3.3 respectively. We examine the solutions for both models for varying levels of immunosuppressant efficiency to determine if our current mathematical model more accurately represents the biological process. The simulations show higher sensitivity to drug dosage with the modified model in Iteration II. We also notice that for the improved model for both higher and lower dosages of immunosuppressant, the kidney function fails, as seen from the rise in creatinine levels (either due to infection or rejection). There are intermediate values of immunosuppressant dosages for which the creatinine levels stay low. Lastly, we pick a low initial viral load as we assume that the transplant recipient does not have an active BKV infection just before and during transplantation. We can see that the modified model solutions are more biologically representative of the true dynamics as described in Table 3.1. In Figure 3.4 we consistently use the following line markers to depict the curves :

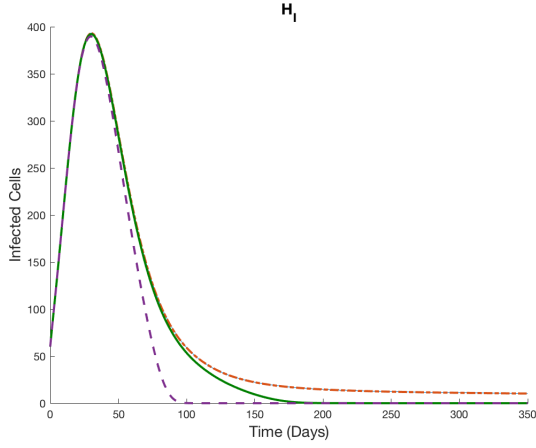
$$\cdots \epsilon_I = 0.8 \quad \text{---} \epsilon_I = 0.6 \quad \text{—} \epsilon_I = 0.4 \quad \text{- -} \epsilon_I = 0.2$$



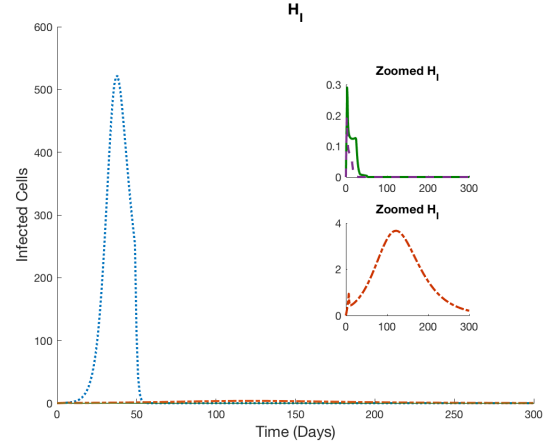
(a) Susceptible cells (Iteration I)



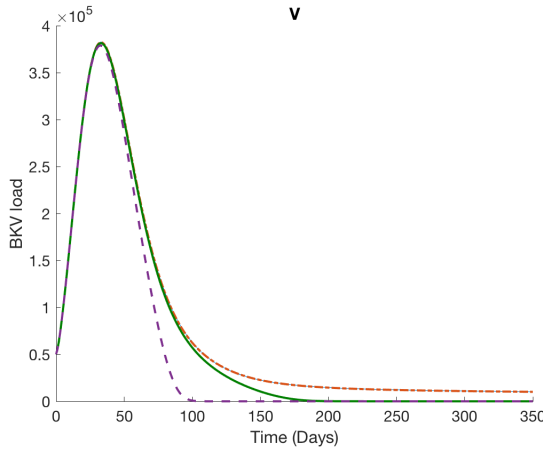
(b) Susceptible cells (Iteration II)



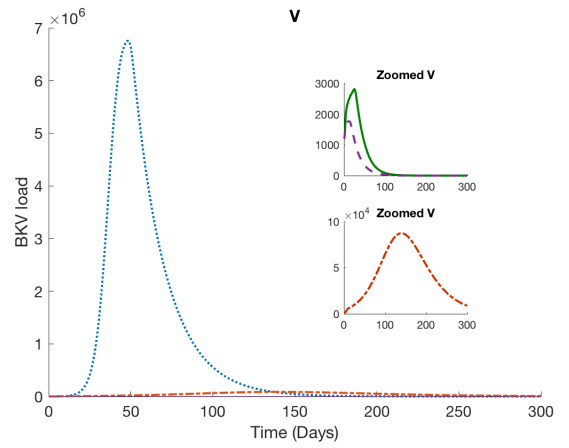
(c) Infected cells (Iteration I)



(d) Infected cells (Iteration II)

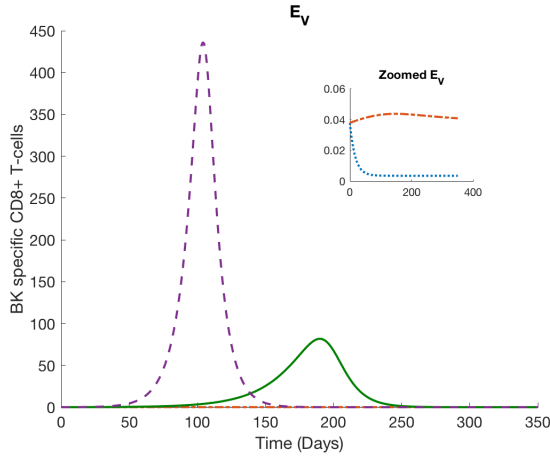


(e) BK Viral load in blood (Iteration I)

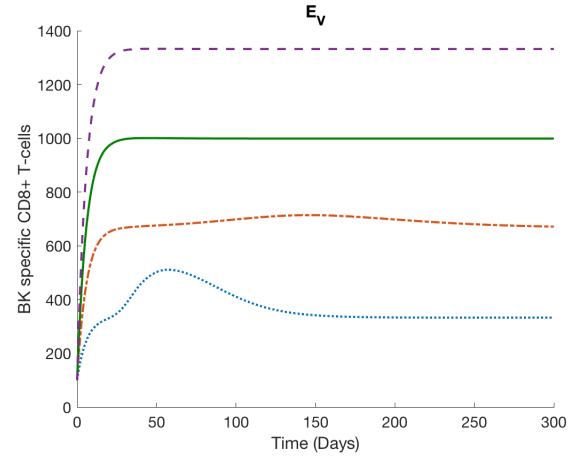


(f) BK Viral load in blood (Iteration II)

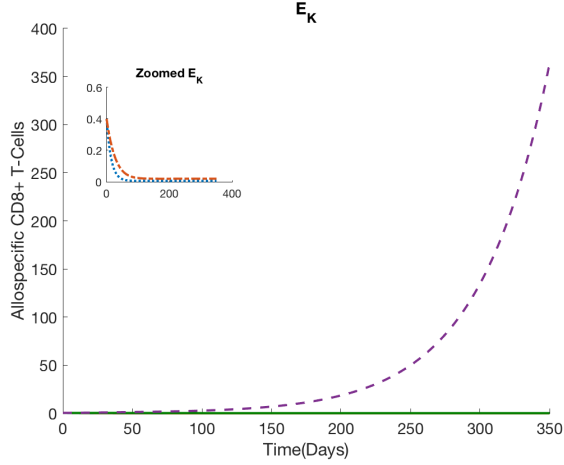
Figure 3.4: Model simulations for both Iteration I and II of modeling for different ϵ_I values.



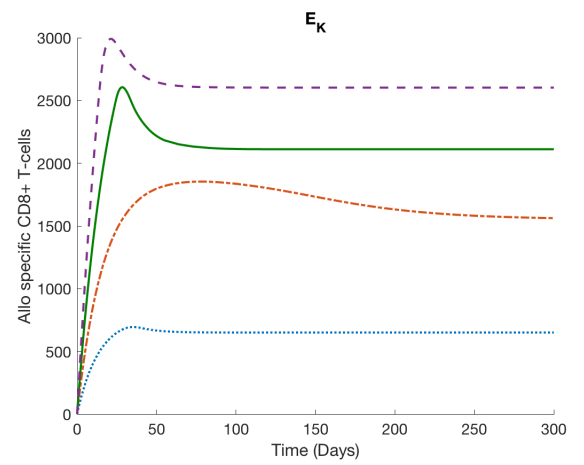
(g) BK specific CD8+ T-cells (Iteration I)



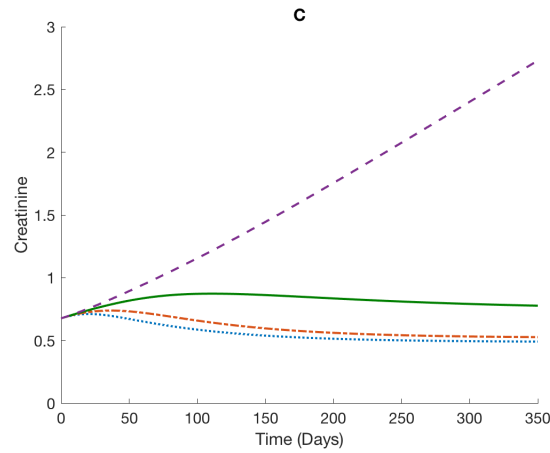
(h) BK specific CD8+ T-cells (Iteration II)



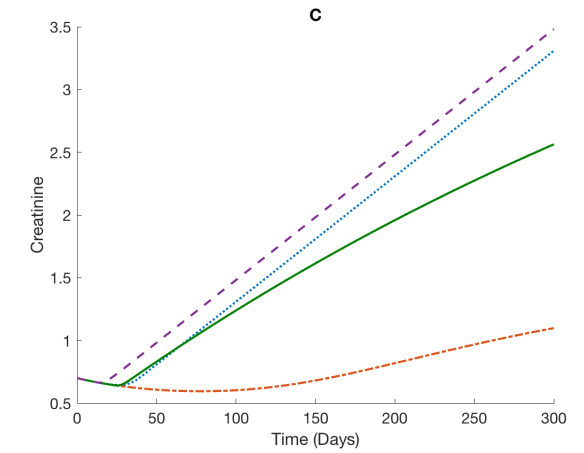
(i) Allospecific CD8+ T-cells (Iteration I)



(j) Allospecific CD8+ T-cells (Iteration II)



(k) Creatinine (Iteration I)



(l) Creatinine (Iteration II)

Figure 3.4 (cont.): Model simulations for both Iteration I and II of modeling for different ϵ_I values.

3.4 Iteration III: Further Prospective Improvements

3.4.1 Mathematical Model

The mathematical model (3.3) presented in Iteration II has an indicator or characteristic function which makes the system of differential equations used to represent the biological system discontinuous. While all our robustness tests (more details of these tests can be seen in Chapter 4) to corroborate the suitability of the model to build a feedback control loop were successful, we still wanted to verify if indeed the discontinuity had little affect on the model outcomes. We replaced each of the discontinuous characteristic functions in model (3.3) (and the associated parameter) with a smooth function as seen below in model (3.5) and Figure 3.5. Recall from Table 3.2 that when using the characteristic functions in model (3.3) the threshold parameter values for E_K^* , V^* and E_V^* were 2500 cells/mL, 1000 copies/mL and 500 cells/mL respectively.

$$\dot{H}_S = -\zeta_1(E_K)H_S E_K - \zeta_2(V)H_S V \quad (3.5a)$$

$$\dot{H}_I = \zeta_2(V)H_S V - \delta_{HI}H_I - \zeta_3(E_V)E_V H_I \quad (3.5b)$$

$$\dot{V} = \rho_V \delta_{HI}H_I - \delta_V V - \zeta_2(V)H_S V \quad (3.5c)$$

$$\dot{E}_V = (1 - \epsilon_I)[\lambda_{EV} + \rho_{EV}(V)E_V] - \delta_{EV}E_V \quad (3.5d)$$

$$\dot{E}_K = (1 - \epsilon_I)[\lambda_{EK} + \rho_{EK}(H_S)E_K] - \delta_{EK}E_K \quad (3.5e)$$

$$\dot{C} = \lambda_C - \delta_C(H_S)C \quad (3.5f)$$

where

$$\zeta_1(E_K) = \frac{\tilde{\beta}}{2} \left(\frac{1 - e^{(-E_K + 2500)}}{1 + e^{(-E_K + 2500)}} + 1 \right), \quad (3.5g)$$

$$\zeta_2(V) = \frac{\beta}{2} \left(\frac{1 - e^{(-V + 1000)}}{1 + e^{(-V + 1000)}} + 1 \right), \quad (3.5h)$$

$$\zeta_3(E_V) = \frac{\delta_{EH}}{2} \left(\frac{1 - e^{(-E_V + 500)}}{1 + e^{(-E_V + 500)}} + 1 \right), \quad (3.5i)$$

and

$$\rho_{EV}(V) = \frac{\bar{\rho}_{EV} V}{V + \kappa_V}, \quad (3.5j)$$

$$\rho_{EK}(H_S) = \frac{\bar{\rho}_{EK} H_S}{H_S + \kappa_{KH}}, \quad (3.5k)$$

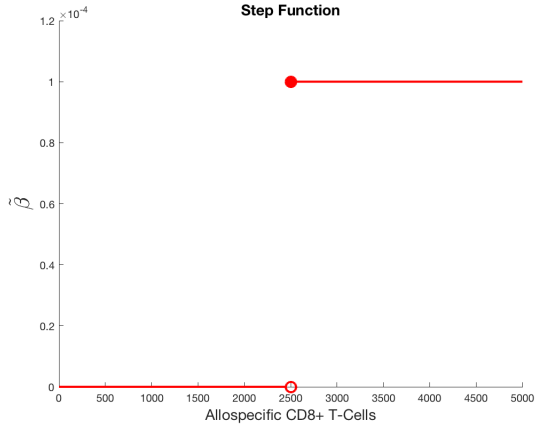
$$\delta_C(H_S) = \delta_{C0} \cdot \frac{H_S}{H_S + \kappa_{CH}}. \quad (3.5l)$$

As before initial conditions are given by,

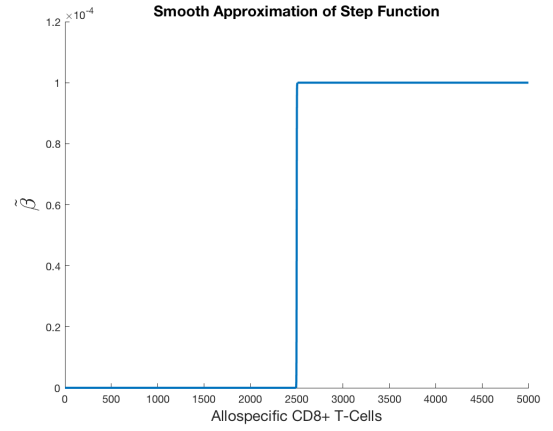
$$(H_S(0), H_I(0), V(0), E_V(0), E_K(0), C(0)) = (H_{S0}, H_{I0}, V_0, E_{V0}, E_{K0}, C_0). \quad (3.5m)$$

3.4.2 Model Analysis and Inference

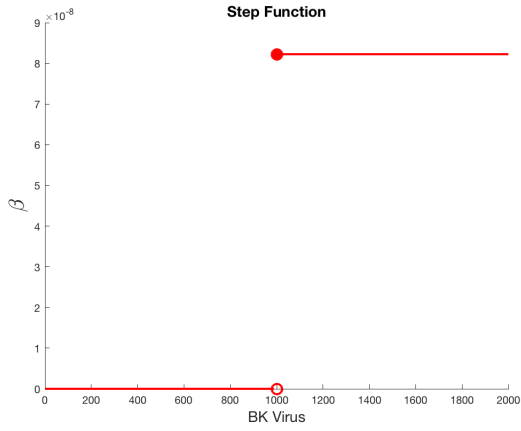
We next use the new model (3.5) to observe model simulations for varying immunosuppressant efficiency levels. Observe in Figure 3.6 the model simulations for Iteration II and Iteration III are identical. We conclude that the use of the characteristic functions in model (3.3) in Iteration II is valid and Iteration III, while a good test need not necessarily be considered as the most improved version of the model (owing to the indistinguishable model dynamics observed in Figure 3.6).



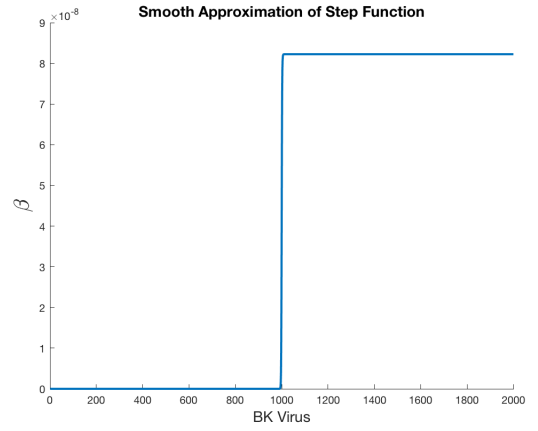
(a) Step function $\tilde{\beta} \chi_{(E_K > E_K^*)}$



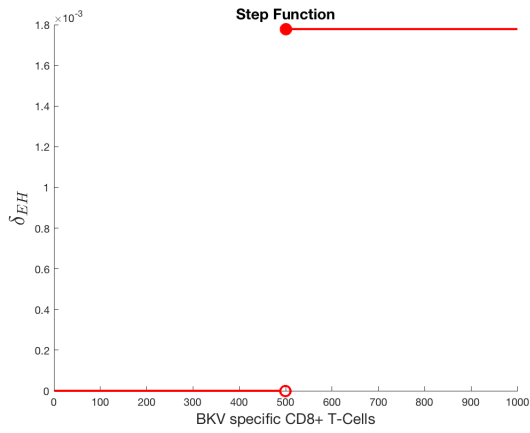
(b) Smooth approximation $\zeta_1(E_K)$



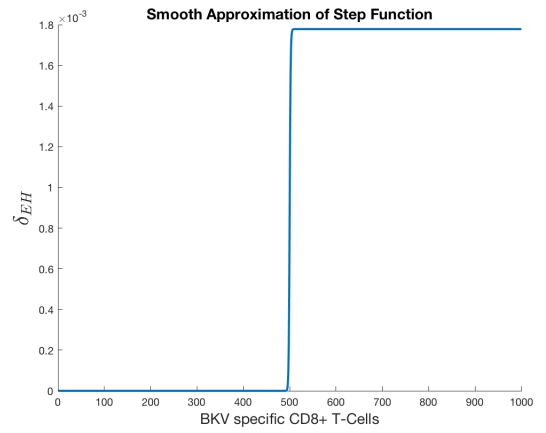
(c) Step function $\beta \chi_{(V > V^*)}$



(d) Smooth approximation $\zeta_2(V)$

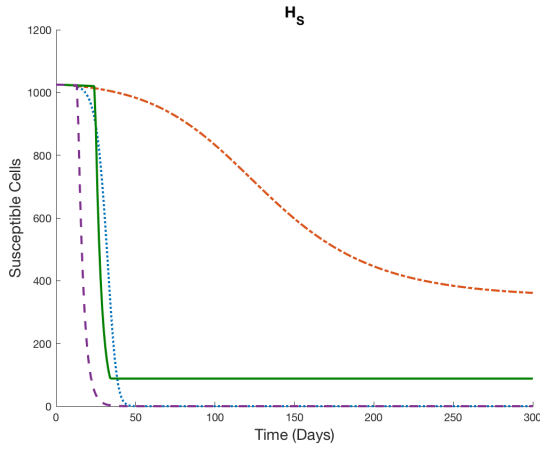


(e) Step function $\delta_{EH} \chi_{(E_V > E_V^*)}$

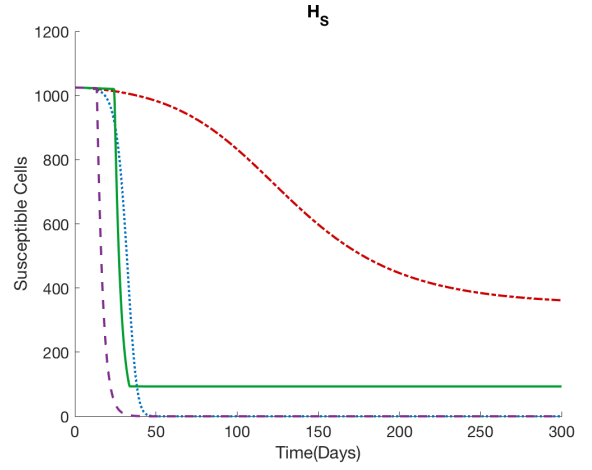


(f) Smooth approximation $\zeta_3(E_V)$

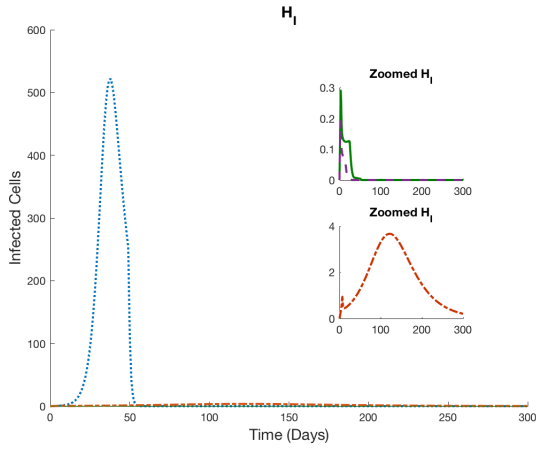
Figure 3.5: Figures showing the step functions used in Iteration II model (3.3) and their smooth approximations used in Iteration III model (3.5).



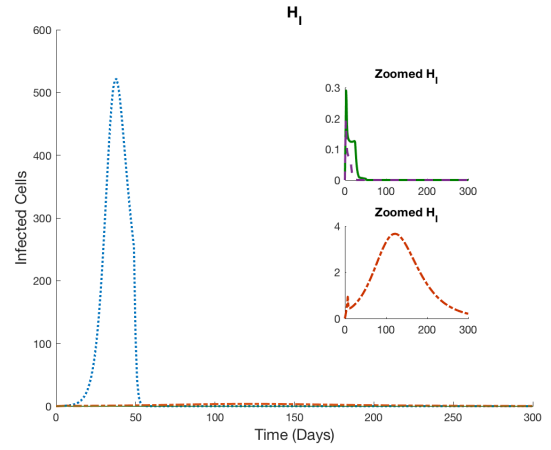
(a) Susceptible cells (Iteration II)



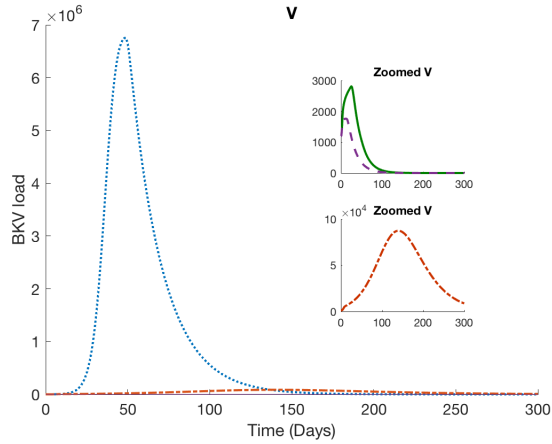
(b) Susceptible cells (Iteration III)



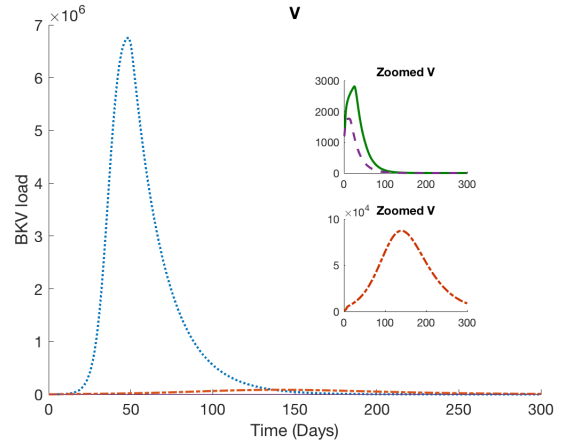
(c) Infected cells (Iteration II)



(d) Infected cells (Iteration III)

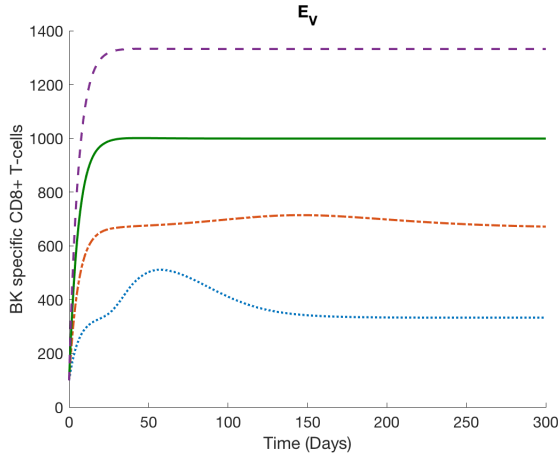


(e) BK Viral load in blood (Iteration II)

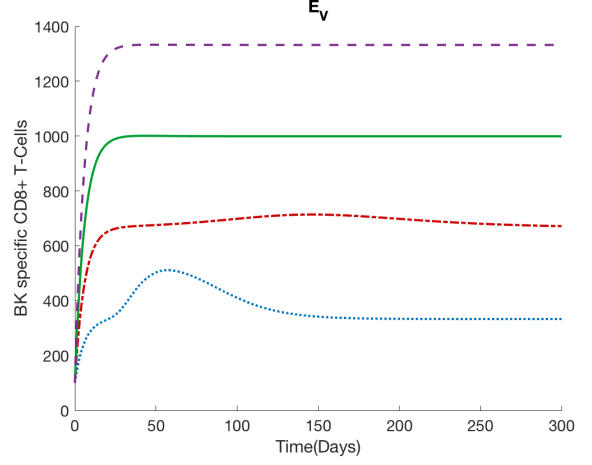


(f) BK Viral load in blood (Iteration III)

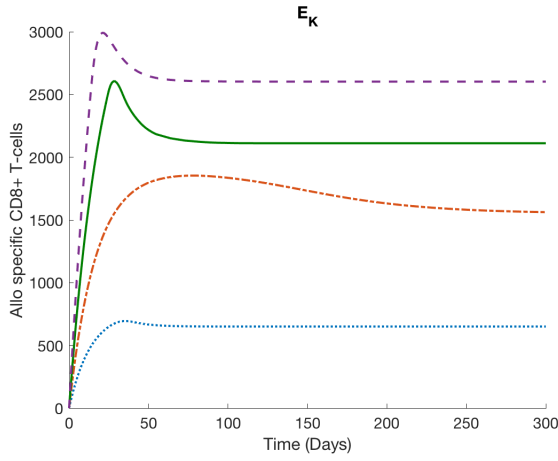
Figure 3.6: Model simulations for Iteration II and III of modeling for different ϵ_I values.



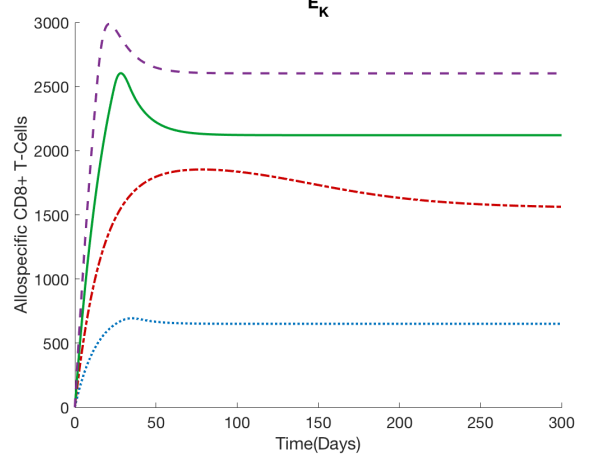
(g) BK specific CD8+ T-cells (Iteration II)



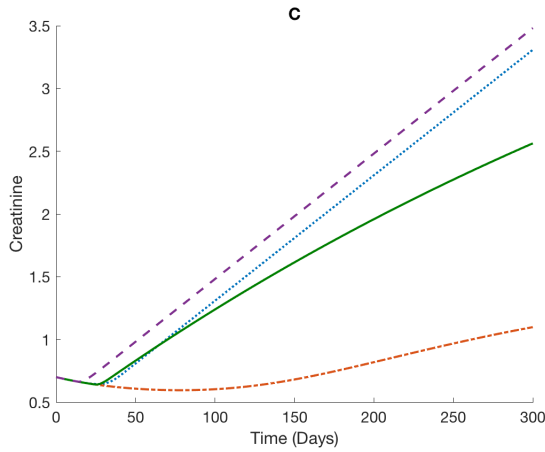
(h) BK specific CD8+ T-cells (Iteration III)



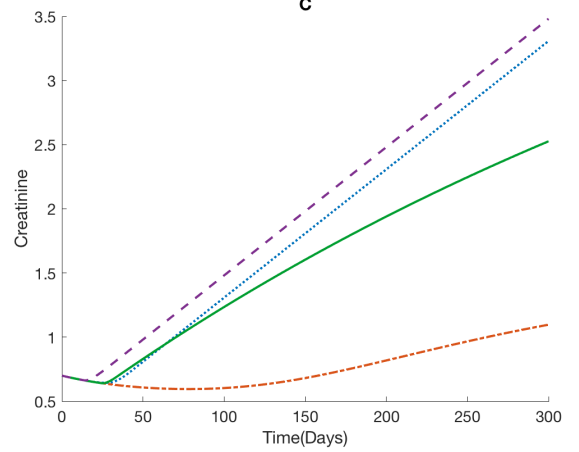
(i) Allospecific CD8+ T-cells (Iteration II)



(j) Allospecific CD8+ T-cells (Iteration III)



(k) Creatinine (Iteration II)



(l) Creatinine (Iteration III)

Figure 3.6 (cont.): Model simulations for Iteration II and III of modeling for different ϵ_I values.

3.5 Discussion

To illustrate the iterative modeling process, we analyze and modify a mathematical model of the immune response in renal transplant recipients infected with BK virus. Motivated by the problem of implementing a control strategy for individual patients, our investigation of the preliminary model indicates a discrepancy between the model and the biological process, prompting another iteration of modeling. The modified model (3.3) now represents the biological process of infection and organ rejection more accurately, as seen in Figure 3.4.

Next to ensure that the discontinuous nature of the characteristic functions in model (3.3) introduced in Iteration II does not hinder model performance, we performed another iteration of modeling. In Iteration III, we approximated the discontinuous characteristic functions with smooth functions. We analyzed the new model's (3.5) solutions for varying levels of immunosuppression and concluded based on the undifferentiated nature of the model plots from Iteration II and Iteration III, that the model in Iteration II is sufficient to design our optimal control upon. Henceforth all simulations in Chapter 4 will be done using model (3.3) from Iteration II of the iterative modeling process.

Future work involves possible further improvements to model (3.3). For instance, as more data (viral load and creatinine measurements) become available we should obtain more confidence in the parameters. Using a diverse data set spanning over several patients one would also be able to conduct further analysis to see if the parameters change substantially for different patients. We are currently exploring the specific effects of individual immunosuppressant drugs on the mechanisms of the immune response by understanding how each of these drugs individually and in combination with others inhibit the functioning of the CD8+ T-cells (some of which is presented in Chapter 5). This might lead to an improvement in the understanding of how immunosuppressed CD8+ T-cells grow and decay in the presence of foreign cells, and in turn might lead to another iteration of modeling.

With this new model (3.3), we will next develop the adaptive optimal control problem to determine the optimal level of immunosuppressive therapy for individual patients to balance over-suppression and under-suppression, as seen in Chapter 4.

OPTIMAL CONTROL OF IMMUNOSUPPRESSANTS IN RENAL TRANSPLANT RECIPIENTS SUSCEPTIBLE TO BKV INFECTION

4.1 Introduction

Personalized medicine investigators attempt to find subgroups of similar patients who may need a different course of action or adaptively treat an individual patient based on their response to treatment [82]. One example, among many, of using mathematical models to enhance personalized treatment is the application of control theory to determine an optimal treatment strategy. There are several related biomedical applications of modeling and control theory [12], such as determining the optimal treatment regimen for cancer patients undergoing chemotherapy in order to minimize tumor density as well as treatment side effects [65]. The authors in [65] also apply control theory to human immunodeficiency virus (HIV)-infected patients undergoing chemotherapy and also to determine the optimal insulin injection level to better regulate blood glucose levels in diabetic patients. Further applications of control theory to HIV therapy strategies are presented in [3, 10,

52]. The authors in [87] present control problems related to cancer therapies, such as antiangiogenic treatments. There are several diseases that require immunosuppressant treatment, besides solid organ transplantation, in which control theory could potentially be beneficial. Allogeneic hematopoietic stem-cell transplant recipients receive immunosuppression therapy to diminish the risk of developing graft-versus-host disease (GvHD), the attack of donor T-cells on the host tissues [90]. Immunosuppression therapy is also one of the current treatment options for autoimmune diseases, which result from the body's immune system attacking its own body [20]. The authors in [7] present a control problem in the context of HCMV-infected kidney transplant recipients. Unlike HCMV infection, the lack of an approved antiviral therapy for symptomatic BKV nephropathy makes the task of carefully monitoring the level of immunosuppression an even more challenging but imperative one.

Quantitative measurements of BK virus in the blood can strongly suggest BK nephropathy, but a graft biopsy with *in situ* hybridization or immunohistochemical techniques is required for a definitive diagnosis. Because there is no proven drug treatment for BK nephropathy, current therapy relies on careful reduction of immunosuppression (with the unavoidable risk of rejection) and options to use intravenous gamma globulin (IVIg) and/or low-dose cidofovir.

While reducing the level of immunosuppression might help keep infections caused by BKV at bay preventing the occurrence of PVAN (which leads to kidney failures), it also makes the immune response stronger. A stronger immune response in turn leads to allograft rejection. Thus the task of achieving the optimum immunosuppression level so that the body reaches the fragile balance between being under-suppressed (and prone to organ rejection) and over-suppressed (and susceptible to infections) is a difficult one (see Figure 3.2 in Chapter 3). In this chapter our goal is to use the improved mathematical model (3.3) developed from Chapter 3 to design an optimal control problem which would find the ideal amount of immunosuppressant dosage required to help achieve both the goal of fighting BKV infection as well as not rejecting the transplant.

The two biomarkers that can be readily measured to ascertain kidney health and infection status from blood plasma samples in a renal transplant patient are the BKV load and creatinine. We will use these two measurements to determine the optimum amount of immunosuppressant dosage.

When designing a control problem there are two broad classifications to consider, an open-loop control and a closed-loop control. An open-loop optimal control is one where the control problem is formulated in a way that the optimization is based on just the initial observations. So for our problem, we would devise a control formulation based on just the initial creatinine and BKV load then predict the optimal amount of immunosuppressant a transplant recipient would need for the rest of his life. As can be expected the body, especially one that had undergone organ transplantation, is an unstable and unpredictable system and making predictions for drug dosage based on just an

initial observation of biomarkers is not a feasible medical strategy.

A closed-loop optimal control problem on the other hand would be more practically suited for our problem as it is formulated in a way that the optimization is based on the most current observations which are updated periodically. Measurements of BKV load and creatinine are usually taken every time a renal transplant patient comes for a check up; such data gets updated routinely with every doctor's visit. Thus using a closed loop optimal control formulation (also called a feedback optimal control formulation), the optimal level of immunosuppression is adjusted at each visit based on the current measurements of viral load and creatinine.

With the model (3.3) from Chapter 3, we formulate in Section 4.3 an optimal feedback control strategy for a simulated transplant recipient susceptible to BKV infection. We first present the model and parameters and then formulate an open loop control to demonstrate the feasibility of eventually designing a feedback control. Next since we want to design a feedback optimal control as opposed to an open loop optimal control we choose to use a Receding Horizon Control (RHC) or Model Predictive Control (MPC) methodology [23, 24]. Implementing a feedback control requires knowledge of all state variables. Since our model is a non-linear dynamical system and we do not have observations for all our model states, we use a state estimation technique such as non-linear filtering to estimate the missing model states. Thus in Section 4.4 we introduce the concept of Non-Linear Kalman Filtering, specifically Continuous Discrete Extended Kalman Filtering as it pertains to our problem. We next present our results which include a cohesive algorithm for optimal control and state estimation to predict the optimal immunosuppression regimen for a transplant recipient in the context of BKV infection. Our last section presents our conclusions.

4.2 Mathematical Model

The current updated BKV model in Chapter 3 which describes the dynamics of the immune response (BKV-specific CD8+T-cells (E_V) and allospecific CD8+T-cells (E_K) that target the kidney) in response to concentrations of susceptible cells (H_S), infected cells (H_I), free BKV (V) and the biomarker serum

creatinine (C) is presented below [73]:

$$\dot{H}_S = -\chi_{(E_K > E_K^*)} \tilde{\beta} H_S E_K - \chi_{(V > V^*)} \beta H_S V \quad (4.1a)$$

$$\dot{H}_I = \chi_{(V > V^*)} \beta H_S V - \delta_{HI} H_I - \chi_{(E_V > E_V^*)} \delta_{EH} E_V H_I \quad (4.1b)$$

$$\dot{V} = \rho_V \delta_{HI} H_I - \delta_V V - \chi_{(V > V^*)} \beta H_S V \quad (4.1c)$$

$$\dot{E}_V = (1 - \epsilon_I) [\lambda_{EV} + \rho_{EV}(V) E_V] - \delta_{EV} E_V \quad (4.1d)$$

$$\dot{E}_K = (1 - \epsilon_I) [\lambda_{EK} + \rho_{EK}(H_S) E_K] - \delta_{EK} E_K \quad (4.1e)$$

$$\dot{C} = \lambda_C - \delta_C(H_S) C, \quad (4.1f)$$

where

$$\rho_{EV}(V) = \frac{\bar{\rho}_{EV} V}{V + \kappa_V}, \quad (4.1g)$$

$$\rho_{EK}(H_S) = \frac{\bar{\rho}_{EK} H_S}{H_S + \kappa_{KH}}, \quad (4.1h)$$

$$\delta_C(H_S) = \delta_{C0} \cdot \frac{H_S}{H_S + \kappa_{CH}}. \quad (4.1i)$$

As described in [73] initial conditions are given by,

$$(H_S(0), H_I(0), V(0), E_V(0), E_K(0), C(0)) = (H_{S0}, H_{I0}, V_0, E_{V0}, E_{K0}, C_0). \quad (4.1j)$$

We define the characteristic or indicator function χ as

$$\chi_{(x > x^*)} = \begin{cases} 1, & \text{for } x > x^* \\ 0, & \text{otherwise.} \end{cases} \quad (4.2)$$

For more details about the detailed mechanisms of the model refer to Chapter 3. The model states, parameters and initial conditions are given in Tables 2.1, 3.2 and 3.3 respectively.

4.3 Optimal Control

Let us suppose we have a non-linear dynamical system

$$\frac{dx(t)}{dt} = f(x(t), u(t), t), \quad x(t_0) = x_0, \quad (4.3)$$

with state variable $x(t) \in \mathcal{R}^n$ and the control input $u(t) \in \mathcal{U} \subset \mathcal{R}^m$ (\mathcal{U} is a control set).

Next let us define a performance index or cost functional associated with the system in (4.3),

$$J(u(t)) = \phi(x(T), T) + \int_{t_0}^T L(x(t), u(t), t) dt, \quad (4.4)$$

where $[t_0, T]$ is our defined time interval. The terminal cost $\phi(x(T), T)$ depends on the final state and time. The running cost or weighting function $L(t, x(t), u(t))$ depends on the intermediate state and control at times in $[t_0, T]$.

The optimal control problem is to find the control $u^*(t)$ on our time interval of interest $[t_0, T]$ that drives the system through the trajectory $x^*(t)$ such that the cost function in (4.4) is minimized and such that the final state constraint function given by $\psi(x(T), T)$ is fixed at zero [57, 66],

$$\psi(x(T), T) = 0. \quad (4.5)$$

For most complex real life problems solving the control problem analytically might not be a feasible option as the closed form solution might not exist. Under such circumstances one must employ numerical methods for solving optimal control problems. Numerical methods can be divided into two broad categories: direct or indirect methods.

In direct methods the optimal control problem is transformed into a nonlinear programming problems or a nonlinear optimization problem which involves the discretization of the original optimal control problem (either just the optimal control or both the optimal control and model states) and then numerically solving using well-established, pre-existing optimization methods. This class of methods is known as direct transcription and is sometimes referred to as "discretize then optimize" [16, 85].

The indirect methods are based on the calculus of variations or the Pontryagin's minimum or maximum principle to derive and solve for the necessary conditions for optimality [57, 66]. The method employs the Hamiltonian function

$$H(x(t), u(t), \lambda(t), t) = L(x(t), u(t), t) + \lambda f(x(t), u(t), t)$$

and the optimal control problem is reduced to the solution of the following system equations given in the form of a two-point boundary value problem (BVP):

$$\text{State Equation : } \dot{x}(t) = H_\lambda(x(t), \lambda(t), u(t), t) = f(x(t), u(t), t)$$

$$\text{Costate/Adjoint Equation : } -\dot{\lambda}(t) = H_x(x(t), \lambda(t), u(t), t)$$

where λ are additional Lagrange multiplier functions.

The boundary conditions are

$$x(t_0) = x_0$$

$$(\phi_x + \psi_x^T \nu - \lambda)^T|_T dx(T) + (\phi_t + \psi_t^T \nu + H)|_T dT = 0 \quad (4.6)$$

where ν is the Lagrange multiplier corresponding to the final state constraint in (4.5). The optimal control is found by,

$$u^*(x(t), \lambda(t), t) = \operatorname{argmin}_{u \in \mathcal{U}} H(x(t), \lambda(t), u(t), t)$$

This problem can be solved analytically for a few simpler models but for most optimization problems it must be solved numerically.

4.3.1 Open Loop Control

An open-loop system is one where the output of the system has no consequence on the input or control because the output is not re-evaluated based on updates on the state inputs. Thus an open-loop control system blindly depends solely on its initial input and fixed path regardless of the final result.

When solving an optimal open loop control problem the goal is to find the input $u^*(t)$ on the time interval $[t_0, T]$ that drives the system along a trajectory $x^*(t)$ such that the cost function in (4.4) is minimized, and such that the final state constraint in (4.5) is satisfied using just an initial input of the model states [57, 66].

For T fixed, solutions of the optimal control problem also solve the following set of differential equations:

$$\text{State Equation : } \dot{x} = H_\lambda = f(x, u, t)$$

$$\text{Costate Equation : } -\dot{\lambda} = H_x = \frac{\partial L}{\partial x} + \lambda \frac{\partial f}{\partial x}$$

$$\text{State initial condition : } x(0) = x_0$$

$$\text{Co-state final condition : } \lambda(T) = (\phi_x + \psi v)|_T.$$

The process of computing an optimal open loop control is given below [23]:

1. Given the initial guess for control and state conditions, solve the state equation forward in time.
2. Given the above computed state and control, solve the costate equation backwards in time.
3. Using the x and λ computed above compute the cost function $J(u)$ and gradient H_u (note that the first order approximation of ∇J is H_u [81]).
4. Update the control using an optimization routine (we use MATLAB's inbuilt solver **fmincon** where we supply the gradient evaluations for greater speed and accuracy).
5. In the optimization routine we want H_u to converge to a certain tolerance or till a maximum number of iterations are reached.
6. Repeat this process at each sampling time point using the current predicted control and model states until time T .

4.3.2 Feedback Loop Control: Receding Horizon Methodology

A closed loop system or a feedback control system is a control system which is similar to open loop except it now incorporates one or more feedback paths between its output and input giving the system a way to re-evaluate itself based on periodically updated input. Unlike the open loop control, a feedback loop has self knowledge of the output and can adapt to any disturbances that occur.

Receding Horizon Control (RHC) is a feedback control formulation which also aims to make use of the computational simplicity of the above mentioned calculus of variations approach. RHC solves an open-loop optimal control problem at each sampling instant for a finite time horizon. Some

factors that need to be considered in this method besides the model and the cost function are the sampling period, the length of the finite time horizon, and the state estimation method to obtain the state at each sampling time point. An advantage of using the RHC methodology is that by solving the open loop control on a finite long time horizon one computes the control far into the future to optimize the present control value (See Figure 4.1).

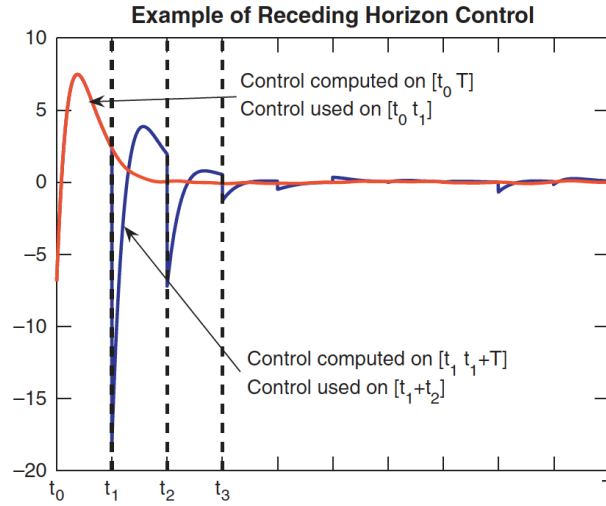


Figure 4.1: A sample plot showing the concept of RHC [24].

We then solve the receding horizon control problem using the following algorithm [24] and diagram of the schematic is also given in Figure 4.2:

1. Given initial condition for state and control, solve the open loop control problem (as described in the previous subsection) on the time interval $[t_i, t_i + t_{ch,i}]$ where $t_{ch,i}$ is the length of the control horizon at time t_i .
2. Using the control defined only on the interval $[t_i, t_{i+1}]$ determine the trajectory of the model states on the same time interval (assumption being we get a data point or “feedback” at t_{i+1}).
3. Use observations for all states if available or a state estimator to determine $x(t_{i+1})$.
4. Repeat step 1 but now over the time interval $[t_{i+1}, t_{i+1} + t_{ch,i+1}]$
5. Continue steps 1 – 4 until T is reached.

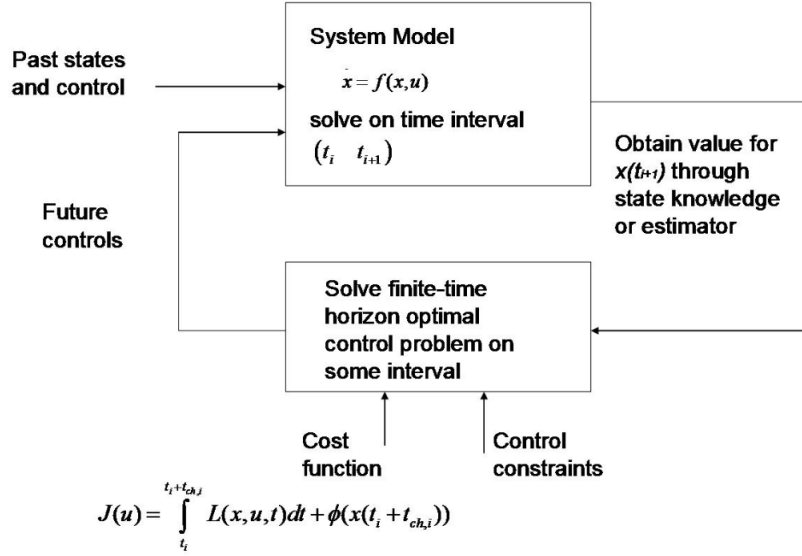


Figure 4.2: A schematic diagram showing the RHC algorithm [24].

4.4 Filtering: Continuous-Discrete Extended Kalman Filter

Filtering is used to combine a set of observations (corrupted with some measurement noise, v) with a model (also corrupted by some noise, w) to obtain an estimate for the true physical system. It provides estimates in real time as data is collected and allows model errors to be taken into account. The Kalman Filter can be extended for nonlinear problems, one such extension is known as the Extended Kalman Filter (EKF). In this case we have a non-linear dynamic model with discrete measurements; hence we use the Continuous-Discrete Extended Kalman Filter.

We use the algorithm described below with the same initial condition x_0 as the open loop control, and the ϵ_I^* obtained and our new data point z_k to obtain the state estimate \hat{x}_k .

We start with the continuous nonlinear dynamic model with discrete data points:

$$\dot{x}(t) = f(x(t), \epsilon_I, t) + g(t)w(t), w(t) \sim N(0, Q) \quad (4.7)$$

$$z_k = h(x(t_k), k) + v_k, v_k \sim N(0, R) \quad (4.8)$$

Here $f(x(t), \epsilon_I, t)$ is our log-scaled version of the model (4.1) and we assumed $g(t) = 1$.

The term $h(x(t_k), k)$ is the observed part of the model solution defined as $h(x(t_k), k) = \begin{bmatrix} V_k \\ C_k \end{bmatrix}$.

Q and R represent the variance of the error in the model and the data respectively. They are chosen after a series of trials and errors. Note if we choose $Q \gg R$, this is because we suspect that there is more noise in the model and then the filter trusts the data more and will fit the data closely. Meanwhile if $R \gg Q$, implying that the data is significantly noisier than the model and then the filter trusts the model more than the data and will fit the model more closely. (See Figures 4.9 and 4.10)

Given below is the algorithm used for state estimation using the Continuous-Discrete Extended Kalman Filter [53, 94]:

Initialization of state and covariance (k = 0):

Initialization of state and covariance (k = 0):

$$\hat{x}_0 = E[x(t_0)]$$

$$P_0 = E[(x(t_0) - E[x(t_0)])(x(t_0) - E[x(t_0)])^T]$$

We chose \hat{x}_0 to be the first data point and $P_0 = \mathbf{I}_6$.

For $k = 1, 2, 3 \dots$

Compute Jacobians:

$$A(x, \epsilon_I, t) = \nabla f(x, \epsilon_I, t), \quad C(x) = \nabla h(x, k).$$

Time Update:

For $t_{k-1} \leq t \leq t_k$, integrate the differential equations

$$\dot{\hat{x}} = f(\hat{x}, \epsilon_I, t).$$

$$\dot{P} = PA^T(\hat{x}, \epsilon_I, t) + A(\hat{x}, \epsilon_I, t)P + gQ^Tg.$$

with the initial conditions $x(t_{k-1}) = x_{k-1}$ and $P(t_{k-1}) = P_{k-1}$ to obtain $\hat{x}_k^- = \hat{x}(t_k)$ and $P_k^- = P(t_k)$

Measurement update:

$$K_k = P_k^- C^T(\hat{x}_k^-) [C(\hat{x}_k^-) P_k^- C^T(\hat{x}_k^-) + R]^{-1}.$$

$$P_k = [I - K_k C(\hat{x}_k^-)] P_k^-.$$

$$\hat{x}_k = \hat{x}_k^- + K_k [z_k - h(\hat{x}_k^-, k)].$$

The above algorithm summarizes the process of state estimation using the Extended Kalman Filter for a continuous model with discrete data points. The time update in the above algorithm is performed over an interval of time when measurement or data will be available. When data becomes available a measurement update will be performed, however in the absence of data the filter can still continue to operate using the time update and the last available state estimate available until data becomes available [42].

4.5 Numerical Results

The results [74] presented in this section are all using a log transformed version of the mathematical model in Equation (4.1) as done in [8] and shown in Chapter 2. Due to sparsity of observational data available, we created simulated data by adding noise to our model solutions as follows:

$$\mathbf{z} = \mathbf{h}(\mathbf{x}) + \eta$$

where the noise $\eta = \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix}$ $\eta_1 \sim N(0, \sigma_1^2)$, $\eta_2 \sim N(0, \sigma_2^2)$ and $\mathbf{h}(\mathbf{x}) = \begin{bmatrix} V \\ C \end{bmatrix}$.

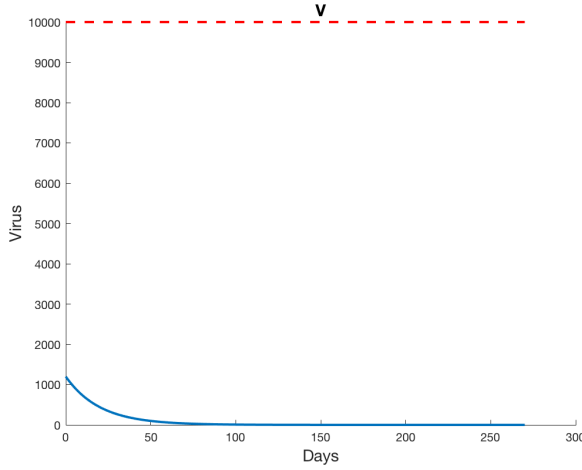
The first subsection presents results for model dynamics if fixed immunosuppressant dosages were prescribed to a transplant recipient. Next we show preliminary simulations for the control problem, starting with open loop simulations to test the sensitivity of the model to the control with respect to varying weights and initial immunosuppression values and feedback control when perfect information is available. We follow with simulations testing the Extended Kalman Filter to ensure it's performance is robust. Lastly we show our results for feedback control with Extended Kalman Filter as our method of choice for state estimation, thus creating an adaptive treatment schedule for renal transplant recipients.

4.5.1 Numerical Results: Fixed Immunosuppressant Dosages

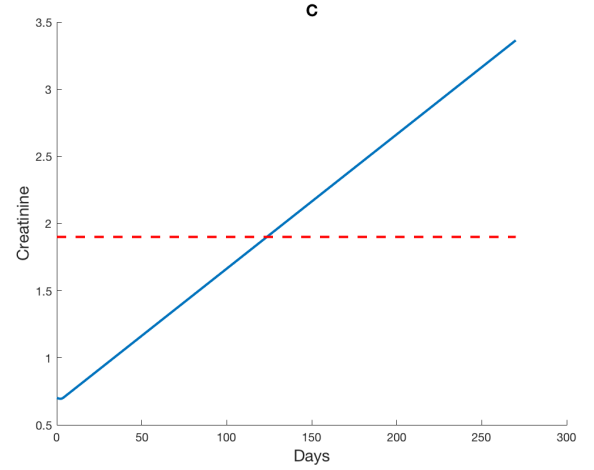
Here we show results for the dynamics of biomarkers of infection (BKV load) and kidney health (Creatinine) when the immunosuppression regimen was not being optimized using an optimal feedback control formulation. We suppose instead one was prescribing predetermined medication to the patients based on the treatment schedule described in [8]. We see in Figure 4.3 that while the viral load was within bounds (minimal viral load of 10,000 copies/ml must be present in plasma for low BK viremia to be detected [28]), the creatinine levels showed that the kidney was undergoing rejection because of the strong allo-specific immune response. As a reference, normal levels of creatinine in the blood are approximately 0.6 to 1.2 mg/dL in adult males and 0.5 to 1.1 mg/dL in

adult females. Usually females have a lower baseline for creatinine levels as they have less muscle mass. Since creatinine is a by product of muscle metabolism, more muscle mass implies higher creatinine levels. Individuals with only one kidney may have a normal creatinine level of about 1.8 mg/dL or 1.9 mg/dL. Since kidney transplant patients are usually individuals with one kidney we will use 1.9 mg/dL as an upper bound baseline for normal creatinine levels [26]. We also define

$$\epsilon_I(t) = \begin{cases} 0.1009 & t \in [0, 21] \\ 0.3658 & t \in (21, 60] \\ 0.5999 & t \in (60, 120] \\ 0.3649 & t \in (120, 450]. \end{cases} \quad (4.9)$$

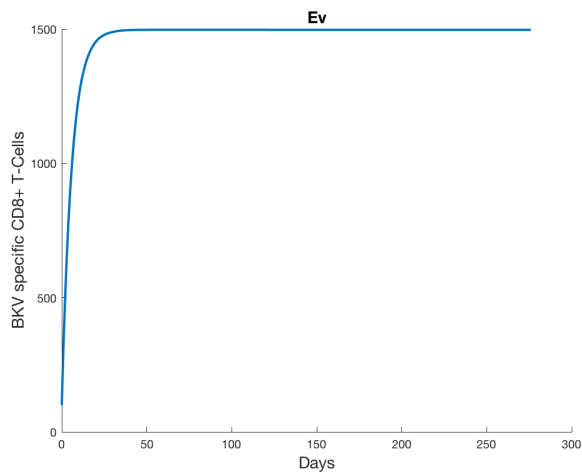


(a) BK Virus

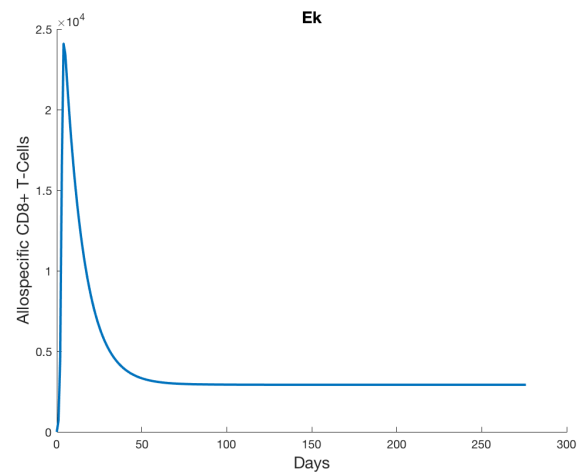


(b) Creatinine

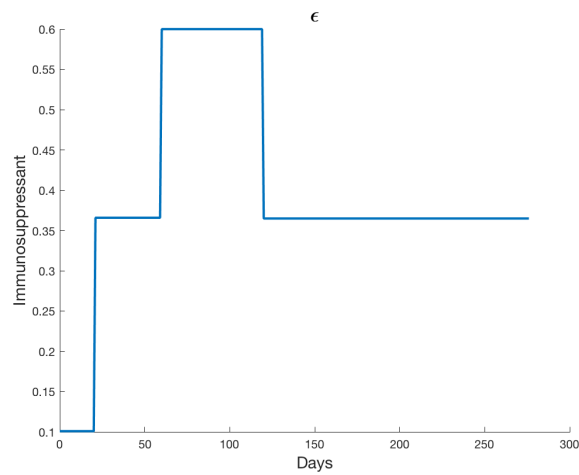
Figure 4.3: Model dynamics when ϵ dosages are fixed. (Red dashed line is the upper bound on healthy biomarker values.)



BKV specific CD8+ T-Cells



Allo-specific CD8+ T-Cells



Immunosuppressant

Figure 4.3 (cont.): Model dynamics when ϵ dosages are fixed. (Red dashed line is the upper bound on healthy biomarker values.)

4.5.2 Numerical Results: Open Loop Control

Our aim is to determine the optimal immunosuppressive drug efficiency over the time interval $[t_0, T]$ such that renal transplant recipients have a functioning kidney free of BKV infection. Since we ascertain from before how far we want to control to be computed, the final time T is fixed. Since we want to drive the viral load and creatinine low but do not want to fix the final outcome, the final state $x(T)$ is free. We define the following cost function

$$J(t) = \int_{t_0}^T \left[W_V V(t)^2 + W_C C(t)^2 \right] dt. \quad (4.10)$$

Note that compared to the cost function we define in general, in Equation (4.4), we do not have a terminal cost in our problem. The formulation of our cost function is under the assumption that sustainable good kidney health under optimal doses of immunosuppression involves minimizing BK viral load in the blood and effective clearance of creatinine from the blood (implying low serum creatinine levels). The weighting terms W_V and W_C adjust how much we want to penalize the cost function for not lowering viral loads and creatinine. Our control here is ϵ_I , the efficiency of the immunosuppressant. Recall that the efficiency has to be a non-negative quantity less than 1. We do not include the control in the cost functional because we do not wish to minimize immunosuppressant dosage. Our only aim is to have an infection-free, healthy kidney irregardless of how much immunosuppressant is needed. Also due to the large differences in magnitude in model states we use a log scaled model (see [8] for details).

We next use the MATLAB optimization solver **fmincon** and give it the cost and the gradient of the Hamiltonian with respect to ϵ_I and performed open loop simulations for 500 days with varying initial guesses of immunosuppressant dosage (for control estimation) ϵ_0 with weights $(W_V, W_C) = (1, 1)$. Our goal before designing the feedback control was to test the robustness of the model in the context of the control problem we are trying to design.

In Figures 4.4 and 4.5 we see the dynamics of our two biomarkers BKV load and Creatinine as we change the initial guess for immunosuppression efficiency when running an open loop control. In Figure 4.4 the control for $\epsilon_0 = 0.5$ looks almost like a straight line with little to none deviation. To investigate it further we chose initial guesses close to $\epsilon_0 = 0.5$. Figure 4.5 confirms that for an open loop control formulation an initial guess of $\epsilon_0 = 0.5$ gives the lowest BKV load and creatinine dynamics seen, hence explaining minimal deviation in the immunosuppression values. Note that ϵ_0 is the initial guess for control estimation and not the initial immunosuppressant dosage recommended to the patient.

Next for initial guess $\epsilon_0 = 0.45$ we observe changes due to varying weights W_V and W_C in BKV and

creatinine dynamics as seen in Figures 4.6 and 4.7. We notice that as we increase the weight for one biomarker the control works harder in lowering it even if it means compromising on minimizing the other. Here again we point out, as seen in Figure 4.7 that on varying W_C , i.e., penalizing to lower creatinine dynamics, we see a dip in creatinine values as we increase W_C soon to be followed by a steady increase later. This is due to the increase in viral load (since we are penalizing heavier on the creatinine than the BKV load), leading to infection and a damaged kidney in the later days.

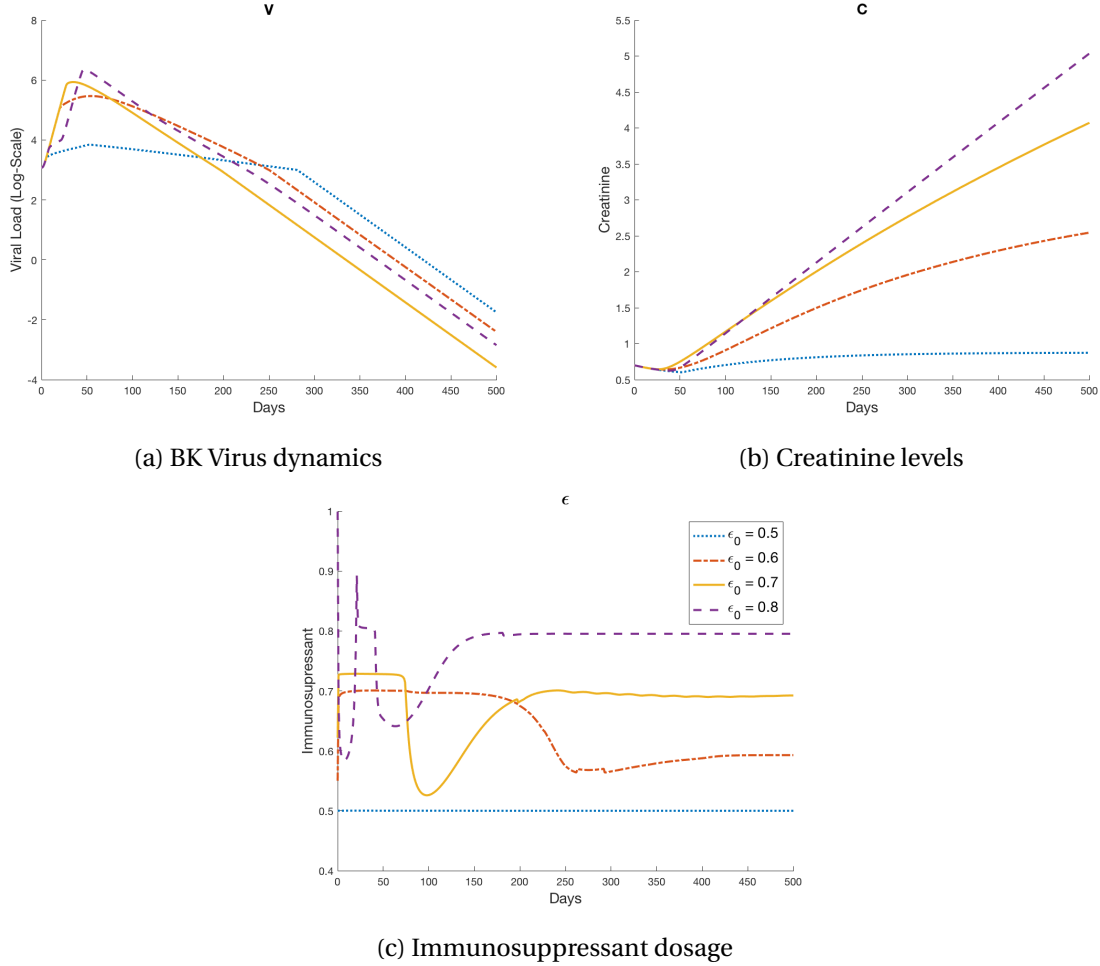


Figure 4.4: Open loop control for weights $(W_V, W_C) = (1, 1)$ with varying ϵ_0 .

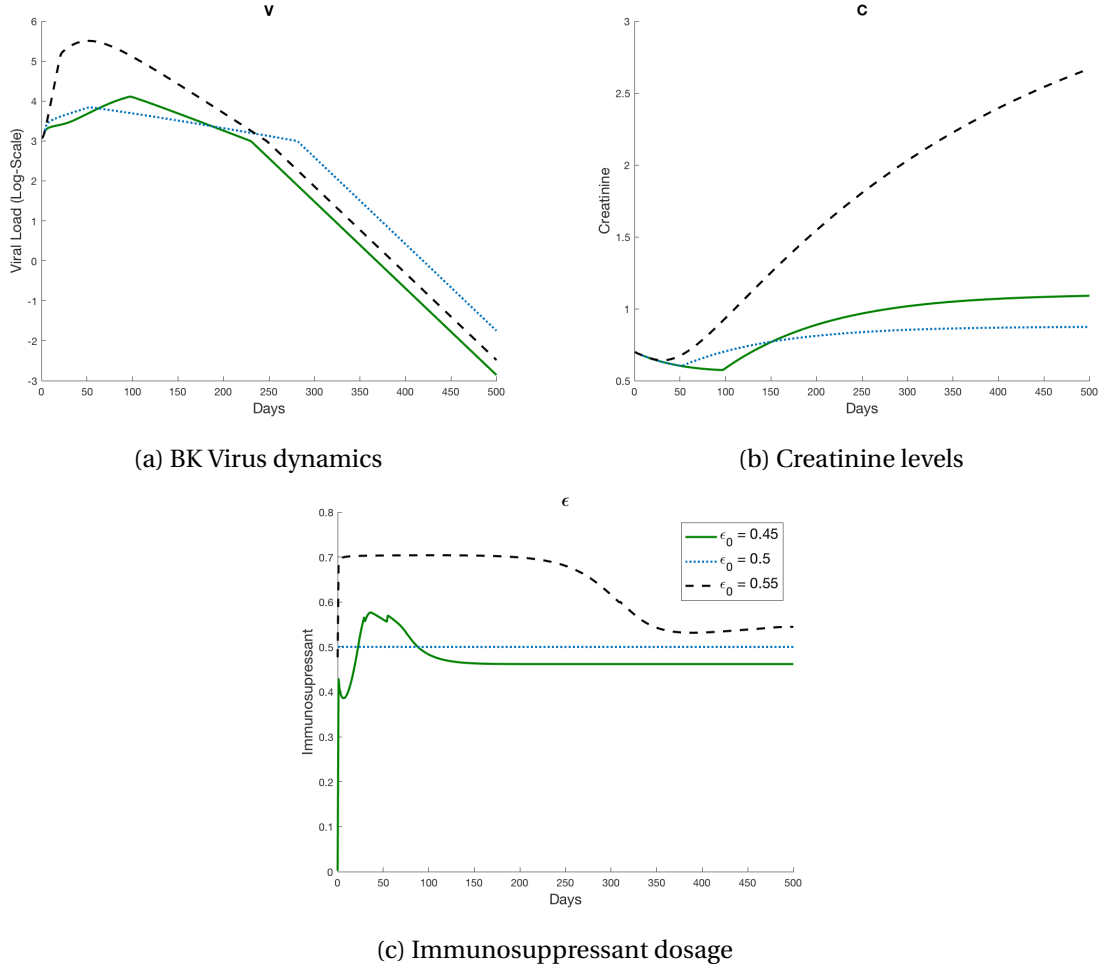
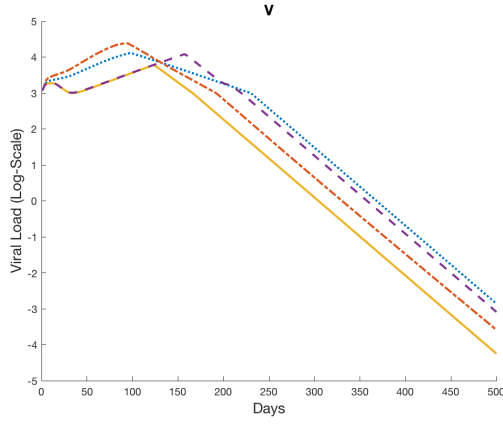
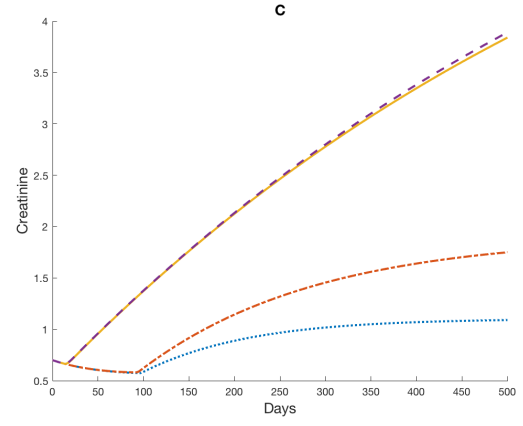


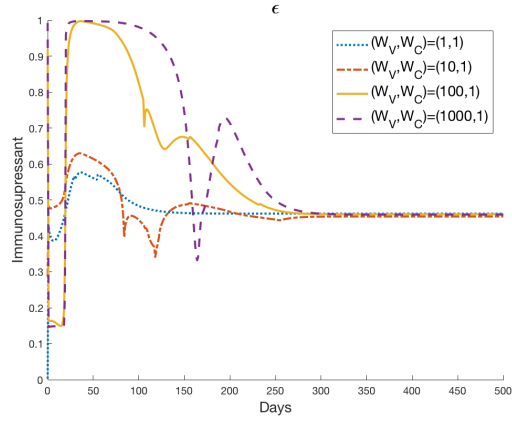
Figure 4.5: Open loop control for weights $(W_V, W_C) = (1, 1)$ with varying ϵ_0 .



(a) BK Virus dynamics

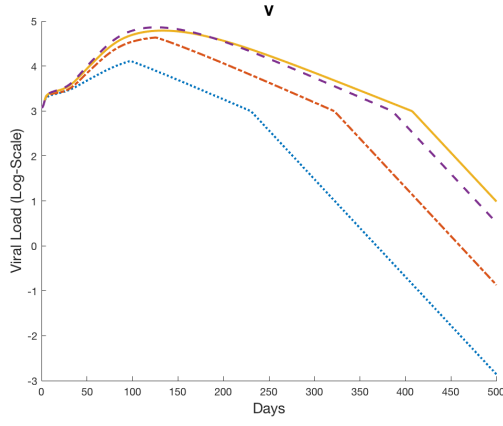


(b) Creatinine levels

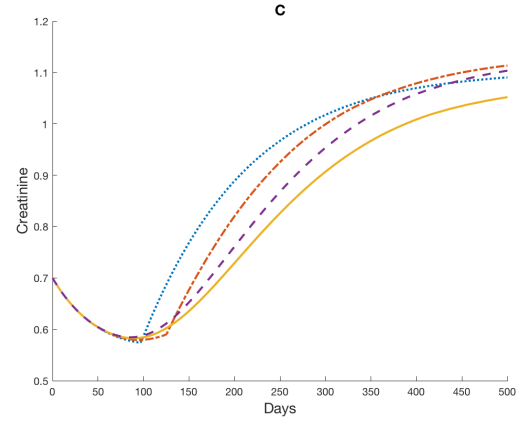


(c) Immunosuppressant dosage

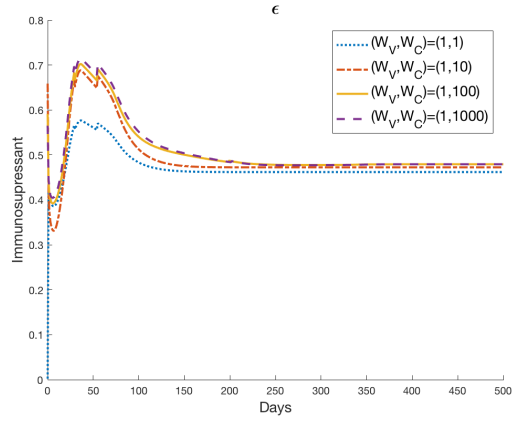
Figure 4.6: Open loop control with varying W_V with initial immunosuppressant, $\epsilon_0 = 0.45$.



(a) BK Virus dynamics



(b) Creatinine levels

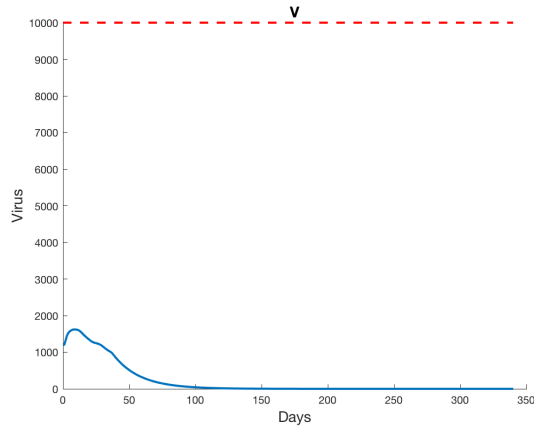


(c) Immunosuppressant dosage

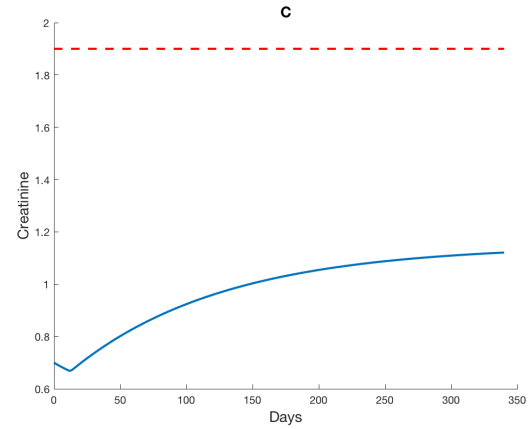
Figure 4.7: Open loop control with varying W_C with initial immunosuppressant, $\epsilon_0 = 0.45$.

4.5.3 Numerical Results: Feedback Control with Perfect Information

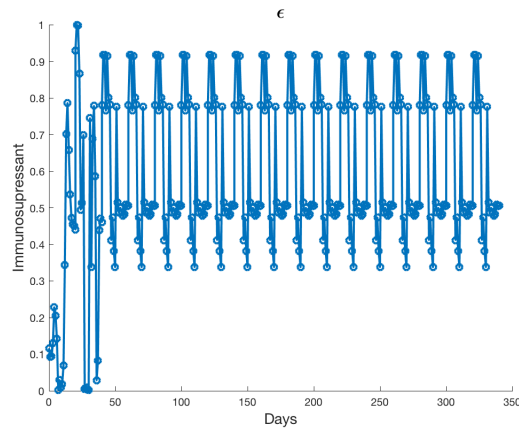
Our eventual aim is to design a feedback control formulation with state estimation to account for the incomplete information of model states one receives during the data collection process. Namely, data as it is currently collected would provide measurements for only BKV load and creatinine. In a step by step build up to our final aim we tested for robustness and compatibility of our model in context of designing a feedback control problem. Hence we first ran simulations for a feedback control problems where perfect information for all states was available during every patient visit, in this case we assume that to be every 20 days. We assume the finite control horizon length $t_{ch,i}$ to be 60 days. We conclude from the low viral and creatinine levels in Figure 4.8 that under the condition of acquiring perfect information, a working adaptive treatment schedule can be built.



(a) BK Virus



(b) Creatinine



(c) Immunosuppressant

Figure 4.8: Optimal immunosuppressant dosage for first 340 days of treatment with an initial guess for control estimation $\epsilon_0 = 0.45$ when perfect information is available on the patient every 20 days. The blue dots are the daily recommended optimal dose. (Red dashed line is the upper bound on healthy biomarker values.)

4.5.4 Numerical Results: Extended Kalman Filter

Before combining Feedback control with Extended Kalman Filter as a state estimation method we wanted to test if our Extended Kalman Filter algorithm is working robustly. Recall that in Equations (4.7) and (4.8), Q and R represent the variance of the error in the model and the data respectively. If we chose $Q \gg R$ we suspect that there is more noise in the model and then the Filter trusts the data more and will fit the data closely. Meanwhile if $R \gg Q$ we expect the data is significantly noisier than the model and then the Filter trusts the model more than the data and will fit the model more closely. We can see that in Figures 4.9 and 4.10. Figure 4.11 shows the current settings where we assume a comparable amount of noise in both our model and data ($R \approx Q$). These variances are picked under the assumption that usually there is insufficient information to conclude which between the model or data is noisier.

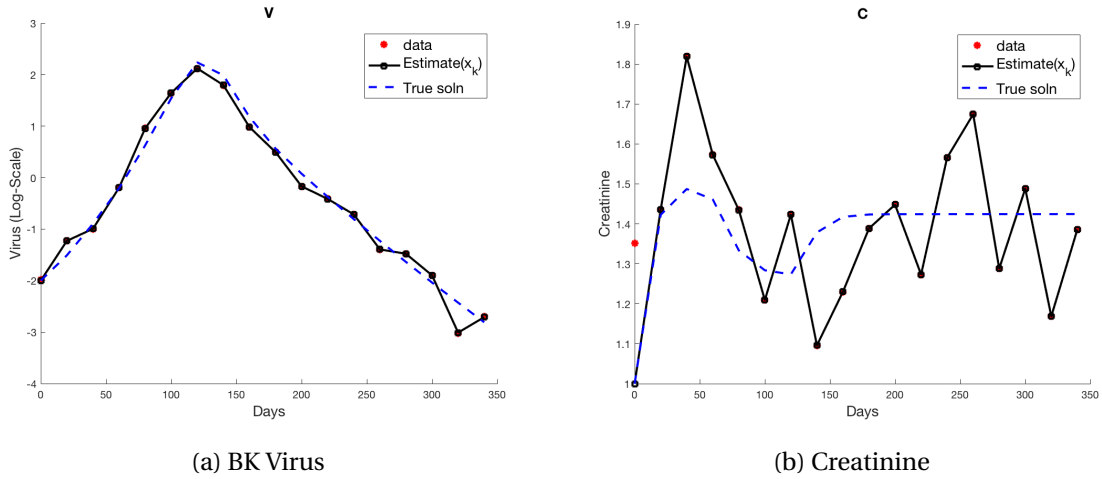
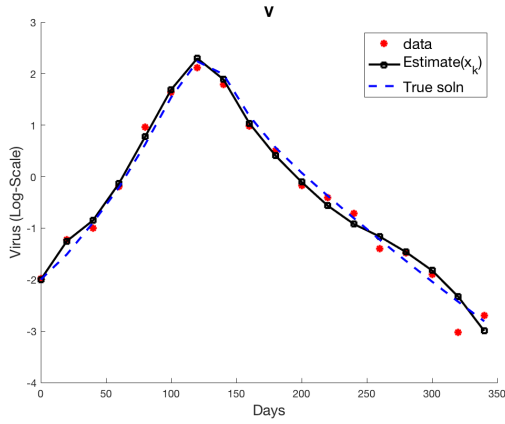
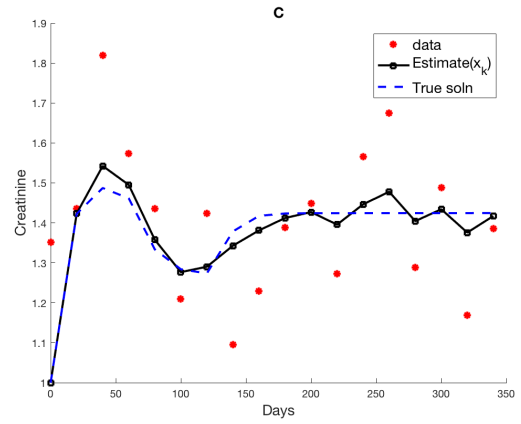


Figure 4.9: State estimation when $Q \gg R$: Filter trusts data more.

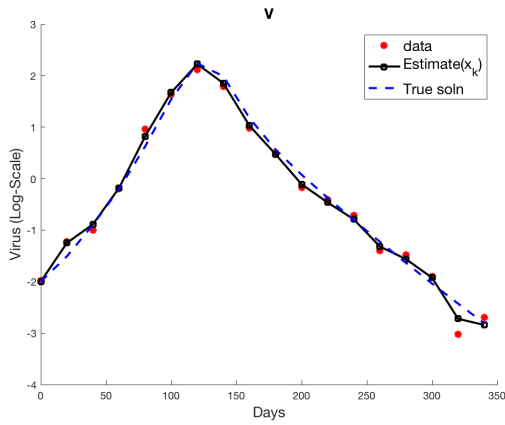


(a) BK Virus

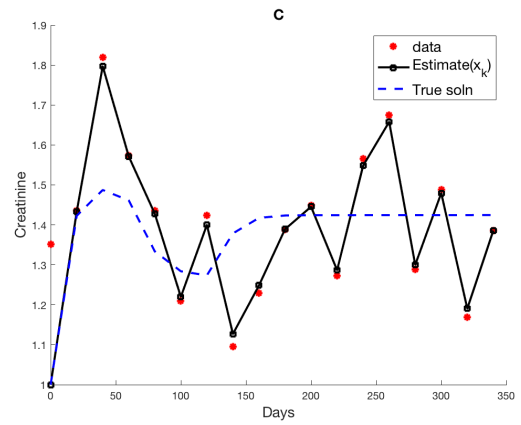


(b) Creatinine

Figure 4.10: State estimation when $R \gg Q$: Filter trusts model more.



(a) BK Virus



(b) Creatinine

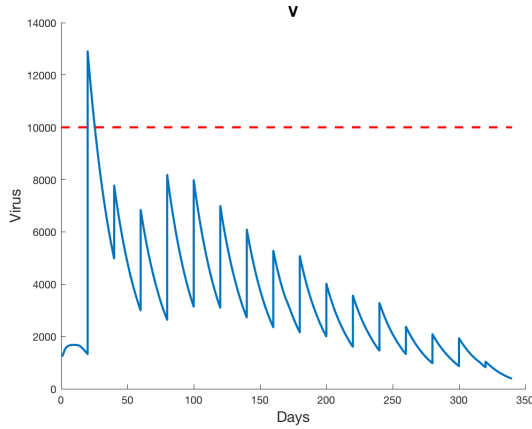
Figure 4.11: State estimation when $R \approx Q$: Current Settings.

4.5.5 Numerical Results: Feedback Control with State Estimation

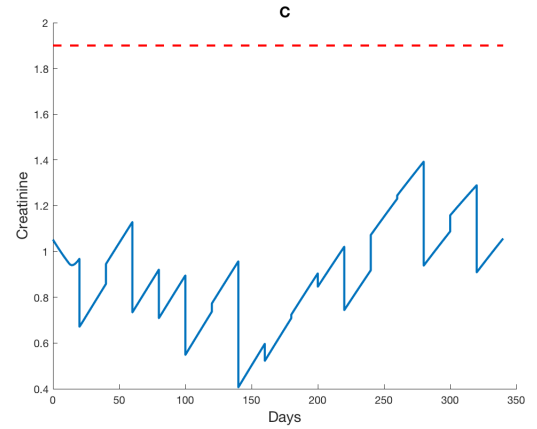
Finally we combine our RHC feedback control methodology and EKF state estimation method to produce an optimal adaptive treatment for a simulated renal transplant recipient. The premise is that the patient is started on an initial immunosuppressant efficiency level at the beginning of treatment and then when they visit the doctor next there are some diagnostic tests performed on them to measure BKV infection and creatinine. This information is then fed back to the RHC algorithm and the remaining model states are estimated using EKF and a new immunosuppressant efficiency is predicted. This efficiency is to be then used to predict dosage until the next patient visit.

Following the below enumerated steps we obtained numerical results for optimal immunosuppressant dosages for the first 340 days of treatment after transplant for $\epsilon_0 = 0.45$ as seen in Figure 4.12. Note that in the model plots, the simulations for immunosuppressive therapy are plotted after the initial treatment ϵ_0 value.

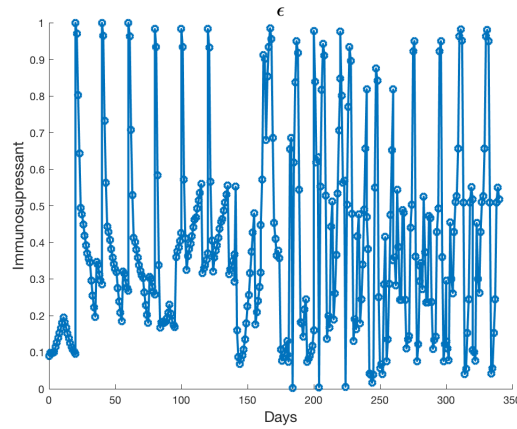
1. Create simulated data as described above choosing a noise level for BKV load and creatinine observations (we picked $\sigma_1 = 0.3$ and $\sigma_2 = 0.15$ respectively).
2. Solve the Receding Horizon Control Problem from $[t_i, t_i + t_{ch,i}]$ (See subsection 4.3.2 for detailed algorithm) and find an ϵ_I^* , $t_{ch,i}$ was chosen to be 200 days while $|t_{i+1} - t_i| = 20$ days.
3. Use the above obtained ϵ_I^* to obtain state estimates for all model states for which we did not have data, using Extended Kalman Filter (See section 4.4 for detailed algorithm). This would then become our new initial condition and repeat steps 2 and 3.



(a) BK Virus



(b) Creatinine



(c) Immunosuppressant

Figure 4.12: Daily optimal immunosuppressant dosages for first 340 days of treatment with an initial control guess $\epsilon_0 = 0.45$ when the patient visited the doctor every 20 days. The blue dots are the daily recommended optimal dose. (Red dashed line is the upper bound on healthy biomarker values.)

4.6 Conclusion

Optimizing drug dosage regimens for immunosuppressed renal transplant recipients is of utmost importance and is a task that is still complicated and difficult to achieve in a lot of cases. This is especially true because the therapeutic index for most of the drugs are very narrow and a small digression from the optimal dosage can very quickly go from beneficial to toxic. Hence finding an individualized treatment schedule for renal transplant recipients is a very pertinent problem to solve. Dosage regimens are highly patient specific and dependent on factors such as patient age, weight, other medications and medical history and there also seems to be a lack of uniformity in the exact immunosuppressive treatment protocols in the United States followed across different organ transplant centers [22, 39].

In this chapter we designed a feedback control algorithm to predict the optimal amount of immunosuppression an individual undergoing a kidney transplant might need. We ran several diagnostics tests to investigate the robustness of the control with respect to the mathematical model as well as the state estimation method. We present results for all the robustness tests. Finally we present and explain the algorithm used to build the adaptive optimal treatment schedule for renal transplant recipients and depict results for the optimal treatment plan for a simulated transplant recipient. In the future we hope to accrue further clinical data and use it to estimate the model parameters with greater confidence. Then use these individual patient specific models in combination with our feedback algorithm to predict optimal treatment schedules. This would be a major step towards incorporating personalized medicine technology in the lives of renal transplant recipients.

The current model (4.1) treats immunosuppressant dosage as drug efficiency which would approximately translate to the percentage of the maximum drug dosage usually prescribed to patients. While modeling immunosuppressive therapy as a unit-less quantity is a good stepping stone in modeling drug dosage, we wish to explore next the amalgamation of drugs that constitute the immunosuppressive therapy to help bridge the gap between efficiency and dosage. Transplant patients are prescribed a cocktail of drugs as part of their therapy. Immunosuppressive drugs can be broken into 3 broad categories: induction drugs, maintenance drugs and reversal drugs used to undo an existing case of rejection [31, 43], a preliminary review of which is provided in Chapter 5. Our next goal is to investigate the combination of most prevalent drugs and their prescribed proportions to either further quantify the relationship between drug efficiency and drug dosage or to model specific kinds of drugs in our existing model itself. Using this new updated model we would then aim to use our adaptive optimal treatment algorithm to optimize drug dosages (instead of efficiencies) as it is applied to individual renal transplant patients. Another possible direction we hope to take is to make the model more representative of the dynamics in the body by incorporating further components of

the human immune system, for example, the helper CD4+ T-cells. Lastly, incorporating specific absorption, distribution, metabolism and excretion mechanisms of the immunosuppressant drugs to build a more physiologically based pharmacokinetic model would be another way to incorporate drug dosage in future efforts.

CHAPTER

5

FUTURE DIRECTIONS

5.1 Dynamics of Immunosuppressive Treatments

While the current mathematical model (3.3) treats the immunosuppressive therapy as a unit-less quantity meant to be a fraction of the total maximum drug dosage to be given to a patient, we wish to explore further the amalgamation of drugs that constitute the immunosuppressive therapy. Immunosuppressive drugs can be broken into 3 broad categories: induction drugs, maintenance drugs and reversal drugs used to undo an existing case of rejection [31, 43]. In this chapter we will focus on induction and maintenance drugs as our mathematical model is based on preventing rejecting and not reversing rejection (which is the case when one would prescribe reversal drugs).

5.1.1 Induction Therapy

Induction therapy which involves a severe amount of immunosuppression is primarily administered to high risk patients pre and post transplantation [31, 43]. Currently induction immunosuppressive drugs contain either polyclonal or monoclonal antibodies. In case of polyclonal antibodies the two antithymocyte globulins are derived by injecting horses (Atgam) or rabbits (Thymoglobulin) with human lymphocytes or thymocytes respectively and then later extracting the immune serums [38, 50]. Polyclonal antibodies target a variety of T-cell markers and cause depletion of the lymphocytes

from the peripheral blood [72]. One short term side effect of Thymoglobulin is cytokine release syndrome which comprises of fever, chills, myalgias(muscle pains), and shortness of breath. In the long term, Thymoglobulin is known to induce a persistent depletion of T-lymphocytes which in turn increases the risk of infections, post-transplantation lymphoproliferative disorder (PTLD), and even in extreme cases autoimmune disease. As an induction agent, antithymocyte globulin (equine) is administered at a dosage of 10-30 mg/kg i.v. for 4-14 days and is generally infused over a period of four to six hours per dose while Thymoglobulin doses range from 1 to 4 mg/kg/day for 3-10 days after transplantation [15, 38, 46, 56, 75].

Some monoclonal antibodies like daclizumab (Zenapax) and basiliximab (Simulect) impair the production and multiplication of lymphocytes and usually do not have too much side effects. Alemtuzumab is another humanized murine monoclonal antibody which is anti-CD52 (a surface protein expressed on T and B lymphocytes, monocytes, and macrophages) [38, 50, 72]. There is no universally accepted dosage for alemtuzumab for induction therapy. Doses of 20-30 mg on the day of transplantation and again on postoperative day 1 or 4 have been proven to work in the past [33, 54, 58, 59, 93]. The side effects of alemtuzumab can be very severe due to its strong depletion of lymphocytes. Some effects include neutropenia (70%), thrombocytopenia (52%), anemia (47%), nausea (54%), vomiting (41%), diarrhea (22%), headache (24%), dysesthesias (15%), dizziness (12%), and autoimmune hemolytic anemia (< 5%) [21, 38, 46]. Due to these adverse effects corticosteroids, acetaminophen, and an antihistamine infusion are often co-administered with the drug to combat the side effects [38].

5.1.2 Maintenance Therapy

Maintenance drugs are anti-rejection medications that are prescribed to patients for a long term.

There are usually 4 classes of maintenance drugs [31, 72]:

- **Calcineurin Inhibitors: Tacrolimus and Cyclosporine**

The calcineurin inhibitors (CNI) impair the expression of several critical cytokine genes that promote T-cell activation. CsA(Cyclosporine) and TAC(Tacrolimus) are two commonly used CNIs, each of which acts as a base immunosuppressant around which additional medications are prescribed to construct the complete immunosuppressant regimen [50, 72]. Since Tacrolimus (Prograf) is mainly absorbed from the small intestine, it has large variability in both inter and intra patient, especially for patients with gastro-intestinal diseases, it has been known to have side effects such as anorexia, nausea, vomiting, diarrhea, abdominal discomfort and frank alopecia(hair loss) [72]. CsA is known to have a narrow therapeutic index

because of which it is pertinent to individualize the CsA dosage schedule for each patient so as to optimize treatment by not only preventing rejection but also avoiding toxicity [71]. CNI are highly nephrotoxic and some of the complications could lead to early post-transplant graft dysfunction, a dose related reversible renal vasoconstriction, chronic interstitial fibrosis, acute microvascular disease, hypertension and electrolyte abnormalities along with sometimes non-renal toxicity (like gastrointestinal toxicity) to name a few [72].

- **Antiproliferative agents: Mycophenolate Mofetil (MMF), Mycophenolate (MPA)**

MMF is a prodrug which is the by product of fermentation of several *Penicillium* species. It has an MPA active compound which is an antimetabolite that blocks growth of T and B cells, prevents antibody formation and the generation of cytotoxic T cells. It can treat an ongoing rejection and hinder the progress of proliferative arteriopathy, a critical lesion in chronic rejection. Gastro-intestinal side effects such as diarrhea, varying degrees of nausea, bloating, dyspepsia, vomiting, gastritis are sometimes observed in patients when using MMF (CellCept) and enteric-coated MPA (Myfortic) [72]. On the slightly brighter side unlike CNIs, nephrotoxicity, neurotoxicity and hepatotoxicity have not been reported with MMF [27, 72].

- **mTOR inhibitor: Sirolimus (Rapamune) and everolimus (Certican)**

Mammalian target of rapamycin (mTOR) inhibitors, sirolimus and everolimus are antibiotic compounds which inhibit mTOR, a key regulatory kinase in the process of cell division [72]. One main advantage mTOR inhibitors have over CNI is that they do not cause nephrotoxicity. Though some other known adverse effects are thrombocytopenia, leukopenia, hypercholesterolemia, stomatitis, diarrhea, and, although rare, interstitial pneumonitis. mTOR inhibitors are metabolized by cytochrome (CYP)-3A4/5 and CYP2C8 enzymes and have a narrow therapeutic index making drug monitoring all the more important [69]. Everolimus is a derivative of rapamycin (sirolimus) but with shorter half life, and higher solubility and bioavailability [69, 72].

Sirolimus is hydrophobic in nature and is rapidly absorbed with an average maximum blood concentration (C_{max}) of $40.5 \pm 22.2 \mu\text{g/L}$ with an orally administered dose of 2.5 mg. The maximum concentration is reached after 2.7 ± 2.1 hours (defined as T_{max} , the time after administration of a drug when the maximum plasma concentration is reached) and is dose dependent [96]. As shown through the experimental study in [95], bioavailability of sirolimus is greatly affected by food intake. The drug was absorbed more slowly when administered after a high-fat meal than when administered after fasting and the oral availability of sirolimus increased uniformly by 35%. The drug also has a high volume of distribution of 7-19 L/kg [68].

Everolimus is less hydrophobic in nature than Sirolimus but is also rapidly absorbed with an

average maximum blood concentration (C_{max}) of $45 \pm 21 \mu\text{g/L}$ when orally administering a dose of 2.5 mg. The maximum concentration is reached after 1.3 ± 0.4 hours (T_{max}) and is dose dependent [77]. Very similar to sirolimus, bioavailability of everolimus is greatly affected by food especially fatty food intake [60]. Interactions with the drug cyclosporine often leads to an altered metabolism and even changes the volume of distribution [61, 69]. For instance suppose if an estimated volume of distribution for a 71-kg patient at steady state is 110 L, when considered in the presence of cyclosporine the volume of distribution is increased by 1.14 L for each kilogram increase in body weight [61]. In the absence of cyclosporine volume of distribution is increased by 148 L [69, 70].

- **Corticosteroids: Prednisone and Prednisolone**

Corticosteroids are an anti-inflammatory drug which also suppresses the immune system by inhibiting all stages of T-cell activation. The most prominent side effects of taking corticosteroids are cosmetic changes, growth impairment, osteonecrosis, osteoporosis, impaired wound healing, cataracts, hyperlipidemia, glucose intolerance, psychological effects [72].

Prednisone and prednisolone are two compounds which are metabolically interconvertible, with the latter being assumed to be the pharmacologically active one. They are both rapidly absorbing drugs with T_{max} between 1 to 3 hours and half lives of prednisone being 3.4 to 3.8 hours while that of prednisolone being a little shorter at 2.1 to 3.5 hours. Pharmacokinetics shown by prednisolone is highly dependent on the dosage, with dose being directly proportional to volume of distribution and plasma clearance. Kinetics of the drug are also dependent on age with a shorter half life in children. Liver disease could increase the fraction unbound and prolong half life as well [80]. Dosage regimen for corticosteroids is very patient specific with some patients being administered high intravenous or oral doses, or a low but steady daily dose or an alternating day maintenance regimen, or a steadily decreasing oral dose spanning over days or weeks [72].

5.1.3 Discussion

From the short review seen in this Chapter about the various dosage regimens and drug combinations prescribed to renal transplant recipients, the three most important conclusions to draw were, first that dosage regimens are highly patient specific and dependent on factors such as patient age, weight, other medications and medical history. Second there also seems to be a lack of uniformity in the exact immunosuppressive treatment protocols in the United States followed across different organ transplant centers [22, 39]. Lastly based on our first and second point, optimizing the drug dosage regimens for immunosuppressed renal transplant recipients is of utmost importance and is a

task that is still complicated and difficult to achieve in a lot of cases especially cause the therapeutic index for most of the drugs are very narrow and a small digression from the optimal dosage can go very quickly from beneficial to toxic. Hence finding an optimal control individualized treatment schedule for renal transplant recipients is a very pertinent problem to solve.

In Chapter 4 we design a adaptive treatment schedule for renal transplant recipients using the model the mathematical model in Chapter 3. However the model (3.3) in Chapter 3 incorporates immunosuppression therapy in the form of efficiency as opposed to dosage. We hope to use our new findings from this chapter to bring changes to the mathematical model (3.3) and include individual standard maintenance and induction drugs dosages and their specific effect on the immune system and kidney. Using this new updated model we then aim to redo the control problem, this time to optimize drug dosages as it is applied to individual renal transplant patients.

5.2 Role of Genetic Variation in the Optimal Control of Immunosuppressants in Renal Transplant Recipients

While formulating the model (3.3) in Chapter 3 we encounter a large variability in certain parameter values which in turn has an effect on the viral steady states achieved. We propose these variations in parameter values (See Table 3.2) could be linked to the genetic variability in immune responses and transplant outcomes. We thus next wish to explore the field of pharmacogenomics as it could be potentially applied to personalized medicine in renal transplant recipients.

5.2.1 Genetic Variation as a Cause of Variability in Immune Responses and Transplant Outcomes

On ranking long-term outcomes in graft survival it was found that for LD (Living donor) grafts, donor age (28%), recipient race (15%), age (15%), transplant year (13%) and recipient sex (11%) play an important role in the variation in results. Whereas for SCD (Standard care donor) allografts - donor age (35%), recipient race (23%), transplant year(15%), recipient sex (8%) and age (5%) [41]. In either case right after donor age the second most influential factor in graft survival is the race of the recipient. This motivates us to look into the role that genetic diversity in individuals (especially caused by race) plays in determining the outcomes of kidney transplants.

Disease causing microorganisms have always played a very significant role in human mortality. Hence it is reasonable to conclude that the human immune system has undergone evolution due to strong selection from pathogens [14, 76]. The first form of humans are considered to be *Homo habilis*,

who evolved in Africa over 2.4 million years ago. As humans started migrating out of Africa they experienced various ranges of environmental factors like climate, diet, and pathogen environment. These played a significant role in the genetic adaptations that followed. As they experienced different selection pressures from different pathogens, their immune response started developing selectively based on the pathogenic environment. Thus humans from different pockets of populations or race show a variation in susceptibility and immune response to diseases. Nédélec *et al.*, [76] in an effort to understand how natural selection is responsible for this genetic variation, tested for the effects of African versus European ancestry on the transcriptional response of primary large phagocytes to live bacterial pathogens such as *Listeria* and *Salmonella*. The authors in [76] state that several genes in the order of magnitude of thousands show differences in transcriptional response to infection based on different populations, for instance populations with African ancestors have shown stronger inflammatory response and reduced intracellular bacterial growth compared to individuals with European ancestry. Thus it is clear that natural selection plays a strong role in differences in immune response due to lineage which are a by product of gene regulation. The authors conclude that even today selective agents continue to shape this genetic diversity which is responsible for the variation in responses to infection. Hélène *et al.*, [83] corroborates on the above findings and uses RNA sequencing to test the transcriptional response of primary large phagocytes not only to live bacterial pathogens but also viral infections. The authors identified a strong transcriptional expression quantitative trait loci (eQTLs) that decreases expression of pro-inflammatory genes in only Europeans. They also identified several cis-eQTLs that are also responsible for the genetic variation in immune response to both viral and bacterial infections within and between populations. They concluded that remnants of Neanderthal DNA might be responsible for responsiveness of the immune system in individuals and is the reason for differences in regulatory responses of genes in Europeans.

Now that genetic variation between immune responses to viral pathogens has been established we narrow down our view to the role of genetic variation in kidney transplants. Bowden [18] provides a review of all the latest research efforts in the genetic studies of renal disease. It states with reasonable confidence through familial studies and comparison of ESRD rates of occurrence for people with different racial backgrounds and ethnicity that the risk of an individual of developing end stage renal disease (ESRD) has a very strong and significant genetic component attached to it. The authors in [40] present that genetic composition of deceased kidney donors have a long-term impact on the survival of the allograft. It also briefly talks about a recent breakthrough which states that lower copy numbers of the complement 4 gene (C4) after receipt of deceased donor kidneys have improved graft survival in transplant recipients. On the other hand, studies [32, 40] have also shown that variation in the gene APOL1 in donor kidneys and individual recipients of African ancestry has led to lower survival rates of allografts compared to their European counterparts.

5.2.2 Renal Pharmacogenomics

Given the role of genes and genetic variation in the outcomes of renal transplants in individuals of different genetic backgrounds, it is only natural to next aim to see if one can tailor medication/ immunosuppressants based on individual genetic makeups. While the feedback control formulated in Chapter 4 would be useful for the adaptive part of this medication process, being able to use genetic information about a patient to decide what immunosuppressants or other drugs to prescribe would definitely contribute to the personalized part (besides individually parametrized models).

Pharmacogenomics is a relatively new field in medicine which combines pharmacology (the science of drugs) and genomics (the study of genes and their functions) to develop effective, safe medications and doses that will be custom made to a person's genetic makeup. It is thus also the study of how genes affect a person's response to drugs [45]. Luttrop *et al.*, [67] talks about the importance of genetic polymorphisms in CKD. Not only would identifying the genetic components that are responsible for the aggravation of CKD improve the process of understanding the causes and outcomes of the disease but also because genetic factors are inherited, using them would bring a certain stability to the prognosis of the disease which it otherwise lacks. Furthermore as mentioned earlier by identifying transplant patients' genetic background, it is possible that a more individualized therapy could be designed. While the feedback optimal control formulation might aid with the dosage of the immunosuppressant, pharmacogenomics would help make the decision of what immunosuppressant would be best suited for the patient based on their genetic make-up.

The review by Kronenberg [62] discusses factors and markers for which there has been evidence collected to support its association with the succession of CKD. The following factors and markers are discussed: asymmetric dimethylarginine, factors involved in calcium-phosphate metabolism, adrenomedullin, A-type natriuretic peptide, N-terminal pro-brain natriuretic peptide, liver-type fatty acid binding protein, kidney injury molecule 1, neutrophil gelatinase-associated lipocalin, apolipoprotein A-IV, adiponectin and some recently identified genetic polymorphisms. However in the end the review does add that substantial amount of verification of these markers and factors needs to be done using epidemiological and experimental data before they can be considered for potential personalized medicine avenues in clinical interventional trials.

Ariadna *et al.*, [79] also focuses on emerging markers to make a predictive association with CKD elevation physiologically, pathologically as well immunologically. Some inhibitors addressed are inhibitors of the renin angiotensin system (including angiotensin-converting enzyme inhibitors and angiotensin II receptor blockers; the vitamin D receptor agonist; salt sensitivity hypertension; and progressive kidney-disease markers with identified genetic polymorphisms). Candidate-gene association studies and genome-wide association studies have assessed the genetic connection

for common renal diseases, including CKD. These studies were conducted on a diverse group of ethnic and racial backgrounds which makes a good case for studying genetic polymorphism with respect to CKD or ESRD. The paper [79] has a table listing all the characteristics and results of the main published studies, there are as many as 13 studies with many sub-studies all involving gene polymorphisms and their effect on renal function. This solidifies our claim that genetic polymorphism plays a substantial role in understanding the immune response in an individual and in turn effects the outcomes for renal transplant recipients. Even though certain diseases like breast cancer have diagnosis and determination of risk factors based on genotyping [86] there is no such concrete definitive DNA-based diagnostics and risk prediction available yet for patients suffering from CKD (Chronic Kidney Disease).

5.2.3 Discussion

Using the mathematical model quantifying the biological relationship of BKV infection and the donor kidney in the presence of an immunosuppressant regime in Chapter 3, we use this model to design a feedback control problem to find the optimum amount of immunosuppressant in renal transplant recipients in Chapter 4.

During the process of improving the original model (3.1) we see that some parameter values have a wide range in literature. Some of these parameters play a crucial role and affect the body's immune response to both the viral load and the kidney in the model (3.3). The existence of such a wide range of parameter values in literature initiates this section to investigate any possible association between these parameter ranges and genetic compositions of individuals. We claim and show through evidence in literature of several studies by other research groups, that there exists a direct connection between genetic diversity (especially that caused by racial and population differences) and the immune response. We also point to studies that specifically show how genetic polymorphism plays a big role in the outcomes of renal transplant procedures. Even though unlike certain diseases, there is no diagnosis and determination of risk factors based on genotyping available yet for patients suffering from CKD, when such genetic markers do become easily determinable, a natural next step after genetic composition based diagnosis for renal transplant recipients (renal pharmacogenomics) would be the regulation and control of these identified biomarkers. The application of control theory to the process of gene regulation as seen in [5] could play a crucial role in shaping the future of renal pharmacogenomics.

CHAPTER

6

CONCLUSION

In this dissertation we show evidence of how mathematical and statistical modeling can be used to model complex biological systems and answer pertinent questions as a result. Specifically, we introduce a model representing the dynamics between the human immune response, a renal graft and BK virus in kidney transplant recipients. Our eventual goal was to use the mathematical model in combination with control theory to develop an optimal treatment plan for renal transplant patients.

In the first chapter, we introduce the problem and show statistical evidence indicating why it is pertinent to solve it. Some of the pressing problems were the high volume of kidney transplants each year and the relatively high number of unfavorable outcomes (renal graft failure) for transplant patients over the course of ten years post-transplant. The recurrent causes for graft failure were either infection or allograft rejection, both of which were often a result of the narrow therapeutic index of immunosuppressant drugs that renal transplant patients were prescribed. With the motivation for modeling, analyzing and eventually addressing this concern, we next present a short survey of previous research on this subject.

An existing model of the immune response in context of BK Virus is first described in Chapter 2. Mathematical model and statistical model misspecification in the context of least squares methodology for this model is investigated using both clinical and simulated data. A difference-based technique is used to determine the statistical model associated with the observation process directly

from data. Now using the chosen statistical model we solve an inverse problem and plot modified residuals. A non-random pattern for the modified residuals would now imply discrepancy in the mathematical modeling process. The two significant advantages of this method were that it was computationally efficient and that one did not need to make tacit assumptions about the correctness of the mathematical model when choosing the statistical model.

With our goal of building an optimal feedback control based treatment, we next, in Chapter 3, set forth analyzing the mathematical model introduced in Chapter 2. Further analysis led us to an iterative process of modeling, when we realized that the current model had certain components which could be improved and made more biologically representative of the true system, the evidence of which is provided in the Chapter 3. We then present the updated model describing the dynamics of BK virus and the immune response in renal transplant recipients under immunosuppressive therapy, and provide evidence in the form of model simulations to support the improvements.

Personalized medicine is a pertinent field aiming to improve patient experience and outcomes with individualized treatments scheduled towards every patient's needs. In Chapter 4, we choose to now use the improved model from Chapter 3 to build an adaptive treatment schedule for renal transplant recipients. We first present numerous robustness tests to authenticate the new model's relevance for building an feedback control formulation. Finally we use Receding Horizon Control Methodology and Extended Kalman Filter State Estimation techniques to build an algorithm which predicts optimal immunosuppressive dosages for renal transplant recipients at every doctor's visit and present results to corroborate. The motivation for using Extended Kalman Filter State Estimation is to provide estimates for model states for which data is not available and also to account for the noise in both the mathematical model and the clinical data. The advantage of using Receding Horizon Control Methodology for our feedback control loop is that this technique avails information from the future to predict optimal solutions for the present.

Lastly, in Chapter 5, literature is presented on renal pharmacogenomics and pharmacodynamics of immunosuppressants (to model specific drug dosage instead of efficiency) and the feasibility of possible future work in personalized medicine for renal transplant recipients in these avenues.

Other avenues of future work contextual to model development specifically would involve acquisition of more clinical data to estimate model parameters with high certainty and incorporating a more elaborate representation of the immune response (for instance including helper CD4+ T-Cells). The work presented in this dissertation has formed a gateway towards development of patient specific models and treatment schedules for individuals who have undergone solid organ transplantations.

BIBLIOGRAPHY

- [1] *2016 Annual Data Report. Scientific Registry of Transplant Recipients.* http://srtr.transplant.hrsa.gov/annual_reports/Default.aspx. Accessed: 2018-03-11.
- [2] M. C. van Aalderen, K. M. Heutinck, C. Huisman, and I. Ten Berge. “BK virus infection in transplant recipients: Clinical manifestations, treatment options and the immune response”. In: *Neth J Med* 70.4 (2012), pp. 172–183.
- [3] B. M. Adams, H. T. Banks, H. Kwon, and H. T. Tran. “Dynamic multidrug therapies for HIV: Optimal and STI control approaches”. In: *Mathematical Biosciences and Engineering* 1.2 (2004), pp. 223–241.
- [4] N. Babel, H.-D. Volk, and P. Reinke. “BK polyomavirus infection and nephropathy: The virus–immune system interplay”. In: *Nature Reviews Nephrology* 7.7 (2011), p. 399.
- [5] F. L. Baldissera, J. E. Cury, and J. Raisch. “A supervisory control theory approach to control gene regulatory networks”. In: *IEEE Transactions on Automatic Control* 61.1 (2016), pp. 18–33.
- [6] H. T. Banks, J. Catenacci, and S. Hu. “Use of difference-based methods to explore statistical and mathematical model discrepancy in inverse problems”. In: *Journal of Inverse and Ill-posed Problems* 24.4 (2016), pp. 413–433.
- [7] H. T. Banks, S. Hu, T. Jang, and H. Kwon. “Modelling and optimal control of immune response of renal transplant recipients”. In: *Journal of Biological Dynamics* 6.2 (2012), pp. 539–567.
- [8] H. T. Banks, S. Hu, K. Link, E. S. Rosenberg, S. Mitsuma, and L. Rosario. “Modelling immune response to BK virus infection and donor kidney in renal transplant recipients”. In: *Inverse Problems in Science and Engineering* 24.1 (2016), pp. 127–152.
- [9] H. T. Banks, S. Hu, and W. C. Thompson. *Modelling and Inverse Problems in the Presence of Uncertainty*. CRC Press, 2014.
- [10] H. T. Banks, H. Kwon, J. Toivanen, and H. T. Tran. “A state-dependent Riccati equation-based estimator approach for HIV feedback control”. In: *Optimal Control Applications and Methods* 27.2 (2006), pp. 93–121.
- [11] H. T. Banks and H. T. Tran. *Mathematical and Experimental Modelling of Physical and Biological Processes*. CRC Press, 2009.
- [12] H. T. Banks. *Modelling and Control in the Biomedical Sciences*. Vol. 6. Springer Science & Business Media, 2013.
- [13] H. T. Banks, R. A. Everett, S. Hu, N. Murad, and H. T. Tran. “Mathematical and statistical model misspecifications in modelling immune response in renal transplant recipients”. In: *Inverse Problems in Science and Engineering* 26.2 (2018), pp. 203–222.

- [14] L. B. Barreiro and L. Quintana-Murci. "From evolutionary genetics to human immunology: How selection shapes host defense genes". In: *Nature Reviews Genetics* 11.1 (2010), p. 17.
- [15] A Beiras-Fernandez, E Thein, and C Hammer. "Induction of immunosuppression with polyclonal antithymocyte globulins: an overview". In: *Exp. Clin. Transplant* 1.2 (2003), pp. 79–84.
- [16] F Biral, E Bertolazzi, and P Bosetti. "Notes on numerical methods for solving optimal control problems". In: *IEEJ Journal of Industry Applications* 5.2 (2016), pp. 154–166.
- [17] D. L. Bohl and D. C. Brennan. "BK virus nephropathy and kidney transplantation". In: *Clinical Journal of the American Society of Nephrology* 2.Supplement 1 (2007), S36–S46.
- [18] D. W. Bowden. "Genetics of kidney disease". In: *Kidney International* 63 (2003), S8–S12.
- [19] B. E. Chambers and R. A. Wingert. "Renal progenitors: Roles in kidney disease and regeneration". In: *World Journal of Stem Cells* 8.11 (2016), p. 367.
- [20] L. Chatenoud. "Precision medicine for autoimmune disease". In: *Nature Biotechnology* 34.9 (2016), p. 930.
- [21] G Ciancio and G. Burke. "Alemtuzumab (Campath-1H) in Kidney Transplantation". In: *American Journal of Transplantation* 8.1 (2008), pp. 15–20.
- [22] J. Colaneri. "An overview of transplant immunosuppression-history, principles, and current practices in kidney transplantation". In: *Nephrology Nursing Journal* 41.6 (2014), p. 549.
- [23] J. David, H. T. Tran, and H. T. Banks. "HIV model analysis under optimal control based treatment strategies". In: *International Journal of Pure and Applied Mathematics* 57 (2009), pp. 357–392.
- [24] J. David, H. T. Tran, and H. T. Banks. "Receding horizon control of HIV". In: *Optimal Control Applications and Methods* 32.6 (2011), pp. 681–699.
- [25] M Davidian and D. M. Giltinan. "Nonlinear models for repeated measurement data: an overview and update". In: *Journal of Agricultural, Biological, and Environmental Statistics* 8.4 (2003), p. 387.
- [26] C. Davis. *Creatinine (Low, High, Blood Test Results Explained)*. https://www.medicinenet.com/creatinine_blood_test/article.htm. Accessed: 2018-03-27.
- [27] C. Dudley, E. Pohanka, H. Riad, J. Dedochova, P. Wijngaard, C. Sutter, H. T. Silva Jr, M. M.C.C. S. Group, et al. "Mycophenolate Mofetil Substitution for Cyclosporine A in Renal Transplant Recipients with Chronic Progressive Allograft Dysfunction: The Creeping Creatinine Study 1". In: *Transplantation* 79.4 (2005), pp. 466–475.

- [28] N. Elfadawy, S. M. Flechner, X. Liu, J. Schold, D. Tian, T. R. Srinivas, E. Poggio, R. Fatica, R. Avery, and S. B. Mossad. “The impact of surveillance and rapid reduction in immunosuppression to control BK virus-related graft injury in kidney transplantation”. In: *Transplant International* 26.8 (2013), pp. 822–832.
- [29] N. K. Foundation. *Glomerular Filtration Rate (GFR)*. <https://www.kidney.org/atoz/content/gfr>. Accessed: 2018-03-27.
- [30] N. K. Foundation. *How to Classify CKD*. <https://www.kidney.org/professionals/explore-your-knowledge/how-to-classify-ckd>. Accessed: 2018-03-11.
- [31] N. K. Foundation. *Immunosuppressants*. <https://www.kidney.org/atoz/content/immuno>. Accessed: 2018-03-27.
- [32] B. I. Freedman, S. O. Pastan, A. K. Israni, D. Schladt, B. A. Julian, M. D. Gautreaux, V. Hauptfeld, R. A. Bray, H. M. Gebel, A. D. Kirk, et al. “APOL1 genotype and kidney transplantation outcomes from deceased African American donors”. In: *Transplantation* 100.1 (2016), p. 194.
- [33] P. Friend, G. Hale, H. Waldmann, S. Gore, S. Thiru, V. Joysey, D. Evans, and R. Calne. “Campath-1M–prophylactic use after kidney transplantation. A randomized controlled clinical trial.” In: *Transplantation* 48.2 (1989), pp. 248–253.
- [34] Y. Fujigaki. “Different modes of renal proximal tubule regeneration in health and disease”. In: *World Journal of Nephrology* 1.4 (2012), p. 92.
- [35] G. A. Funk, R. Gosert, P. Comoli, F. Ginevri, and H. H. Hirsch. “Polyomavirus BK replication dynamics in vivo and in silico to predict cytopathology and viral clearance in kidney transplants”. In: *American Journal of Transplantation* 8.11 (2008), pp. 2368–2377.
- [36] G. A. Funk and H. H. Hirsch. “From plasma BK viral load to allograft damage: Rule of thumb for estimating the intrarenal cytopathic wear”. In: *Clinical Infectious Diseases* 49.6 (2009), pp. 989–990.
- [37] G. A. Funk, J. Steiger, and H. H. Hirsch. “Rapid dynamics of polyomavirus type BK in renal transplant recipients”. In: *The Journal of Infectious Diseases* 193.1 (2006), pp. 80–87.
- [38] S. Gabardi, S. T. Martin, K. L. Roberts, and M. Grafals. “Induction immunosuppressive therapies in renal transplantation”. In: *American Journal of Health-System Pharmacy* 68.3 (2011), pp. 211–218.
- [39] R. Gaston. “Immunosuppressive Therapy: The Scientific Basis and Clinical Practice of Immunosuppressive Therapy in the Management of Transplant Recipients”. In: *Extending Medicare Coverage for Preventive and Other Services (Appendix D, Part I)* (2000). Editors: M. J. Field, R. L. Lawrence and L. Zwanziger, US, Washington (DC), 2000.
- [40] M. D. Gautreaux and B. I. Freedman. “Genotypic variation and outcomes in kidney transplantation: Donor and recipient effects”. In: *Kidney international* 84.3 (2013), pp. 431–433.

- [41] D. W. Gjertson. "Explainable variation in renal transplant outcomes: A comparison of standard and expanded criteria donors." In: *Clinical Transplants* (2004), pp. 303–314.
- [42] M. S. Grewal and A. P. Andrews. *Kalman Filtering: Theory and Practice Using MATLAB*, Wiley, 2001.
- [43] P. F. Halloran. "Immunosuppressive drugs for kidney transplantation". In: *New England Journal of Medicine* 351.26 (2004), pp. 2715–2729.
- [44] N. I.o. D. National Institutes of Health, Digestive, and K. Diseases. *2015 USRDS annual data report: Epidemiology of Kidney Disease in the United States*. United States Renal Data System.
- [45] U. D. of Health and H. Services. *What is pharmacogenomics?* <https://ghr.nlm.nih.gov/primer/genomicresearch/pharmacogenomics>. Accessed: 2018-03-28.
- [46] T. Healthcare. *Micromedex® Healthcare Series [internet database]*. 2008.
- [47] H. H. Hirsch, C. B. Drachenberg, J. Steiger, and E. Ramos. "Polyomavirus-associated nephropathy in renal transplantation: critical issues of screening and management". In: (2013).
- [48] H. H. Hirsch, W. Knowles, M. Dickenmann, J. Passweg, T. Klimkait, M. J. Mihatsch, and J. Steiger. "Prospective study of polyomavirus type BK replication and nephropathy in renal-transplant recipients". In: *New England Journal of Medicine* 347.7 (2002), pp. 488–496.
- [49] H. S. C. Institute. "The Kidney Repair Shop". In: *Stem Cell Lines* 3.2 (2008).
- [50] N. Issa and W. E. Braun. "Immunosuppression for renal transplant patients and common medical problems in renal transplantation". In: *Current Clinical Medicine* (2009).
- [51] A. James and R. B. Mannon. "The Cost of Transplant Immunosuppressant Therapy: Is This Sustainable?" In: *Current Transplantation Reports* 2.2 (2015), pp. 113–121.
- [52] T. S. Jang, J. Kim, H.-D. Kwon, and J. Lee. "Hybrid on-off controls for an HIV model based on a linear control problem". In: *Journal of the Korean Mathematical Society* 52.3 (2015), pp. 469–487.
- [53] A. H. Jazwinski. *Stochastic Processes and Filtering Theory*. Courier Corporation, 2007.
- [54] M. J. Keating, S. O'Brien, and A. Ferrajoli. "Alemtuzumab: A novel monoclonal antibody". In: *Expert Opinion on Biological Therapy* 1.6 (2001), pp. 1059–1065.
- [55] G. M. Kepler, H. T. Banks, M. Davidian, and E. S. Rosenberg. "A model for HCMV infection in immunosuppressed patients". In: *Mathematical and Computer Modelling* 49.7-8 (2009), pp. 1653–1663.
- [56] A. D. Kirk. "Induction immunosuppression". In: *Transplantation* 82.5 (2006), pp. 593–602.

- [57] D. E. Kirk. *Optimal Control Theory: An Introduction*. Courier Corporation, 2012.
- [58] S. J. Knechtle, L. A. Fernandez, J. D. Pirsch, B. N. Becker, L. T. Chin, Y. T. Becker, J. S. Odorico, A. M. D’alessandro, and H. W. Sollinger. “Campath-1H in renal transplantation: The University of Wisconsin experience”. In: *Surgery* 136.4 (2004), pp. 754–760.
- [59] S. J. Knechtle, J. D. Pirsch, J. H. Fechner, B. N. Becker, A. Friedl, R. B. Colvin, L. K. Lebeck, L. T. Chin, Y. T. Becker, J. S. Odorico, et al. “Campath-1H Induction Plus Rapamycin Monotherapy for Renal Transplantation: Results of a Pilot Study”. In: *American Journal of Transplantation* 3.6 (2003), pp. 722–730.
- [60] J. M. Kovarik, S. Hartmann, J. Figueiredo, C. Rordorf, G. Golor, A. Lison, K. Budde, and H. H. Neumayer. “Effect of food on everolimus absorption: quantification in healthy subjects and a confirmatory screening in patients with renal transplants”. In: *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy* 22.2 (2002), pp. 154–159.
- [61] J. M. Kovarik, C.-H. Hsu, L. McMahon, S. Berthier, and C. Rordorf. “Population pharmacokinetics of everolimus in de novo renal transplant patients: impact of ethnicity and comedications”. In: *Clinical Pharmacology & Therapeutics* 70.3 (2001), pp. 247–254.
- [62] F. Kronenberg. “Emerging risk factors and markers of chronic kidney disease progression”. In: *Nature Reviews Nephrology* 5.12 (2009), p. 677.
- [63] Y. N.H.H.M. C. Department of Laboratory Medicine. “Quantitative BK virus DNA PCR for Diagnosis in Compromised Hosts.” In: *Clinical Virology Laboratory Newsletter* 14.4 (2005).
- [64] Y. N.H.H.M. C. Department of Laboratory Medicine. “BK virus Nephropathy, Genome Variability and the Pitfalls of PCR Surveillance.” In: *Clinical Virology Laboratory Newsletter* 20.1 (2011).
- [65] S. Lenhart and J. T. Workman. *Optimal Control Applied to Biological Models*. CRC Press, 2007.
- [66] F. L. Lewis, D. Vrabie, and V. L. Syrmos. *Optimal Control*. John Wiley & Sons, 2012.
- [67] K. Luttrupp, P. Stenvinkel, J. J. Carrero, R. Pecoits-Filho, B. Lindholm, and L. Nordfors. “Understanding the role of genetic polymorphisms in chronic kidney disease”. In: *Pediatric Nephrology* 23.11 (2008), pp. 1941–1949.
- [68] H.-U. Meier-Kriesche and B. Kaplan. “Toxicity and efficacy of sirolimus: relationship to whole-blood concentrations”. In: *Clinical Therapeutics* 22 (2000), B93–B100.
- [69] D. J. A. Moes, H.-J. Guchelaar, and J. W. de Fijter. “Sirolimus and everolimus in kidney transplantation”. In: *Drug Discovery Today* 20.10 (2015), pp. 1243–1249.
- [70] D. J. A. Moes, R. R. Press, J. den Hartigh, T. van der Straaten, J. W. de Fijter, and H.-J. Guchelaar. “Population pharmacokinetics and pharmacogenetics of everolimus in renal transplant patients”. In: *Clinical Pharmacokinetics* 51.7 (2012), pp. 467–480.

- [71] R. G. Morris. "Cyclosporin therapeutic drug monitoring-an established service revisited". In: *The Clinical Biochemist Reviews* 24.2 (2003), p. 33.
- [72] A. Muntean and M. Lucan. "Immunosuppression in kidney transplantation". In: *Clujul Medical* 86.3 (2013), p. 177.
- [73] N. Murad, R. A. Everett, H. T. Tran, E. S. Rosenberg, and H. T. Banks. "Immunosuppressant Treatment Dynamics in Renal Transplant Recipients: An Iterative Modelling Approach". In: *CRSC Technical Report, CRSC-TR17-23* (2017). Discrete and Continuous Dynamical Systems Series B, (Manuscript submitted for publication.)
- [74] N. Murad, H. T. Tran, and H. T. Banks. "Optimal Control of Immunosuppressants in Renal Transplant Recipients Susceptible to BKV Infection". In: *CRSC Technical Report, CRSC-TR18-02* (2018). Optimal Control, Applications and Methods, (Manuscript submitted for publication.)
- [75] B. Nashan. "Antibody induction therapy in renal transplant patients receiving calcineurin-inhibitor immunosuppressive regimens". In: *BioDrugs* 19.1 (2005), pp. 39–46.
- [76] Y. Nédélec, J. Sanz, G. Baharian, Z. A. Szpiech, A. Pacis, A. Dumaine, J.-C. Grenier, A. Freiman, A. J. Sams, S. Hebert, et al. "Genetic ancestry and natural selection drive population differences in immune responses to pathogens". In: *Cell* 167.3 (2016), pp. 657–669.
- [77] H. Neumayer, K Paradis, A Korn, C Jean, L Fritsche, K Budde, M Winkler, V Kliem, R Pichlmayr, I. Hauser, et al. "Entry-into-human study with the novel immunosuppressant SDZ RAD in stable renal transplant recipients". In: *British Journal of Clinical Pharmacology* 48.5 (1999), p. 694.
- [78] *Organ Procurement and Transplantation Network*. <http://optn.transplant.hrsa.gov/>. Accessed: 2018-03-11.
- [79] A. Padullés, I. Rama, I. Llaudó, and N. Lloberas. "Developments in renal pharmacogenomics and applications in chronic kidney disease". In: *Pharmacogenomics and Personalized Medicine* 7 (2014), p. 251.
- [80] M. Pickup. "Clinical pharmacokinetics of prednisone and prednisolone". In: *Clinical Pharmacokinetics* 4.2 (1979), pp. 111–128.
- [81] E. Polak. *Computational Methods in Optimization: A Unified Approach*. Vol. 77. Academic press, 1971.
- [82] M. Prague, D. Commenges, and R. Thiébaud. "Dynamical models of biomarkers and clinical progression for personalized medicine: The HIV context". In: *Advanced Drug Delivery Reviews* 65.7 (2013), pp. 954–965.
- [83] H. Quach, M. Rotival, J. Pothlichet, Y.-H. E. Loh, M. Dannemann, N. Zidane, G. Laval, E. Patin, C. Harmant, M. Lopez, et al. "Genetic adaptation and neandertal admixture shaped the immune system of human populations". In: *Cell* 167.3 (2016), pp. 643–656.

- [84] P Randhawa and D. Brennan. “BK virus infection in transplant recipients: An overview and update”. In: *American Journal of Transplantation* 6.9 (2006), pp. 2000–2005.
- [85] A. V. Rao. “A survey of numerical methods for optimal control”. In: *Advances in the Astronautical Sciences* 135.1 (2009), pp. 497–528.
- [86] M. Robson and K. Offit. “Management of an inherited predisposition to breast cancer”. In: *New England Journal of Medicine* 357.2 (2007), pp. 154–162.
- [87] H. Schättler and U. Ledzewicz. *Optimal Control for Mathematical Models of Cancer Therapies*. Springer, 2015.
- [88] G. Seber and C. Wild. *Nonlinear Regression*. John Wiley & Sons, 2003.
- [89] J Sellares, D. De Freitas, M Mengel, J Reeve, G Einecke, B Sis, L. Hidalgo, K Famulski, A Matas, and P. Halloran. “Understanding the causes of kidney transplant failure: The dominant role of antibody-mediated rejection and nonadherence”. In: *American Journal of Transplantation* 12.2 (2012), pp. 388–399.
- [90] W. D. Shlomchik. “Graft-versus-host disease”. In: *Nature Reviews Immunology* 7.5 (2007), p. 340.
- [91] L. E. Tushla. *When a transplant fails*. https://www.kidney.org/transplantation/transaction/TC/summer09/TCsm09_TransplantFails. Accessed: 2018-03-11.
- [92] *United Network for Organ Sharing*. <https://unos.org/data/>. Accessed: 2018-03-11.
- [93] C. J. Watson, J. A. Bradley, P. J. Friend, J. Firth, C. J. Taylor, J. R. Bradley, K. G. Smith, S. Thiru, N. V. Jamieson, G. Hale, et al. “Alemtuzumab (CAMPATH 1H) induction therapy in cadaveric kidney transplantation efficacy and safety at five years”. In: *American Journal of Transplantation* 5.6 (2005), pp. 1347–1353.
- [94] L. Xie, D. Popa, and F. L. Lewis. *Optimal and Robust Estimation: With an Introduction to Stochastic Control Theory*. CRC press, 2007.
- [95] J. J. Zimmerman, G. M. Ferron, H.-K. Lim, and V. Parker. “The Effect of a High-Fat Meal on the Oral Bioavailability of the Immunosuppressant Sirolimus (Rapamycin)”. In: *The Journal of Clinical Pharmacology* 39.11 (1999), pp. 1155–1161.
- [96] J. J. Zimmerman and B. D. Kahan. “Pharmacokinetics of sirolimus in stable renal transplant patients after multiple oral dose administration”. In: *The Journal of Clinical Pharmacology* 37.5 (1997), pp. 405–415.

APPENDIX

APPENDIX

A

MATHEMATICAL AND STATISTICAL MODEL MISSPECIFICATION IN MODELING IMMUNE RESPONSE IN RENAL TRANSPLANT RECIPIENTS

A.1 Derivation of Modified Pseudo Errors

Consider a general N -dimensional dynamical system with parameter vector \mathbf{q} ,

$$\begin{aligned}\frac{d\mathbf{x}}{dt}(t) &= \mathbf{g}(t, \mathbf{x}(t); \mathbf{q}), \\ \mathbf{x}(t_0) &= \mathbf{x}_0,\end{aligned}$$

with an m dimensional observation process

$$f(t; \boldsymbol{\theta}) = \mathcal{C} \mathbf{x}(t; \boldsymbol{\theta}),$$

where $\theta = (q^\top, \hat{x}_0^\top)^\top$ is the vector of parameters along with the initial conditions to be estimated and \mathcal{C} is the $m \times N$ observation matrix.

Consider the following statistical model for observable k ,

$$Y_i^k = f_k(t_i^k; \theta_0) + f_k(t_i^k; \theta_0)^{\gamma_k} \mathcal{E}_i^k, \quad i = 1, 2, \dots, n_k,$$

where $\gamma_k \geq 0$ and the vector $\theta_0 \in \Omega$ is the “true” or nominal parameter set. The $n_k \times 1$ random error vector \mathcal{E}^k is assumed to be independent and identically distributed (*i.i.d*) with mean zero and $\text{Var}(\mathcal{E}_i^k) = \sigma_{0k}^2$. The corresponding realization is,

$$y_i^k = f_k(t_i^k; \theta_0) + f_k(t_i^k; \theta_0)^{\gamma_k} \epsilon_i^k, \quad i = 1, 2, \dots, n_k. \quad (\text{A.1})$$

Note that (A.1) can be rewritten as

$$f_k(t_i^k; \theta_0) = y_i^k - f_k(t_i^k; \theta_0)^{\gamma_k} \epsilon_i^k. \quad (\text{A.2})$$

The observation error is represented by the term $f_k(t_i^k; \theta_0)^{\gamma_k} \epsilon_i^k$, which is proportional to the size of the observations and quantifies the discrepancy between the model and the data.

We define

$$\hat{\epsilon}_i^k \equiv |f_k(t_i^k; \theta_0)|^{\gamma_k} \epsilon_i^k. \quad (\text{A.3})$$

We use difference based methods to calculate this approximation of observation error (pseudo measurement errors) by the following

$$\hat{\epsilon}_i^k = \begin{cases} \frac{1}{\sqrt{2}}(y_{i+1}^k - y_i^k) & \text{for } i = 1 \\ \frac{1}{\sqrt{6}}(y_{i-1}^k - 2y_i^k + y_{i+1}^k) & \text{for } i = 2, \dots, n_{k-1} \\ \frac{1}{\sqrt{2}}(y_i^k - y_{i-1}^k) & \text{for } i = n_k. \end{cases}$$

Substituting the approximation (A.3) in (A.2) we obtain

$$f_k(t_i^k; \theta_0) \approx y_i^k - \hat{\epsilon}_i^k. \quad (\text{A.4})$$

Using (A.4), we can rewrite (A.3) as

$$\begin{aligned}\epsilon_i^k &\approx \frac{\hat{\epsilon}_i^k}{|f_k(t_i^k; \theta_0)|^{\gamma_k}} \\ &\approx \frac{\hat{\epsilon}_i^k}{|y_i^k - \hat{\epsilon}_i^k|^{\gamma_k}}.\end{aligned}$$

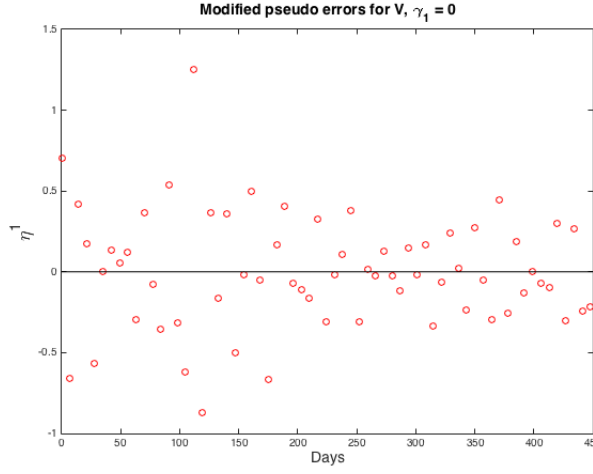
Thus we define modified pseudo errors given by the following

$$\eta_i^k = \frac{\hat{\epsilon}_i^k}{|y_i^k - \hat{\epsilon}_i^k|^{\gamma_k}}.$$

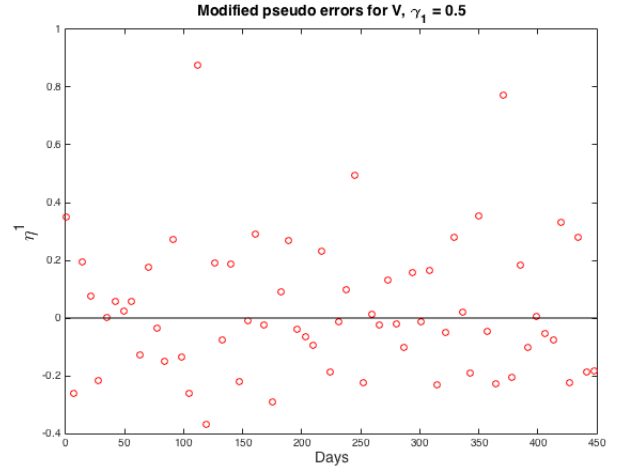
A.2 Simulated Data from the Original Model (3.1)

We apply the difference-based method to the simulated data set (2.6) created using the original model (3.1) to determine the correct γ value. The modified pseudo errors for the viral load with various γ_1 values are given in Figure A.1. As expected, $\gamma_1 = 0.5$ produces the desired scatter plot whereas $\gamma_1 = 0$ and $\gamma_1 = 1$ produce undesired megaphone shapes. Figure A.2a contains the desired modified pseudo errors for creatinine levels with $\gamma_2 = 0$. While Figure A.2b with $\gamma_2 = 1.5$ is similar, it is not quite as symmetric and has larger modified pseudo error values.

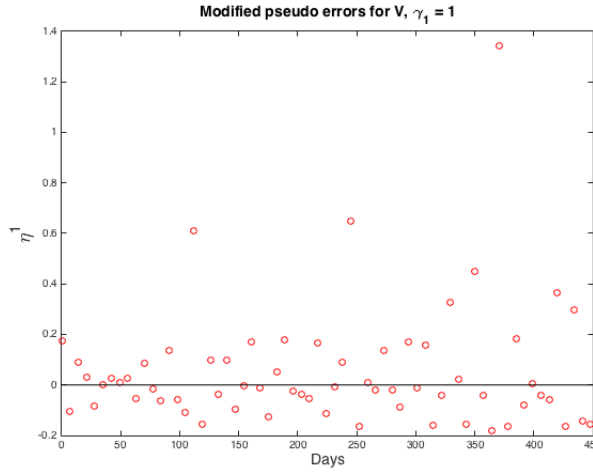
We then solve the inverse problem using the original model (3.1) with $\gamma = (0.5, 0)$ and plot the modified residuals to verify there is no mathematical model misspecification. The model solutions and corresponding modified residuals are plotted in Figure A.3 and Figure A.4. As expected, the model solutions fit the data well and the corresponding modified residuals appear to form a uniform band around the horizontal axis. The five estimated parameters are, $[\log_{10} \beta, \log_{10} \bar{\rho}_{EV}, \log_{10} \delta_{EV}, \log_{10} \delta_{EK}, \log_{10} \bar{\rho}_{EK}] = [7.0735, -0.6008, -0.9628, -0.9948, -0.7836]$.



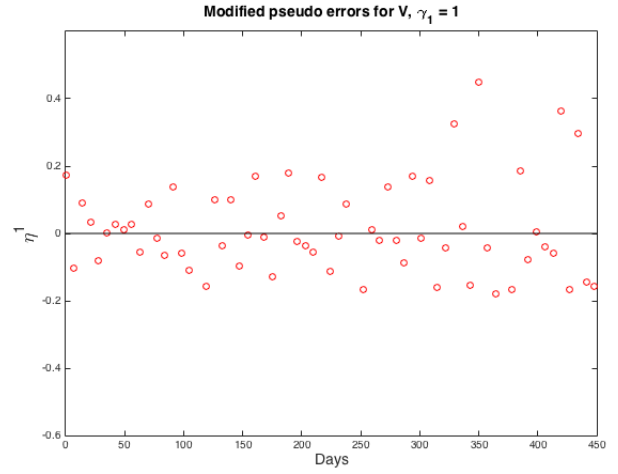
(a) $\gamma_1 = 0$



(b) $\gamma_1 = 0.5$

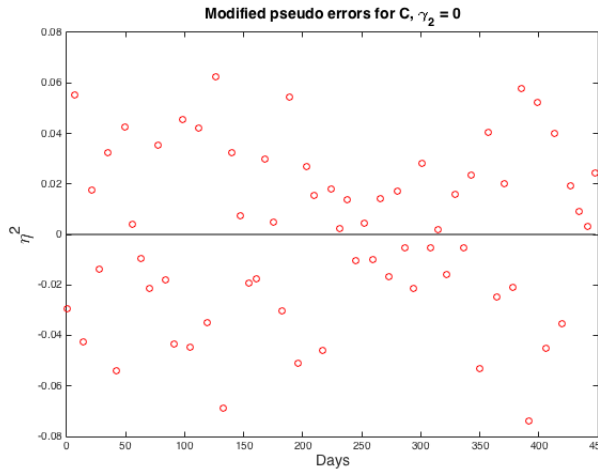


(c) $\gamma_1 = 1$

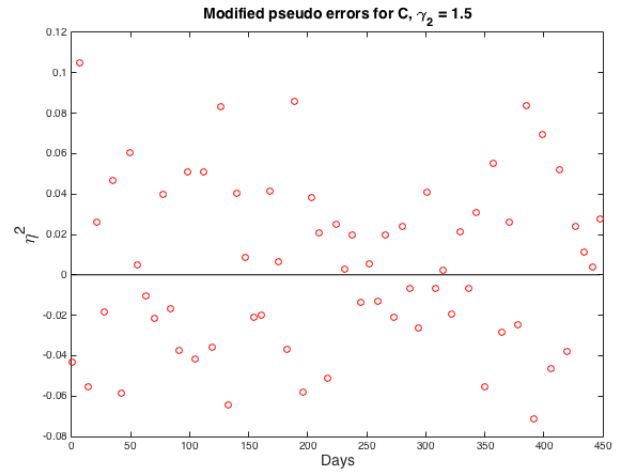


(d) Enlarged image, $\gamma_1 = 1$

Figure A.1: Simulated viral load modified pseudo errors vs. time for various γ_1 values.

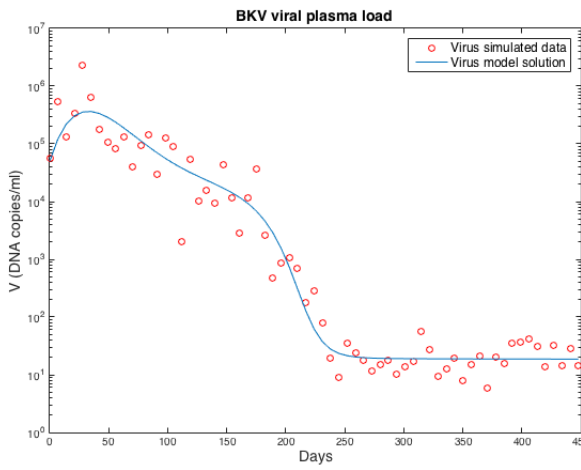


(a) $\gamma_2 = 0$

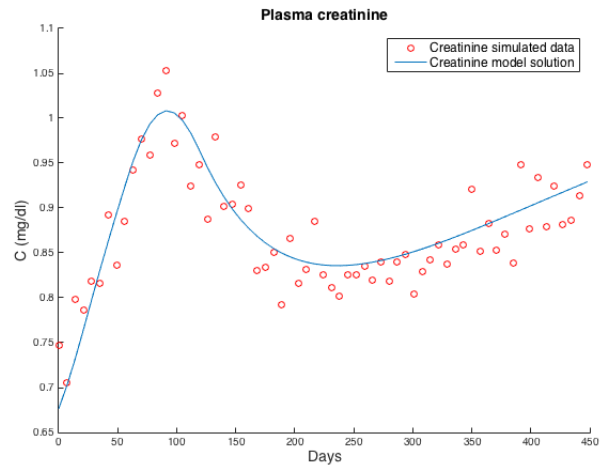


(b) $\gamma_2 = 1.5$

Figure A.2: Simulated creatinine modified pseudo errors vs. time for various γ_2 values.

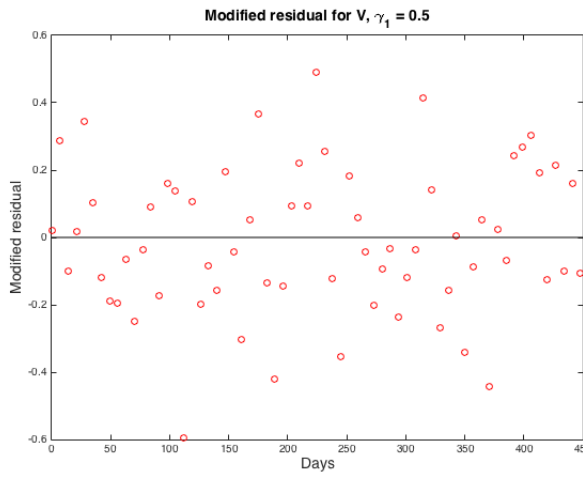


(a) V model solution and simulated data

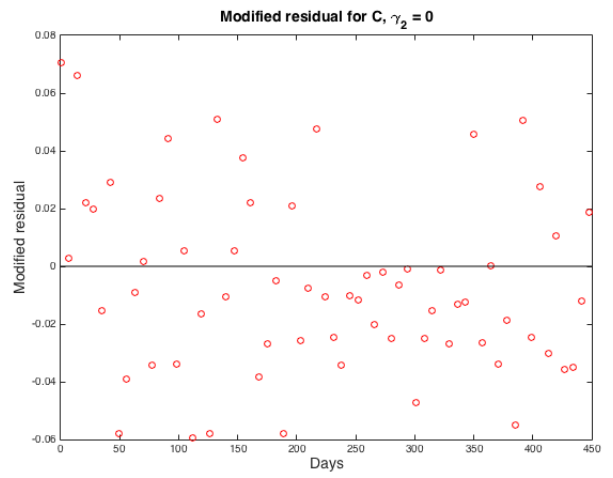


(b) C model solution and simulated data

Figure A.3: Inverse problem model (3.1) solution and simulated data with $\gamma = (0.5, 0)$.



(a) Modified residuals for V



(b) Modified residuals for C

Figure A.4: Modified residuals for V and C with $\gamma = (0.5, 0)$.

A.3 Simulated Data from the Simpler Model (2.7)

We apply the difference-based method to the simulated data set (2.6) generated using the simpler model (2.7) to determine the correct γ value. The randomness in the modified pseudo errors for both the viral load with $\gamma_1 = 0$ (Figure A.5a) and the creatinine levels with $\gamma_2 = 0.7$ (Figure A.5b) reiterate that the difference-based method works as expected.

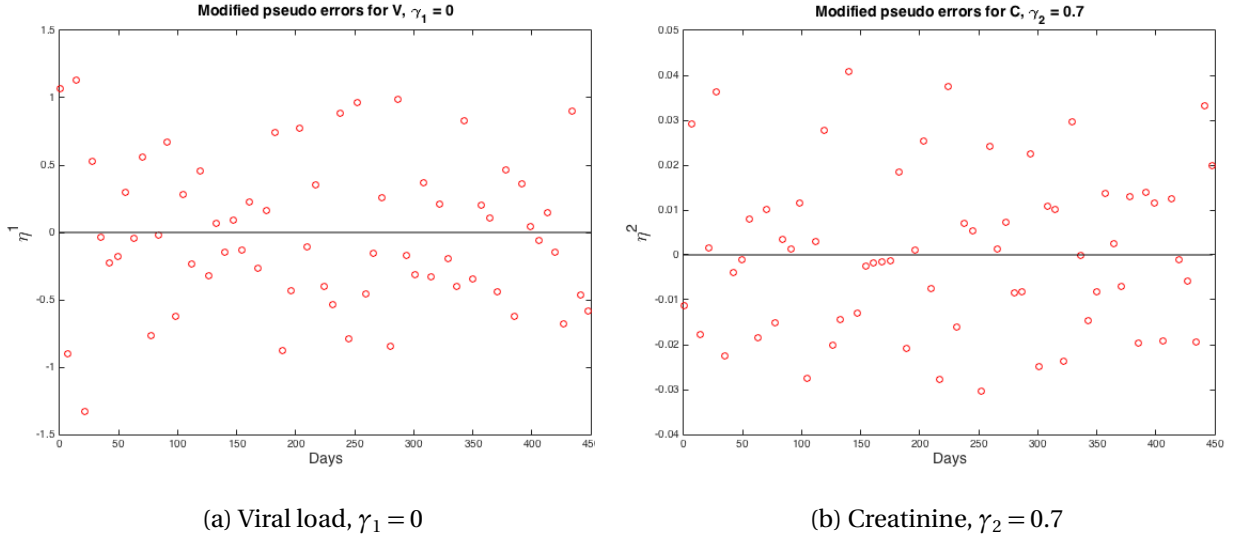


Figure A.5: Modified pseudo errors for the viral load and creatinine with $\gamma = (0, 0.7)$.