Abstract

ADAMS, BRIAN MICHAEL. Non-parametric Parameter Estimation and Clinical Data Fitting with a Model of HIV Infection. (Under the direction of H. Thomas Banks.)

The focus of this dissertation is to develop a combined mathematical and statistical modeling approach for analyzing clinical data from an HIV acute infection study. We amalgamate two existing models from the literature to create a nonlinear differential equation model of in-host infection dynamics that is capable of predicting sustained low-level viral loads and multiple stable equilibria. Using this example system of differential equations we demonstrate two contrasting parameter identification problem formulations for estimating the distribution of model parameters across a population: the first at the individual patient level and the second directly at the population level itself. In the latter case one leverages data from all patients to estimate a probability density function representing the distribution. We discuss well-posedness and computational implementation for such inverse problems. Directly estimating the distribution in this way may offer computational advantages over estimating parameters for individual patients.

In the context of the model, we implement the Expectation Maximization (EM) Algorithm for maximum likelihood estimation to handle patient measurements censored by assay resolution limits. This censored data method is beneficial since with it

we do not arbitrarily assign values for measurements below the limit of detection, but rather compute their expected value based on the dynamics model and conditioned on the knowledge that they are censored. In addition, in both inverse problem contexts (estimating a vector of parameters for a single patient and the distribution of a parameter across all patients) we develop and apply methods for estimating variability of the resulting parameter estimates by using sensitivity analysis to calculate confidence intervals.

We validate each of the methods with simulated data and demonstrate typical results. Finally we present results for the application of the methods to actual clinical data and give examples of conclusions that one might draw from them. This model fitting approach may help clinicians better understand patient behaviors and notably, could alert them to the expected long-term trend for a particular patient.

NON-PARAMETRIC PARAMETER ESTIMATION AND CLINICAL DATA FITTING WITH A MODEL OF HIV INFECTION

BY

BRIAN M. ADAMS

A DISSERTATION SUBMITTED TO THE GRADUATE FACULTY OF

NORTH CAROLINA STATE UNIVERSITY

IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

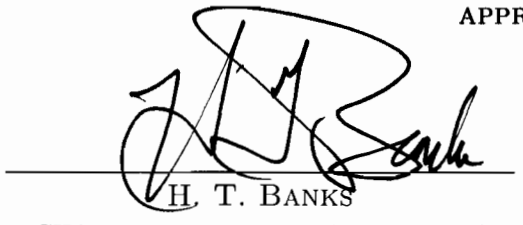
DOCTOR OF PHILOSOPHY

APPLIED MATHEMATICS, COMPUTATIONAL MATHEMATICS CONCENTRATION

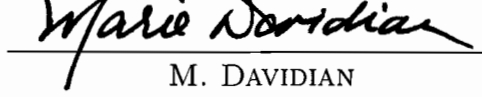
RALEIGH, NORTH CAROLINA

JULY 2005

APPROVED BY:



H. T. BANKS
CHAIR OF ADVISORY COMMITTEE



M. DAVIDIAN

R. H. MARTIN

H. T. TRAN

Biography

Brian Michael Adams, born October 10, 1977, in Bridgeport, Connecticut, grew up in the rural Vermont towns of Underhill and Jeffersonville. He was primarily educated in the public schools of those towns including Lamoille Union High School, where his love of science and mathematics became evident and interests in technical theatre and music flourished. Community and family experiences complemented his traditional academic learning. Through Scouting, working in a local hardware store, and assisting his parents in restoring a late 1800s farmhouse, he developed skills and sensibilities he values immensely to this day. He graduated Salutatorian from Lamoille Union in June 1995, and, not feeling too geographically adventurous, chose to attend Saint Michael's College in Colchester, Vermont.

Some have suggested that Brian's undergraduate career consisted mainly of learning about and playing music, supporting the campus computing community as a systems and network specialist, participating in church activities and community service projects, and engineering live sound for concerts. While the author does not refute these allegations, he did spend considerable time studying mathematics, computer science and humanities. Brian earned a Bachelor of Science degree in Mathematics with minors in both Music and Computer Science and certification to teach secondary school mathematics. He graduated Valedictorian in May 1999, mathematically richer, significantly more aware of cultural and social justice concerns, and desiring to expe-

rience the world in a more adventurous manner.

Responding to his adventurous itch, Brian served in Spring 2000 as an Americorps volunteer in New Orleans, Louisiana, teaching inner-city middle schoolers religion and computer skills. He concurrently installed a data, voice, and video network that he would maintain for four more years. Aware that a lifetime in the Crescent City was not for him, Brian next tried to satiate his curiosity as to how he could utilize all the previously-learned mathematics.

In Summer 2000 the author settled in Raleigh, NC, where he quickly learned about living in a bigger city and to appreciate cooking and eating good food and traveling to various places for conferences. His graduate career started at North Carolina State University with the support of a U.S. Department of Education GAANN Computational Science Fellowship. After completing initial coursework in applied mathematics and a summer internship at Fred Hutchinson Cancer Research Center, he received his Master's degree in December 2000. Delighted to be more aware of the various uses of mathematics and compelled to study biological applications, Brian continued research in mathematical biology under the direction and support of Dr. H.T. Banks. Dr. Banks advised him on a minor project on population models for aphids, and then through completion of his Ph.D. project on modeling HIV infection in Summer 2005. He earned a Ph.D. in Computational and Applied Mathematics to be awarded December 2005.

Brian has accepted a research position at Sandia National Laboratories in Albuquerque, New Mexico. He looks forward to the adventures that await him there.

Acknowledgments

Many extraordinary people influenced my growth through course work and research collaboration at NC State. Since recruiting me to the applied math program five years ago, Dr. H.T. Banks has served as my teacher, research advisor, and mentor. He exposed me to the breadth of mathematics and sciences through classes, personal dialog, and conference travel, and helped me settle on research in math biology. Dr. Banks took an interest in my mathematical and personal pursuits and, at crucial times, reminded me of the importance of balance between the two. I am grateful for his selfless commitment to our work and constant encouragement and support.

The members of my advisory committee have played important roles as well; I appreciate the classes they each taught and time they spent discussing my research. Dr. Hien Tran taught an excellent class in mathematical modeling and shared in lengthy discussions on our HIV model and biological modeling and control in general. Dr. Marie Davidian spent hours meeting and preparing notes to help me understand statistical concerns and approaches. Two classes and a qualifying examination with Dr. Bob Martin convinced me of the beauty of functional analysis. He clarified analysis details in this dissertation and mentored me in the Preparing the Professoriate program.

Numerous others contributed to the HIV modeling project. I initially learned about HIV infection pathways and current related mathematical research with the

help of David Bortz and Stacy Beun. Sarah Holte and Tim Randolph served as mentors during my internship in Seattle; their collaboration and hospitality made working and living there a joy. More recent collaborations have been with Shannon Wynne, Yanyuan Ma, Hee-dae Kwon, Jari Toivanen, Shuhua Hu, and Sarah Grove. I thank Grace Kepler for helpful discussions and candid advice, and Eric Rosenberg for providing clinical data and sharing in discussions on HIV immunology.

This work would not be possible without mathematical foundations and I am grateful to the professors who have influenced me over the years. Professors in the St. Michael's College mathematics department were instrumental in starting me in this field. At NC State, Tim Kelley and Michael Shearer particularly influenced me through their courses. For mathematical, personal, and computer-related discussions, especially early-on in graduate school, I thank fellow students Brian Ball, Kristy Coffey, Todd Coffey, Nathan Gibson, Katie (Kavanagh) Fowler, Rachel Levy, Brian Lewis, Jordan Massad, Jim Nealis, and Bob Wieman. The staff in our department has been wonderfully supportive; Rory Schnell and Brenda Currin in particular helped me through administrative hassles and listened patiently many times.

My friends and family have kept me focused on my work and helped me relax when necessary. Long-term friends from high school and college, especially Adam Kropelin, Susan Hebert, and Casey Reever supported me at various stages of the process. Here in Raleigh I recognize all my (also mathematical) friends, including Jason Osborne, Jill Reese, Jeremy Scott, Brandy Benedict (including for suggestions on this manuscript), and Teresa Selee. The many people with whom I have played music over the years deserve special recognition for helping preserve my sanity. Finally, huge thanks go to my eternally supportive and encouraging parents, Carl and Barbara Adams.

My graduate studies, including this research, have been supported financially

by the NC State University Mathematics Department and Center for Research in Scientific Computation, as well as the U.S. Department of Education GAANN fellowship program. This dissertation work was also supported in part by the Joint DMS/NIGMS Initiative to Support Research in the Area of Mathematical Biology under grant 1R01GM67299-01, and benefited from facilities at the Statistical and Applied Mathematical Sciences Institute, which is funded by NSF under grant DMS-0112069.

Table of Contents

List of Tables

List of Figures

Notation

Model State Variables and Data	
ODE model state vector (unscaled)	$\bar{x}$
ODE model state vector ($\log_{10}$ scaled)	x
state observer matrix $\in \mathbb{R}^{m \times n}$	$\mathcal{O}$
observed unscaled state vector	$\bar{z}$
observed $\log_{10}$ scaled state vector	z
data (unscaled clinical or simulated)	$\bar{y}$
data ($\log_{10}$ scaled clinical or simulated)	y, w
time (days)	t (or t_s^{ij})
parameter vector	q
fixed parameters	$\grave{q}$
differential equation dynamics	$g(t, x; q), h(t, x; q)$
assay censor limits (unscaled)	$\bar{L}1, \bar{L}2$
assay censor limits (log-scaled)	$L, L1, L2$
Dimensions	
size of full ODE model state vector	n
size of observed ODE model state vector	m
number of time points (possibly patient-dependent)	N (or N^j, N_s^j)
number of patients	N_P
number of spline intervals	N_S
dimension of estimated parameter vector	p
Indices	
time point index	i
patient index	j
spline indices	k, l
state vector index	s

Probability and Statistics

admissible parameter space	Q
standard error	ν
variance	σ^2
covariance matrix	Σ
cumulative distribution function	$F(q)$
probability density function	$f(q)$
set of functions	$\mathcal{F}$
a probability distribution	P
probability space (set of distributions)	$\mathcal{P}$
expected value	E
matrix operator for expected value	$\mathcal{E}$
indicator function	χ
likelihood function	$\mathcal{L}$

Other

spline "hat" basis function	$\phi_k(q)$
spline coefficients	d_k
spline knots	q_k
Jacobian matrix	$\mathcal{J}$
cost criterion	J
regularization parameter	β_R
rational numbers	$\mathbb{Q}$
natural numbers $\{1, 2, 3, \ldots\}$	$\mathbb{N}$

- Optimal estimates resulting from parameter estimation procedures will be denoted by $(\hat{\cdot})$ or $(\cdot)^*$.
- $u(t)$ denotes a control input (treatment function), which we assume has a piecewise linear form and so is specified by times when treatment changes.

Chapter 1

Introduction

Human Immunodeficiency Virus (HIV) is a retrovirus that infects T-helper cells of the immune system and is the causative agent for Acquired Immune Deficiency Syndrome (AIDS). HIV and AIDS are among the world's most serious public health concerns, affecting people of all demographics worldwide, with some regions impacted disproportionately. As of 2003, an estimated 38 million HIV-infected individuals are living worldwide, with approximately two-thirds in Africa, where 2.2 million people died from opportunistic infections related to AIDS in 2003 (UNAIDS 2004 Report on the Global HIV/AIDS Epidemic [3]).

Despite many successful public health and clinical interventions since the first identification of HIV-positive patients in 1981, there remains no cure and the HIV/AIDS epidemic continues to grow. In 2003, 4.8 million people became newly infected with HIV, with over half of new cases occurring in youth ages 15–24. This is despite the fact that effective transmission prevention strategies exist. A possible factor for continued spread in industrialized countries is behavior resulting from the myth that antiretroviral drugs, which often successfully suppress virus and improve patient quality of life, constitute a cure for HIV infection. Infection rates continue to rise around the world,

with the fastest expansions of the epidemic occurring in Asia and Eastern Europe. Thus, developing effective methods for prevention of transmission and related public health education campaigns remains crucial.

While antiretroviral drugs are widely available in the United States and Western Europe, their cost and side effects may make their use challenging. In developing nations, UNAIDS estimates that only 7% of the infected population has access to antiretroviral drugs. Access to treatment for and education about this disease remain serious human rights issues around the world. In all geographies, ever-improving strategies are needed for efficient and appropriate use of drug therapy.

The epidemiology of HIV and public health issues like transmission (inter-host dynamics) are important to study. As important to investigate are the effective use and improvement of antiretroviral drugs, which depend on understanding viral behavior within each host, including pathways of infection and effects of drugs. Understanding intra-host viral and immune system pathways depends on knowledge from various biological areas including physiology and immunology. Mathematical models can aid in quantifying dynamic physiologic and immunologic processes and correlating the scientific knowledge of these processes with observed patient behavior.

It is believed that the acute and early phases of HIV infection provide crucial information about immune responses and viral dynamics. In particular, long-term viral set points and speed of progression to AIDS may possibly be understood by studying these key periods. Motivated by clinical study data from patients observed during the crucial acute infection phase and beyond, we develop a combined mathematical and statistical approach to modeling HIV infection in this dissertation. We use both simulated (virtual) data and clinical data to demonstrate the methods and kinds of conclusions one may draw from them.

Several patients for whom we have clinical data underwent therapy interruptions.

Some of these drug holidays were unprescribed or single interruptions, while others were structured treatment interruptions (STIs) according to a study protocol. STIs are being explored as an alternative to continuous therapy with antiretrovirals and in addition to offering the benefit of reduced side effects may also serve to boost HIV-specific immune responses. We therefore incorporate STI protocols in our mathematical models. A good overview of the concept of STI and its applicability in various phases of HIV infection can be found in [35].

In Chapter 2 we overview the clinical acute HIV infection study and the methods used to gather the data analyzed in this dissertation. We describe the characteristics of the data set, including the treatment regimens undergone by various patients, and set goals for understanding it.

Chapter 3 begins with a survey of mathematical models of in-host HIV infection dynamics and illustrates the various disease features and pathways one might wish to model. We then describe the particular system of differential equations used to model HIV infection in our work and discuss its properties and solvability. We present a statistical model to describe the relationship between the differential equation model and the observed data and explain how it will be used to generate simulated data. Finally, we use sensitivity equations to determine which dynamic parameters most influence model solutions and to compute confidence intervals on parameter estimates.

In Chapter 4 we present two contrasting approaches to inverse problems with multiple longitudinal data sets, including one in which the distribution of model parameters across the population is determined directly. Theory for well-posedness of this method is then presented. We explain the statistical methods used to construct confidence intervals from the sensitivity equations and the Expectation Maximization Algorithm used to handle patient data below the limit of measurement detection. We conclude with a discussion of the numerical algorithms and approaches used to solve

the parameter identification problems. The inverse problem for a distribution presented in Chapter 4 can offer substantial computational advantage over estimating parameters individually for each patient in a data set or over more complex hierarchical methods where parameters and distribution parameters are estimated for each patient.

Before applying the mathematical methods to clinical data, we test them on simulated data designed to represent the clinical data sets. In Chapter 5 we present results of these experiments and discuss some of the strengths and weaknesses of the censored data and probability distribution estimation methods.

Following the exploration with simulated data, we examine results for applying the methods to clinical data in Chapter 6. Model fitting results can offer surprising insight into patient behavior – we discuss examples here. To conclude the dissertation, in Chapter 7 we summarize key results and offer pointers to issues requiring further investigation.

Chapter 2

Clinical Data and Desired Outcomes

The data for our study come from over 100 adults with symptomatic acute or early HIV-1 infection. These subjects were enrolled in a study based at Massachusetts General Hospital and associated regional centers and followed for varying lengths of time between 1996 and 2004. The study cohort is unique in that its members were all identified soon after initial infection, making its data particularly useful for understanding early viral dynamics and related immune responses. A principal goal of the clinical study is to assess the potential immunologic consequences of early treatment initiation, including preservation of HIV-specific $CD4^+$ T-cells, extent of latent reservoir development, and homogeneity of viral population. The researchers strive to understand the role of early immune responses in long-term viral suppression.

Clinical and demographic data were collected at the time of study enrollment and blood draw assays of $CD4^+$ T-lymphocyte count and RNA viral load performed at roughly monthly follow-up visits. Viral load was quantified with Reverse Transcriptase-Polymerase Chain Reaction (RT-PCR) methods using the commercially available

HIV-1 Roche Amplicor or Chiron Quantiplex assay, yielding measurements in viral RNA copies per milliliter (ml). The standard assay has a linear range of 400 to 750,000 copies/ml, while the ultra-sensitive assay, 50 to 100,000 copies/ml. The latter is typically employed when a measurement is below the 400 copies/ml limit of the standard assay, as is often the case for a patient successfully suppressing virus. Standard flow cytometry methods were employed to obtain total plasma $CD4^+$ T-lymphocyte counts per microliter (μl) [32].

Nearly all subjects in the study underwent combination therapy with three or more antiretroviral drugs, although the precise regimen varied from patient to patient as dictated by the treating physician. Fourteen of the subjects underwent structured treatment interruptions according to a study protocol, including patients with identification numbers 2, 4, 5, 6, 10, 13, 14, and 16 for whom immune responses were assessed during interruption [48]. Several others simply discontinued drugs at various points. Table 2.1 summarizes the data for all 118 patients in the data set, including the clinical identification number assigned to the patient, number of longitudinal viral load and $CD4^+$ measurements, the total length of time from presentation to last observation, total number of days on and off treatment, and the number of periods (of any length) the patient was off and on therapy. Blank entries indicate patients for whom there were no observations. The number of treatment interruptions varies drastically over the population and some patient records include an initial brief off-treatment phase after presentation, but before therapy commenced. In some cases, the sum of days on and days off exceeds the total days. This is because total days indicates the time from the start of study to last measurement of viral load or CD4 count; for some patients, data indicating treatment status extends beyond this period.

The treatment protocols and overall length of observation for each of the 118 patients are depicted in Figures 2.1 (patients 1–59) and 2.2 (patients 60–118). In

these schematics, thicker lines denote on-treatment periods and the thinner lines, off-treatment. The seventeen patients with no markers had fewer than two measurements and will not be considered in our work. Since we fit a complex dynamic model to these data, we restrict attention to the 59 patients with at least ten viral load and ten CD4 measurements (the patients marked with an asterisk in Table 2.1) and denote this set of patients by PS59.

The distribution of percentage of time spent off treatment by patients in PS59 is shown by histograms in Figure 2.3. The left panel includes frequency for all 59 patients, while the right panel focuses on the 28 patients who spent 10–90% time off treatment. The total number of patients in each range is summarized in Table 2.2. A total of 39 patients spend less than 20% time on drug holiday, with 31 spending less than 10% time on holiday.

Some aspects of the mathematical model later considered are more readily validated in the context of treatment protocols with a balance between time on and time off treatment. Therefore, to validate mathematical methods, we later consider the treatment schedules and observation times of patients spending 30–70% time off treatment. This set of eighteen patients consists of those numbered 2, 4, 5, 6, 9, 10, 13, 14, 15, 16, 23, 24, 26, 27, 33, 46, 47, and 76, and we denote it by PS18. The members of PS18 each have at least fourteen measurements per state and they will serve as model or virtual patients for algorithm testing when we generate simulated data based on their schedules and observation times.

Due to the linear range limits described above, the clinical viral load assays effectively have lower and upper limits of quantification. The upper limit is typically readily handled by repeatedly diluting the sample until the resulting viral load measurement is in range and then scaling. The lower limit, or left censor point, however, directly influences the observed data. When a data point is left-censored (below the

lower limit of detection), the only available knowledge is that the true measurement is between zero and the limit of detection $\bar{L}\star$ for the assay. Those at hand have two limits of detection, $\bar{L}1 = 400$ copies/ml for the standard and $\bar{L}2 = 50$ copies/ml for the ultra-sensitive assay. These are illustrated in sample data shown in Figure 2.4, where censored data points are those appearing identically on the drawn censor lines $\bar{L}1 = 400, \bar{L}2 = 50$. A statistical methodology for handling this type of censored data is described later in Section 4.3.2.

The observation times and intervals vary substantially between patients. The sample data in Figure 2.4 also reveal that observations of viral load and CD4 may not have been made at the same time points, so in general for patient number j we have $CD4^+$ T-cell data pairs $(t_1^{ij}, y_1^{ij}), i = 1, \ldots, N_1^j$ and (potentially different) viral RNA data pairs $(t_2^{ij}, y_2^{ij}), i = 1, \ldots, N_2^j$.

We have several goals for working with this clinical data:

1. Describe the data with a mathematical model of time-varying infection dynamics. Leverage data to calibrate the model by estimating model dynamic parameters. Determine if the model can predict long-term T-cell preservation versus decline.

2. Use the data-calibrated model to extrapolate beyond the observation period to determine consequences of various treatment schemes.

3. Use the calibrated model to correlate differences in model parameters with clinically observed differences among patients. For example, one might ask, "Are there model parameters that can predict rapid versus long-term non-progression to AIDS over the course of HIV infection?"

4. Use the calibrated model to determine novel optimal treatment schemes. A

group of collaborators is considering both open-loop [2] and feedback control [11] in the context of the model proposed in the next section. Ultimately these methodologies should suggest better treatment schemes for potential clinical investigation.

Table 2.1: Summary of patient data, ordered by clinical identification number. Includes number of measurements, duration of observation and time on versus off treatment. Asterisks (*) indicate patients with ten or more viral load and ten or more CD4 measurements. (Data are not available for patients with blank entries.)

pat num	num VL	num CD4	total days	days on/off	periods on/off	pat num	num VL	num CD4	total days	days on/off	periods on/off
1*	102	84	1527	1316/211	4/3	60*	19	18	746	720/26	1/1
2*	107	82	1966	902/1064	2/2	61*	14	14	749	748/1	1/1
3*	76	61	1943	1589/354	3/2	62	8	8	741	721/20	1/1
4*	154	107	1919	1248/671	4/4	63*	16	17	846	714/132	2/2
5*	158	115	2061	1067/994	4/4	64*	23	15	539	534/5	1/1
6*	143	111	1839	923/916	4/5	65*	18	17	755	728/27	1/1
7*	23	22	1932	1924/8	1/1	66*	14	13	552	497/55	3/3
8*	34	33	1672	1668/4	1/1	67	9	3	427	421/6	1/1
9*	32	32	1626	1112/514	2/3	68	6	5	185	174/11	1/1
10*	73	63	1711	582/1129	1/1	69*	14	13	394	398/31	1/1
11	9	8	384	379/5	1/1	70*	19	12	423	363/60	1/2
12*	24	19	1575	1540/35	2/1	71					
13*	64	55	914	537/377	3/3	72	10	7	1213	1159/54	2/2
14*	136	91	1637	659/978	3/3	73	12	1	428	421/7	1/1
15*	46	46	1659	932/727	1/1	74	5	6	440	433/7	1/1
16*	77	57	2228	1337/891	2/2	75*	16	15	549	521/28	3/3
17	11	7	1658	1441/217	1/1	76*	14	14	532	220/315	2/2
18*	32	30	1545	1545/0	1/0	77	7	1	441	422/19	1/1
19*	21	19	1430	1416/14	1/1	78	18	2	418	413/5	1/1
20*	29	27	1581	1469/112	1/2	79					
21*	38	36	1433	1412/21	1/1	80	4	3	78	51/28	1/2
22	8	7	194	179/15	1/1	81*	11	10	425	419/6	1/1
23*	37	36	1505	671/834	4/5	82*	11	11	448	416/32	1/1
24*	36	35	1436	841/595	4/3	83					
25*	83	60	1412	1255/157	4/4	84*	16	15	461	461/0	1/0
26*	100	72	1434	754/680	3/4	85	9	8	363	336/27	1/1
27*	36	35	1379	591/788	2/2	86	4	7	203	0/203	0/1
28	9	8	363	359/4	1/1	87	9	8	1289	1289/0	1/0
29*	34	34	1024	1017/7	1/1	88	8	4	412	270/142	1/2
30*	16	13	841	837/4	1/1	89					
31*	30	30	1256	1228/28	2/2	90	5	4	809	283/652	2/3
32*	33	33	1230	1209/21	1/1	91	5	5	245	0/245	0/1
33*	75	52	1302	658/644	4/4	92					
34*	24	23	1174	1173/1	1/1	93					
35	10	9	484	483/1	1/1	94*	12	11	352	322/30	1/1
36*	33	31	1167	1161/6	1/1	95	4	3	55	40/15	1/1
37*	25	25	1146	1139/7	1/1	96	6	3	332	10/322	1/2
38						97					
39*	29	28	1023	910/113	3/3	98					
40	9	1	328	328/0	1/0	99	3	0	147	0/147	0/1
41*	22	21	717	940/29	2/1	100	7	7	215	215/0	1/0
42*	30	30	1218	1170/48	2/1	101	10	9	273	270/3	1/1
43*	28	29	1134	1060/74	1/1	102	7	7	177	173/4	1/1
44	6	4	994	980/14	1/1	103	8	7	218	203/15	1/1
45*	46	28	499	418/81	2/2	104	6	6	121	121/0	1/0
46*	100	55	1004	496/508	3/3	105	5	4	160	146/14	1/1
47*	23	23	1002	496/506	1/2	106	4	3	157	146/11	1/1
48	5	1	161	154/7	1/1	107	7	7	189	189/0	1/0
49						108					
50*	17	12	141	108/33	1/1	109					
51*	10	10	2043	1519/524	1/2	110					
52*	20	19	708	674/34	1/1	111	1	1	0	40/0	1/0
53						112					
54*	25	25	878	868/10	1/1	113	5	4	94	83/11	1/1
55*	14	14	806	748/58	1/1	114	7	5	122	115/7	1/1
56	10	9	738	701/37	1/1	115					
57						116	6	4	77	63/14	1/1
58*	11	10	671	594/77	1/1	117	3	2	36	21/15	1/1
59	6	6	106	27/79	1/1	118	2	2	151	37/117	2/1

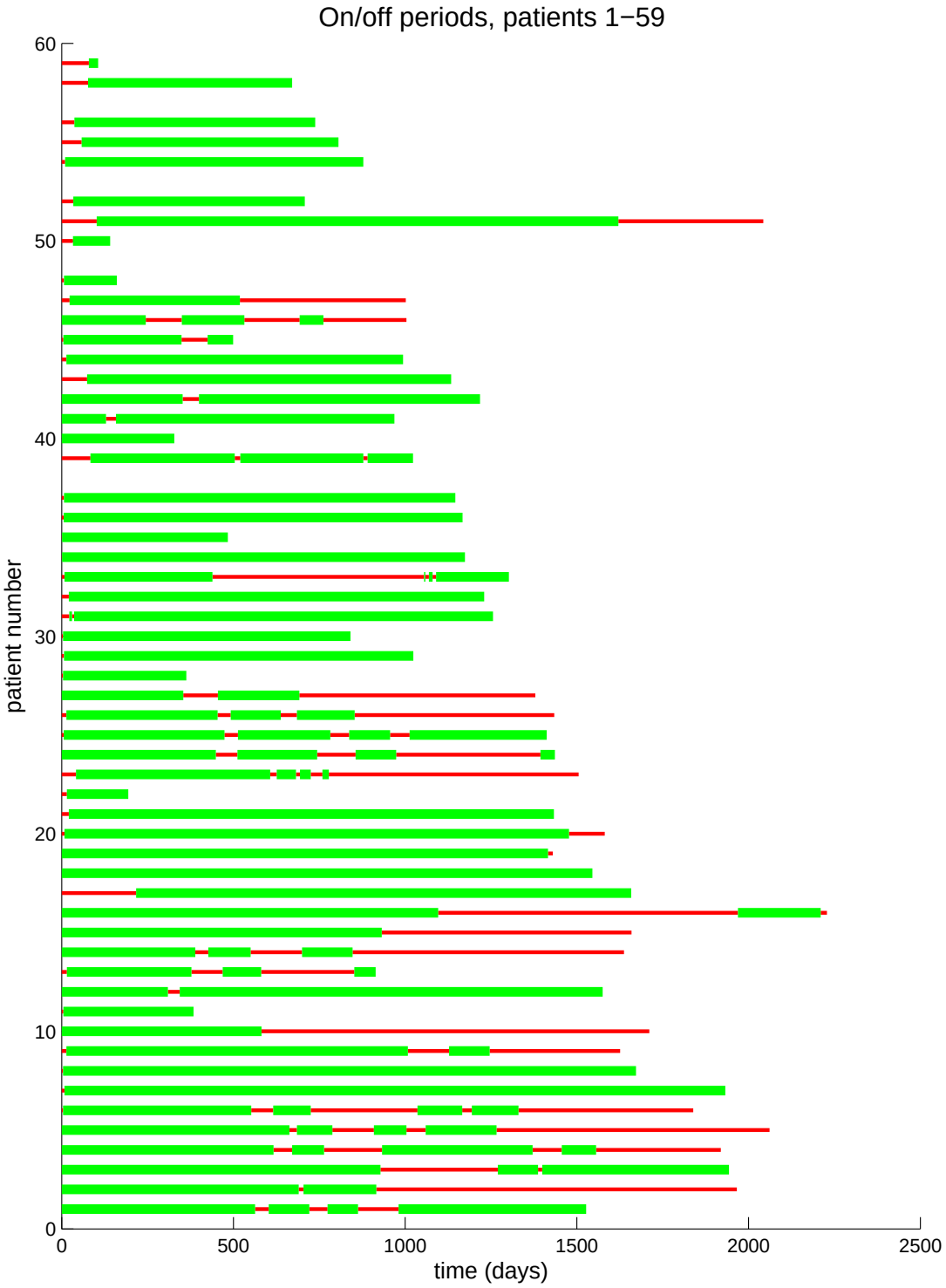


Figure 2.1: Treatment protocols and observation periods for patients 1–59. Thick green lines denote on-treatment periods whereas thin red lines denote off-treatment.

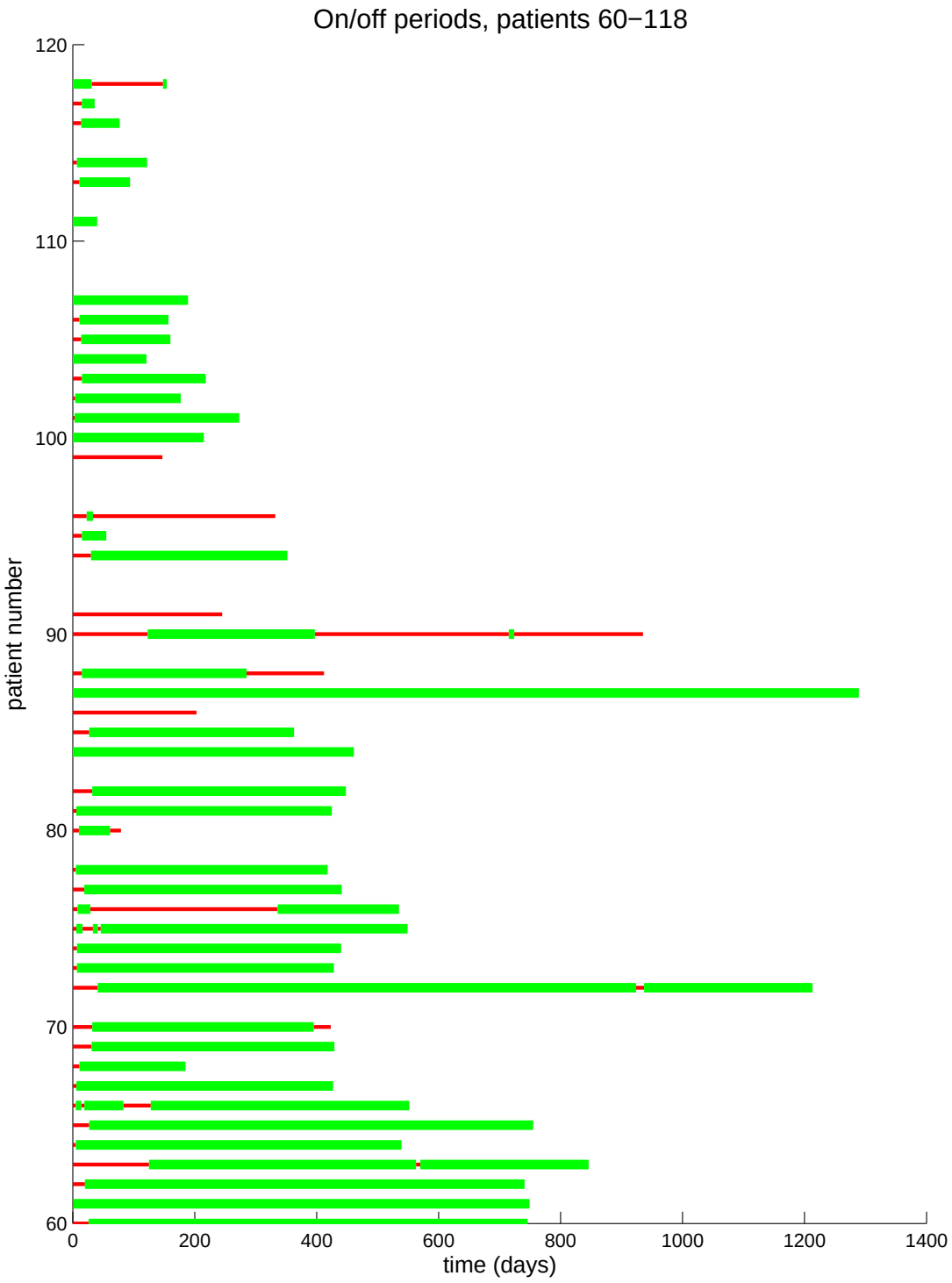


Figure 2.2: Treatment protocols and observation periods for patients 60–118. Thick green lines denote on-treatment periods whereas thin red lines denote off-treatment.

Table 2.2: Number of patients in various ranges of percentage time spent off treatment.

percent time off treatment	number of patients
40–60	14
30–70	18
20–80	20
10–90	28
0–100	59

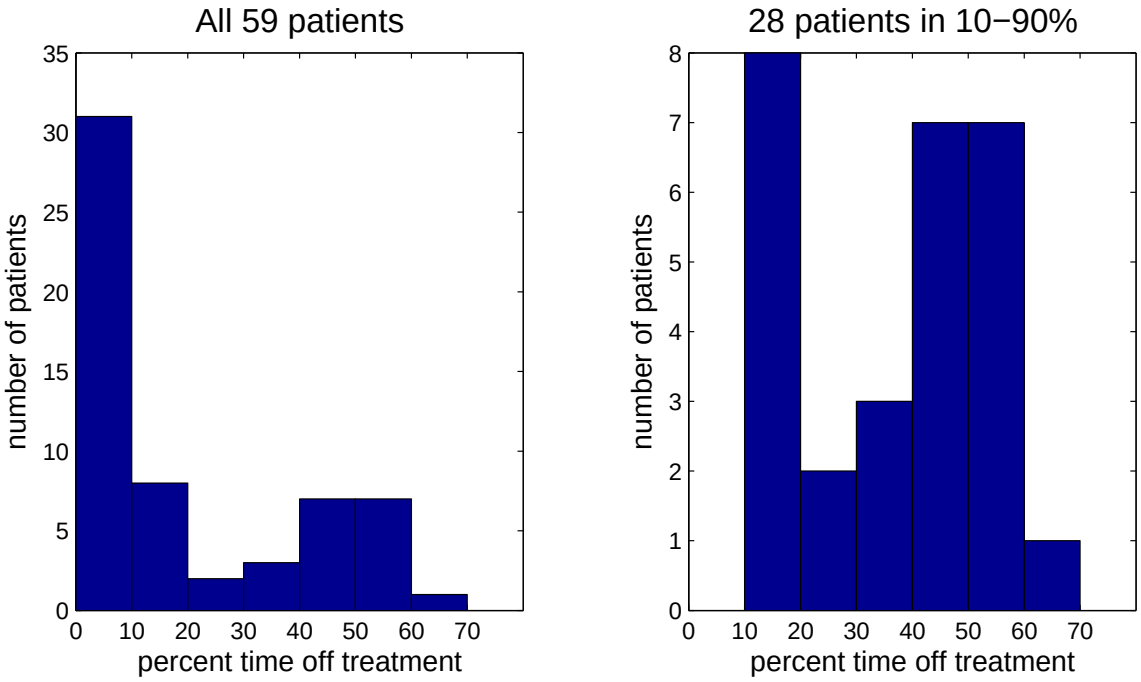


Figure 2.3: Frequency plots of percentage of time off treatment. Left panel includes all 59 patients, whereas right panel only 28 patients spending 10–90% time off treatment.

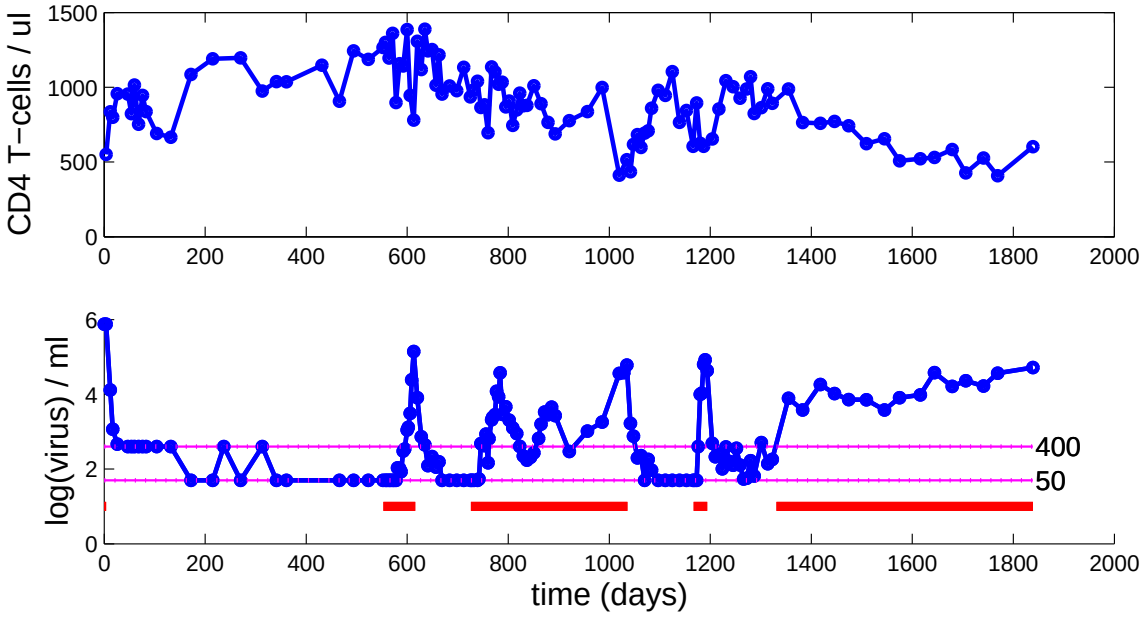


Figure 2.4: Sample patient CD4$^+$ T-cell and viral load data, including censor points (lines at $\bar{L}1 = 400, \bar{L}2 = 50$) for viral load, and periods of therapy interruption (bars below data).

Chapter 3

Mathematical Models for HIV Infection

3.1 Survey of Existing Models

In modeling HIV infection one must typically choose only a critical subset of the many possible biological compartments and interactions. Moreover, scale is important in that one must decide whether to model at the micro level, e.g., of viral RNA, or more at the systemic level. Our focus is on compartmental models in which each compartment corresponds to a type of cell population throughout the body. We do not attempt to provide a comprehensive survey of the extensive collection of mathematical models used for HIV infection. Rather, we refer the reader to one of the excellent survey articles already published (see, for example, [17] and [45]). We provide a brief overview of some important developments here.

Investigations of the kinetics of virus and $CD4^+$ T-cell populations using mathematical models with data from patients undergoing Highly Active Anti-Retroviral Therapy (HAART) support the theory of very rapid and constant turnover of the

viral and infected cell populations in all individuals studied; see, for example, the work of Ho, et al. [30], Wei, et al. [53], and Perelson, et al. [46]. This contrasts with researchers' previous assumptions that the stable viral and $CD4^+$ T-cell concentrations seen during the period of clinical latency of chronic HIV infection were due to the absence of any significant viral replication. The studies by Ho, Wei and Perelson indicate that both the viral and infected cell populations are turning over rapidly and continuously. Further work by Perelson, et al. [44] revealed a second population of longer-lived infected cells contributing to the population of viral RNA. Since these reports, numerous groups have used mathematical models to estimate decay rates for infected cell populations [36, 38, 41, 42, 56]. In Section 3.2 we present a model that can predict the observed persistent low-level replication of virus and includes multiple infected cell populations.

The early linear models developed in [53, 30, 46, 44] are approximations to more realistic nonlinear models for viral and infected cell decay, and thus are applicable only over short periods of time, most likely on the order of days. While these linear models have been extremely useful in characterizing short-term dynamics of HIV infection after therapy, several researchers have attempted to use these models to estimate time to eradication of virus from individuals. Such predictions involve periods of time which extend beyond that which is appropriate for approximation of the nonlinear dynamics by a linear model.

To model data over longer periods of time and make predictions about long-term outcomes, nonlinear mathematical models are necessary. In addition to the unrealistic simplifying assumptions that make it difficult for linear models to accurately describe long-term HIV infection dynamics, factors that could play an important role in dynamic disease outcomes may be omitted in linear models. For example, several authors have raised the question as to whether or not these linear mathematical

models have adequately described the decay of compartments relevant to HIV infection dynamics. The authors of [15] argue that more complex nonlinear models are needed to accurately describe long-term viral decay. In [36] the authors point out that the biphasic pattern, which has been attributed to two populations of infected cells, could be the result of exponential decay of a single population of infected cells with decreasing exponent over time. This phenomenon is well-known in population biology, and is often referred to as density-dependent decay.

Viral production by cells infected with HIV depends on the "age" (e.g., time since infection) of the infected cells, and there are several different biological aspects of this age dependence. Intra-cellular delay due to viral reverse transcription, integration, transcription, and virion formation is described by Mittler [39], extending the work of Perelson, et al., [44]. Mittler allows intra-cellular delay to vary across cells, and estimates these delays to be more significant than the pharmacological delays associated with drug absorption. Recent efforts [7] with *in vitro* data suggest the importance of modeling these distributed delays with some care. Incorporation of this variability of delays into models may lead to improved estimates of the half-life of free virus from short-term clinical data on patients undergoing HAART.

Since the qualitative behavior of a dynamical system is determined by its underlying parameters, knowledge of the bifurcation properties of the system is important for understanding the associated characteristics of the biological system described by the model. If the range of model parameters for a population is such that dramatically different outcomes are predicted for different individuals, bifurcation values for different parameters could suggest target interventions for continued successful treatment. For example, loss of stability of the zero steady state for viral load could be reversed by treatments affecting the parameters which influence this stability. In addition, variability in initial conditions, which one can consider as parameters in the model,

can lead to trajectories of the system lying in different regions of attraction, i.e., different initial population sizes can lead to dramatically different qualitative outcomes in a nonlinear model. Such situations are described in [49, 54] where the authors model structured treatment interruption (STI). The models in these reports and the model discussed in Section 3.2 all have multiple equilibria. Different equilibria describe the success or failure of the immune system to control infection, and the initial conditions and parameters of the system determine which equilibrium is realized. Careful qualitative analysis of mathematical models that describe HIV infection dynamics can contribute to understanding of fundamental qualitative features of infection, and possibly suggest targets for improved disease monitoring and/or treatment.

In using any of these complex models, one would like to understand the identifiability of parameters, i.e., which parameters can be successfully estimated and what type of data is necessary to do so. In Chapter 4 we examine estimating parameters in our example nonlinear model of HIV infection dynamics. In doing so, we describe two inverse problem formulations and corresponding computational approaches, together with a sensitivity analysis method which leads to estimation of standard errors on parameter estimates arising from the inverse problems process.

3.2 Proposed Model: Features and Analysis

The breadth of models discussed in Section 3.1 makes clear the assortment of HIV infection mechanisms and features that may be important to model. In order to fit the clinical data at hand and make realistic predictions, a few key features are essential in a dynamic model of HIV progression in an individual. The model proposed here is therefore a "typical" model of HIV infection that includes those certain features and will serve to demonstrate the mathematical and statistical methodologies presented

in the remainder of this dissertation. Given our goal of method validation and the existence of survey papers noted earlier, we do not include here discussions of all the important aspects of HIV modeling. Further, we do not purport to have a good model of HIV progression and treatment; rather we propose a model that behaves reasonably and contains some of the more desirable features one could expect.

We consider a compartmental ordinary differential equation (ODE) model for in-host HIV infection dynamics. A wide variety of such ODE models have been proposed to describe various aspects of the dynamics. The most basic of these models typically include two or three of the key dynamic compartments: virus, uninfected target cells, and infected cells. The proposed model includes all three of these physiologically-relevant compartments since they are direct players in the infection process. However infected and uninfected $CD4^+$ T-cells in patients are not typically measured separately, so these compartments must be aggregated for purposes of model fitting. In addition, the documented importance of the immune system in responding to HIV infection (and especially its apparent crucial role during structured treatment interruptions) strongly motivates the inclusion of at least one model compartment representing immune response to the pathogen. We therefore propose a model that includes some measure of cytotoxic T-lymphocyte (CTL) $CD8^+$ response to HIV infection. While the presently available data do not directly quantify the presence of HIV-specific CTLs, these immune responders are important for control of infected cells and may eventually be correlated to available epitope-challenge data.

Since the model will be considered in the context of time-variable treatment strategies, it should be capable of incorporating the action of commonly-used reverse transcriptase inhibitors (RTIs) and protease inhibitors (PIs). Inclusion of the latter usually implies inclusion of separate compartments for infectious free virus and virus rendered non-infectious by the PI. Given the predominance of HAART in the form

of drug cocktails combining two or more drugs, the model should behave reasonably when simulating such multi-drug therapy.

As Callaway and Perelson [17] point out in their 2002 review paper, a reasonable model of HIV infection predicts a non-zero steady-state viral load even in the presence of effective drug therapy. Patients subjected to drug therapy often successfully suppress virus for a long time, potentially at undetectable levels. However, some reservoir or mechanism exists which almost invariably causes the virus to grow out to detectable levels upon removal of drug therapy. Hence one does not expect incorporation of drug therapy in the model, at a sensible efficacy, to drive the viral load to zero, but rather reduce it considerably, perhaps below the assay limits of detection. The authors of [17] analyze models typically employed to describe HIV infection and show that many do not exhibit a reasonable relationship between drug efficacy and predicted viral load. In such models, a very slight change in drug efficacy can yield a drastic change in the predicted viral load set point. This has important consequences for the control problem of determining treatment strategies: a successful model must exhibit reasonable sensitivity of the viral load equilibrium to treatment efficacy.

Kepler and Perelson [34] offer one potential quantitative explanation for maintenance of a low steady-state viral load. They propose a drug sanctuary compartmental model which includes physiologically distinct compartments for blood plasma and other tissue (e.g., brain tissue or lymph nodes) where drug effectiveness is reduced. This reflects the belief that HIV may replicate in body reservoirs such as the central nervous system where replication rates and drug penetration are different from plasma. Since clinical observations most commonly only include data from plasma, this can explain the undetectable viral load in many patients. In the present work, we consider a similar model, adapted from equation (5.3) in Callaway and Perelson. As shown in Figure 3.1, it includes two co-circulating populations of target

cells, potentially representing CD4$^+$ T-lymphocytes (T_1) and macrophages or other HIV-targeted cells (T_2). The two cell populations may have different activation requirements or susceptibility to drug therapy. A summary of model compartments and notation is provided in Table 3.1.

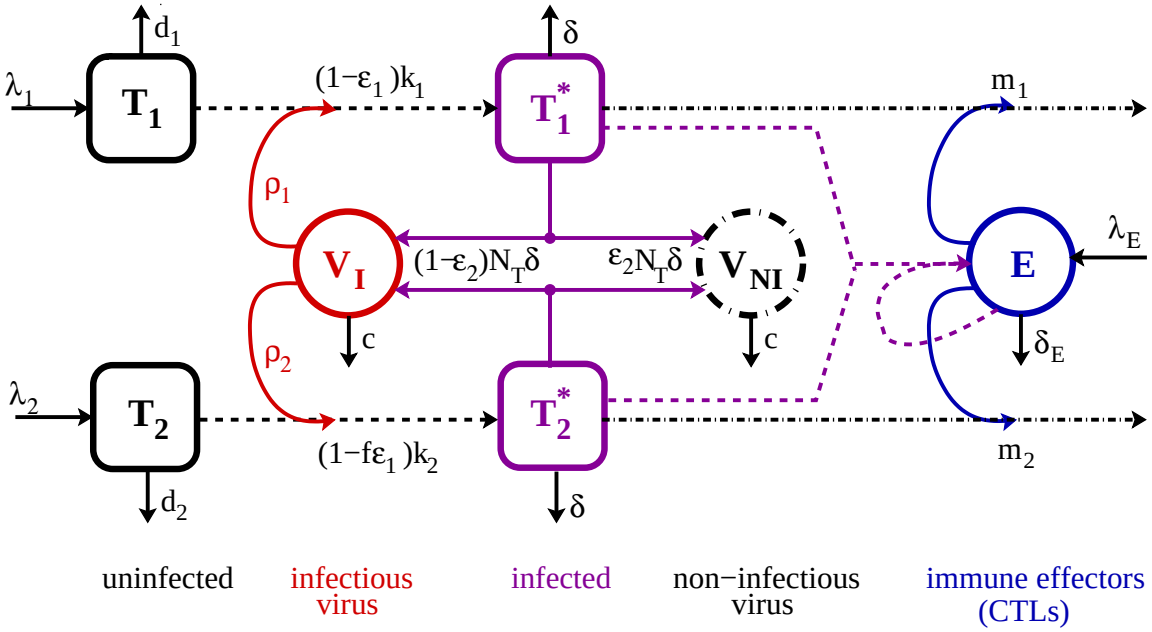


Figure 3.1: Schematic of compartmental HIV infection dynamics model. Only key pathways are indicated in the schematic – for further details, see the system of differential equations (3.1) below.

Table 3.1: State variables used in HIV model

variable	description	units
T_1	uninfected type 1 target cells (e.g. CD4$^+$ T-cells)	cells / μl
T_2	uninfected type 2 target cells (e.g. macrophages)	cells / μl
T_1^*	infected type 1 cells	cells / μl
T_2^*	infected type 2 cells	cells / μl
V_I	infectious free virus	RNA copies / ml
V_{NI}	non-infectious free virus	RNA copies / ml
E	cytotoxic T-lymphocytes	cells / μl

The corresponding system of differential equations is principally based on the Callaway–Perelson model, but includes an immune response compartment and dy-

namics as suggested by Bonhoeffer, et. al. [49]. This compartment, denoted by E, represents CTLs which lyse and thus kill infected cells. The adapted system of ODEs is given by

$$\dot{T}_1 = \lambda_1 - d_1 T_1 - (1 - \bar{\epsilon}_1(t))\, k_1 V_I T_1 \tag{3.1a}$$

$$\dot{T}_2 = \lambda_2 - d_2 T_2 - (1 - f\bar{\epsilon}_1(t)) k_2 V_I T_2 \tag{3.1b}$$

$$\dot{T}_1^* = (1 - \bar{\epsilon}_1(t)) k_1 V_I T_1 - \delta T_1^* - m_1 E T_1^* \tag{3.1c}$$

$$\dot{T}_2^* = (1 - f\bar{\epsilon}_1(t)) k_2 V_I T_2 - \delta T_2^* - m_2 E T_2^* \tag{3.1d}$$

$$\dot{V}_I = (1 - \bar{\epsilon}_2(t)) 10^3 N_T \delta (T_1^* + T_2^*) - c V_I \tag{3.1e}$$

$$- (1 - \bar{\epsilon}_1(t)) \rho_1 10^3 k_1 T_1 V_I - (1 - f\bar{\epsilon}_1(t)) \rho_2 10^3 k_2 T_2 V_I$$

$$\dot{V}_{NI} = \bar{\epsilon}_2(t) 10^3 N_T \delta (T_1^* + T_2^*) - c V_{NI} \tag{3.1f}$$

$$\dot{E} = \lambda_E + \frac{b_E (T_1^* + T_2^*)}{(T_1^* + T_2^*) + K_b} E - \frac{d_E (T_1^* + T_2^*)}{(T_1^* + T_2^*) + K_d} E - \delta_E E, \tag{3.1g}$$

together with an initial condition vector

$$[T_1(0), T_1^*(0), T_2(0), T_2^*(0), V_I(0), V_{NI}(0), E(0)]^T .$$

Here the factors 10^3 are introduced to convert between microliter and milliliter scales, preserving the units from some of the published papers. In this dynamical system, the treatment factors $\bar{\epsilon}_1(t) = \epsilon_1 u(t)$ and $\bar{\epsilon}_2(t) = \epsilon_2 u(t)$ represent the effective treatment impact, consisting of efficacy factors ϵ_1, ϵ_2 and a time-dependent treatment function $0 \leq u(t) \leq 1$ representing HAART drug level, where $u(t) = 0$ is fully off and $u(t) = 1$, fully on. See Figure 3.2 for a sample time-varying treatment protocol representing structured therapy interruption. The relative effectiveness of RTIs is modeled by ϵ_1 and that of PIs by ϵ_2. Since HIV treatment is nearly always administered as

combination therapy, we do not consider the possibility of monotherapy, even for a limited period of time, though this could be implemented by considering separate treatment functions $u_1(t), u_2(t)$.

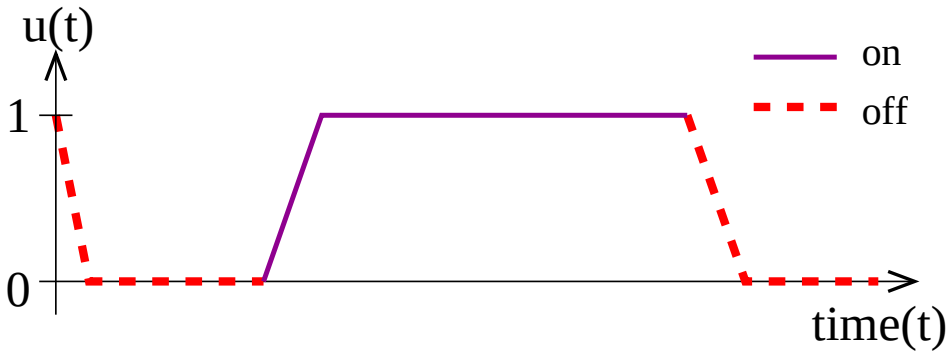


Figure 3.2: Sample control input (treatment protocol) $u(t)$ representing structured treatment interruption. This is a schematic in that interruption periods need not be periodic and one might assume more smooth ramp functions for the absorption and dissipation of the drug.

As is common in models of HIV infection, infected cells T_i^* result from encounters between uninfected target cells T_i and infectious free virus V_I in a well-mixed environment. As noted above, this model involves two co-circulating populations of target cells, perhaps representing $CD4^+$ T-lymphocytes (T_1) and macrophages (T_2). The natural infection rate k_i may differ between the two populations, which could account for believed differences in activation rates between lymphocytes and macrophages. The drug efficacy parameter ϵ_1 represents an RTI that blocks new infections and is potentially more effective in population 1 (T_1, T_1^*) than in population 2 (T_2, T_2^*), where the efficacy is $f\epsilon_1$, with $f \in [0, 1]$. The differences in infection rates and treatment efficacy help create a low, but non-zero, infected cell steady state for T_2^*, which is commensurate with the idea that macrophages may be an important source of virus after T-cell depletion. The populations of uninfected target cells T_1 and T_2 may have different source rates λ_i and natural death rates d_i.

For our efforts here we assume that both target cell types have the same death rate δ, though this could be readily generalized as well. Infected cells T_1^*, T_2^* may be removed from the system via either natural death or by the action of immune effector cells E described below.

To preserve simplicity in the model, we omit the chronically infected cell compartments proposed in the original Callaway–Perelson model. The important qualitative behaviors seem preserved in the model we propose and specifically modeling this feature is not essential to our present work. In particular, the existence of a low steady-state viral load equilibrium and sensitivity of the viral load equilibrium to the drug efficacy is obtained with or without such compartments. We note that while removing the chronically infected compartments does not affect the sensitivity to treatment, the addition of immune response terms does, as discussed below.

Free virus particles are produced by both types of infected cells, which we assume produce virus at the same rate (again this could be readily generalized to account for different productivity). In the Callaway–Perelson model, virus only leaves the V_I compartment via natural death at rate c; there is no removal term in the $\dot{V}_I$ equation representing loss of virus due to infection of a cell. One potential justification for this omission is offered by Nelson and Perelson [45] (page 10) who suggest that this term can be omitted since the term $k_i T_i V_I$ is small in comparison to cV in the typical HIV-infected patient. They further assert that if T_i is approximately constant, one can absorb the loss of virus due to infection into the cV_I term, thus making it account for all clearance processes.

While the arguments offered by Nelson and Perelson could justify the exclusion of the virus removal term, we investigate situations where treatment is interrupted abruptly, potentially effecting drastic changes in all of the cell populations under consideration. We therefore include the term $\left[(1-\epsilon)\rho_1 k_1 T_1 + (1-f\epsilon)\rho_2 k_2 T_2\right] V_I$ in

the $\dot{V}_I$ equation to account for the removal of free virus that takes place when free virions infect a T_1 or T_2 cell. We make the simplifying assumption $\rho_i = 1$, i.e., one free virus particle is responsible for each new infection. This could easily be adapted for multiple virus particles being responsible for each new infection by choosing $\rho_i > 1$.

The action of a PI, which causes infected cells to produce non-infectious virus V_{NI} is modeled by $\bar{\epsilon}_2$. Tracking non-infectious virus is important since the clinically-measured viral load data for patients includes total free virus (infectious V_I and non-infectious V_{NI}). Model fits to data are therefore to the sum $V_I + V_{NI}$. However, see the discussion below regarding the decoupling of this compartment from the remainder of the ODE system.

Finally, the immune effectors E (CTLs), are produced in response to the presence of infected cells and existing immune effectors. The immune response assumed here is similar to that suggested by Bonhoeffer, et al., in their 2000 paper [49], with a Michaelis-Menten type saturation nonlinearity. The infected cell-dependent death term in the immune response represents immune system impairment "at high virus load". In [49] the authors demonstrate that a model with this immune reponse structure and a latently infected cell compartment can exhibit transfer between "healthy" and "unhealthy" stable steady states via STI, making it a good candidate for our investigation. (Indeed the same is true for the model (3.1) considered here.) We add a source term λ_E to create a non-zero off-treatment steady state for E, rather than explicitly modeling immune memory. While immune effectors are not inherently present in the absence of pathogen, they persist at low levels during infection. We note that other immune responses models, such as those considered by Wodarz-Nowak [54] or Nowak-Bangham [43] could be substituted if desired. However, the latter does not appear to admit multiple stable off-treatment steady states.

The immune response we model is that of cytotoxic T-lymphocytes. CTL act by

lysing infected cells, causing them to explode. Thus they remove infected cells from the system in the equations for $\dot{T}_1^*$ and $\dot{T}_2^*$, at rates m_1 and m_2, respectively. Unlike interferons, they do not directly target free virus, so there is no interaction term with the virus compartment. As with any immune system responders, we suspect that CTL sometimes mistarget or misidentify receptors and thus kill healthy cells or misidentify self versus antigen, but for simplicity, we do not model that here.

While the model (3.1) explicitly includes a V_{NI} compartment

$$\dot{V}_{NI} = \bar{\epsilon}_2(t)10^3 N_T \delta(T_1^* + T_2^*) - cV_{NI},$$

it serves only as a collection compartment and does not couple with any of the other model dynamics. Explicitly including this compartment and its dynamics complicates the linear analysis of the ODE system's stability since it introduces a zero eigenvalue. It can be explicitly solved using variation of parameters to obtain the necessary quantity V_{NI} to use in model fitting as follows.

Defining $G(t) = \bar{\epsilon}_2 10^3 N_T \delta\left(T_1^*(t) + T_2^*(t)\right)$ (so G(t) depends on the solution to the remainder of the ODE system (3.1)), the solution to (3.1f) is given by

$$V_{NI}(t) = V_{NI}(0)e^{-ct} + \int_0^t e^{-c(t-s)} G(s) ds. \tag{3.2}$$

Given a solution to the remaining equations in the ODE system, one can compute $V_{NI}(t)$ for purposes of model fitting to the clinically observed quantity $V = V_I + V_{NI}$. For analysis, we therefore consider the V_{NI} compartment as dependent on the other dynamics.

Throughout the remainder of this dissertation, $\bar{x}$ will denote the vector of solutions

to the ODE system (3.1):

$$\bar{x}(t) = \begin{bmatrix} T_1(t) & T_1^*(t) & T_2(t) & T_2^*(t) & V_I(t) & V_{NI}(t) & E(t) \end{bmatrix}^T, \tag{3.3}$$

where components 1–4 of $\bar{x}$ are on a cells/μl scale, 5 and 6 (corresponding to V_I and V_{NI}) on a copies/ml scale, and 7 on a cells/μl scale. The differential equation model (3.1) can therefore be summarized by

$$\frac{d\,\bar{x}}{d\,t} = \bar{g}(t, \bar{x}; q),$$

with q denoting model dynamic parameters and $\bar{g}$ the vector of derivatives. Model fits will be to the base-10 logarithm of these quantities ($x = \log_{10} \bar{x}$) and in general, as explained in the notation section, variables with an overbar will denote an unscaled quantity and those without, $\log_{10}$-transformed variables.

3.2.1 Sample model parameters, steady states, and treatment efficacy

The model (3.1) contains numerous parameters that must be assigned values before simulations can be carried out. In specifying model parameters, to the greatest extent possible we employ values similar to those reported or justified in the literature. The parameters indicated in Table 3.2 are principally extracted from the Callaway–Perelson [17] and Bonhoeffer, et al., [49] papers.

Callaway and Perelson point out that several model parameters are not available from human or animal data. They choose the parameters $\lambda_1, k_1, \lambda_2,$ and k_2 such that several conditions on target cell and viral load equilibria are satisfied for their model. These conditions are not precisely satisfied by our model which has no chronic cell

Table 3.2: Parameters used in model (3.1). Those in the top section of the table are taken directly from Callaway and Perelson [17]. Parameters in the bottom section of the table are taken from Bonhoeffer, et. al. [49], with slight adjustments. The superscripts * denote parameters the authors indicated were estimated from human data and ** denote those estimated from macaque data.

parameter	**value**	**units**	**description**
λ_1	10	$\frac{cells}{ul \cdot day}$	target cell type 1 production (source) rate
d_1	0.01^{**}	$\frac{1}{day}$	target cell type 1 death rate
ϵ_1	$\in [0,1]$	–	RTI treatment efficacy
k_1	8.0×10^{-7}	$\frac{ml}{virions \cdot day}$	population 1 infection rate
λ_2	0.03198	$\frac{cells}{ul \cdot day}$	target cell type 2 production (source) rate
d_2	0.01^{**}	$\frac{1}{day}$	target cell type 2 death rate
f	0.34 ($\in [0,1]$)	–	treatment efficacy reduction in population 2
k_2	1×10^{-4}	$\frac{ml}{virions \cdot day}$	population 2 infection rate
δ	0.7^*	$\frac{1}{day}$	infected cell death rate
m_1	0.01	$\frac{ul}{cells \cdot day}$	immune-induced clearance rate for pop. 1
m_2	0.01	$\frac{ul}{cells \cdot day}$	immune-induced clearance rate for pop. 2
ϵ_2	$\in [0,1]$	–	PI treatment efficacy
N_T	100^*	$\frac{virions}{cell}$	virions produced per infected cell
c	13^*	$\frac{1}{day}$	virus natural death rate
ρ_1	1	$\frac{virions}{cell}$	average number virions infecting a type 1 cell
ρ_2	1	$\frac{virions}{cell}$	average number virions infecting a type 2 cell
λ_E	0.001	$\frac{cells}{ul \cdot day}$	immune effector production (source) rate
b_E	0.3	$\frac{1}{day}$	maximum birth rate for immune effectors
K_b	0.1	$\frac{cells}{ul}$	saturation constant for immune effector birth
d_E	0.25	$\frac{1}{day}$	maximum death rate for immune effectors
K_d	0.5	$\frac{cells}{ul}$	saturation constant for immune effector death
δ_E	0.1^*	$\frac{1}{day}$	natural death rate for immune effectors

compartment and an added immune response. However, the conditions are closely approximated by the model's behavior (partially due to multiple stable equilibria) and we believe the parameters could be adjusted to obtain the same qualitative behavior.

In general, immune response parameters are not well-known and are thus frequently chosen to demonstrate model behavior in simulations. The parameters m_1 and m_2 represent cytopathicity of the immune effectors. Their common value $m = m_1 = m_2 = 0.01$ was taken from [17] where the authors note that the value was

suggested originally by Nowak and Bangham [43]. The parameters in the $\dot{E}$ equation (3.1g) are also slightly adjusted from published values to demonstrate the possibility of multiple stable equilibria for the model.

In order to understand the possible behaviors of the model under treatment interruptions, we examine the possible equilibria to which the model dynamics might converge. These do not depend on the dynamics for the non-infectious virus, so we define the reduced vector $x_R = [\bar{x}_1, \ldots, \bar{x}_5, \bar{x}_7]^T$ and $g_R(t, x_R; q)$, the corresponding model dynamics (derivatives). Model steady states (equilibria) result from solving the system of six algebraic equations $g_R(t, x_R; q) = 0$ for the steady state values $\tilde{x}_R$, i.e., $\tilde{x}_R$ are the state values where derivatives are all zero. By setting $\dot{V}_{NI} = 0$ in (3.1f) and substituting values of T_1^*, T_2^*, and V_I from $\tilde{x}_R$, the corresponding V_{NI} equilibrium follows.

To assess the stability of the equilibria, calculated steady state values $\tilde{x}_R$ may then be substituted for x_R in the Jacobian matrix

$$\mathcal{J}(x_R; q) = \frac{\partial\, g_R(t, x_R; q)}{\partial\, x_R} =$$

$$\begin{pmatrix}
-d_1 - \bar{k}_1 V_I & 0 & 0 & 0 & -\bar{k}_1 T_1 & 0 \\
0 & -d_2 - \bar{k}_2 V_I & 0 & 0 & -\bar{k}_2 T_2 & 0 \\
\bar{k}_1 V_I & 0 & -\delta - m_1 E & 0 & \bar{k}_1 T_1 & -m_1 T_1^* \\
0 & \bar{k}_2 V_I & 0 & -\delta - m_2 E & \bar{k}_2 T_2 & -m_2 T_2^* \\
-\bar{k}_1 10^3 V_I & -\bar{k}_2 10^3 V_I & (1-\bar{\epsilon}_2) 10^3 N_T \delta & (1-\bar{\epsilon}_2) 10^3 N_T \delta & \mathcal{J}_{5,5} & 0 \\
0 & 0 & \mathcal{J}_{6,3} & \mathcal{J}_{6,4} & 0 & \mathcal{J}_{6,6}
\end{pmatrix},$$

where

$$\begin{aligned}
\bar{k}_1 &= (1-\bar{\epsilon}_1)k_1, \\
\bar{k}_2 &= (1-f\bar{\epsilon}_1)k_2, \\
\mathcal{J}_{5,5} &= -c-(1-\bar{\epsilon}_1)10^3 k_1 T_1 - (1-f\bar{\epsilon}_1)10^3 k_2 T_2, \\
\mathcal{J}_{6,3} = \mathcal{J}_{6,4} &= \frac{b_E K_b E}{\left(T_1^* + T_2^* + K_b\right)^2} - \frac{d_E K_d E}{\left(T_1^* + T_2^* + K_d\right)^2}, \text{ and} \\
\mathcal{J}_{6,6} &= \frac{b_E\left(T_1^* + T_2^*\right)}{\left(T_1^* + T_2^* + K_b\right)} - \frac{d_E\left(T_1^* + T_2^*\right)}{\left(T_1^* + T_2^* + K_d\right)} - \delta_E.
\end{aligned}$$

As is well-known from linearization theory, the eigenvalues of $\mathcal{J}(\tilde{x}_R; q)$ yield information about the local stability of the substituted equilibrium $\tilde{x}_R$. In the local stability results below, stable means locally asymptotically stable, or all eigenvalues having strictly negative real part. While it is possible to analytically determine the steady states in closed form, the expressions are unwieldy and not terribly informative. We therefore computationally explore the model steady states and stability for the set of parameters q indicated in Table 3.2.

Given the specified parameters, in the absence of therapy the model exhibits several steady states (calculated with Maplesoft's Maple 8 and) shown in Table 3.3. Equilibrium $EQ1$ represents the uninfected patient, with healthy T-cell counts and no virus or infected cells present. (While a healthy person's T-cell count varies naturally over time, a typical measurement would be 1000 cells/μl.) When the system is in this state, introduction of a small amount of virus causes the system to converge to $EQ2$, where healthy target cells are substantially depleted and a dangerously high viral load is present. The system also exhibits an additional stable equilibrium $EQ3$, where a strong immune response has developed, successfully controlling the virus and consequently restoring CD4$^+$ target cell (T_1) help. The immune response

variable incorporated here does not directly correspond to quantities easily measured in the lab, since HIV-specific immune responses are quantified for specific proteins or epitopes and here we have a more general summative response. While calculations

Table 3.3: Off treatment ($\epsilon = 0$) steady states (copies/ml) for model (3.1). Non-physical steady states have been omitted.

	Off Treatment ($\epsilon_1 = \epsilon_2 = 0$)			$\epsilon_1 = 1$	
	$EQ1$	$EQ2$	$EQ3$	f=0	f=0.5
T_1	1000	163.57	967.84	1000	1000
T_2	3.198	0.005	0.6205	1.3135	2.6267
T_1^*	0	11.945	0.0760	0	0
T_2^*	0	0.0456	0.0061	0.0269	0.0082
V_I	0	63919	415.38	143.45	43.497
V_{NI}	0	0	0	0	0
E	0.01	0.0235	353.11	0.0203	0.0123
local stab.	unstable	stable	stable	stable	stable

reveal the model may exhibit as many as nine steady states, several are non-physical (negative or complex equilibrium values), and the remaining unreported physical equilibrium is never locally stable. Other models with multiple stable steady states have been proposed for the STI scenario including one by Wodarz–Nowak [54].

Considering the model under RTI monotherapy ($\epsilon_2 = 0$), we turn to the criteria on the steady states suggested by Callaway and Perelson. When $\epsilon_1 = 1$ and $f = 0$, they desire a viral load of 100. In this scenario, our model exhibits a viral load of $V = 143$ copies/ml, which is on that order. The entire high virus equilibrium is shown in Table 3.3. When $\epsilon_1 = 1$ and $f = 0.5$, they desire eradication of the virus. Our model does not precisely attain that, though the viral steady state ($V = 43$ copies/ml) is below the limit of detection for most assays currently in use. (Again, see the table for the entire equilibrium.) In both these cases, the reported equilibrium is the only stable physical equilibrium. The persistence of low levels of virus, even while under treatment, is a key feature of HIV models, since rapid viral rebound is

typically observed upon the withdrawal of drugs. Also, by increasing the secondary compartment treatment efficacy factor to $f \geq 0.59$, the virus is eradicated because the uninfected steady state becomes stable.

The topmost plot in Figure 3.3 ($\epsilon_2 = 0$) depicts the sensitivity of the viral load equilibrium $V = V_I + V_{NI}$ to drug efficacy ϵ_1. The introduction of the immune response terms causes a discontinuity in this curve as stability is exchanged between viral dominant and immune-dominant equilibria. This is in contrast to the sensitivity curve of the original Callaway–Perelson model which is continuous across the range of drug efficacies. However, this may be reasonable: a certain amount of drug may be necessary to augment the immune system to counteract effects of the virus. The low (nonzero) steady-state viral load is maintained across the entire range of reasonable drug efficacies. The model also exhibits approximately a 1–2 log drop in viral load across the range of drug efficacies, which is typical for monotherapy with RTIs.

The remaining plots in Figure 3.3 demonstrate how the viral equilibria vary for other fixed values of ϵ_2 (corresponding to various levels of protease inhibition). If $\epsilon_2 = 1$ the system has entirely different dynamics as the infectious virus is eradicated; this case is not explored. In fact as shown in Figure 3.6, for a region of ϵ_2, including $\epsilon_2 > 0.88$, the uninfected $EQ1$ is solely stable, so this is not a reasonable range of ϵ_2 to simulate persistent low-level infection.

Figure 3.4 includes plots of the sensitivity of the stable viral load equilibria to the efficacy of protease inhibitors for various fixed RTI efficacies. In Figure 3.5 we consider the sensitivity of the viral load equilibrium as the effectiveness of both drugs vary in relationship. The plots contain the following, in order: $\epsilon_2 = \frac{1}{2}\epsilon_1$, $\epsilon_2 = \epsilon_1$, and $\epsilon_2 = 2\epsilon_1$. Considering the same parameters as above, in Figure 3.6 we summarize regions in treatment parameter space where the various equilibria are stable. For large enough values of a combination of ϵ_1 and ϵ_2 only the uninfected equilibrium is stable.

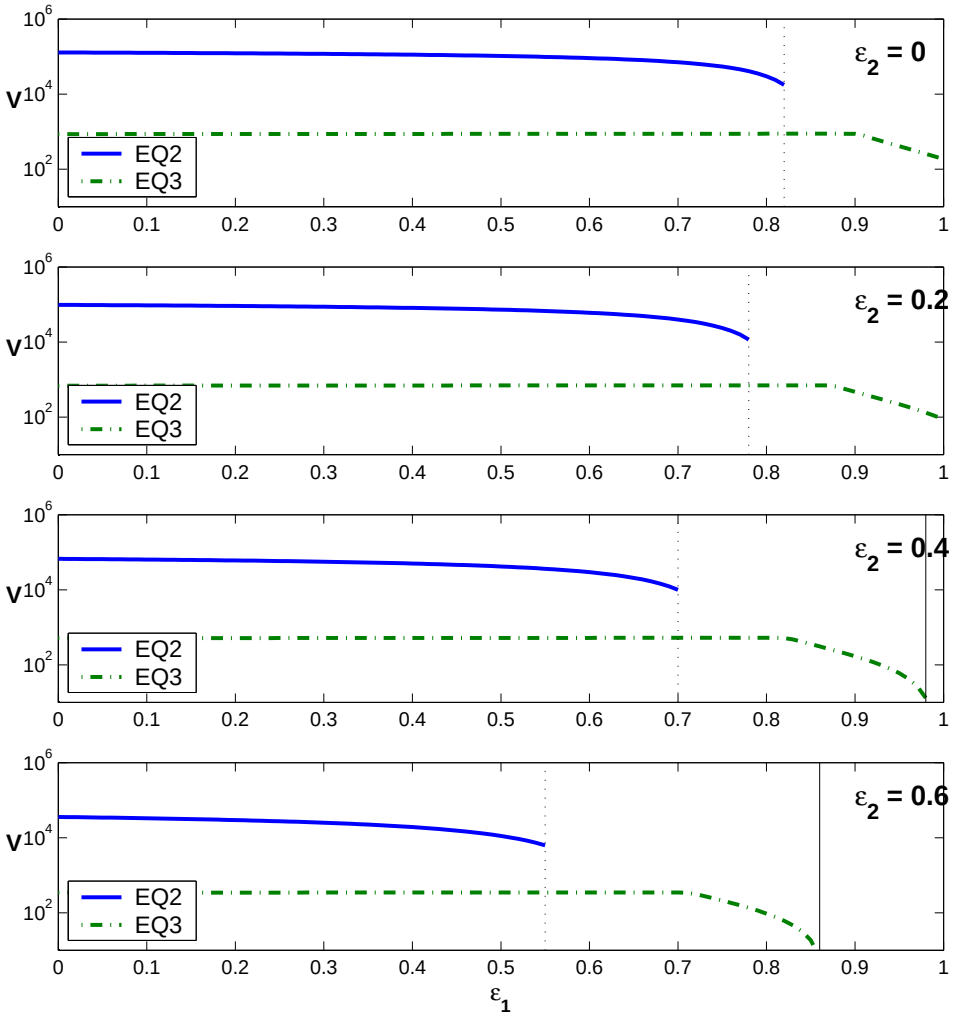


Figure 3.3: Sensitivity of stable viral load equilibria (total virus $V = V_I + V_{NI}$) to drug efficacy ϵ_1 for various fixed values of ϵ_2. The dashed vertical line separates the region of bi-stability from that with a unique equilibrium and the solid line separates the unique equilibrium region from the region where only the uninfected equilibrium is stable.

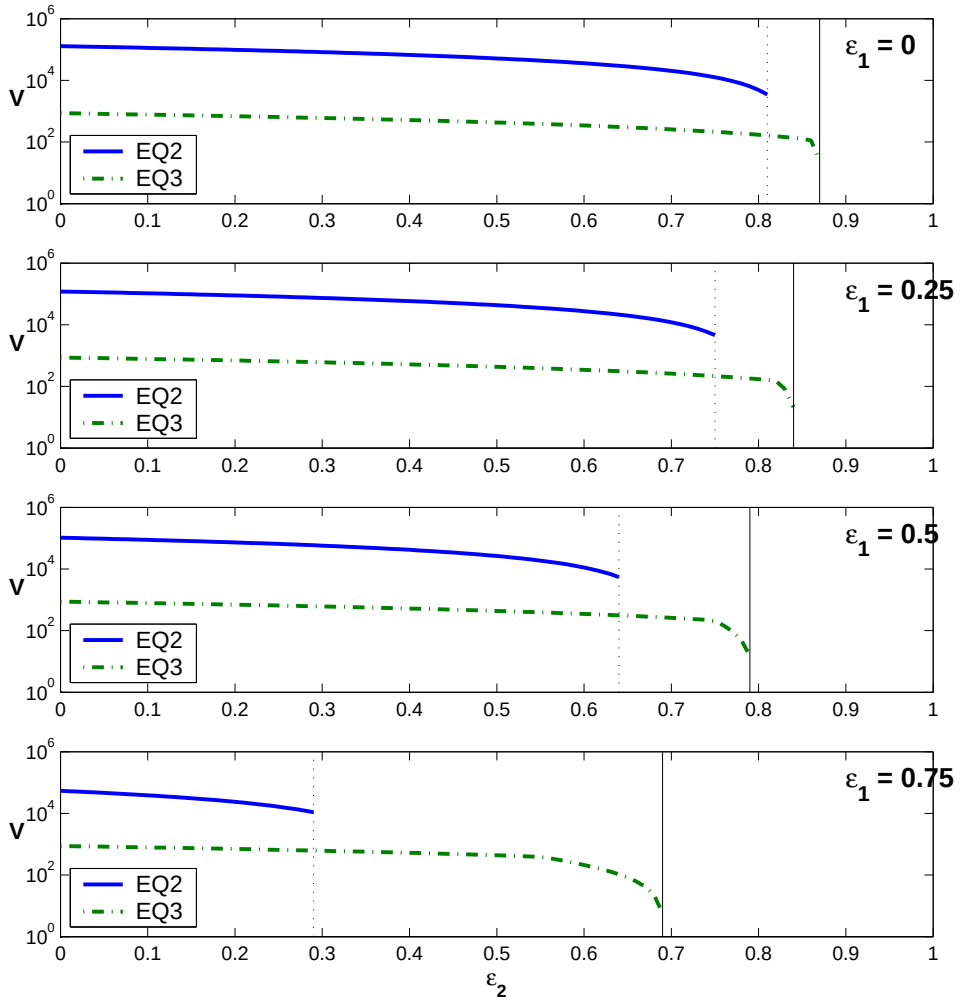


Figure 3.4: Sensitivity of stable viral load equilibria (total virus $V = V_I + V_{NI}$) to drug efficacy ϵ_2 for various fixed values of ϵ_1.

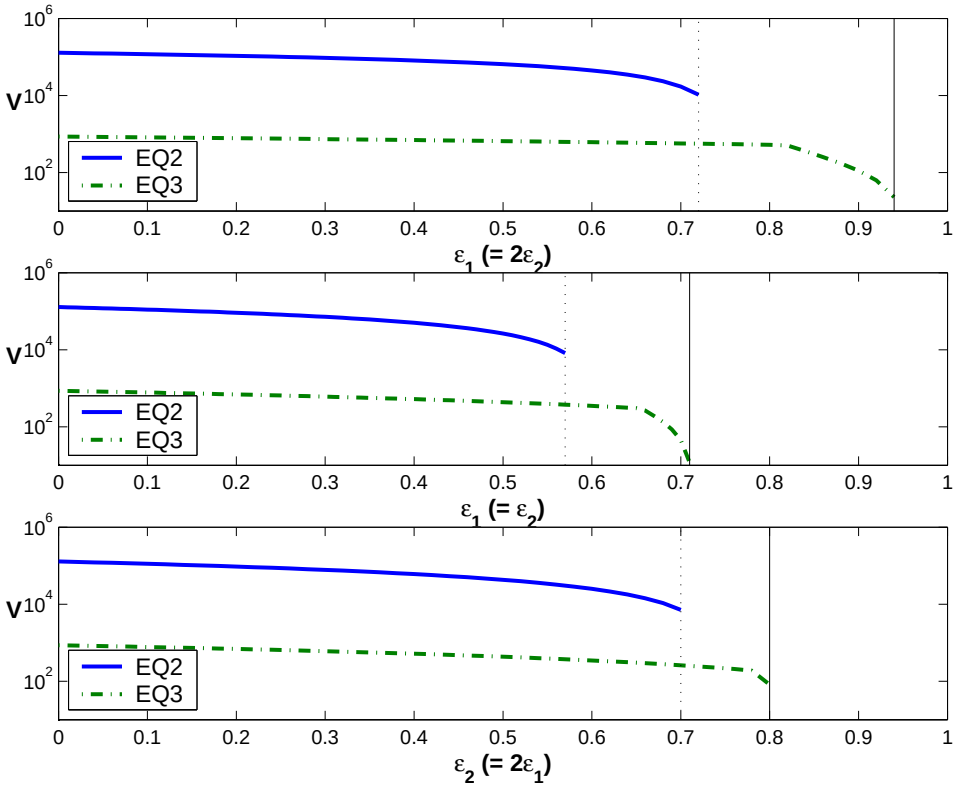


Figure 3.5: Sensitivity of stable viral load equilibria (total virus $V = V_I + V_{NI}$) to drug efficacy ϵ_1, ϵ_2 for various relationships between the two parameters. The relationships are shown on the lines in Figure 3.6.

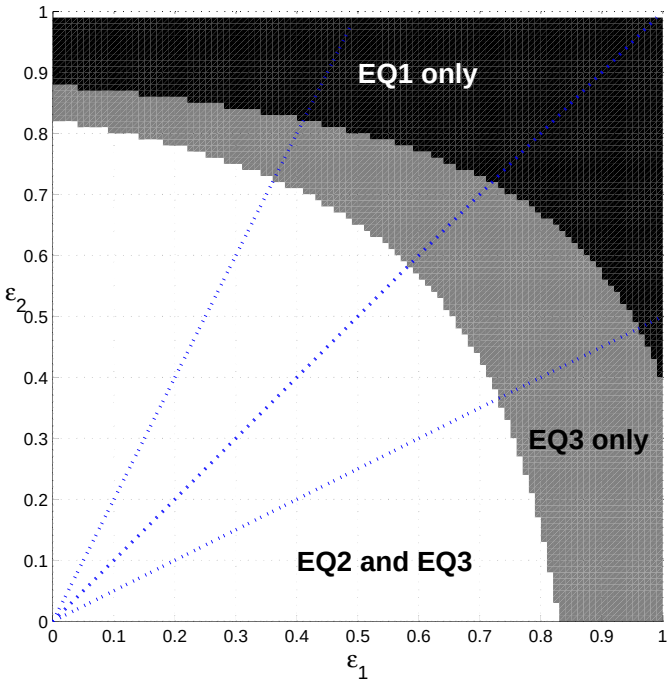


Figure 3.6: Regions of stability for equilibria given various values of treatment efficacies ϵ_1 and ϵ_2. Note that while the specific values of the steady states vary with these parameters as well, $EQ1$ refers to the disease free-equilibrium, $EQ2$ to the viral dominant, and $EQ3$ to the immune dominant. The lines indicate the linear relationships of parameters ϵ_1, ϵ_2 considered in Figure 3.5.

3.2.2 Existence and computation of model solution

Existence of solution to ODE system

Theory for existence and uniqueness of solutions to a differential equation system $\dot{\bar{x}} = \bar{g}(t, \bar{x}; q)$ typically appeals to Lipschitz continuity in $\bar{x}$ of the function $\bar{g}(t, \bar{x}; q)$. However the model dynamics for $\bar{g}$ given by the right side of (3.1) do not satisfy such a Lipschitz condition in $\bar{x}$, due to product nonlinearities such as $k_1 T_1 V_I$. If these approximate terms are replaced by ones that saturate out as cell populations grow large (as is more biologically reasonable since one does not expect rates to grow without bound as the population increases), the function $\bar{g}(\bar{x})$ is piecewise differentiable and satisfies a global Lipschitz condition. The typical theory for solutions to ODEs then applies to guarantee the existence of a unique solution.

To determine a global Lipschitz bound on the dynamics we proceed as in [16] and first rewrite the system as

$$\frac{d\bar{x}}{dt} = S + \mathcal{L}(t)\bar{x} + h(t, \bar{x}),$$

where $S = [\lambda_1, \lambda_2, 0, 0, 0, 0, \lambda_E]^T$ includes the source terms,

$$\mathcal{L}(t)\bar{x} = \begin{bmatrix} -d_1 & & & & & & \\ & -d_2 & & & & & \\ & & -\delta & & & & \\ & & & -\delta & & & \\ & & (1-\bar{\epsilon}_2(t))10^3 N_T \delta & (1-\bar{\epsilon}_2(t))10^3 N_T \delta & -c & & \\ & & \bar{\epsilon}_2(t)10^3 N_T \delta & \bar{\epsilon}_2(t)10^3 N_T \delta & & -c & \\ & & & & & & -\delta_E \end{bmatrix}$$

is the component of the system that is linear in the state, and the nonlinearities are incorporated in

$$h(t,\bar{x}) = \begin{bmatrix} -\left(1-\bar{\epsilon}_1(t)\right)k_1\bar{x}_1\bar{x}_5 \\ -(1-f\bar{\epsilon}_1(t))k_2\bar{x}_2\bar{x}_5 \\ (1-\bar{\epsilon}_1(t))k_1\bar{x}_1\bar{x}_5 - m_1\bar{x}_3\bar{x}_7 \\ (1-f\bar{\epsilon}_1(t))k_2\bar{x}_2\bar{x}_5 - m_2\bar{x}_4\bar{x}_7 \\ -(1-\bar{\epsilon}_1(t))\rho_1 10^3 k_1\bar{x}_1\bar{x}_5 - (1-f\bar{\epsilon}_1(t))\rho_2 10^3 k_2\bar{x}_2\bar{x}_5 \\ 0 \\ \frac{b_E(\bar{x}_3+\bar{x}_4)}{(\bar{x}_3+\bar{x}_4)+K_b}\bar{x}_7 - \frac{d_E(\bar{x}_3+\bar{x}_4)}{(\bar{x}_3+\bar{x}_4)+K_d}\bar{x}_7 \end{bmatrix}.$$

In the nonlinear term $h(t,\bar{x})$ for $s=1,2$, the potentially unbounded terms $k_s\bar{x}_s\bar{x}_5$ can be replaced by $k_s\bar{x}_s\bar{x}_5 = \bar{k}_s^s(\bar{x}_s)\bar{k}_s^5(\bar{x}_5)$ and the terms $m_s\bar{x}_{s+2}\bar{x}_7$ replaced by $m_s\bar{x}_{s+2}\bar{x}_7 = \bar{m}_s^{s+2}(\bar{x}_{s+2})\bar{m}_s^7(\bar{x}_7)$, where the right sides of these functions saturate out. For example for $s=1,2$ let

$$\bar{k}_s^s(\bar{x}_s) = \begin{cases} 0 & \bar{x}_s < 0 \\ \sqrt{k_s}\bar{x}_s & 0 \le \bar{x}_s \le \bar{x}_s^M \\ \sqrt{k_s}\bar{x}_s^M & \bar{x}_s^M < \bar{x}_s \end{cases}; \qquad \bar{k}_s^5(\bar{x}_5) = \begin{cases} 0 & \bar{x}_5 < 0 \\ \sqrt{k_s}\bar{x}_5 & 0 \le \bar{x}_5 \le \bar{x}_5^M \\ \sqrt{k_s}\bar{x}_5^M & \bar{x}_5^M < \bar{x}_5 \end{cases};$$

$$\bar{m}_s^{s+2}(\bar{x}_{s+2}) = \begin{cases} 0 & \bar{x}_{s+2} < 0 \\ \sqrt{m_s}\bar{x}_{s+2} & 0 \le \bar{x}_{s+2} \le \bar{x}_{s+2}^M \\ \sqrt{m_s}\bar{x}_{s+2}^M & \bar{x}_{s+2}^M < \bar{x}_{s+2} \end{cases}; \text{ and } \quad \bar{m}_s^7(\bar{x}_7) = \begin{cases} 0 & \bar{x}_7 < 0 \\ \sqrt{m_s}\bar{x}_7 & 0 \le \bar{x}_7 \le \bar{x}_7^M \\ \sqrt{m_s}\bar{x}_7^M & \bar{x}_7^M < \bar{x}_7 \end{cases}.$$

Note that each of these new rate functions are globally bounded and piecewise differentiable, e.g., $\bar{k}_s^s \le \sqrt{k_s}\bar{x}_s^M$ and $(\bar{k}_s^s)'(\bar{x}_s) \le \sqrt{k_s}$. Similarly, let $\bar{b}_E = \sqrt{b_E}, \bar{d}_E = \sqrt{d_E}$

and define the bounded and piecewise differentiable terms

$$\bar{b}_E^7(\bar{x}_7) = \begin{cases} 0 & \bar{x}_7 < 0 \\ \sqrt{b_E}\bar{x}_7 & 0 \le \bar{x}_7 \le \bar{x}_7^M \\ \sqrt{b_E}\bar{x}_7^M & \bar{x}_7^M < \bar{x}_7 \end{cases}, \text{ and } \quad \bar{d}_E^7(\bar{x}_7) = \begin{cases} 0 & \bar{x}_7 < 0 \\ \sqrt{d_E}\bar{x}_7 & 0 \le \bar{x}_7 \le \bar{x}_7^M \\ \sqrt{d_E}\bar{x}_7^M & \bar{x}_7^M < \bar{x}_7 \end{cases}$$

to replace the unbounded terms in the dynamics for $\bar{x}_7$.

Now take $h(t, \bar{x})$ together with these substitutions and define a new, saturated nonlinear term:

$$h^S(t,\bar{x}) = \begin{bmatrix} -\left(1-\bar{\epsilon}_1(t)\right)\bar{k}_1^1(\bar{x}_1)\bar{k}_1^5(\bar{x}_5) \\ -(1-f\bar{\epsilon}_1(t))\bar{k}_2^2(\bar{x}_2)\bar{k}_2^5(\bar{x}_5) \\ (1-\bar{\epsilon}_1(t))\bar{k}_1^1(\bar{x}_1)\bar{k}_1^5(\bar{x}_5) - \bar{m}_1^3(\bar{x}_3)\bar{m}_1^7(\bar{x}_7) \\ (1-f\bar{\epsilon}_1(t))\bar{k}_2^2(\bar{x}_2)\bar{k}_2^5(\bar{x}_5) - \bar{m}_2^4(\bar{x}_4)\bar{m}_2^7(\bar{x}_7) \\ -(1-\bar{\epsilon}_1(t))\rho_1 10^3\bar{k}_1^1(\bar{x}_1)\bar{k}_1^5(\bar{x}_5) - (1-f\bar{\epsilon}_1(t))\rho_2 10^3\bar{k}_2^2(\bar{x}_2)\bar{k}_2^5(\bar{x}_5) \\ 0 \\ \frac{\bar{b}_E(\bar{x}_3+\bar{x}_4)}{(\bar{x}_3+\bar{x}_4)+K_b}\bar{b}_E^7(\bar{x}_7) - \frac{\bar{d}_E(\bar{x}_3+\bar{x}_4)}{(\bar{x}_3+\bar{x}_4)+K_d}\bar{d}_E^7(\bar{x}_7) \end{bmatrix}.$$

Therefore when the states $\bar{x}$ are in the set $\Omega = \left\{\bar{x} \in \mathbb{R}^7 | \bar{x}_s \le \bar{x}_s^M, s = 1, \ldots, 5, 7\right\}$, $h^S(t,\bar{x}) = h(t,\bar{x})$ (the original system), but on $\mathbb{R}^7 \backslash \Omega$ one or more rates saturate.

The rates $\frac{\bar{b}_E(\bar{x}_3+\bar{x}_4)}{(\bar{x}_3+\bar{x}_4)+K_b}$ and $\frac{\bar{d}_E(\bar{x}_3+\bar{x}_4)}{(\bar{x}_3+\bar{x}_4)+K_d}$ are also bounded and have bounded derivatives with respect to $\bar{x}_3$ and $\bar{x}_4$. For example,

$$B(\bar{x}_3+\bar{x}_4) = \frac{\bar{b}_E(\bar{x}_3+\bar{x}_4)}{(\bar{x}_3+\bar{x}_4)+K_b} \le \bar{b}_E$$

and, taking a derivative with respect to either variable,

$$B'(\bar{x}_3 + \bar{x}_4) = \frac{\bar{b}_E K_b}{(\bar{x}_3 + \bar{x}_4 + K_b)^2} \leq \frac{\bar{b}_E}{K_b}.$$

The bounds on the states and derivatives directly imply that the 7×7 derivative matrix $D_{\bar{x}} h(t, \bar{x})$ is bounded:

$$\left\| D_{\bar{x}} h(t, \bar{x}) \right\|_\infty < \infty.$$

We can therefore prove global existence and uniqueness of a solution to the modified system

$$\frac{d\bar{x}}{dt} = \bar{g}^S(t, \bar{x}) = S + \mathcal{L}(t)\bar{x} + h^S(t, \bar{x})$$

since the right side satisfies a global Lipschitz condition:

$$\begin{aligned} \left| \bar{g}^S(t, \phi) - \bar{g}^S(t, \psi) \right| &= \left| \mathcal{L}(t)(\phi - \psi) + h^S(t, \phi) - h^s(t, \psi) \right| \\ &\leq \|\mathcal{L}(t)\| |\phi - \psi| + \left\| D_{\bar{x}} h(t, \bar{x}) \right\|_\infty |\phi - \psi| \\ &= K_L |\phi - \psi|. \end{aligned}$$

Standard theory for ordinary differential equation systems then guarantees the existence of a unique solution (see for example Theorem 3.1 in Robinson [47]).

Remark: While we showed existence and uniqueness of a solution to the more biologically realistic problem involving saturation terms, we compute using the original system (3.1). In simulations we never encountered a problem of the states (and consequently derivatives) growing unbounded and therefore did not have to substitute the saturated dynamics.

Numerical simulation of ODE system

In solving the HIV dynamics system numerically we substitute a log-transformed system. This resolved a problem of states becoming unrealistically negative during solution due to round-off error: nonnegative solutions of this model should stay so throughout integration. It also enables efficient handling of unrealistic cases where states get infinitesimally small during integration due to parameters selected by optimization algorithms. Using the transformation $x = \log_{10}(\bar{x})$, with the original system $\dot{\bar{x}}_i = \bar{g}_i(t, \bar{x}; q)$ we obtain the system

$$\frac{dx_i}{dt} = \frac{10^{-x_i}}{\ln(10)} \bar{g}_i(t, 10^x; q), \; i = 1, \ldots, 5, 7, \tag{3.4}$$

which is the log-transformed analog of the reduced system $\dot{x}_R = g_R(t, x_R; q)$ discussed above, i.e., for all states except V_{NI}.

Given a vector of model dynamics parameters q and specified initial conditions $\bar{x}(0)$, we calculate numerical solutions for the model using variable-order adaptive BDF-based integrators. The solvers used for results in this dissertation are Matlab's `ode15s` [50] and Lawrence Livermore's `LSODE` [29] with relative error tolerance 10^{-9}. The latter is written in Fortran and thus enabled more rapid computation which substantially helped in both optimization problems here and Monte Carlo simulations reported in [9].

Having obtained the model solution $x(t)$ to (3.4) and therefore $\bar{x}(t)$, we use its information on $T_1^*(t), T_2^*(t)$, and $V_I(t)$ to integrate (3.2) using composite sixteen-point Gaussian quadrature. If T_F denotes the final time (day) at which a solution is desired, then for $t = 1, \ldots, T_F$ we integrate on one day subintervals $[t-1, t]$:

$$V_{NI}(t) = V_{NI}(t-1)e^c + \int_{t-1}^{t} e^{-c(t-s)} G(s) ds,$$

using sixteen Gauss points to evaluate the integral in the second term. We take this approach rather than integrating the full seven-state ODE system because, in the absence of PI treatment ($\bar{\epsilon}_2(t) \equiv 0$), exponential decay makes it impractical to solve for V_{NI} on the logarithmic scale.

A simulation of early infection from the model is shown in Figure 3.7. Simulation is started near steady state $EQ1$, with the addition of one viral copy per ml ($V = 1$ copy/ml) and a corresponding small amount of infected cells ($\bar{x}^0 = [1000, 3.198, 1 \times 10^{-4}, 1 \times 10^{-6}, 1, 1, 0.01]^T$). Upon infection, the virus replicates to a peak before converging in damped oscillations to equilibrium. There is a delayed initial immune response to the presence of the infected cells, but a sustained and vigorous immune response does not develop. The higher infection rate in the T_2 population is evident in its more dramatic decline than the T_1 cell population. The plot for V includes non-infectious virus $V = V_I + V_{NI}$; however the simulation presumes no treatment, so the V_{NI} population is quickly extinguished.

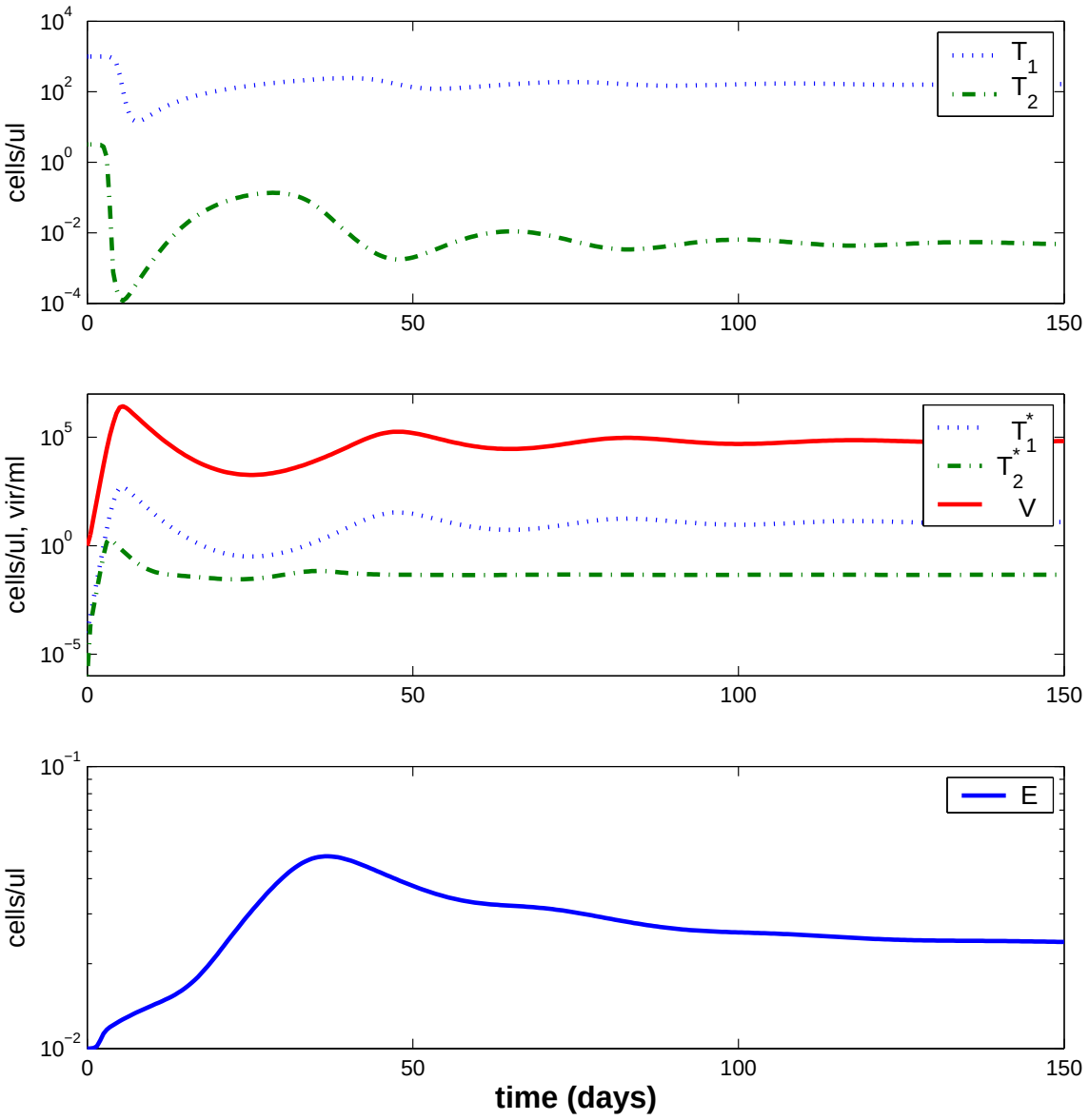


Figure 3.7: Simulation of early infection scenario using HIV model (3.1). Note varying logarithmic scales in different subplots.

3.3 Error Model and Simulated Data

In order to generate simulated data and perform inference on estimated parameters, one must make assumptions about the structure of errors in the model and process. These errors are not relative to the model states directly, but rather the experimentally-observed components of them. Applying an observation operator $\mathcal{O}$ to the vector of unscaled model solutions $\bar{x}(t^{ij})$ yields the vector of observed model states $\bar{z}(t^{ij}) = \mathcal{O}\bar{x}(t^{ij})$ to be fit to data. For example the clinical data includes at most total CD4$^+$ T-cell count $(T_1 + T_1^*)$, free virus $(V + V_{NI})$, and immune response (E), although we do not fit E data. The corresponding observer matrix $\mathcal{O} \in \mathbb{R}^{m\times n}$ that produces the $m = 2$ outputs for this system would be

$$\mathcal{O} = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \end{pmatrix}.$$

We assume observation errors for a particular patient, at times t^{ij}, are independent for each time. As is typical for assay data, we further assume that the unscaled observations $\bar{y}_s^{ij}$ for patient j and state $s = 1, \ldots, m$, are distributed lognormally such that

$$E(\bar{y}_s^{ij}) = \bar{z}_s(t_s^{ij}; q^j), \text{ and} \tag{3.5a}$$

$$\mathrm{Var}(\bar{y}_s^{ij}) = \bar{\sigma}_s^2 \left\{\bar{z}_s(t_s^{ij}; q^j)\right\}^2. \tag{3.5b}$$

To be precise, one should distinguish between random variables $\bar{Y}_s^{ij}$ and their realizations $\bar{y}_s^{ij}$, but in most cases the meaning is clear from context and we will not make this distinction. Equations (3.5) comprise a constant coefficient of variation model, where $\bar{\sigma} = \frac{\sqrt{\mathrm{Var}(\bar{y})}}{E(\bar{y})}$ is the coefficient of variation (ratio of standard deviation

to mean). In this case, the variance in the observations is proportional to the square of the mean response and is equivalent to an adjusted lognormal distribution of the data about the mean response. In particular, defining $y_s^{ij} = \log_{10} \bar{y}_s^{ij}$, the log-scaled y_s^{ij} are distributed normally: $\mathcal{N}(\zeta_s^{ij}, \sigma_s^2)$, where

$$\zeta_s^{ij} = \log_{10} \bar{z}_s^{ij} - \frac{\log_{10}(\bar{\sigma}_s^2 + 1)}{2}, \text{ and} \tag{3.6a}$$

$$\sigma_s^2 = \log_{10}\left(\bar{\sigma}_s^2 + 1\right) \log_{10}(e). \tag{3.6b}$$

We first verify a similar result involving natural logarithms using the moment generating function for the normal distribution.

Proposition 3.1. *If a random variable $W \sim \mathcal{N}(m, s^2)$, where*

$$m = \ln \mu - \frac{\ln(\sigma^2 + 1)}{2} \quad \textit{and} \quad s^2 = \ln\left(\sigma^2 + 1\right),$$

then $\bar{W} = e^W$ is such that

$$E(\bar{W}) = \mu \quad \textit{and} \quad \operatorname{Var} \bar{W} = \sigma^2 \mu^2.$$

Proof. From the moment generating function of the normal distribution $M(t) = E\left(e^{tW}\right) = e^{mt + s^2 t^2/2}$, it immediately follows that

$$E(\bar{W}) = E(e^W) = M(1) = e^{m + s^2/2} \text{ and}$$

$$E(\bar{W}^2) = E(e^{2W}) = M(2) = e^{2m + 2s^2}.$$

Therefore,

$$\begin{aligned} E(\bar{W}) &= e^{m+s^2/2} = e^{\ln\mu - \frac{\ln(\sigma^2+1)}{2} + \frac{\ln(\sigma^2+1)}{2}} = \mu \text{ , and} \\ \mathrm{Var}(\bar{W}) &= E(\bar{W}^2) - \{E(\bar{W})\}^2 \\ &= e^{2m+2s^2} - \left(e^{m+s^2/2}\right)^2 \\ &= \left(e^{s^2} - 1\right) e^{2m+s^2} \\ &= \sigma^2\mu^2. \end{aligned}$$

□

We now verify the equivalence of (3.5) and (3.6), where the logarithms are base-10 rather than base e.

Proposition 3.2. *Let $Y \sim \mathcal{N}(\zeta, \sigma^2)$, where ζ and σ^2 are as in* (3.6) *(indices suppressed). Then $\bar{Y} = 10^Y$ is such that the variance model* (3.5) *holds.*

Proof. Define $W := \ln(10)Y$ (and note that $\bar{Y} := 10^Y = e^{\ln(10)Y} = e^W$). Then, defining $m = \ln(10)\zeta$ and $s^2 = \ln(10)^2\sigma^2$, W is such that $W \sim \mathcal{N}(m, s^2)$. Directly applying Proposition 3.1 to W yields

$$\begin{aligned} E(10^Y) = E(e^W) &= \bar{z}, \text{ and} \\ \mathrm{Var}(10^Y) = \mathrm{Var}(e^W) &= \bar{\sigma}^2 \{\bar{z}\}^2 . \end{aligned}$$

□

We therefore use the equivalence of these two formulations and generate simulated data by sampling y_s^{ij} from $\mathcal{N}(\zeta_s^{ij}, \sigma_s^2)$ (where ζ_s^{ij} depends on the solution to the

differential equation model) and then exponentiating to obtain the data:

$$\bar{y}_s^{ij} = 10^{y_s^{ij}}.$$

We generate simulated data with coefficient of variation $\bar{\sigma}_s = 0.2$ or $\bar{\sigma}_s = 0.3$ for each state, which are in the range observed in practice for viral load and CD4 count assays [51, 55].

3.4 Sensitivity Computations

Computing sensitivity of observed model outputs to dynamic parameters both yields information about identifiability and helps construct the relationship between estimated parameters and error estimates (confidence intervals) in an inverse problem process. In the case of our ordinary differential equation model, semi-relative sensitivities can be computed explicitly by differentiating the dynamical system with respect to one or more parameters of interest and then integrating the newly-formed block system in time. For example, to determine the relative sensitivity of observed model outputs $z_s(t_s^{ij}; q)$ with respect to parameters q_k, one must compute

$$\frac{\partial z_s(t_s^{ij}; q)}{\partial q_k} \cdot q_k \tag{3.7}$$

which scales the sensitivity by the magnitude of the parameter considered [6]. Note that a parameter q_k in this case might be a model dynamic parameter or an initial condition.

Since the solution of the ODE model is better posed on the logarithmic scale, we determine the sensitivities using the log scale model dynamics, and then transform to relate them to the sensitivities of the observed states on the log scale (which we

fit to data). Denoting the system of log-scaled dynamics from (3.4) by

$$\frac{dx}{dt} = h(t, x; q) \tag{3.8}$$

$$x(0) = x^0 \tag{3.9}$$

we formally differentiate with respect to each q_k and interchange the order of the time and parameter derivatives [6, 23, 25]. In the case of r parameters and n model state variables, we thus obtain an $(n \times r)$-dimensional system of differential equations for the sensitivities $x_q(t; q) = \frac{\partial x}{\partial q}(t; q)$, where q is the vector of parameters considered:

$$\frac{d}{dt}\left(\frac{\partial x}{\partial q}(t)\right) = \frac{\partial h}{\partial x}\frac{\partial x}{\partial q}(t) + \frac{\partial h}{\partial q} \tag{3.10}$$

with initial condition

$$\frac{\partial x}{\partial q}(0) = \frac{\partial x^0}{\partial q}. \tag{3.11}$$

The initial condition matrix (3.11) has zero entries (as the initial conditions are independent of the model dynamic parameters) except when one or more initial conditions are included in the vector of parameters of interest. In the latter case, the sensitivity initial condition for a model state with respect to its own initial condition is 1. Note that $\frac{\partial h}{\partial x}$ is the Jacobian of the ODE system, and $\frac{\partial h}{\partial q}$ is similarly a matrix containing the derivatives of the right side with respect to the parameters considered.

We solve the system {(3.10),(3.11)} for $x_q(t; q)$ by coupling it with the original differential equation system to obtain an $(nr+n)$-dimensional system which we again solve numerically with the Matlab solver `ode15s`. More specifically, since the sensitivity of V_{NI} with respect to any parameter does not depend on V_{NI} itself, we couple states $[1, 2, 3, 4, 5, 7]$ from (3.4) (omitting V_{NI}) to the sensitivity equations (3.10) and integrate, then perform Gaussian quadrature as before to obtain model solutions for

V_{NI}. (To visit this sensitivity computation method in another context, see [1].)

Since we obtain sensitivity solutions on a log scale, we have to be careful to transform before applying the observation operator. As before let $x = \log_{10} \bar{x}$. Then we can compute the sensitivity of the unscaled model solutions by

$$\frac{\partial \bar{x}_s}{\partial q} = \ln(10) 10^{x_s} \frac{\partial x_s}{\partial q}.$$

Since the observation process is linear in these unscaled states and the derivative is a linear operator, we can now apply the observer to determine the sensitivity of the log of the observed model states as follows:

$$\begin{aligned} \frac{\partial z_s}{\partial q}(t;q) &= \frac{\partial}{\partial q} \log_{10} \left(\mathcal{O}_s \bar{x}(t;q) \right) && (3.12) \\ &= \log_{10}(e) \frac{\mathcal{O}_s \left(\frac{\partial \bar{x}}{\partial q}(t;q) \right)}{\mathcal{O}_s \left(\bar{x}_i(t;q) \right)}. && (3.13) \end{aligned}$$

These can then be scaled by the relative size of the various parameters q_k considered.

This process yields sensitivity information as a function of time over the interval of integration considered which will be used directly to compute confidence intervals. To understand overall which model dynamic parameters and initial conditions most influence the outputs of the system we take the L^2 norm of these over time. In particular, given treatment schedule, observation times, and parameters for virtual patient j, we compute the influence of each parameter q_k on CD4 (z_1) and viral RNA (z_2) response over time:

$$\sqrt{\frac{1}{N_1^j} \sum_{i=1}^{N_1^j} \left| \frac{\partial z_1}{\partial q_k}(t_1^{ij}; q^j) q_k \right|^2}, \quad \sqrt{\frac{1}{N_2^j} \sum_{i=1}^{N_2^j} \left| \frac{\partial z_2}{\partial q_k}(t_2^{ij}; q^j) q_k \right|^2}.$$

For the cases considered here, using the sup norm yields similar results.

We examine results for relative sensitivity for CD4 ($T_1 + T_1^*$) and viral RNA (V_I+V_{NI}) to each of the 20 parameters and seven initial conditions. The conditions for these computations are based on the sample set of eighteen patients PS18, described in Chapter 2.

1. For Scenario 1, we set the treatment schedules $u^j(t)$ and observation times t_s^{ij} for each patient based on the actual record for patient j in the clinical data. However, we fix the model parameter values $q^j, j = 1, \ldots, 18$ for all patients at the values in Table 3.2 and use initial condition $\bar{x}^0 = [600, 50, 60, 50, 900000, 10, 20]$ to simulate viral peak in early infection.

2. For Scenario 2, we again consider the treatment schemes and observation times from the clinical data sets, however now the model parameter values $q^j, j = 1, \ldots, 18$ and initial condition $\bar{x}^{0j}$ for each patient were set to those from preliminary estimates based on clinical data.

Table 3.4 contains the semi-relative sensitivities for each scenario and parameter. Results for each scenario in the tables are based on averaging across the eighteen virtual patients considered. Perhaps more interesting is Table 3.5 where for each scenario the parameters are ranked in increasing order of semi-relative sensitivity. The model outputs are most sensitive to the parameters V_I^0, T_1^0, ϵ_2, and V_{NI}^0 and secondarily to the parameters $\lambda_1, \epsilon_1, k_1, T_1^{*0}$, and c. These results are dependent on the relative magnitude of the parameters considered and therefore vary between the two scenarios. This sensitivity analysis will inform the inverse problem process, as one should not expect to estimate a parameter to which the model solutions are insensitive.

Other approaches exist for estimating solutions to sensitivity equations and may be necessary if the sensitivity system (3.10) proves difficult to integrate. Finite-

difference approximations may provide a reasonable estimate to the sensitivities. An example using forward-difference derivative approximations in the context of an electromagnetics model is presented in [26].

Table 3.4: Semi-relative sensitivities of CD4 and viral RNA with respect to each of 27 parameters in two scenarios.

	scenario 1		scenario 2	
param	CD4	viral RNA	CD4	viral RNA
λ_1	2.12e+00	3.82e+01	3.29e-01	1.09e+00
d_1	1.75e+00	3.23e+01	3.05e-01	9.57e-01
ϵ_1	4.17e+00	7.78e+01	1.06e-02	4.63e+00
k_1	1.98e+00	3.43e+01	9.77e-02	1.07e+00
λ_2	2.53e-01	6.24e+00	1.84e-02	5.03e+00
d_2	7.32e-02	1.78e+00	1.09e-02	3.35e+00
f	1.25e-01	2.81e+00	4.58e-03	2.73e+00
k_2	4.05e-01	8.99e+00	9.11e-03	4.54e+00
δ	2.09e-01	2.92e+00	3.46e-02	6.82e-01
m_1	1.09e-02	2.44e-01	9.05e-04	2.89e-02
m_2	5.10e-03	9.44e-02	2.35e-04	5.54e-02
ϵ_2	1.06e+00	7.54e+02	5.13e-02	1.22e+03
N_T	1.51e+00	3.08e+01	1.09e-02	5.94e+00
c	1.64e+00	3.76e+02	1.01e-01	1.90e+02
λ_E	3.22e-02	6.26e-01	1.11e-03	7.48e-02
b_E	2.65e+00	5.13e+01	1.84e-02	1.17e+00
K_b	7.08e-01	1.39e+01	1.57e-03	2.36e-01
d_E	1.39e+00	2.70e+01	6.40e-03	1.97e-01
K_d	1.53e+00	2.95e+01	1.45e-02	1.07e+00
δ_E	1.22e+00	2.38e+01	1.29e-02	1.01e+00
T_1^0	2.96e+02	5.09e+03	1.78e+02	4.25e+02
T_2^0	1.79e+00	3.27e+01	7.72e-02	5.98e+00
T_1^{*0}	2.85e+00	4.84e+01	1.80e-01	1.47e+00
T_2^{*0}	1.83e+00	3.36e+01	5.11e-04	1.45e-01
V_I^0	1.30e+04	3.56e+05	2.91e+03	2.29e+05
V_{NI}^0	0.00e+00	2.60e+00	0.00e+00	1.68e+03
E^0	1.67e+00	3.19e+01	2.24e-05	1.74e-02

Table 3.5: Parameters ranked by semi-relative sensitivity for CD4 and viral RNA in two scenarios.

	scenario 1		scenario 2	
rank	CD4	viral RNA	CD4	viral RNA
1	25 (V_I^0)	25 (V_I^0)	25 (V_I^0)	25 (V_I^0)
2	21 (T_1^0)	21 (T_1^0)	21 (T_1^0)	26 (V_{NI}^0)
3	3 (ϵ_1)	12 (ϵ_2)	1 (λ_1)	12 (ϵ_2)
4	23 (T_1^{*0})	14 (c)	2 (d_1)	21 (T_1^0)
5	16 (b_E)	3 (ϵ_1)	23 (T_1^{*0})	14 (c)
6	1 (λ_1)	16 (b_E)	14 (c)	22 (T_2^0)(N_T)
7	4 (k_1)	23 (T_1^{*0})	4 (k_1)	13 (N_T)
8	24 (T_2^{*0})	1 (λ_1)	22 (T_2^0)	5 (λ_2)
9	22 (T_2^0)	4 (k_1)	12 (ϵ_2)	3 (ϵ_1)
10	2 (d_1)	24 (T_2^{*0})	9 (δ)	8 (k_2)
11	27 (E^0)	22 (T_2^0)	5 (λ_2)	6 (d_2)
12	14 (c)	2 (d_1)	16 (b_E)	7 (f)
13	19 (K_d)	27 (E^0)	19 (K_d)	23 (T_1^{*0})
14	13 (N_T)	13 (N_T)	20 (δ_E)	16 (b_E)
15	18 (d_E)	19 (K_d)	6 (d_2)	1 (λ_1)
16	20 (δ_E)	18 (d_E)	13 (N_T)	4 (k_1)
17	12 (ϵ_2)	20 (δ_E)	3 (ϵ_1)	19 (K_d)
18	17 (K_b)	17 (K_b)	8 (k_2)	20 (δ_E)
19	8 (k_2)	8 (k_2)	18 (d_E)	2 (d_1)
20	5 (λ_2)	5 (λ_2)	7 (f)	9 (δ)
21	9 (δ)	9 (δ)	17 (K_b)	17 (K_b)
22	7 (f)	7 (f)	15 (λ_E)	18 (d_E)
23	6 (d_2)	26 (V_{NI}^0)	10 (m_1)	24 (T_2^{*0})
24	15 (λ_E)	6 (d_2)	24 (T_2^{*0})	15 (λ_E)
25	10 (m_1)	15 (λ_E)	11 (m_2)	11 (m_2)
26	11 (m_2)	10 (m_1)	27 (E^0)	10 (m_1)
27	26 (V_{NI}^0)	11 (m_2)	26 (V_{NI}^0)	27 (E^0)

Chapter 4

Parameter Identification (Inverse) Problem

We wish to use the HIV model (3.1) to describe clinical data, make predictions, and suggest better treatment schemes. To do this more realistically, the model should be calibrated to patient data by estimating its parameters. We explore methods for doing so in the context of data from one or more patients.

As noted in the data description, in performing an inverse problem we do not have the luxury of observing the full vector of model states at each measurement time. Given an observation operator appropriate for the data, let $\bar{x}$ denote native model solutions and $\bar{z} = \mathcal{O}\bar{x}$, observed model solutions. Recall that the number of observations might vary from patient to patient, so for each patient $j = 1, \ldots, N_P$, we have times $\{t_1^{ij}, i = 1, \ldots N_1^j\}$ for CD4 measurements $y_1(t)$ and $\{t_2^{ij}, i = 1, \ldots N_2^j\}$ for viral RNA measurements $y_2(t)$. We fit the model using the base-10 logarithm of these quantities: $x = \log_{10} \bar{x}, z = \log_{10} \bar{z}, y = \log_{10} \bar{y}$.

4.1 Inverse Problem Formulations

Consider the two inverse problem formulations given by (4.1) and (4.2) below. The first employs data from a single patient in order to estimate one or more parameters (q). In this case, for each fixed patient j, the goal is to fit the ODE model to his data by minimizing the cost criterion

$$q^{*j} = \arg\min_{q \in Q} J(q) = \sum_{s=1}^{m} \frac{1}{N_s^j} \sum_{i=1}^{N_s^j} \left| z_s(t_s^{ij}; q) - y_s^{ij} \right|^2 \tag{4.1}$$

over an admissible parameter space $Q \subset \mathbb{R}^p$ to obtain optimal estimates. This is the typical nonlinear least squares formulation, where $J(q)$ depends through z on the solution to the nonlinear system of differential equations. Many sampling- and gradient-based methods are available to iteratively solve (4.1) for q^{*j} (see Kelley [33] and references therein).

In the second inverse problem method one estimates a distribution of parameters using data from multiple patients simultaneously. For this case, assume each patient can be identified with a set of "true" parameters q^{j0} and that these are therefore realizations of a random variable which comes from a distribution P on the admissible parameter space Q. So P describes the distribution of the parameter values q^{j0} across the population.

Given the probability space $\mathcal{P}(Q)$ of all probability distributions on the space Q in which the parameters of interest q live, and some admissible subset $\tilde{\mathcal{P}}(Q) \subset \mathcal{P}(Q)$ solve for

$$P^* = \arg\min_{P \in \tilde{\mathcal{P}}} J(P) = \frac{1}{N_P} \sum_{j=1}^{N_P} \sum_{s=1}^{m} \frac{1}{N_s^j} \sum_{i=1}^{N_s^j} \left| E\left[z(t_s^{ij}; q) | P(q) \right] - y_s^{ij} \right|^2. \tag{4.2}$$

This is a more abstract (in particular infinite-dimensional) problem, and the space $\tilde{\mathcal{P}}$ over which the minimization takes place to determine P^* must be carefully specified to determine well-posedness and stability of the inverse problem and certainly in order to compute a solution numerically.

In general the solution to the differential equation model and consequently the expected value in (4.2) may vary between patients based on factors other than observation times. In addition to patient-dependent observation times, the model solutions $z_s(\cdot; q)$ may depend on patient-specific treatment schemes $u^j(t)$. Further, some subset of model parameters $\grave{q}^j$ may be fixed, while others q^j are estimated. For this most general case, the model solutions look like

$$z_s(t_s^{ij}; \grave{q}^j; u^j(t); q), s = 1, \ldots, m$$

and we estimate either a vector of model parameters q or their distribution P as described. Some of this more general notation will be suppressed in sections of what follows.

One could consider more involved inverse problem schemas where parameters are estimated for each individual as well as distributions and errors across the population. These hierarchical statistical methods are widely used and have the advantage of being able to incorporate a high degree of structure, but are not considered here. The reader is referred to the discussion in [20] for more information. The methods considered here may offer some of the same information at reduced cost and complexity.

The single-patient inverse problem formulation (4.1) is well-suited to computation (minimization over a subset $Q \subset \mathbb{R}^p$). However, when considering (4.2) computationally, one cannot minimize over the space $\mathcal{P}(Q)$ of all probability distributions on Q so we consider sets $\tilde{\mathcal{P}}(Q)$ over which this is more feasible. Examples of $\tilde{\mathcal{P}}(Q) \subset \mathcal{P}(Q)$

include:

1. *Parameterized families of distributions.* For example, the set of normal distributions $\tilde{\mathcal{P}}(Q) = \{P \in \mathcal{P}(Q) | P \sim \mathcal{N}(q^0, \Sigma)\}$ is parameterized by the components of the mean vector q^0 and covariance matrix Σ, thus inducing a finite-dimensional inverse problem for these components. Similarly, other parametric distributions may be employed, e.g., mixtures of normals, lognormals, gammas, etc.

2. *Nonparametric point masses.* Banks and Bihari [4] demonstrate that a finite linear combination of Dirac measures ("point masses") approximate generic elements of $\mathcal{P}(Q)$. For example in the case of a univariate distribution, choose sets of nodes $Q^{N_m} = \left\{q_k^{N_m}\right\}_{k=0}^{N_m}, N_m = 1, 2, \ldots, \infty,$, such that each $Q^{N_m} \subset Q$ and their union $\cup_{N_m=1}^{\infty} Q^{N_m}$ is dense in Q. Then the set of distributions

$$\tilde{\mathcal{P}}(Q) = \mathcal{P}^{N_m}(Q) = \left\{ P \in \mathcal{P}(Q) | P = \sum_{k=1}^{N_m} w_k \delta_{q_k^{N_m}}, q_k^{N_m} \in Q^{N_m}, w_k \in \mathbb{Q}, \sum w_k = 1 \right\},$$

   where δ_q is the Dirac measure with atom at q, can be used to approximate $\mathcal{P}(Q)$ and is finite dimensional, yielding a tractable inverse problem in the rational weights w_k.

3. *Distributions with densities.* A third possibility involves approximation using finite linear combinations of linear splines to represent the associated probability density functions as developed in Banks–Pinter [13]. This entails restricting $\mathcal{P}(Q)$ to a class of distributions that are absolutely continuous with $L^2(Q)$ density. We use this approach here.

Direct estimation of the distribution of one or more model parameters across a population using (4.2) offers benefits over the single patient inverse problem (4.1).

In the multiple-patient context it is possible to leverage data from all patients simultaneously to obtain an estimate of the distributional form. This is possible, though potentially more computationally intensive, with the single patient inverse problems or through hierarchical methods. One could estimate parameters for each patient (samples from the true parameter distribution) and then use them to construct histograms or other representations of the joint distributions. In addition, as in the case of the formulations considered shortly, if only a small subset of parameters vary across the population with the rest of the model parameters fixed at known values, the method may offer substantial computational advantage since a basis of model solutions can be computed ahead of time, before the distribution is estimated. Further, while parameterized distributions (1.) are often used to describe population characteristics, the formulations (2.) and (3.) allow more flexible fitting to population characteristics and may help avoid distribution mis-specification as discussed in [12].

We now overview the particular finite-dimensional concretization of (4.2) considered in this dissertation. Presently, we are interested in using data to estimate the distribution of a single parameter. We consider the situation where the vector $\grave{q}^j$ of all but one of the model parameters is fixed for each patient and estimate a distribution for the remaining parameter q. We minimize over a subset of $\mathcal{P}(Q)$ consisting of probability measures that have L^2 probability density functions (pdfs) $f(q)$. We discretize the pdfs with piecewise linear spline approximations, reducing the inverse problem to finding spline coefficients, similar to the problem of finding Dirac weights in (2.) above. In this case, the inverse problem reduces to a quadratic programming problem.

Given a set of functions $\mathcal{F}$ to be specified later, consider $\mathcal{F}^{N_S} \subset \mathcal{F} \subset L^2(Q)$, where $\mathcal{F}^{N_S}$ consists of piecewise linear approximations to functions in $\mathcal{F}$. We solve the inverse problem (4.2) by minimizing over the set of probability measures with

piecewise linear density functions

$$\mathcal{P}^{N_S}(Q) = \left\{ F \in \mathcal{P}(Q) | F = \int f, f \in \mathcal{F}^{N_S} \right\}.$$

In the next section we explore the well-posedness of inverse problems in this context.

For q in the interval $[q_L, q_U]$ consider a partition $q_L = q_0 < q_1 < q_2 < \ldots < q_{N_S} = q_U$ as shown in Figure 4.1. Then

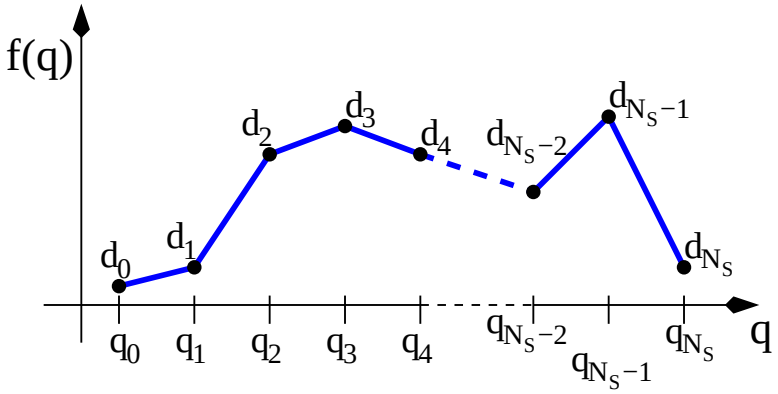


Figure 4.1: Example of piecewise linear spline approximation to density function $f(q)$.

$$f(q) \approx f^{N_S}(q) = \sum_{k=0}^{N_S} d_k \phi_k(q)$$

approximates a generic density $f(q) \in \mathcal{F}$, where $\phi_k(q)$ denotes the piecewise linear "hat" basis function centered at node q_k. Now, since we represent probability distributions $P(q)$ by their densities $f(q)$, the expected value in the cost criterion depends

on the them. Letting $E^{ij}(f) = E\left[z^{ij}(q)|f(q)\right] = E\left[z(t^{ij};\dot{q}^j;u^j(t);q)|f(q)\right]$,

$$\begin{aligned} E^{ij}(f) = \int_Q z^{ij}(q)f(q)dq &\approx \int_Q z^{ij}(q)\sum_{k=0}^{N_S} d_k\phi_k(q)dq \\ &= \sum_{k=0}^{N_S} d_k \int_Q z^{ij}(q)\phi_k(q)dq \\ &= \sum_{k=0}^{N_S} d_k \mathcal{S}_k^{ij}. \end{aligned}$$

The vector of integrals $[\mathcal{S}_k^{ij} = \int_Q z^{ij}(q)\phi_k(q)dq], k = 0,\ldots,N_S$ can be computed explicitly once the parameter space Q has been fixed. The cost function can now be rewritten

$$\begin{aligned} N_P\, J(f) &= \sum_{j=1}^{N_P}\frac{1}{N^j}\sum_{i=1}^{N^j}\left|\sum_{k=0}^{N_S} d_k\mathcal{S}_k^{ij} - y^{ij}\right|^2 \\ &= \sum_{j=1}^{N_P}\frac{1}{N^j}\sum_{i=1}^{N^j}\left[d^T\mathcal{S}^{ij}\left(\mathcal{S}^{ij}\right)^T d - 2d^T\mathcal{S}^{ij}y^{ij} + \left(y^{ij}\right)^2\right] \\ &= d^T\left[\sum_{j=1}^{N_P}\frac{1}{N^j}\sum_{i=1}^{N^j}\mathcal{S}^{ij}\left(\mathcal{S}^{ij}\right)^T\right]d - 2d^T\sum_{j=1}^{N_P}\frac{1}{N^j}\sum_{i=1}^{N^j}\mathcal{S}^{ij}y^{ij} \\ &\quad + \sum_{j=1}^{N_P}\frac{1}{N^j}\sum_{i=1}^{N^j}\left(y^{ij}\right)^2. \end{aligned} \tag{4.3}$$

(For clarity we suppress the state index s, although its sum can be readily included.) This can be summarized as

$$J(f(d)) = d^T Ad - 2d^T b + c,$$

which reveals that the minimization problem is a quadratic programming problem in the spline coefficients d_k, where the coefficients $A, b,$ and c contain information from

model basis functions and patient data as shown in (4.3). In order to ensure that the estimated coefficients determine a probability density function, we require them to be nonnegative and that the estimated pdf $f^{N_S}(q)$ integrate to one, yielding a constrained optimization problem:

$$\min_d J(d) = d^T A d - 2d^T b + c, \text{ subject to} \tag{4.4a}$$

$$1 = \sum_{k=0}^{N_S-1} \frac{d_k + d_{k+1}}{2} \Delta q_k \tag{4.4b}$$

$$d_k \geq 0, k = 0, 1, \ldots, N_S. \tag{4.4c}$$

Note that even though the matrices $A^{ij} = \mathcal{S}^{ij} \left(\mathcal{S}^{ij}\right)^T$ are rank one for each pair (i, j), their sum is not necessarily so.

The inverse problem as stated in (4.4) is potentially very ill-conditioned. We also explore a similar (improved) inverse problem that incorporates a penalty for irregularity of the density function $f(q)$. Specifically, we augment the cost function with a Tikhonov regularization-like term that penalizes the L^2 norm of the derivative of f:

$$J(f) = \frac{1}{N_P} \sum_{j=1}^{N_P} \frac{1}{N^j} \sum_{i=1}^{N^j} \left| E^{ij}(f) - y^{ij} \right|^2 + \beta_R \left\| \frac{df}{dq} \right\|^2_{L^2(Q)},$$

where the regularization parameter β_R is a design parameter that controls trade-off between model misfit and smoothness of f. Though not smooth, a piecewise approximation to the derivative of $f(q)$ suffices for our purposes. The derivative on subinterval k can be approximated by $\frac{d_{k+1}-d_k}{\Delta q_k}, k = 0, \ldots, N_S - 1$ and therefore

$$\beta_R \left\| \frac{df}{dq} \right\|^2_{L^2} = \beta_R \int_Q \left| \frac{df}{dq}(q) \right|^2 dq \approx \beta_R \sum_{k=0}^{N_S} \left(\frac{d_{k+1} - d_k}{\Delta q_k} \right)^2 \Delta q_k \tag{4.5}$$

$$= \beta_R d^T A_R d, \tag{4.6}$$

where

$$A_R = \begin{bmatrix} -1 & & & \\ 1 & -1 & & \\ & \ddots & \ddots & \\ & & 1 & -1 \\ & & & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\Delta q_0} & & & \\ & \frac{1}{\Delta q_1} & & \\ & & \ddots & \\ & & & \frac{1}{\Delta q_{N_S-2}} \\ & & & & \frac{1}{\Delta q_{N_S-1}} \end{bmatrix} \begin{bmatrix} -1 & 1 & & & \\ & -1 & 1 & & \\ & & \ddots & \ddots & \\ & & & -1 & 1 \end{bmatrix}.$$

The modified quadratic programming problem objective function is

$$J(f(d)) = d^T \left(A + \beta_R A_R\right) d - 2d^T b + c.$$

4.2 Analysis of Inverse Problems

The single-patient inverse problem is typical nonlinear least squares and as such, its analysis has been thoroughly discussed in other contexts. We do not consider that case here, except later when issues with censored data arise.

In the case of the second, probability-distribution-based inverse problem, the cost functional (4.2) is minimized over a space of probability measures. To clarify the well-posedness and stability [10] of this inverse problem in the finite-dimensional computational framework, we follow a progression, drawing from the previous work of Banks–Pinter [13] and Banks–Bihari [4]:

- Define the Prohorov ρ metric to measure convergence of probability measures and demonstrate that given an appropriate set of functions $\mathcal{F}$ a set

$$\mathcal{P}_{\mathcal{F}} = \left\{ F \in \mathcal{P}(Q) | F = \int f, f \in \mathcal{F} \right\}$$

over which the minimization may take place is compact in this metric.

- Show that elements of this set $\mathcal{P}_{\mathcal{F}}$, characterized by their probability density functions f, can be approximated by piecewise linear splines from the set

$$\mathcal{P}^{N_S}(Q) = \left\{ F \in \mathcal{P}(Q) | F = \int f, f \in \mathcal{F}^{N_S} \right\}.$$

- Use results from Banks–Bihari to demonstrate that the inverse problem has a minimizer and is stable.

4.2.1 Minimization over a ρ compact set

In order to address the well-posedness of the inverse problem posed over probability distributions, one needs a measure of convergence of distributions. The Prohorov metric induces a topology on the probability distribution functions and thus, a means to measure distance between distributions.

Definition 4.1. *Prohorov metric ρ: If (Q, d) is a complete metric space, and for any closed subset $S \subset Q$ and $\epsilon > 0$ we define $S^\epsilon = \{q \in Q | d(\tilde{q}, q) < \epsilon, \tilde{q} \in S\}$, the Prohorov metric $\rho : \mathcal{P}(Q) \times \mathcal{P}(Q) \to \mathbb{R}^+$ is defined by*

$$\rho(P_1, P_2) \equiv \inf \{\epsilon > 0 : P_1[S] \leq P_2[S^\epsilon] + \epsilon, S \text{ closed}, S \subset Q\} \tag{4.7}$$

It is well known that

1. $(\mathcal{P}(Q), \rho)$ is a complete metric space; and
2. if Q is compact then $(\mathcal{P}(Q), \rho)$ is a compact metric space.

The following equivalent characterizations of the Prohorov metric [14] may be more intuitive and in particular, we will soon employ *(ii.)*.

Theorem 4.1. *Given $P_k, P \in \mathcal{P}(Q)$, the following convergence statements are equivalent:*

(i.) $\rho(P_k, P) \to 0$*;*

(ii.) $\int_Q g dP_k(q) \to \int_Q g dP(q)$ *for all bounded, uniformly continuous functions* $g : Q \to \mathbb{R}$*;*

(iii.) $P_k[A] \to P[A]$ *for all Borel sets* $A \subset Q$ *with* $P[\partial A] = 0$*.*

We consider minimizing our most general inverse problem cost function over the (infinite-dimensional) set $\mathcal{P}_{\mathcal{F}}(Q)$ of probability measures on the admissible parameter space Q that have underlying densities f taken from a subset $\mathcal{F}$ of $L^2(Q)$:

$$\mathcal{P}_{\mathcal{F}}(Q) = \left\{ F \in \mathcal{P}(Q) | F = \int f, f \in \mathcal{F} \right\}. \tag{4.8}$$

Here $\mathcal{P}(Q)$ denotes the set of *all* probability distributions on Q. The following theorem gives a condition for $\mathcal{P}_{\mathcal{F}}(Q)$ to be compact in $\mathcal{P}(Q)$, and we offer a paraphrased version of its proof.

Theorem 4.2. *(Banks–Pinter [13] Theorem 5.1) If $\mathcal{F}$ is a weakly compact subset of $L^2(Q)$, with Q compact, then $\mathcal{P}_{\mathcal{F}}(Q)$ as defined in* (4.8) *above is a ρ-compact subset of $(\mathcal{P}(Q), \rho)$.*

Proof. Take a sequence $\{F_n\} \subset \mathcal{P}_{\mathcal{F}}(Q)$, so each element of the sequence is identified with a density $f_n \in \mathcal{F}$, i.e. $F_n = \int f_n dq$. Weak compactness of $\mathcal{F}$ implies that $\mathcal{F}$ is norm bounded (V.6.1 in Dunford–Schwartz [22]) and therefore $\{f_n\} \subset \mathcal{F}$ is weakly sequentially compact (every sequence has a weakly convergent subsequence). In this case, there exists an $f \in \mathcal{F}$ and a subsequence $\{f_{n_k}\}$ of $\{f_n\}$ such that $f_{n_k} \rightharpoonup f$ with

$\|f_{n_k}\|_{L^2} \leq M$. Since this limit holds for any test function in $L^2(Q)$, it does specifically for a function $g \in C(Q)$ (bounded continuous functions on Q):

$$\int_Q g f_{n_k} dx \to \int_Q g f dx$$

and therefore $F_{n_k} \to F = \int f ds$ in the ρ metric by Theorem 4.1, so $\mathcal{P}_{\mathcal{F}}(Q)$ is sequentially compact.

We now show that the set $\mathcal{P}_{\mathcal{F}}(Q)$ is closed in the ρ metric by examining a sequence $\{F_n\} \subset \mathcal{P}_{\mathcal{F}}(Q)$ such that $F_n \to F \in \mathcal{P}(Q)$ and must show that $F \in \mathcal{P}_{\mathcal{F}}(Q)$. Again, each $F_n = \int f_n ds$ and since convergence in Prohorov metric is equivalent to weak* convergence in $\mathcal{P}(Q)$,

$$\int_Q g f_n dx \to \int_Q g dF \tag{4.9}$$

for every $g \in C(Q)$. As above, since $\{f_n\} \subset \mathcal{F}$ (which is weakly compact) there exists $\{f_{n_k}\}$ and $f \in \mathcal{F}$ such that

$$\int_Q g f_{n_k} dx \to \int_Q g f dx$$

for all $g \in C(Q)$. Now, f_{n_k} is bounded in $L^2(Q)$ and $C(Q)$ is dense in $L^2(Q)$, so

$$\int_Q h f_{n_k} dx \to \int_Q h f dx$$

for all $h \in L^2(Q)$. (To see this (N, ϵ)-style use $hf_{n_k} - g_l f_{n_k} + g_l f_{n_k} - g_l f + g_l f - hf$ for $\|g_l - h\|_{L^2} \to 0$.) Therefore by uniqueness of limits and (4.9),

$$\int_Q \tilde{g} dF = \int_Q \tilde{g} f dx \ \ \forall \ \ \tilde{g} \in L^2(Q);$$

that is, F is absolutely continuous with $F' = f$, so $F \in \mathcal{P}_{\mathcal{F}}(Q)$, so $\mathcal{P}_{\mathcal{F}}(Q)$ is compact

in $(\mathcal{P}(Q), \rho)$. □

Closed balls are an example of a weakly compact set in $L^2(Q)$. Indeed, closed, bounded, convex sets completely characterize weakly compact sets in the reflexive space $L^2(Q)$. This follows from two theorems in Dunford–Schwartz [22]: from II.3.28, a set in a reflexive space is weakly sequentially compact if and only if it is bounded; and from V.3.13, a convex subset of a locally convex linear topological space is weakly closed if and only if it is closed.

4.2.2 Approximation by piecewise linear splines

In order to show method stability in the manner of Banks–Bihari [4], we first establish that convergence of piecewise linear densities f^{N_S} to a continuous density f implies convergence of the corresponding distributions in the Prohorov metric.

Theorem 4.3. *Let $F \in \mathcal{P}(Q)$, where $F = \int f$, and define a series of piecewise linear spline approximations to the density f,*

$$f(q) \approx f^{N_S}(q) = \sum_{k=1}^{N_S} d_k^{N_S} \phi_k^{N_S}(q),$$

such that $f^{N_S} \to f$ in $L^2(Q)$ as $N_S \to \infty$. Then $F^{N_S} = \int f^{N_S}$ are such that $\rho(F^{N_S}, F) \to 0$.

Proof. Strong convergence of f^{N_S} to f implies weak convergence in $L^2(Q)$, so

$$\int_Q g f^{N_S} dq \to \int_Q g f dq$$

for all $g \in L^2(Q)$ and consequently for all $g \in C(Q)$, so for $F^{N_S} = \int f^{N_S}$,

$$\rho(F^{N_S}, F) \to 0.$$

□

Therefore, letting

$$\mathcal{F}^{N_S} = \left\{ h \in L^2(Q) | h(q) = \sum_{k=1}^{N_S} d_k^{N_S} \phi_k^{N_S}(q) \right\},$$

($L^2(Q)$ functions with piecewise linear N_S representations), the set

$$\mathcal{P}_D(Q) = \left\{ F \in \mathcal{P}(Q) | F = \int f, f \in \cup_1^\infty \mathcal{F}^{N_S} \right\} = \cup_{N=1}^\infty \mathcal{P}^{N_S}(Q)$$

is dense in $\mathcal{P}_\mathcal{F}(Q)$ in the ρ metric (i.e., $f^N \to f$). We can use elements of

$$\mathcal{P}^{N_S}(Q) = \left\{ F \in \mathcal{P}(Q) | F = \int f, f \in \mathcal{F}^{N_S} \right\}$$

to approximate elements of $\mathcal{P}_\mathcal{F}(Q)$ for computational purposes. Note that since it is finite dimensional, $\mathcal{P}^{N_S}(Q)$ is a closed subset of $\mathcal{P}(Q)$. One can see this by intersecting the closed set of piecewise linear N_S functions with the closed set of densities f that integrate to one.

4.2.3 Existence of a minimizer and method stability

For a fixed set of data y, we wish to minimize our cost criterion

$$\min_{P \in \mathcal{P}_\mathcal{F}(Q)} J(P) = \sum_{i=1}^{N^j} \left| \int_Q z^{ij}(q) dP(q) - y^{ij} \right|^2$$

over the set of distributions $\mathcal{P}_{\mathcal{F}}(Q)$ (we again omit the state index for clarity) and denote the set of minimizers of this problem by $P^*(y)$. We also define a series of approximate problems based on the piecewise linear splines described above. The cost functions look like

$$\min_{P_{N_S} \in \mathcal{P}^{N_S}(Q)} J(P_{N_S}) = \sum_{i=1}^{N^j} \left| \int_Q z^{ij}(q) dP_{N_S}(q) - y^{ij} \right|^2,$$

where $P^*_{N_S}(y)$ denotes the set of minimizers to this approximate problem. As above, we assume Q is compact and $\mathcal{F}$ is weakly compact.

We wish to show existence of a minimizer and method stability for this approximation scheme. Method stability in this case means that $dist(P^*_{N_S}, P^*) \to 0$, where $dist$ denotes the usual Hausdorff distance between the two sets: $dist(A, B) = \inf\{\rho(P_1, P_2) : P_1 \in A, P_2 \in B\}$.

Solutions $z^{ij}(q)$ to the system of differential equations depend continuously on the parameters q, i.e., $q \to z^{ij}(q)$ is a continuous map from Q into $\mathbb{R}$ for each i, j pair (and therefore bounded since a continuous map on a compact set). This follows from standard results for differential equations, appealing to continuity of the function $\bar{g}(t, \bar{x}, q)$ with respect to q and continuity of the composite linear observer and log-scaling map from $\bar{x}$ to z. Now, when a generic sequence $P_{N_S} \to P$ in $\mathcal{P}(Q)$, by the

equivalence of convergence in Prohorov metric with weak* convergence in $\mathcal{P}(Q)$,

$$\begin{aligned}
\lim_{N\to\infty} J(P_{N_S}) &= \lim_{N_S\to\infty} \sum_{i=1}^{N^j} \left| E\left[z^{ij}(q)|P_{N_S}\right) \right] - y^{ij}\Big|^2 && (4.10a)\\
&= \lim_{N_S\to\infty} \sum_{i=1}^{N^j} \left| \int_Q z^{ij}(q) dP_{N_S}(q) - y^{ij} \right|^2 && (4.10b)\\
&= \sum_{i=1}^{N^j} \left| \int_Q z^{ij}(q) dP(q) - y^{ij} \right|^2 && (4.10c)\\
&= \sum_{i=1}^{N^j} \left| E\left[y^{ij}|P\right) \right] - y^{ij}\Big|^2 && (4.10d)\\
&= J(P).
\end{aligned}$$

So when $P_N \to P$, the approximate *problems* "converge" to the exact problem, i.e., $J(P) : \mathcal{P}(Q) \to \mathbb{R}$ is a continuous map. When there is an underlying representative density $f_{N_S} \to f$ the same is true since $dP_{N_S}(q) = f_{N_S}(q)$.

Since J is continuous on the compact set $\mathcal{P}_{\mathcal{F}}(Q)$, there exists at least one minimizer $P^* \in P^*(y)$ to the infinite dimensional problem. Since the subset $\mathcal{P}^{N_S}(Q)$ is a closed subset of the compact set $\mathcal{P}(Q)$ it is therefore compact under the ρ metric. The cost J is continuous on this compact set, so there exists at least one minimizer $P^*_{N_S}$ to the approximate problem. Let $P^*_{N_S}(y)$ denote the set of minimizers of $J(P)$ over $\mathcal{P}^{N_S}(Q)$.

Let $\left\{P^*_{N_S}\right\}, N_S \in \mathbb{N}$ be a sequence of minimizers in $\mathcal{P}_{\mathcal{F}}(Q)$, where each $P^*_{N_S} \in P^*_{N_S}(y)$. Then by compactness of $\mathcal{P}_{\mathcal{F}}(Q)$, there exists a ρ-convergent subsequence $P^*_{N_{S_l}}$ such that

$$\lim_{N_{S_l}\to\infty} P^*_{N_{S_l}} = \tilde{P} \in \mathcal{P}_{\mathcal{F}}(Q). \tag{4.11}$$

Note that for any $P_{N_{S_l}} \in \mathcal{P}^{N_{S_l}}(Q)$, at the minimizer,

$$J(P^*_{N_{S_l}}) \leq J(P_{N_{S_l}}).$$

Also it follows from the definition of the cost criterion $J(P)$, (4.10a), and (4.11), that

$$\begin{aligned}\lim_{N_{S_l}\to\infty} J(P^*_{N_{S_l}}) &= \lim_{N_{S_l}\to\infty} \sum_{i=1}^{N^j} \left| E\left[z^{ij}(q)|P^*_{N_{S_l}}\right] - y^{ij}\right|^2 \\ &= \sum_{i=1}^{N^j} \left| E\left[z^{ij}(q)|\tilde{P}\right] - y^{ij}\right|^2 \\ &= J(\tilde{P}).\end{aligned}$$

Now we involve the convergence of the approximations. Let $P \in \mathcal{P}_{\mathcal{F}}(Q)$. Since $\mathcal{P}_D$ as defined above is dense in $\mathcal{P}_{\mathcal{F}}(Q)$, there exists a sequence $\left\{P_{N_{S_l}}\right\}, P_{N_{S_l}} \in P^{N_{S_l}}(Q)$ such that $P_{N_{S_l}} \to P$ as $N_{S_l} \to \infty$. Thus it follows that

$$\begin{aligned}\lim_{N_{S_l}\to\infty} J(P_{N_{S_l}}) &= \lim_{N_{S_l}\to\infty} \sum_{i=1}^{N^j} \left| E\left[z^{ij}(q)|P_{N_{S_l}}\right] - y^{ij}\right|^2 \\ &= \sum_{i=1}^{N^j} \left| E\left[z^{ij}(q)|P\right] - y^{ij}\right|^2 \\ &= J(P),\end{aligned}$$

and putting these results all together:

$$\begin{array}{ccc} J(P^*_{N_{S_l}}) & \leq & J(P_{N_{S_l}}) \\ \downarrow & & \downarrow \\ J(\tilde{P}) & & J(P). \end{array}$$

Thus $\tilde{P}$ is a minimizer of $J(P)$ over $P \in \mathcal{P}_{\mathcal{F}}(Q)$; that is, $\tilde{P} \in P^*(y)$.

Summary: Any sequence $P^*_{N_S}$ in $P^*_{N_S}(y)$ has a subsequence $P^*_{N_{S_l}}$ such that $P^*_{N_{S_l}} \to \tilde{P} \in P^*(y)$. Therefore $dist(P^*_{N_{S_l}}(y), P^*(y)) \to 0$ as $N_{S_l} \to \infty$. Each finite-dimensional problem has a minimizer and these can be used to approximate the minimizers of the

infinite-dimensional problem.

The argument presented here presumes a single fixed data set y; however, method stability requires that if a sequence of data sets $\{y^k\}$ converge ($y^k \to y$ as $k \to \infty$), the corresponding minimizers converge: $P^*_{N_S}(y^k) \to P^*(y)$ as both $k \to \infty$ and $N_S \to \infty$. This means the solutions depend continuously on the data as well and that the problems are method stable. The arguments in that more detailed case similarly appeal to the continuity of J with respect to the data and are summarized in [8], with details in [4].

4.3 Statistical Theory and Methods

4.3.1 Confidence intervals

In performing an inverse problem via ordinary least squares to determine an estimate of either a vector $q \in \mathbb{R}^p$ of parameters or the probability distribution $f(q)$ for a particular parameter (in our case parameterized by coefficients d_k), it is vital to understand the uncertainty of the process used to obtain that estimate. In particular, in the case of estimating a finite set of parameters q, we wish to determine confidence intervals $q_l \pm \tilde{\nu}_l$, where $\tilde{\nu}_l$ is based on a calculated standard error for parameter component l. The method presented here is based on that described in Davidian and Giltinan [20] and Carroll and Ruppert [18]. Other examples in the context of inverse problems with nonlinear models can be seen in [1] and [5]. A derivation of these results using linearization can be found in the latter.

We first focus on estimating a Euclidean vector of parameters q in the context of a single patient inverse problem and therefore temporarily suppress the patient index j. Consider the two-output case where the observed time/data pairs correspond to total

CD4 count $(t_1^i, y_1^i), i = 1, \ldots, N_1$, and total viral load count $(t_2^i, y_2^i), i = 1, \ldots, N_2$, where the subscript indexes the state. As before, we correspondingly subscript the log-scaled model solutions z_s, $s = 1, 2$. Recall that in the single patient inverse problem where we fit two states to data on a logarithmic scale, the problem is

$$q^* = \arg\min_{q \in Q} J(q) = \sum_{i=1}^{N_1} \left| z_1(t_1^i; q) - y_1^i \right|^2 + \sum_{i=1}^{N_2} \left| z_2(t_2^i; q) - y_2^i \right|^2 .$$

(We do not scale by the number of time points in this derivation or in computation as the numbers of measurements for each state are comparable.) The assumption for the variance model (3.5) implies that the log-scaled observations y_s^i are normally distributed. Adding the assumption that longitudinal measurements are independent and taking into account the inverse problem process used to obtain the optimal estimate q^*, we have from asymptotic statistical theory that

$$q^* \sim \mathcal{N}\left(q^0, \Sigma\right),$$

where q^0 denotes the true underlying mean parameter vector (see, e.g., page 14 in Carroll and Ruppert [18]). The covariance matrix Σ incorporates the following intermediate results. We combine the model responses into a vector

$$Z(q) = \left[z_1(t_1^1; q), z_1(t_1^2; q), \ldots, z_1(t_1^{N_1}; q), z_2(t_2^1; q), z_2(t_2^2; q), \ldots, z_2(t_2^{N_2}; q) \right]^T$$

and let Z_q denote the $(N_1 + N_2) \times p$ matrix of partial derivatives of Z with respect to each estimated parameter:

$$[Z_q]_{il} = \left[\frac{\partial Z_i}{\partial q_l} \right].$$

Note that each component of Z_q depends on the solution to the sensitivity equations

discussed in Section 3.4. Finally since on the log scale the variance for each state is assumed constant and observation independent, the diagonal weighting matrix is

$$W(\sigma) = \begin{bmatrix} \sigma_1^2 I_{N_1} & \\ & \sigma_2^2 I_{N_2} \end{bmatrix},$$

and finally,

$$\Sigma(q^0, \sigma) = \left\{ Z_q(q^0)^T Z_q(q^0) \right\}^{-1} Z_q(q^0)^T W(\sigma) Z_q(q^0) \left\{ Z_q(q^0)^T W(q^0, \sigma) Z_q(q^0) \right\}^{-1}$$

As in Banks–Bokil [5], this formulation of Σ essentially describes the effect of the inverse problem process on the errors in the observed data by accounting for the sensitivity of the model solution with respect to estimated parameters.

In practice, when working with clinical data, the true values of the mean q^0 and variances σ_s^2 are unknown and must be approximated as part of the inverse problem process. Following standard statistical practice, we substitute the estimated parameters q^* for q^0 and an unbiased estimator of the variance $\hat{\sigma}_s^2 = \frac{1}{N_s - p} \sum_{i=1}^{N_s} |z_s(t_s^i; q^*) - y_s^i|^2$ for σ_s. Standard errors for parameter component l are then obtained by taking $\nu_l = \sqrt{(\Sigma(q^*, \hat{\sigma}_s)_{ll})}$. These in turn can be used to compute confidence intervals at the $(1 - \alpha)$ level (often $(1 - \alpha) = 0.95$ for 95% confidence intervals) for each parameter component,

$$\left[q_l^* - t_{1-\alpha/2} \nu_l(q^*), q_l^* + t_{1-\alpha/2} \nu_l(q^*) \right], l = 1, \ldots, p,$$

indicating that $(1 - \alpha) \times 100\%$ of intervals constructed through this process would cover the true value of the parameter q^0. Here $t_{(1-\alpha/2)}$ denotes the value such that Student's t-distribution with $N_1 + N_2 - p$ degrees of freedom has a tail probability of $1 - \alpha/2$.

Assuming that estimates of variance σ_s^2 in the data are available, this calculation applies to the case of density estimation as well. Here, the unknown parameters are the components of the vector d_k and we use the above results to construct confidence intervals at the spline knots q_k. A schematic of such piecewise linear confidence bands

$$f^- = \sum_{k=0}^{N_S} (d_k^* - \bar{\nu}_k)\phi_k(q) \leq f^* \leq \sum_{k=0}^{N_S} (d_k^* + 2\bar{\nu}_k)\phi_k(q) = f^+ \tag{4.12}$$

is shown in Figure 4.2. These bands only provide an estimate of confidence at the

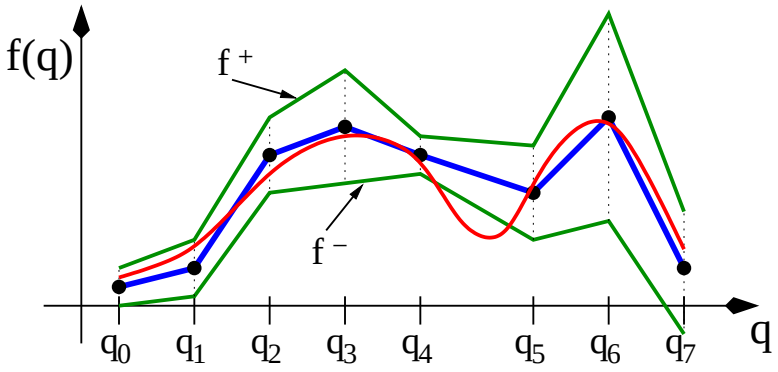


Figure 4.2: Example of spline approximation to density function $f(q)$ including nodal confidence intervals and an example of a true density not contained in them.

nodes and as illustrated, the true density function might never be contained in them.

Note that in the density-estimation case, model solutions from the differential equation system are a fixed basis and the operator $E^{ij}(d) : \mathbb{R}^p \to \mathbb{R}$ is linear, so estimates of standard errors are exact, rather than linear approximations via sensitivity equations as above. The matrix Z_q in the results above is replaced by derivatives of $E^{ij}(f) \approx E^{ij}(d)$ with respect to d_k. The entries of the matrix no longer require sensitivity computations, but rather are given explicitly by

$$[Z_q]_{il} = \frac{\partial E^{ij}(d)}{\partial d_l} = \int_Q z^{ij}(q)\phi_l(q)dq.$$

Since the inverse problem process for probability density estimation may include a regularization term it must be accounted for when determining standard errors. The cost function, dependent on a vector of spline coefficients d is

$$J(f(d)) = \frac{1}{N_P}\sum_{j=1}^{N_P}\sum_{s=1}^{2}\frac{1}{N_s^j}\sum_{i=1}^{N^j}\left|E_s^{ij}(f(d)) - y_s^{ij}\right|^2 + \beta_R\left\|\frac{df}{dq}\right\|_{L^2}^2,$$

or equivalently,

$$J(f(d)) = \frac{1}{2}\left(\mathcal{E}d - \tilde{y}\right)^T\left(\mathcal{E}d - \tilde{y}\right) + \beta_R d^T A_R d.$$

Here $\tilde{y}$ denotes the scaled vector of data: $\tilde{y}_s^{ij} = y_s^{ij}\sqrt{\frac{2}{N_P N_s^j}}$ and $\mathcal{E}$ the similarly scaled linear operator that maps the spline coefficients to expected values by integrating against model basis functions. Specifically, each row of $\mathcal{E}$ is comprised of entries $\mathcal{E}_s^{ij} = \mathcal{S}_s^{ij}\sqrt{\frac{2}{N_P N_s^j}}$, where $\mathcal{S}_s^{ij}$ is the row vector of model solutions integrated against spline basis functions (as defined in Section 4.1) for a particular (patient j, state s, time point t_s^{ij}) triple.

If d^0 denotes the coefficients in the approximation of the true underlying density f^0, let $d = d^0 + \Delta d$. As is done similarly in Banks–Bokil [5], we derive an ordinary least squares estimating equation for Δd and use it to determine variance of estimates $\mathrm{Var}(d^*) = \mathrm{Var}(d^0 + \Delta d) = \mathrm{Var}(\Delta d)$.

Denoting the error in fit to data by $D_e = \tilde{y} - \mathcal{E}d^0$ and writing $L_R^T L_R = 2\beta_R A_R$, we have

$$\begin{aligned}
\Delta d_{OLS}^* &= \min_{\Delta d}\frac{1}{2}\left[\left(\mathcal{E}(d^0+\Delta d) - \tilde{y}\right)^T\left(\mathcal{E}(d^0+\Delta d) - \tilde{y}\right) + (d^0+\Delta d)^T L_R^T L_R (d^0+\Delta d)\right] \\
&= \min_{\Delta d}\frac{1}{2}\left[\left(\mathcal{E}\Delta d - D_e\right)^T\left(\mathcal{E}\Delta d - D_e\right) + (d^0+\Delta d)^T L_R^T L_R (d^0+\Delta d)\right] \\
&= \min_{\Delta d}\frac{1}{2}\left[\Delta d^T\mathcal{E}^T\mathcal{E}\Delta d - 2D_e^T\mathcal{E}\Delta d + D_e^T D_e + \right. \\
&\qquad \left. \Delta d^T L_R^T L_R \Delta d + 2{d^0}^T L_R^T L_R \Delta d + {d^0}^T L_R^T L_R d^0\right].
\end{aligned}$$

Differentiating the right side of this with respect to Δd and setting equal to zero yields the optimality condition for Δd^*_{OLS}:

$$\mathcal{E}^T\mathcal{E}\left(\Delta d^*_{OLS}\right) - \mathcal{E}^T D_e + L_R^T L_R\left(\Delta d^*_{OLS}\right) + L_R^T L_R d^0 = 0,$$

and therefore the estimator

$$\Delta d^*_{OLS} = \left(\mathcal{E}^T\mathcal{E} + L_R^T L_R\right)^{-1}\left(\mathcal{E}^T D_e - L_R^T L_R d^0\right).$$

The terms involving d^0 are exact and $\operatorname{Var} D_e = \operatorname{Var}(\tilde{y})$. Therefore

$$\begin{aligned}
\operatorname{Cov}(d^*_{OLS}) &= \operatorname{Cov}\left\{\left(\mathcal{E}^T\mathcal{E} + L_R^T L_R\right)^{-1}\mathcal{E}^T D_e\right\} \\
&= \left(\mathcal{E}^T\mathcal{E} + L_R^T L_R\right)^{-1}\mathcal{E}^T \operatorname{Cov}(\tilde{y})\mathcal{E}\left(\mathcal{E}^T\mathcal{E} + L_R^T L_R\right)^{-1} \\
&= \left(\mathcal{E}^T\mathcal{E} + L_R^T L_R\right)^{-1}\mathcal{E}^T D(\sigma_1,\sigma_2)\mathcal{E}\left(\mathcal{E}^T\mathcal{E} + L_R^T L_R\right)^{-1} \\
&= \Sigma_R, \qquad (4.13)
\end{aligned}$$

and the standard errors, as before, are $\nu_l = \sqrt{\Sigma_{R_{ll}}}$. The diagonal matrix D consists of entries σ_1^2 and σ_2^2 arranged to respect the ordering of the vector $\tilde{y}$ and matrix $\mathcal{E}$, where the σ_s^2 are estimates of the variance for fitting the expected value to the data.

Some care is necessary when computing estimates of variance in the density estimation case as we do not fit the differential equation model directly to each patient's data, where a sum of squared residuals would yield information about the variance in the particular patient's data. In this context, even when assuming each patient's longitudinal data are independent for each time, one should consider a more structured variance model that accounts for inter-patient correlation. As a first step, however, we consider the expected value (conditioned on the model and probability distribu-

tion) the "model" and seek a common estimate of variance among all patients. In this situation, the sum of squared residuals gives a preliminary estimate of variance to be used in computing standard errors:

$$\hat{\sigma}_s^2 = \frac{1}{N_P} \sum_{j=1}^{N_P} \frac{1}{N_s^j} \sum_{i=1}^{N_s^j} \left| E_s^{ij}(f(d)) - y_s^{ij} \right|^2 . \tag{4.14}$$

4.3.2 Censored data methodology

In this section we focus on handling the censored data points for a particular patient. Unscaled measurements of viral load (second observed component, $\bar{y}_2^i$) are censored when below the limit of detection, at either $\bar{L}1 = 400$ or $\bar{L}2 = 50$. According to our presumed error model (3.5), the log-scaled measurements

$$y_2^i \sim \mathcal{N}(\zeta_2(t^i; q), \sigma_2^2)$$

where ζ and σ are given by (3.6). Denote the log-scaled censor points by $L1 = \log_{10} \bar{L}1$, $L2 = \log_{10} \bar{L}2$. For censored data points, the available knowledge is that the observed value $y_2^i \leq L^i$, where L^i denotes the relevant censor point ($L^i = L1$ or $L^i = L2$) at time t^i.

In this context we observe pairs $(w^i, \chi^i), i = 1, \ldots, N$, where

$$\begin{aligned} w^i &= \begin{cases} y_2^i & \text{if } y_2^i > L^i \\ L^i & \text{if } y_2^i \leq L^i \end{cases} \\ \chi^i &= I_{(y_2^i > L^i)}, \end{aligned}$$

and $I_{(\cdot)}$ is the indicator function. Defining

$$\phi(\xi) = \frac{1}{\sqrt{2\pi}} e^{-\xi^2/2} \qquad \text{(standard normal pdf), and}$$
$$\Phi(\xi) = \int_{-\infty}^{\xi} \phi(s)ds \qquad \text{(standard normal cdf),}$$

the viral load portion of the likelihood function for (q, σ_2) given the observations w^i is

$$\bar{\mathcal{L}}(q, \sigma_2) = \prod_{i=1}^{N} \left[\frac{1}{\sigma_2} \phi\left(\frac{w^i - \zeta_2^i}{\sigma_2} \right) \right]^{\chi^i} \left[\Phi\left(\frac{w^i - \zeta_2^i}{\sigma_2} \right) \right]^{1-\chi^i},$$

where the first term accounts for the probability of observing w^i given that it is uncensored and the second term the probability that the observation is in the interval $(-\infty, L^i)$ when censored. This is using a truncated normal distribution for the censored measurements. The log-likelihood is

$$\begin{aligned} & \mathcal{L}(q, \sigma_2) \\ = & \sum_{i=1}^{N} \left(\chi^i \left[\log \phi\left(\frac{w^i - \zeta_2^i}{\sigma_2} \right) - \log \sigma_2 \right] + (1 - \chi^i) \left[\log \Phi\left(\frac{w^i - \zeta_2^i}{\sigma_2} \right) \right] \right) \\ = & \sum_{i=1}^{N} \left(\chi^i \left[\log \phi\left(\frac{y_2^i - \zeta_2^i}{\sigma_2} \right) - \log \sigma_2 \right] + (1 - \chi^i) \left[\log \Phi\left(\frac{L^i - \zeta_2^i}{\sigma_2} \right) \right] \right), \end{aligned} \tag{4.15}$$

which we maximize to estimate q and σ_2. This is analogous to the typical log likelihood estimator in the absence of a limit of detection, where

$$\mathcal{L}(q, \sigma_2) = -\frac{N}{2} \log 2\pi - N \log \sigma_2 - \sum_{i=1}^{N} \frac{(y_2^i - \zeta_2^i)^2}{2\sigma_2^2}. \tag{4.16}$$

However, while maximizing (4.16) in the parameters q is equivalent to minimizing the sum of squared residuals $\sum_{i=1}^{N} (y_2^i - \zeta_2^i)^2$ (typical least squares) and the estimation of q and σ_2 decouple, maximizing (4.15) is not as simple, since a joint estimation

of q and σ_2 must be performed. Maximizing (4.15) is possible with the Expectation Maximization (EM) algorithm [21, 37], which iteratively improves the estimates of q and σ_2.

First, with the assumptions about distributions made above, let $\xi^i = \frac{L^i - \zeta_2^i}{\sigma_2}$ and $\Lambda(\xi^i) = \frac{\phi(\xi^i)}{\Phi(\xi^i)}$ and use properties of a truncated normal distribution to obtain

$$\begin{aligned} E\left[y_2^i | y_2^i \leq L\right] &= \zeta_2^i - \sigma_2 \Lambda(\xi^i), \text{ and} \\ E\left[(y_2^i)^2 | y_2^i \leq L\right] &= (\zeta_2^i)^2 - 2\sigma_2 \zeta_2^i \Lambda(\xi^i) - \sigma_2^2 \xi^i \Lambda(\xi^i) + \sigma_2^2. \end{aligned}$$

These can be used to update the data points and estimate of squared residuals for the second observed state by the following

$$\begin{aligned} \tilde{y}^i &= \chi^i y_2^i + (1 - \chi^i) E\left[y_2^i | y_2^i \leq L^i\right] \\ &= \chi^i y_2^i + (1 - \chi^i)\left[\zeta_2^i - \sigma_2 \Lambda(\xi^i)\right] \end{aligned} \tag{4.17}$$

and

$$\begin{aligned} \tilde{r}^i &= \chi^i E\left[(y_2^i - \zeta_2^i)^2\right] + (1 - \chi^i) E\left[(y_2^i - \zeta_2^i)^2 | y_2^i \leq L\right] \\ &= \chi^i (y_2^i - \zeta_2^i)^2 + (1 - \chi^i)\left\{E\left[(y_2^i)^2 | y_2^i \leq L^i\right] - 2\zeta_2^i E\left[(y_2^i) | y_2^i \leq L^i\right] + (\zeta_2^i)^2\right\} \\ &= \chi^i (y_2^i - \zeta_2^i)^2 + (1 - \chi^i)\sigma_2^2\left[1 - \xi^i \Lambda(\xi^i)\right]. \end{aligned} \tag{4.18}$$

We proceed with the EM Algorithm.

Algorithm 4.1. *Expectation Maximization (EM) Algorithm*

Step 1 (Initialize) Create adjusted data $\tilde{y}^i$ by replacing censored y_2^i values (those for which $\chi^i = 0$) by $L^i/2$, and use ordinary least squares to estimate $\hat{q}^{(0)}$ using both CD4 data y_1^i and viral RNA data $\tilde{y}^i$ (which includes replaced censored

values). Get an initial estimate for σ_2^2 *from*

$$(\hat{\sigma}_2^{(0)})^2 = \frac{1}{N_2}\sum_{i=1}^{N_2}\left|\tilde{y}^i - \zeta_2(t_2^i;\hat{q}^{(0)})\right|^2.$$

Set $k = 0$.

Step 2 Define $\hat{\zeta}_2^{i(k)} = \zeta_2(t^i;\hat{q}^{(k)})$ *and* $\hat{\xi}^{i(k)} = \frac{L^i - \hat{\zeta}_2^{i(k)}}{\hat{\sigma}_2^{(k)}}$ *and update the data and residuals by*

$$\tilde{y}^{i(k)} = \chi^i y_2^i + (1-\chi^i)\left[\hat{\zeta}_2^{i(k)} - \hat{\sigma}_2^{(k)}\Lambda(\hat{\xi}^{i(k)})\right] \quad (4.19)$$

$$\tilde{r}^{i(k)} = \chi^i(y_2^i - \hat{\zeta}_2^{i(k)})^2 + (1-\chi^i)(\hat{\sigma}_2^{(k)})^2\left[1 - \hat{\xi}^{i(k)}\Lambda(\hat{\xi}^{i(k)})\right]. \quad (4.20)$$

Step 3 Update the estimates to $\hat{q}^{(k+1)}, \hat{\sigma}_2^{(k+1)}$ *by performing ordinary least squares minimization in* q

$$\hat{q}^{(k+1)} = \arg\min_q \sum_{i=1}^{N_1}\left|y_1^i - \zeta_1(t_1^i;q)\right|^2 + \sum_{i=1}^{N_2}\left|\tilde{y}^{i(k)} - \zeta_2(t_2^i;q)\right|^2$$

and computing

$$(\hat{\sigma}_2^{(k+1)})^2 = \frac{1}{N_2}\sum_{i=1}^{N_2}\tilde{r}^{i(k)}.$$

If relative changes in $\hat{q}$ *and* $\hat{\sigma}$ *are small, terminate. Otherwise set* $k = k+1$ *and go to Step 2.*

This iterative process yields estimates of the parameters, variance, and expected values of the data at times where censored observations were recorded. Note that the model on the log scale is given by

$$\zeta^{ij} = \log_{10}\bar{z}^{ij} - \frac{\log_{10}(\bar{\sigma}^2+1)}{2},$$

dependent on the variance $\bar{\sigma}^2$, which we iteratively estimate. However, the second term $\frac{\log_{10}(\bar{\sigma}^2+1)}{2}$ for reasonable values of $\bar{\sigma}$ is small in comparison to $\log_{10} \bar{z}^{ij}$, so as in the inverse problems above we disregard it and substitute the log-scaled $z^{ij} = \log_{10} \bar{z}^{ij}$ directly for ζ^{ij} when performing these calculations.

4.4 Computational Methods

In the case of individual patient estimates, using the censored data methodology, a combination of sampling and gradient-based methods is employed. Currently for each of the single patients in PS59, we do the following:

1. Estimate all model parameters and initial conditions using 100,000 function evaluations with the `DIRECT` sampling algorithm. We use the Matlab implementation by Finkel [24].

2. Use Matlab's `lsqnonlin` to refine initial parameter estimates with relative function and iterate tolerance 10^{-8}. This algorithm is a subspace trust region method and is based on the interior-reflective Newton method described in [19].

3. Employ the censored data algorithm to iteratively update the estimates of parameters, variance, and censored data values, using `lsqnonlin` at each step.

Solution of the quadratic programming problem is via the medium-scale algorithm in the Matlab optimization toolbox function `quadprog`. The medium-scale algorithm (employed due to the constraints on this problem) is an active set projection method [27]. In all cases, we presume no knowledge of the true probability density function being estimated, so we use a uniform distribution on $[q_L, q_U]$ as the initial iterate for the optimizer. Moreover we tested the algorithm with various initial iterates for the

spline coefficient vector, and the quadratic programming algorithm always converged to the same optimal vector, suggesting that the approach is robust to choice of initial iterate.

Chapter 5

Method Validation with Simulated Data

Before applying inverse problem methods to clinical data we validate them on simulated data to better understand their behavior. Section 5.1 contains results for estimating probability density functions from simulated data generated under three different scenarios. We demonstrate the effect of varying the number of patients, varying the number of splines used in approximation, and applying regularization. We then compare results for simulated patients who have nearly identical underlying dynamics parameters to a cohort in which each patient's behavior is governed by one of eighteen sets of dynamic parameters. In Section 5.2 we validate censored data methods on simulated data.

5.1 Estimation of PDFs and Confidence Bands

To test the probability density function estimation methods, we choose a probability distribution f^0 to represent the "truth" of how a particular parameter q will be

distributed across a virtual (simulated) patient population, while the remaining model parameters $\grave{q}$ remain fixed. We generate virtual patients by sampling parameters $\hat{q}$ from this true distribution using Matlab's random number generators `rand` and `randn`. We then create corresponding model solutions $z_s(t_s^{ij}; \grave{q}; \hat{q})$, and add noise with coefficient of variation $\bar{\sigma} = 0.2$ to yield data with representative measurement error as described previously. Then, presuming no prior knowledge of the form of the underlying true density, we apply the inverse problem method (4.2) to the data in hopes of recovering the true density. That is, we use a uniform initial iterate for the quadratic programming problem. If some estimate or presumption of the distribution were available, it could alternatively be used as an initial iterate.

We consider three types of virtual patient cohort:

1. **Uniform simulated cohort:** Members of the first simulated patient cohort all have the same treatment protocol, observation times, initial conditions, and, with the exception of the parameter distributed across the population, model dynamic parameters. The treatment regimen and observation times are based on Patient 4 who has a total of 154 viral load and 107 CD4 measurements, four on and four off treatment periods, and a total observation period of 1919 days. Model parameters are set as in Table 3.2, except for efficacy values $\epsilon_1 = 0.7, \epsilon_2 = 0.4$, and the initial condition $\bar{x}^0 = [600, 50, 60, 50, 900000, 10, 20]^T$ to simulate acute infection.

2. **Treatment-varied simulated cohort:** Members of the second simulated patient cohort will still have the same underlying dynamics parameters and initial condition as described for the first. However in this set of simulated data each patient will be assigned one of eighteen different time-varying treatment schemes $\{u^j(t)\}_{j=1}^{18}$ and sampling schemes for CD4 (t_1^{ij}) and viral load (t_2^{ij}). In partic-

ular, we consider regimens and sampling times from the set PS18 of patients with on/off treatment ratios between 30–70%. (PS18 is described in the clinical data description in Chapter 2.) From these we generate simulated data corresponding to 128 sets for each of the 18 treatment/observation schemes for a total of 2304 virtual patients, each dependent on a parameter q of interest sampled from a true distribution f^0 as described above.

In this case the expected value in the cost criterion explicitly depends on the patient:

$$E\left[x(t_s^{ij}, u^j(t); \dot{q}, \bar{x}^0; q)|f(q)\right],$$

where the parameters $\dot{q}$ and initial condition x^0 are the same for each patient, the times t_s^{ij} and treatment $u^j(t)$ depend on the data structure for a patient from PS18, and as before q is assumed to vary across the population.

3. **Dynamics-varied simulated cohort:** As in (2.), each of the 2304 members of the third cohort will again be assigned the characteristics of one of the eighteen members of PS18. In this case, however, each of the eighteen base patients will be taken with his own model parameters and initial conditions. Now the expected value depends on several factors according to patient j:

$$E\left[x(t_s^{ij}, u^j(t), \dot{q}^j, \bar{x}^{0j}; q)|f(q)\right].$$

This scenario is the most similar to the clinical data as each patient is likely to have different behavior. Indeed, the parameters assigned to each of the eighteen patients resulted from initial efforts fitting the model to their clinical data and therefore approximately represent the variability in patient dynamics.

As examples, we consider in turn the parameters $d_1, k_1, N_T,$ and c, letting q denote

one of these distributed across the population, while the rest $\grave{q}$ of the 27 model parameters and initial conditions remain fixed. We assume for each q, $\log_{10} q$ is distributed either normally or according to a mixture of normals (bimodally) across the population. Initial parameter estimates from clinical data motivate the choice of a lognormal distribution as several estimated parameters exhibit that characteristic distribution. The maximal ranges considered for $\log_{10} q$ are denoted $[q_L, q_U]$ and prescribed in Table 5.1.

Table 5.1: Ranges prescribed for generating distributions of parameters.

parameter q	range ($\log_{10}(q)$) $[q_L, q_U]$
d_1	[-3, -0.9]
k_1	[-7, -5.5]
N_T	[1.8, 2.6]
c	[0.7, 1.3]

Setting $q_{span} = q_U - q_L$, we consider the following distributions for $\log_{10}(q)$:

- *Normal:* mean centered in the range interval and standard deviation $\frac{1}{8}$ the interval
$$\mathcal{N}\left(\frac{q_L + q_U}{2}, \left(\frac{q_{span}}{8}\right)^2\right);$$
- *Bimodal:* a one-third, two-thirds mixture of two normals with different means and standard deviation:
$$\frac{1}{3}\mathcal{N}\left(\frac{2q_L + q_U}{3}, \left(\frac{q_{span}}{12}\right)^2\right) + \frac{2}{3}\mathcal{N}\left(\frac{q_L + 2q_U}{3}, \left(\frac{q_{span}}{20}\right)^2\right).$$

A diagram of the probability density functions for these distributions on the range $[-1, 1]$ is shown in Figure 5.1.

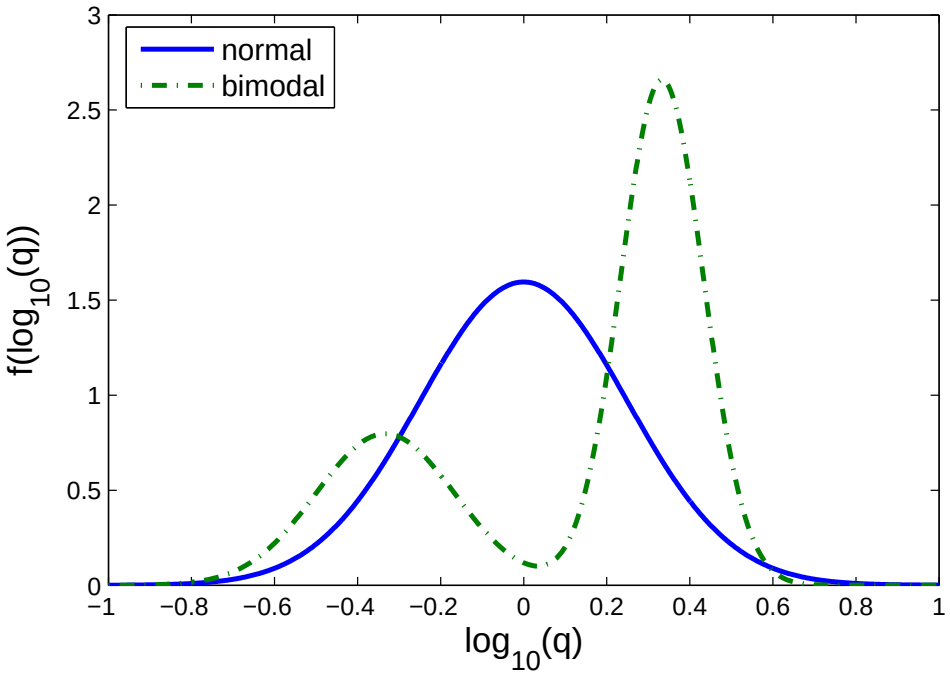


Figure 5.1: Sample normal and bimodal densities f on $\log_{10}(q) \in [-1, 1]$.

5.1.1 Uniform simulated cohort

We first use the data from the uniform simulated patient cohort to test the multi-patient inverse problem for determining the density f in the simplest case where the expected values of model solutions have no dependence on virtual patient. We wish to understand the effect of various factors on the inverse problem process, including number of patients sampled, number of splines used in the approximation of the density f, and regularization.

We begin by examining results for estimating distributions of the viral infectivity parameter k_1. For $N_S = 8$ splines, Figure 5.2 shows the estimated densities in the normal scenario for various numbers of sampled patients. As the number of patients increases, the L_1 norm of the error is reduced and the qualitative fit improves. In contrast, we observe from Figure 5.3 that in the case of the bimodal density the approximation with eight splines is not as good, even with 1024 sampled patients.

In this case, it seems more splines would be necessary to capture the detail of the distribution. However, even using only $N_S = 8$, the method returns an estimate that clearly suggests that the underlying density is bimodal.

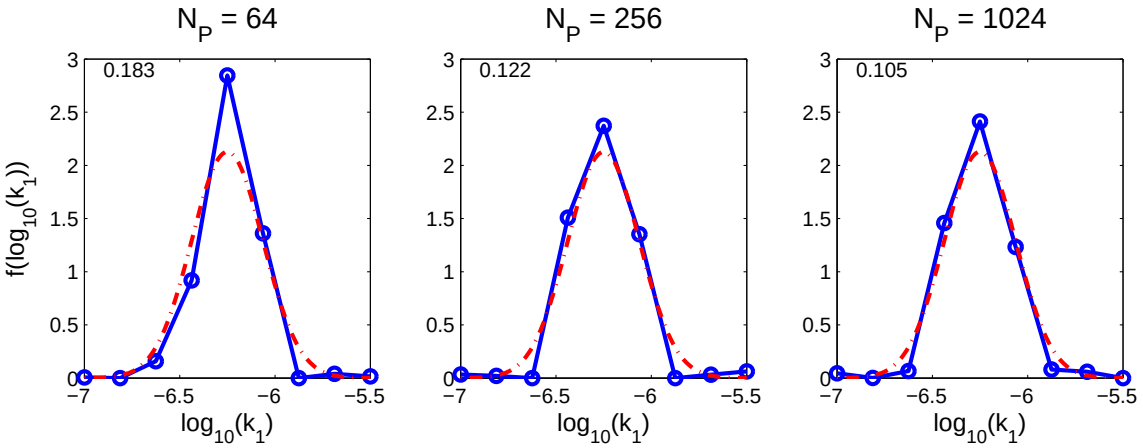


Figure 5.2: Uniform simulated cohort: True (red dash-dot line) and estimated spline-parametrized (blue solid line) probability density functions f for normal distribution of k_1 with $N_S = 8$ splines for various numbers of patients N_P. The L_1 norm of the error is indicated in the upper left corner of each plot.

In Figures 5.4 and 5.5, we fix the number of virtual patients at 1024 and increase the number of approximating splines in hopes of improving the approximation. In the normal case, increasing to $N_S = 16$ offers slight improvement, but a subsequent increase to $N_S = 32$ results in a poorer fit to the data and erratic approximation. In the bimodal case we see satisfactory improvement by increasing from eight to sixteen splines, but again severe oscillations are observed in increasing to 32. The poor fit with larger numbers of splines is related to ill-conditioning of the matrix A in the quadratic programming problem and can be alleviated by applying regularization.

The incorporation of a small nonzero regularization term $\beta_R = 1e-4$ resolves the observed oscillations. Even this minor adjustment stabilizes the condition number of the quadratic programming problem as revealed in Table 5.2, allowing reliable approximation with a greater number of splines. Results with the regularization term

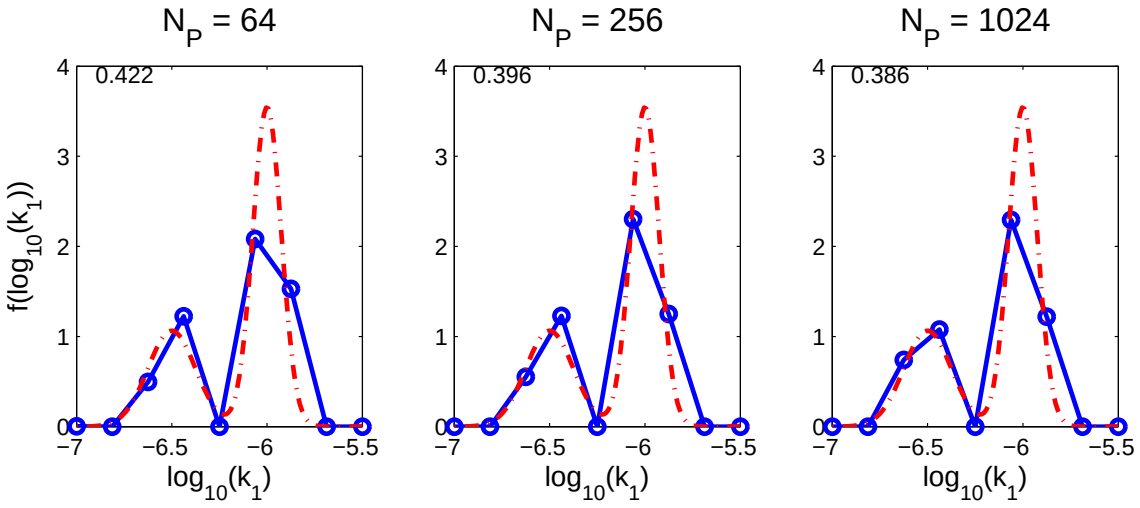


Figure 5.3: Uniform simulated cohort: True (red dash-dot line) and estimated spline-parametrized (blue solid line) probability density functions f for bimodal distribution of k_1 with $N_S = 8$ splines for various numbers of patients N_P. The L_1 norm of the error is indicated in the upper left of each plot.

for the same normal and bimodal cases appear in Figures 5.6 and 5.7, respectively. Application of regularization reduces the L_1 norm of the error, especially for larger numbers of splines.

Table 5.2: Influence of regularization with $\beta_R = 10^{-4}$ on two-norm condition number κ_2 of matrix $A + \beta_R A_R$ in quadratic programming problem. As defined previously, A is the matrix dependent on the model solutions and $\beta_R A_R$ is the regularization component.

N_S	$\kappa_2(A)$	$\kappa_2(A + \beta_R A_R)$
8	2.6041e+04	2.5836e+04
16	5.5955e+05	3.1947e+05
32	3.0632e+08	2.7232e+05

In addition to depending on the number of splines used, the choice of regularization parameter β_R is highly problem-dependent and can be as much art as science. While there exist methods for choosing β_R, including traditional L-curve or zero-crossing methods (discussed in [31]) and Generalized Cross-Validation or Generalized Maximum Likelihood (see [40] for an overview), for the simulated data results here

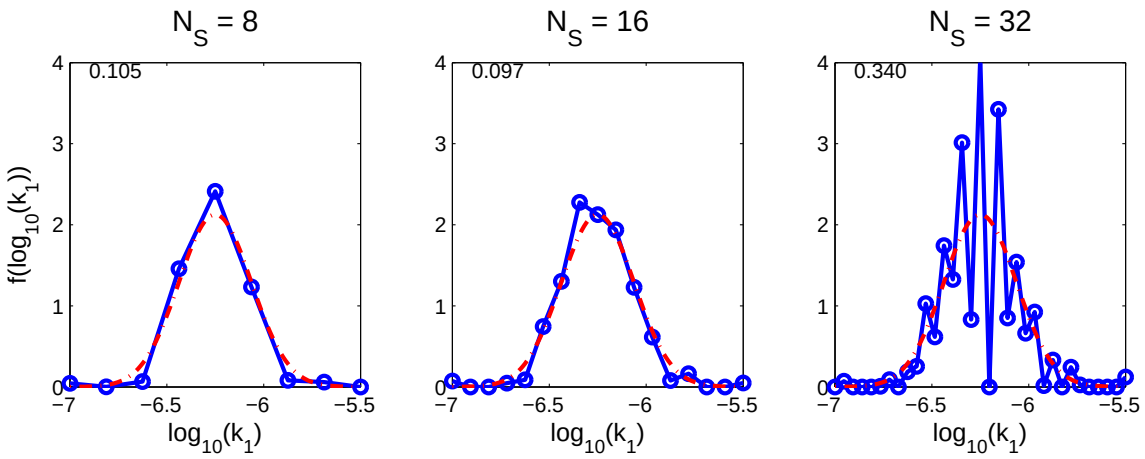


Figure 5.4: Uniform simulated cohort: True (red dash-dot line) and estimated spline-parametrized (blue solid line) probability density functions f for normal distribution of k_1 with $N_P = 1024$ patients for various numbers of splines N_S.

we simply choose a value sufficiently large enough to remove severe oscillations. The possible dangers in selecting extreme values of the parameter are demonstrated in Figure 5.8: when too small, not enough emphasis is placed on smoothing and erratic behavior remains, but when too large, fit to data is sacrificed in the interest of smoothness.

We conclude by examining some sample results for the other model parameters considered. Data from only 64 virtual patients is employed, as this is realistic for the kind of data we might hope to collect. Figures 5.9, 5.10, and 5.11 demonstrate that reasonable fits to distributions for a variety of model parameters are possible even with limited data. Regularization parameters were adjusted slightly as conditioning varied between data sets.

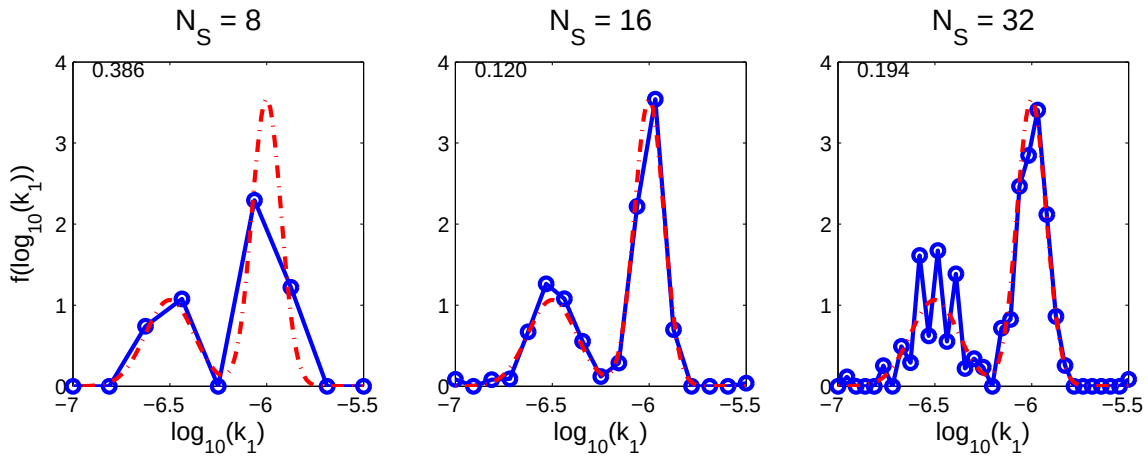


Figure 5.5: Uniform simulated cohort: True (red dash-dot line) and estimated spline-parametrized (blue solid line) probability density functions f for bimodal distribution of k_1 with $N_P = 1024$ patients for various numbers of splines N_S.

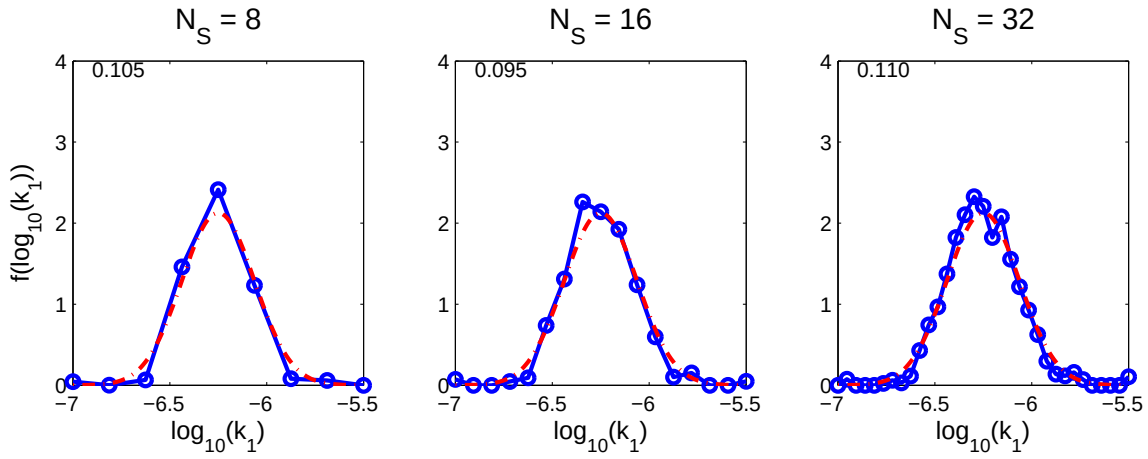


Figure 5.6: Uniform simulated cohort: True (red dash-dot line) and estimated spline-parametrized (blue solid line) probability density functions f for normal distribution of k_1 with $N_P = 1024$ patients for various numbers of splines N_S with regularization $\beta_R = 1e-4$.

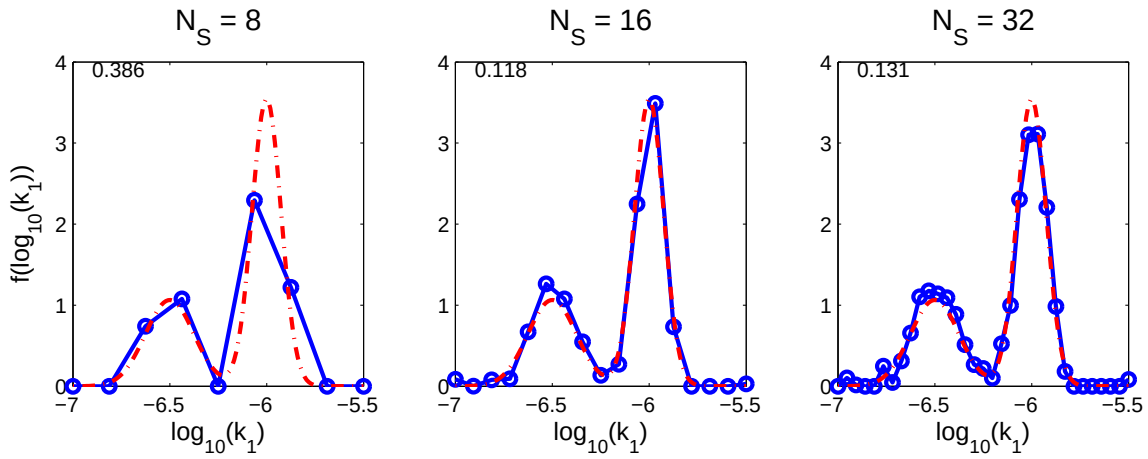


Figure 5.7: Uniform simulated cohort: True (red dash-dot line) and estimated spline-parametrized (blue solid line) probability density functions f for bimodal distribution of k_1 with $N_P = 1024$ patients for various numbers of splines with regularization $\beta_R = 1e-4$.

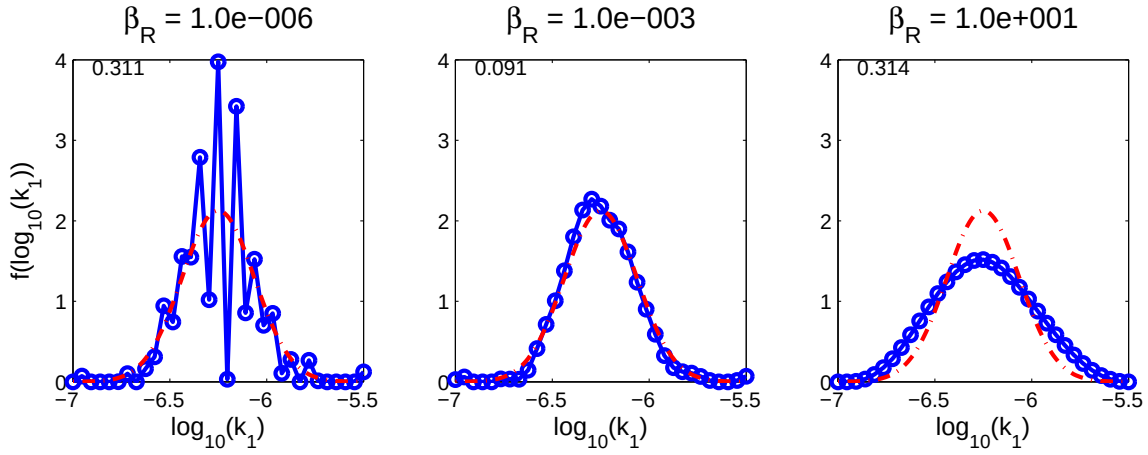


Figure 5.8: Influence of choice of regularization parameter β_R on estimated spline-parametrized (blue solid line) probability density functions f for normal distribution with $N_P = 1024, N_S = 32$.

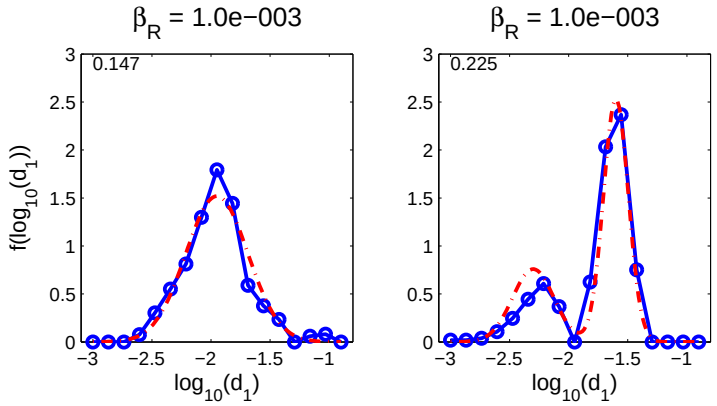


Figure 5.9: Uniform simulated cohort: True (red dash-dot line) and estimated spline-parametrized (blue solid line) probability density functions f for normal (left) and bimodal (right) distribution of d_1 with $N_P = 64, N_S = 16$, with regularization $\beta_R = 1e-3$.

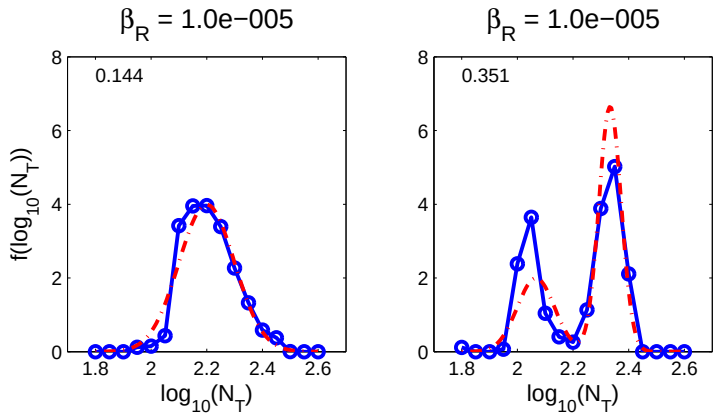


Figure 5.10: Uniform simulated cohort: True (red dash-dot line) and estimated spline-parametrized (blue solid line) probability density functions f for normal (left) and bimodal (right) distribution of N_T with $N_P = 64, N_S = 16$, with regularization $\beta_R = 1e-5$.

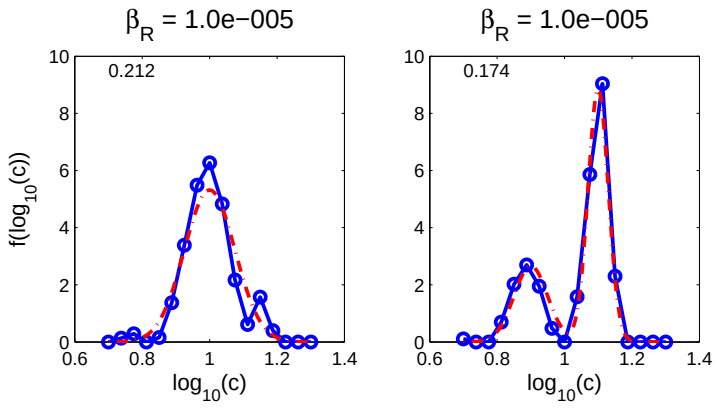


Figure 5.11: Uniform simulated cohort: True (red dash-dot line) and estimated spline-parametrized (blue solid line) probability density functions f for normal (left) and bimodal (right) distribution of c with $N_P = 64, N_S = 16$, with regularization $\beta_R = 1e-5$.

5.1.2 Treatment-varied simulated cohort

In this section we consider the second patient cohort, where patients are each assigned one of eighteen treatment/observation schemes, but each have the same underlying dynamics parameters and initial conditions.

Results obtained are similar for each of the parameters considered, and sample results for the parameter k_1 distributed bimodally with eight and sixteen splines and varying numbers of patients are shown in Figures 5.12 and 5.13, respectively.

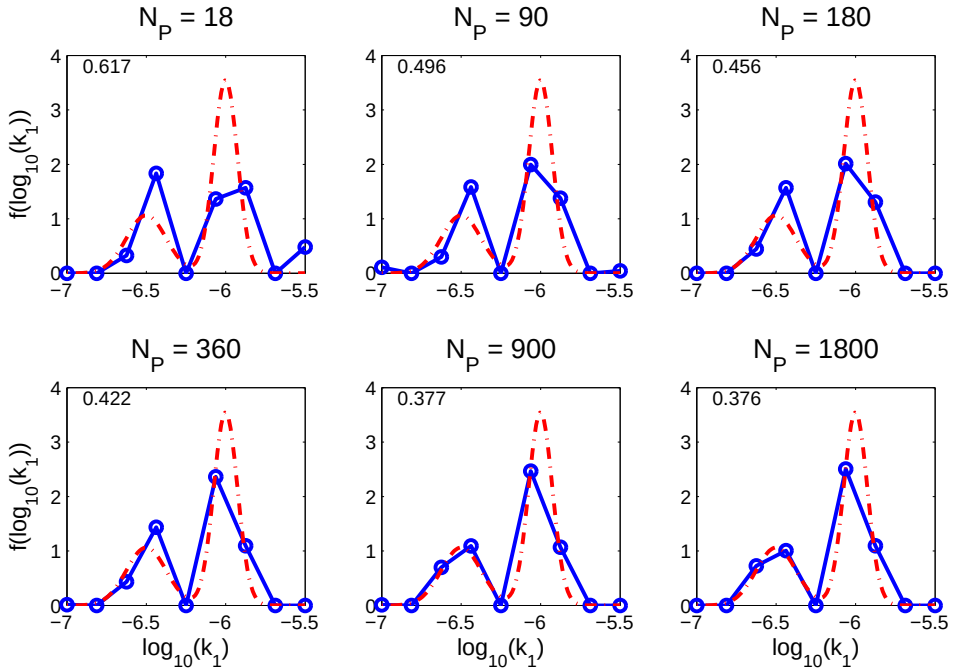


Figure 5.12: Treatment-varied simulated cohort: True (red dash-dot line) and estimated spline-parametrized (blue solid line) probability density functions f for bimodal distribution (right) of k_1 with $N_S = 8$, $\beta_R = 0$. L_1 norm of the error is again indicated in the upper left of each plot.

Since this inverse problem method fits the expected value of the model conditioned on the distribution of a model parameter to observed data, one must be cautious if there is more than one density function f that gives rise to a similar expected value. It is possible that two probability distributions, even with different means, will cause

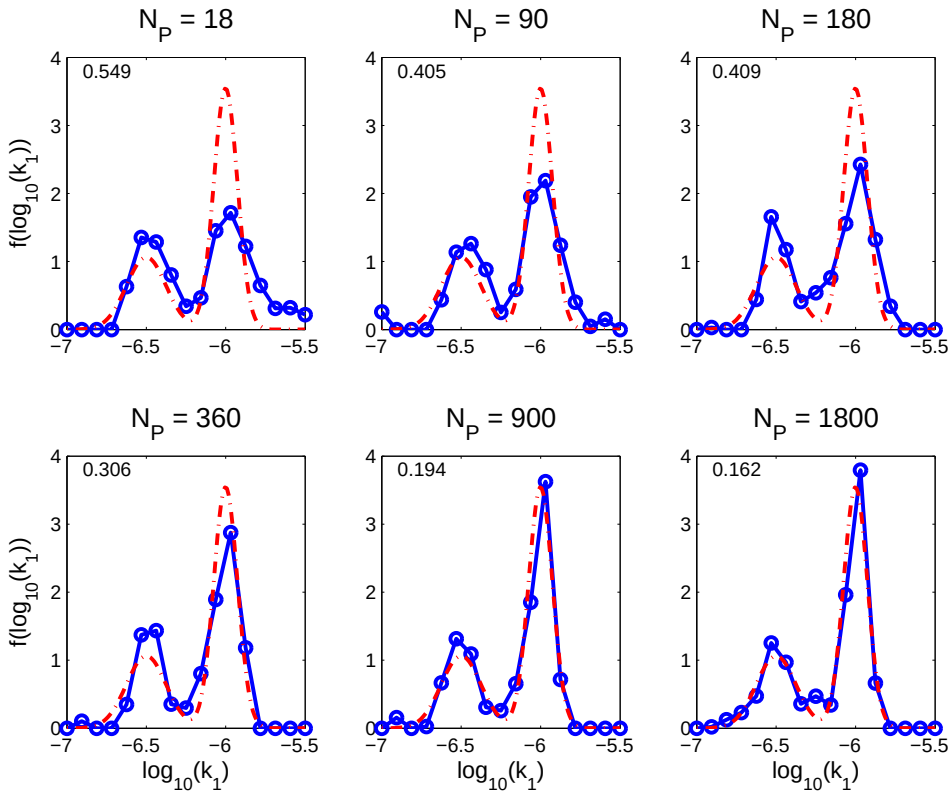


Figure 5.13: Treatment-varied simulated cohort: True (red dash-dot line) and estimated spline-parametrized (blue solid line) probability density functions f for bimodal distribution (right) of k_1 with $N_S = 16$, $\beta_R = 5e-3$.

the expected value $E^{ij}(f)$ to be the same. As an example, we consider adjusting the bimodal distribution of the examples above so that the expected value of the model, given the bimodal distribution, is similar to that when conditioned on the normal distribution.

In particular, considering the parameter d_1, the variance for the modes of the bimodal distribution remain fixed, but the second mean $\mu_2 = -1.6$ is adjusted to $\mu_2 = -1.78$ to yield an expected value of the model similar to that for the normal density. The resulting bimodal density and previous normal are shown in Figure 5.14. Results for estimating the bimodal pdf f, given data generated from a bimodal distribution are shown in Figures 5.15 and 5.16 for eight and sixteen splines, respectively. These

examples make clear the danger of limited data: with only eighteen patients, despite rich time series data, we cannot reliably estimate the distribution and indeed might conclude that it is normal. While this example was created in an *ad hoc* manner, it is possible there exist situations where, regardless of the number of observations, one cannot differentiate between two densities.

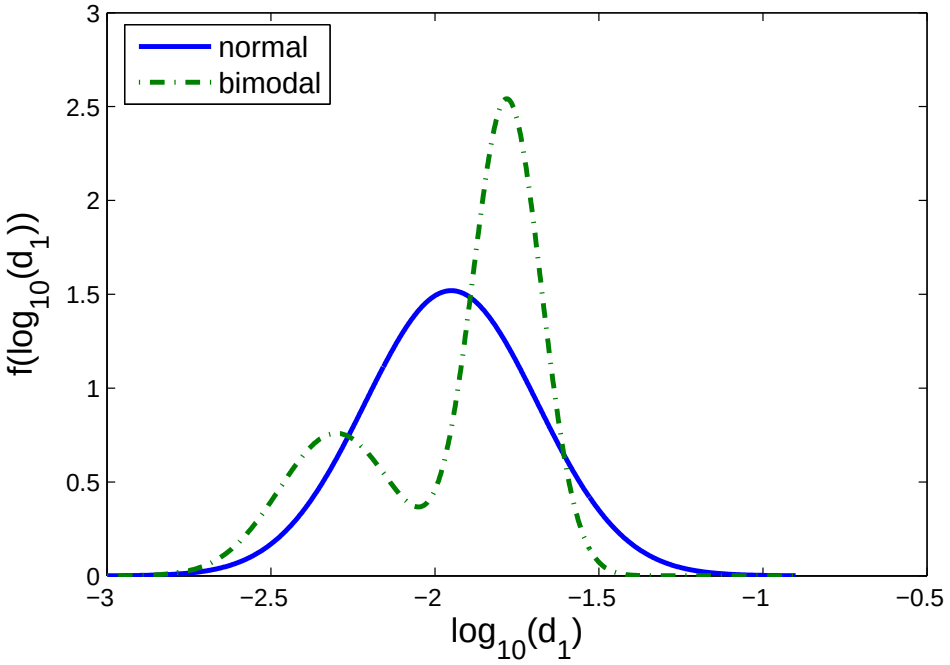


Figure 5.14: Sample normal and adjusted bimodal distributions for d_1.

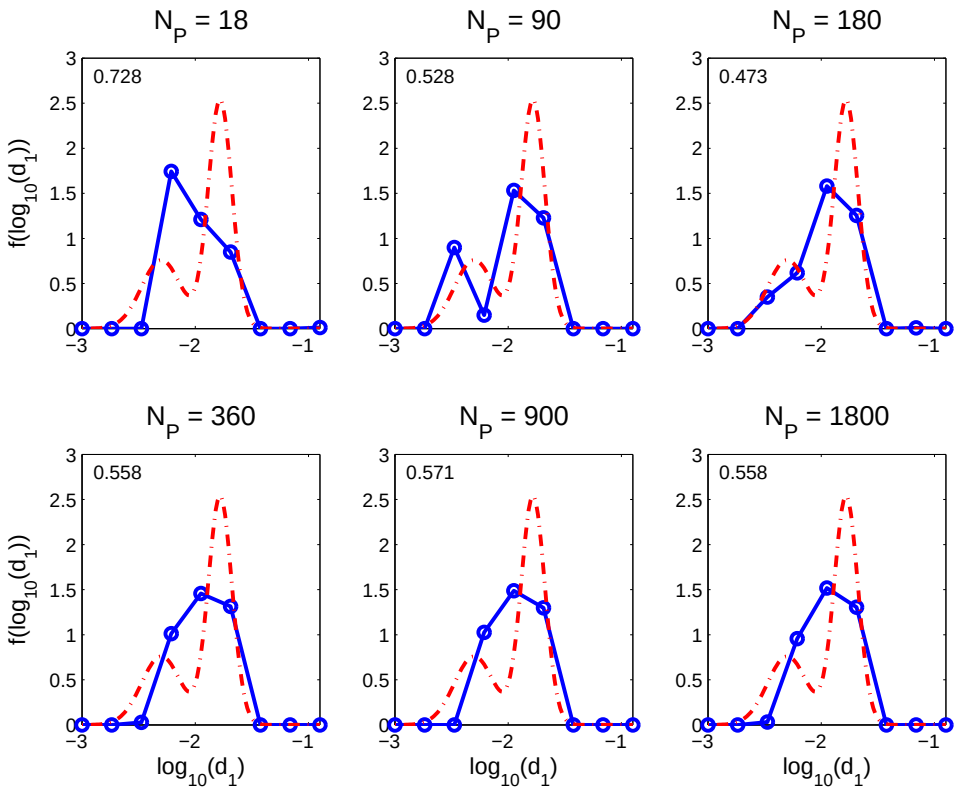


Figure 5.15: Treatment-varied simulated cohort: Fit to bimodal distribution when expected value is close to that of normal, $N_S = 8$. True (red dash-dot line) and estimated spline-parametrized (blue solid line) probability density functions f for bimodal distribution (right) of d_1 with $\beta_R = 0$.

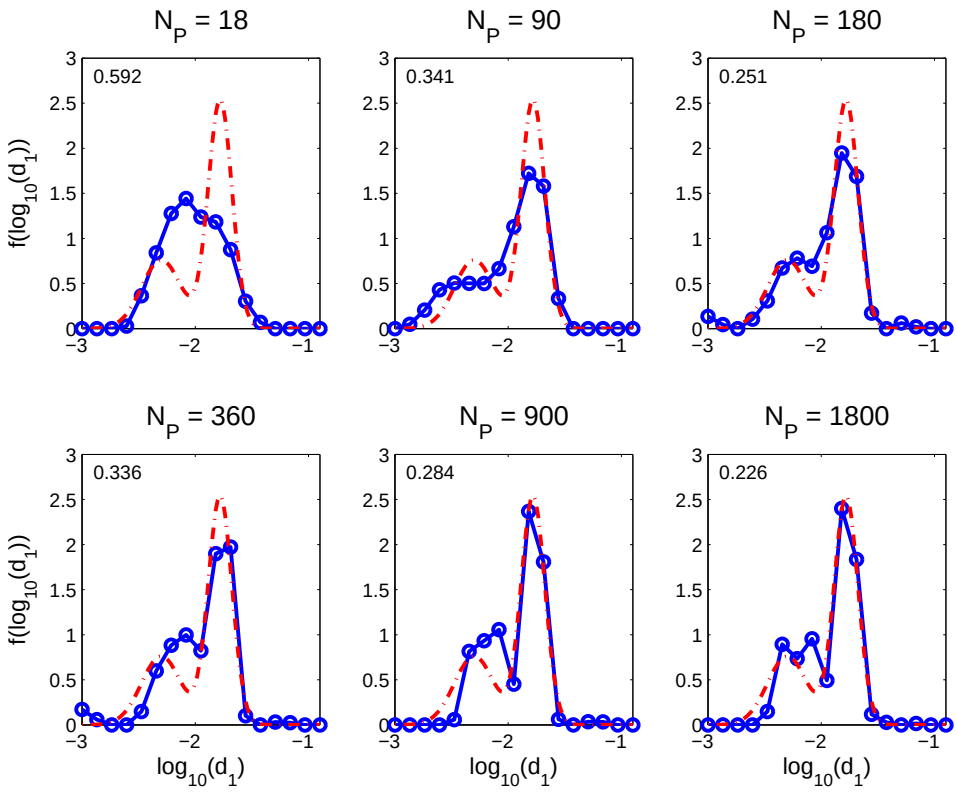


Figure 5.16: Treatment-varied simulated cohort: Fit to bimodal distribution when expected value is close to that of normal, $N_S = 16$. True (red dash-dot line) and estimated spline-parametrized (blue solid line) probability density functions f for bimodal distribution (right) of d_1 with $\beta_R = 0$.

5.1.3 Dynamics-varied simulated cohort

Finally we consider an example of the most general case where patients are assigned one of eighteen sets of parameters, initial conditions and treatment and observation schemes. It is essential that the methods work in this scenario before applying them to clinical data.

The method performs well in this most general scenario too. One sample set of fits is presented in Figure 5.17 for estimating a bimodal distribution of the parameter N_T with sixteen splines. For comparison, the top row is from scenario 2 with patients taken from the treatment-varied cohort. The bottom row contains results for fitting to data where the dynamic parameters and initial conditions also vary across the patients in PS18 from which the virtual patients were generated. Note that the choice of regularization parameter may be a factor in the difference between these examples.

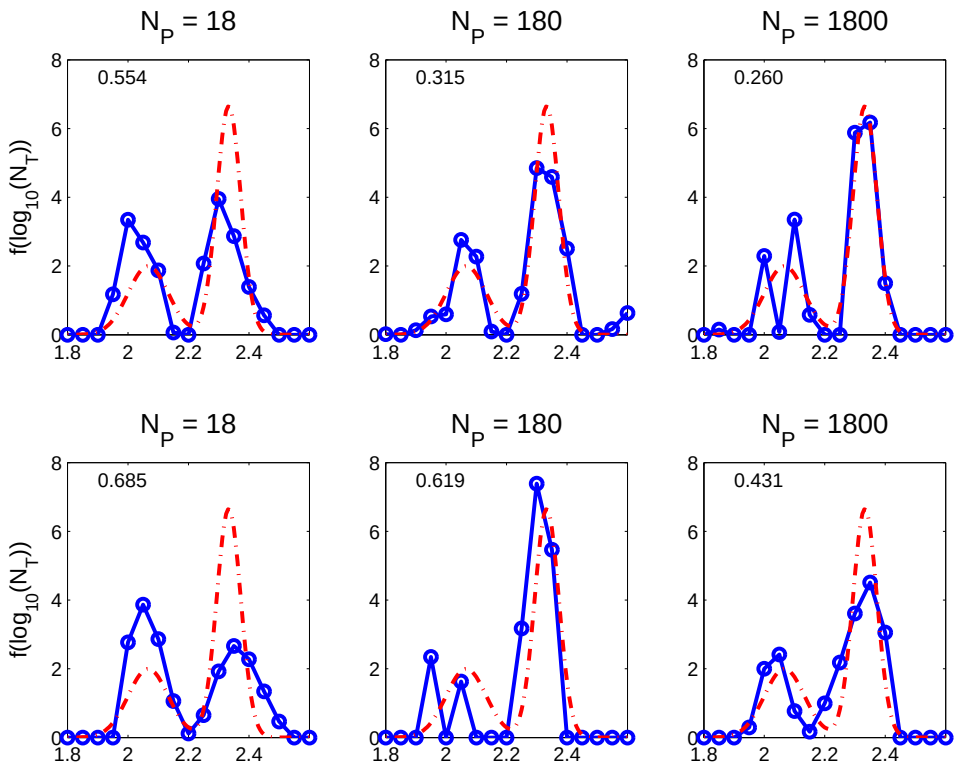


Figure 5.17: Top: estimate with only treatment and census times varying per patient. Bottom: all parameters vary across the eighteen baseline patients. True (red dash-dot line) and estimated spline-parametrized (blue solid line) probability density functions f for bimodal distribution (right) of N_T with $N_S = 16$, $\beta_R = 5e-5$. L_1 norm of the error is indicated in the upper left of each plot.

5.1.4 Confidence intervals

While the examples above demonstrate the ability of the inverse problem process to reliably estimate densities using piecewise linear approximations, one would like some idea of the certainty of the estimates. The formulation (4.12) above gives a framework for constructing piecewise linear confidence bands on the estimated density functions $f(q)$. The examples in this section are all constructed from the scenario described in section 5.1.1, and 95% confidence intervals at the nodes are shown.

We saw previously that regularization is an important aid in estimating density functions. Figure 5.18 contains results for estimating a normal distribution with various numbers of splines, demonstrating that the same is true for constructing confidence intervals. The same ill-conditioning that makes it difficult to recover the true density function is revealed in the progressively enormous confidence intervals. Figure

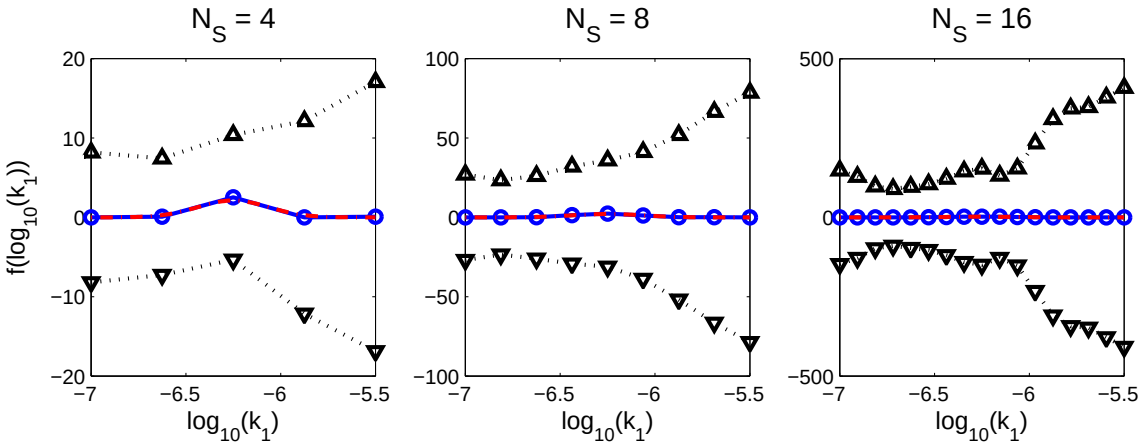


Figure 5.18: Uniform simulated cohort: Confidence intervals (dotted lines) for estimates (solid lines) of normal distribution of k_1, with $N_S = 4, 8, 16$; $\beta_R = 0$.

5.19 on the other hand shows the same scenario, with the addition of a regularization term $\beta_R = 0.01$.

We consider a bimodal example in Figure 5.20. Despite incorporation of regularization, the confidence intervals indicate no reasonable certainty as to what the

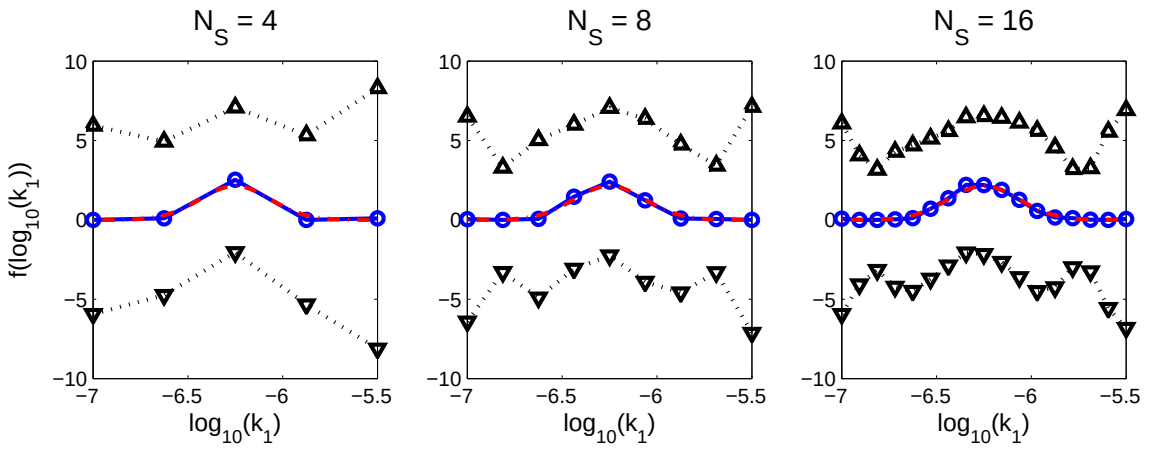


Figure 5.19: Uniform simulated cohort: Confidence intervals (dotted lines) for estimates (solid lines) of normal distribution of k_1, with $N_S = 4, 8, 16$; $\beta_R = 0.01$.

underlying distribution is, as any number of uni- or multi-modal densities could reliably fit within them. The emphasizes the importance of constructing confidence intervals to assess results. In contrast, consider Figure 5.21, where we show results for the parameter d_1 in two situations. In these cases, we have high confidence in the form of the distributions.

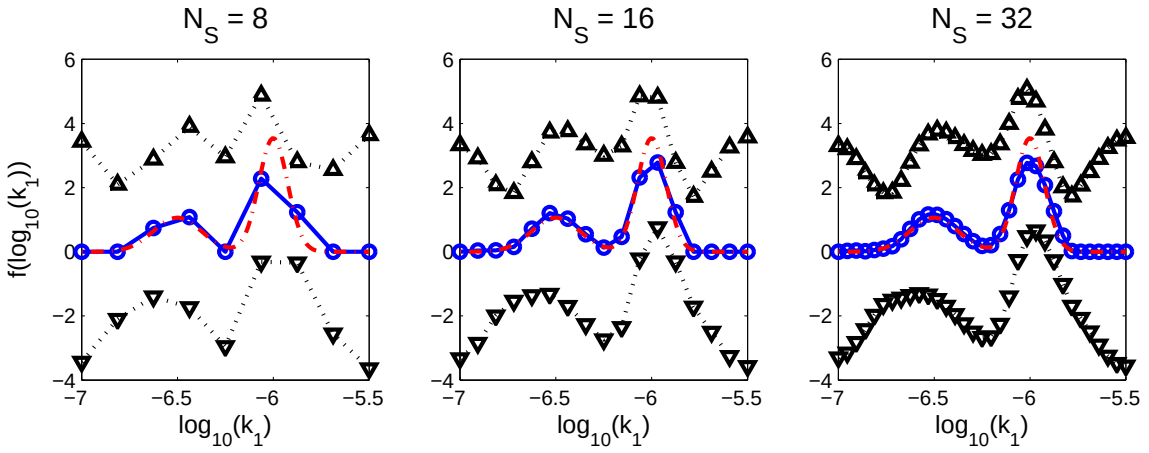


Figure 5.20: Uniform simulated cohort: Confidence intervals (dotted lines) for estimates (solid lines) of bimodal distribution of k_1, with $N_S = 4, 8, 16$; $\beta_R = 0.0001$.

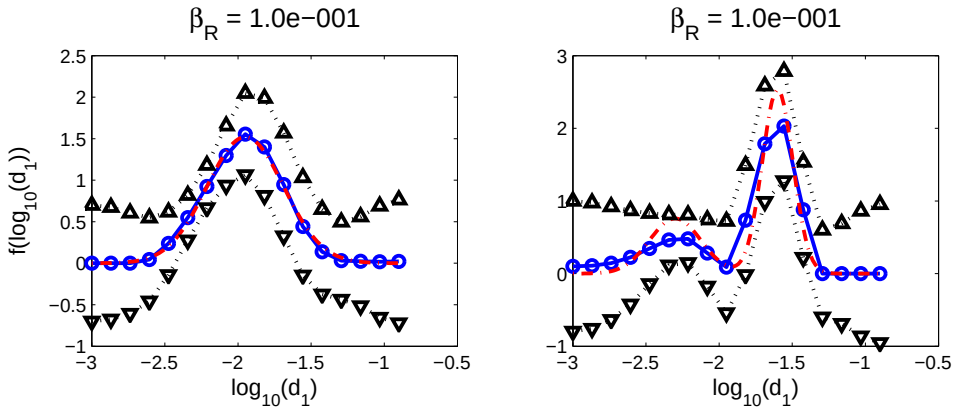


Figure 5.21: Uniform simulated cohort: Confidence intervals (dotted lines) for estimates (solid lines) of normal and bimodal distribution of d_1, with $N_S = 16$; $\beta_R = 0.01$.

5.2 Testing of Censored Data Methods

As with the methods for testing density estimation in the previous section, we test the censored data methods using the treatment regimen and observation schedule of Patient 4. We consider the baseline parameters established in Table 3.2 and generate simulated data at the desired observation times for cases with coefficient of variation $\bar{\sigma} = 0.2$ and 0.3, for two noise levels. To emulate the censor point of the assays we set two-thirds of the observations less than 400 copies per ml to the 400 ml censor point and one-third of the censored measurements to 50 copies per ml to emulate the use of the ultra-sensitive assay with its lower limit of detection. The additional detail provided by having measurements from the ultra-sensitive assay helps substantially with model fitting.

We proceed to test the algorithm by estimating eight sets of four parameters chosen from the six most sensitive $(\lambda_1, \epsilon_1, k_1, \epsilon_2, c, b_E)$, holding the remainder of the parameters and initial conditions at their known true values. As described in section 4.3.2, measurements censored at limit L are first set to $L/2$ for an initial least squares estimation, then the iterative censored data algorithm proceeds as described, with a relative stop tolerance on changes in parameter components and variance estimate of 10^{-6}. This convergence goal demands varying numbers of iterations depending on the data set considered.

For each data set we initiate the inverse problem algorithm with the true parameters q^0 from which the data were generated, as well as with positive and negative 1%, 5%, and 10% perturbations of the parameter vector q^0. Since the coefficient of variation for simulated data is known and small ($\bar{\sigma}_1 = \bar{\sigma}_2 = 0.2, 0.3$), the term $\frac{\log_{10}(\bar{\sigma}^2+1)}{2}$ is small and we therefore fit the base-10 logarithm of the model response ($\log_{10} \bar{z}^{ij}$) directly to the base-10 logarithm of the data. This is in contrast to fitting ζ^{ij} (which

includes the small correction term) to y^{ij} as derived when describing the error model above.

Tables 5.3 and 5.4 contain results for the two coefficients of variation. They include percent relative error between the estimated and known true value

$$\frac{|q_l^* - q_l^0|}{|q_l^0|} \times 100\%$$

for each parameter considered, number of iterations the censored data algorithm took to converge, the cost J at the true and optimal parameters, and the estimates of coefficient of variation returned by the algorithm. In this simulated data case, we report cost values J corresponding to the model fit to the *uncensored* noisy data. In a clinical data setting we would not directly observe data points below the censor point to make this comparison, but they more closely represent the true underlying dynamics and by using them the cost criterion J gives an estimate of how well the censored data algorithm can predict those original dynamics.

We observe that the censored data algorithm converges in a reasonable number of iterations for all the cases considered. When working with the clinical data, many more iterations are often necessary. Note that since the simulated data used for these experiments include random noise, we do not expect to exactly recover the true values of the parameters. Even when starting the algorithm at the known true values q^0 (no perturbation), the best fit to data may be achieved for nearby parameters. Therefore even in the zero perturbation case, we anticipate some small deviation from the true parameters.

The inverse problem algorithm produces estimates of variance σ^2 on the transformed scale. From (3.6) we can transform these back to find the original coefficient

of variation

$$\bar{\sigma}^2 = 10^{\frac{\sigma^2}{\log 10(e)}} - 1$$

from which the data were generated. These are the values reported in the tables. The values estimated for variance in the CD4 data, i.e., $\hat{\sigma}_1$, are reasonable, given the true values from which the data came. However, the censored data algorithm seems to consistently overestimate the coefficient of variation $\hat{\sigma}_2$ for the viral load. To understand this, we refer back to (4.18) which characterizes the squared residuals $\tilde{r}^i$ in the censored data context. Estimates of $\hat{\sigma}_2$ are based on sums of these squared residuals. For censored data points these are calculated using $E\left[(y^i - \zeta^i(q))^2 | y^i < L^i\right]$, i.e., conditioned on the knowledge of the censor point. We would not expect the estimate of variance to be the same as in the perfect information case. Indeed, when an exact model solution (one evaluated using the true parameters) is used to calculate values of $\hat{\sigma}_2$ for these two scenarios, we find 0.3595 and 0.3921. Therefore, the estimates in the tables seem reasonable.

Table 5.3: Parameter relative error and statistics for testing censored data algorithm with $\bar{\sigma} = 0.2$ for various perturbations of known true value q^0 and various sets of four parameters.

		perturbation						
set	statistic	0.90	0.95	0.99	1.00	1.01	1.05	1.10
1	ϵ_1 rel. err.	0.68	0.81	0.90	0.68	2.79	1.42	1.09
	ϵ_2 rel. err.	1.24	1.47	2.01	1.39	2.20	3.45	2.37
	c rel. err.	1.40	1.45	1.23	1.34	4.07	1.12	1.28
	b_E rel. err.	0.17	0.09	0.02	0.19	0.74	0.29	0.11
	numits	20	10	10	11	9	13	15
	$J(q_{true})$	0.014	0.014	0.014	0.014	0.014	0.014	0.014
	$J(q_{est})$	0.020	0.020	0.020	0.020	0.118	0.021	0.020
	$\hat{\sigma}_1$	0.206	0.206	0.206	0.206	0.211	0.206	0.206
	$\hat{\sigma}_2$	0.38	0.38	0.38	0.38	0.96	0.37	0.37
2	ϵ_1 rel. err.	0.47	1.52	0.90	0.28	1.65	1.28	4.07
	k_1 rel. err.	2.97	0.77	0.36	1.55	1.40	4.76	5.66
	ϵ_2 rel. err.	2.59	4.55	2.41	0.89	2.20	4.35	2.63
	c rel. err.	4.16	0.15	1.37	2.13	2.59	5.15	0.20
	numits	19	14	10	10	15	16	11
	$J(q_{true})$	0.014	0.014	0.014	0.014	0.014	0.014	0.014
	$J(q_{est})$	0.021	0.022	0.022	0.022	0.082	0.022	0.128
	$\hat{\sigma}_1$	0.206	0.206	0.206	0.206	0.209	0.205	0.211
	$\hat{\sigma}_2$	0.37	0.38	0.37	0.37	0.80	0.37	0.98
3	ϵ_1 rel. err.	2.28	1.28	0.75	0.51	1.02	0.87	1.16
	k_1 rel. err.	0.76	0.32	0.21	0.11	0.30	0.15	0.49
	ϵ_2 rel. err.	6.18	4.39	3.25	2.77	3.77	3.58	3.96
	b_E rel. err.	0.60	0.10	0.20	0.32	0.05	0.10	0.04
	numits	21	10	11	10	10	9	16
	$J(q_{true})$	0.014	0.014	0.014	0.014	0.014	0.014	0.014
	$J(q_{est})$	0.021	0.021	0.020	0.020	0.020	0.020	0.020
	$\hat{\sigma}_1$	0.206	0.205	0.205	0.205	0.205	0.205	0.205
	$\hat{\sigma}_2$	0.37	0.38	0.38	0.38	0.38	0.38	0.38
4	λ_1 rel. err.	4.09	4.60	4.90	5.37	5.64	5.65	5.28
	ϵ_1 rel. err.	0.40	0.25	0.07	0.22	0.23	0.17	0.13
	c rel. err.	4.27	4.42	4.30	4.84	5.03	4.92	4.64
	b_E rel. err.	1.43	1.50	1.50	1.80	1.92	1.89	1.69
	numits	20	15	10	10	13	8	10
	$J(q_{true})$	0.014	0.014	0.014	0.014	0.014	0.014	0.014
	$J(q_{est})$	0.025	0.026	0.027	0.029	0.029	0.029	0.028
	$\hat{\sigma}_1$	0.211	0.212	0.212	0.213	0.214	0.214	0.213
	$\hat{\sigma}_2$	0.35	0.35	0.35	0.35	0.35	0.35	0.35

		perturbation						
set	statistic	0.90	0.95	0.99	1.00	1.01	1.05	1.10
5	λ_1 rel. err.	4.77	3.94	3.83	2.94	3.21	4.41	4.19
	ϵ_1 rel. err.	0.06	0.08	0.20	0.22	0.27	0.29	0.30
	k_1 rel. err.	3.84	2.51	2.01	0.89	1.10	2.46	2.23
	c rel. err.	1.73	1.95	2.11	2.30	2.24	2.04	2.00
	numits	19	10	11	10	18	16	10
	$J(q_{true})$	0.014	0.014	0.014	0.014	0.014	0.014	0.014
	$J(q_{est})$	0.023	0.023	0.024	0.024	0.024	0.025	0.024
	$\hat{\sigma}_1$	0.214	0.212	0.211	0.209	0.210	0.212	0.212
	$\hat{\sigma}_2$	0.36	0.36	0.36	0.36	0.36	0.36	0.36
6	λ_1 rel. err.	23.30	5.61	4.86	4.39	4.24	5.67	5.11
	ϵ_2 rel. err.	38.57	2.09	0.97	0.57	0.22	1.32	1.01
	c rel. err.	16.31	5.98	4.79	4.25	3.94	5.52	4.66
	b_E rel. err.	3.03	1.81	1.46	1.23	1.16	1.81	1.93
	numits	17	21	10	12	13	9	10
	$J(q_{true})$	0.014	0.014	0.014	0.014	0.014	0.014	0.014
	$J(q_{est})$	1.047	0.027	0.026	0.025	0.025	0.028	0.032
	$\hat{\sigma}_1$	0.328	0.215	0.212	0.211	0.211	0.214	0.215
	$\hat{\sigma}_2$	1.80	0.34	0.35	0.35	0.35	0.34	0.39
7	λ_1 rel. err.	6.47	3.20	2.72	2.90	3.10	2.06	3.34
	k_1 rel. err.	5.54	1.78	1.19	1.38	1.54	0.32	1.78
	ϵ_2 rel. err.	0.55	0.11	0.36	0.38	0.39	0.31	0.26
	c rel. err.	1.17	2.20	2.49	2.51	2.58	2.66	2.49
	numits	35	10	10	9	10	10	12
	$J(q_{true})$	0.014	0.014	0.014	0.014	0.014	0.014	0.014
	$J(q_{est})$	0.024	0.023	0.022	0.023	0.023	0.023	0.024
	$\hat{\sigma}_1$	0.219	0.210	0.210	0.210	0.211	0.209	0.211
	$\hat{\sigma}_2$	0.35	0.36	0.36	0.36	0.36	0.36	0.36
8	λ_1 rel. err.	8.80	6.15	5.84	5.04	4.55	5.19	4.76
	k_1 rel. err.	4.96	1.37	0.51	0.25	0.74	0.69	0.25
	c rel. err.	3.08	4.04	4.42	4.50	4.53	3.61	3.65
	b_E rel. err.	1.28	1.47	1.66	1.56	1.52	1.68	1.65
	numits	21	11	10	14	10	12	9
	$J(q_{true})$	0.014	0.014	0.014	0.014	0.014	0.014	0.014
	$J(q_{est})$	0.028	0.028	0.029	0.029	0.028	0.032	0.032
	$\hat{\sigma}_1$	0.225	0.215	0.214	0.212	0.211	0.214	0.213
	$\hat{\sigma}_2$	0.34	0.34	0.35	0.35	0.35	0.39	0.39

Table 5.4: Parameter relative error and statistics for testing censored data algorithm with $\bar{\sigma} = 0.3$ for various perturbations of known true value q^0 and various sets of four parameters.

		perturbation						
set	statistic	0.90	0.95	0.99	1.00	1.01	1.05	1.10
1	ϵ_1 rel. err.	0.21	0.50	0.13	0.30	4.59	0.86	0.39
	ϵ_2 rel. err.	0.76	2.77	1.93	2.46	2.91	3.99	0.81
	c rel. err.	0.00	0.18	0.19	0.26	7.32	0.39	0.25
	b_E rel. err.	0.60	0.15	0.41	0.30	1.69	0.05	0.76
	numits	25	13	11	10	11	20	9
	$J(q_{true})$	0.031	0.031	0.031	0.031	0.031	0.031	0.031
	$J(q_{est})$	0.031	0.032	0.032	0.032	0.120	0.032	0.032
	$\hat{\sigma}_1$	0.256	0.256	0.256	0.256	0.266	0.256	0.257
	$\hat{\sigma}_2$	0.41	0.41	0.41	0.41	0.94	0.41	0.41
2	ϵ_1 rel. err.	6.48	0.68	0.18	0.27	3.16	1.42	2.11
	k_1 rel. err.	16.25	0.40	1.33	2.15	3.10	4.58	5.98
	ϵ_2 rel. err.	23.37	3.80	2.19	0.75	3.49	2.89	5.56
	c rel. err.	14.46	0.18	0.68	1.40	4.12	3.50	4.91
	numits	31	12	9	10	9	18	10
	$J(q_{true})$	0.031	0.031	0.031	0.031	0.031	0.031	0.031
	$J(q_{est})$	0.033	0.033	0.033	0.032	0.129	0.033	0.032
	$\hat{\sigma}_1$	0.258	0.257	0.257	0.257	0.264	0.257	0.257
	$\hat{\sigma}_2$	0.41	0.41	0.41	0.41	0.97	0.40	0.40
3	ϵ_1 rel. err.	0.94	0.89	0.21	0.20	0.47	0.76	1.02
	k_1 rel. err.	0.92	0.58	0.91	0.92	1.08	1.05	1.17
	ϵ_2 rel. err.	0.23	3.93	1.74	1.74	1.24	0.54	0.02
	b_E rel. err.	0.75	0.30	0.35	0.33	0.45	0.67	0.78
	numits	21	9	12	9	10	10	9
	$J(q_{true})$	0.031	0.031	0.031	0.031	0.031	0.031	0.031
	$J(q_{est})$	0.031	0.032	0.032	0.032	0.032	0.032	0.032
	$\hat{\sigma}_1$	0.257	0.256	0.257	0.257	0.257	0.257	0.257
	$\hat{\sigma}_2$	0.41	0.41	0.41	0.41	0.41	0.41	0.41
4	λ_1 rel. err.	1.37	1.84	1.98	2.01	2.29	2.67	2.59
	ϵ_1 rel. err.	0.38	0.51	0.60	0.57	0.60	0.51	0.49
	c rel. err.	1.06	1.20	1.16	1.23	1.40	1.77	1.75
	b_E rel. err.	0.13	0.06	0.08	0.05	0.01	0.25	0.23
	numits	23	10	10	10	10	14	10
	$J(q_{true})$	0.031	0.031	0.031	0.031	0.031	0.031	0.031
	$J(q_{est})$	0.032	0.032	0.032	0.032	0.033	0.033	0.033
	$\hat{\sigma}_1$	0.259	0.260	0.260	0.260	0.261	0.262	0.262
	$\hat{\sigma}_2$	0.41	0.40	0.40	0.40	0.40	0.40	0.40

		perturbation						
set	statistic	0.90	0.95	0.99	1.00	1.01	1.05	1.10
5	λ_1 rel. err.	1.44	0.30	1.19	0.42	0.55	1.60	0.28
	ϵ_1 rel. err.	0.37	3.29	0.64	0.62	0.60	3.91	0.60
	k_1 rel. err.	4.12	2.19	4.63	2.45	2.24	2.93	3.33
	c rel. err.	1.99	2.83	2.18	1.88	1.85	2.26	2.01
	numits	24	9	14	19	18	12	24
	$J(q_{true})$	0.031	0.031	0.031	0.031	0.031	0.031	0.031
	$J(q_{est})$	0.032	0.124	0.034	0.033	0.033	0.129	0.033
	$\hat{\sigma}_1$	0.255	0.263	0.256	0.257	0.258	0.265	0.256
	$\hat{\sigma}_2$	0.41	0.99	0.41	0.40	0.40	0.97	0.41
6	λ_1 rel. err.	20.99	1.71	1.64	1.41	1.71	1.50	1.35
	ϵ_2 rel. err.	33.43	1.11	1.50	1.62	1.24	1.50	1.85
	c rel. err.	16.76	1.20	0.93	0.71	1.13	0.85	0.54
	b_E rel. err.	2.27	0.23	0.18	0.08	0.21	0.12	0.05
	numits	11	30	10	10	10	10	9
	$J(q_{true})$	0.031	0.031	0.031	0.031	0.031	0.031	0.031
	$J(q_{est})$	1.052	0.032	0.033	0.032	0.032	0.032	0.033
	$\hat{\sigma}_1$	0.321	0.260	0.259	0.259	0.260	0.259	0.259
	$\hat{\sigma}_2$	1.64	0.41	0.41	0.41	0.41	0.41	0.41
7	λ_1 rel. err.	1.24	0.34	0.14	0.54	0.45	0.07	0.65
	k_1 rel. err.	0.01	1.21	1.46	0.96	1.07	1.79	2.44
	ϵ_2 rel. err.	1.17	1.30	1.47	1.45	1.33	0.93	0.89
	c rel. err.	0.94	1.11	1.03	0.98	1.06	1.51	1.52
	numits	27	11	10	13	10	10	10
	$J(q_{true})$	0.031	0.031	0.031	0.031	0.031	0.031	0.031
	$J(q_{est})$	0.032	0.032	0.033	0.033	0.032	0.033	0.032
	$\hat{\sigma}_1$	0.259	0.257	0.257	0.257	0.257	0.257	0.256
	$\hat{\sigma}_2$	0.41	0.41	0.41	0.41	0.41	0.41	0.41
8	λ_1 rel. err.	0.68	0.76	0.85	1.43	1.38	1.01	2.65
	k_1 rel. err.	2.05	2.58	2.46	1.95	1.64	2.09	2.50
	c rel. err.	2.52	2.97	2.96	2.99	2.77	2.79	1.46
	b_E rel. err.	0.56	0.74	0.73	0.81	0.69	0.70	1.05
	numits	25	10	10	10	10	10	9
	$J(q_{true})$	0.031	0.031	0.031	0.031	0.031	0.031	0.031
	$J(q_{est})$	0.032	0.032	0.032	0.033	0.032	0.032	0.152
	$\hat{\sigma}_1$	0.258	0.258	0.258	0.259	0.259	0.258	0.270
	$\hat{\sigma}_2$	0.41	0.41	0.41	0.41	0.41	0.41	1.11

We examine some plots for the case $\bar{\sigma} = 0.2$. Figure 5.22 contains the model fit to data for the fourth parameter set with a 1.05 perturbation as initial iterate. The fit to data is not identical to the true model values, but is certainly reasonable. Even when the relative errors in the parameters are moderately large (up to 10%) the model fits remain plausible.

There are many examples in the table where a large perturbation of the initial iterate yields poorer results than a smaller one. Plots that show the corresponding fits to data for the seventh parameter set are in Figures 5.23 for 0.9 perturbation and 5.24 for 1.01 perturbation. These remind us that the parameters that yield a good fit to data are not necessarily unique – adjusting one can compensate for another.

Finally an example of failure is shown in Figure 5.25 for a 0.90 perturbation of the sixth parameter set. The fits to the two states are substantially worse and the table indicates 3% – 40% relative error on the estimated parameters. The algorithm cannot estimate the true values of parameters given this large of an initial deviation from the true values.

One reassurance that the censored data method is working is that the log likelihood discussed above should be increasing with each iteration of the method. A plot of log likelihood vs. iteration number is shown for parameter set 2 in Figure 5.27 for perturbations 0.9, 0.95, and 0.99. With a few exceptions, the likelihood is monotonically increasing. While theory guarantees monotonicity under certain assumptions, these are likely not met in this situation where the least squares minimization at each step of the censored data algorithm involves a numerical optimization algorithm. In turn the optimization algorithm at each step requires one or more numerical solutions to the differential equation system and may converge to a local minimum. This is particularly evident in the rightmost subplot (0.99 perturbation) where the likelihood is changing very slightly due to good initial iterate.

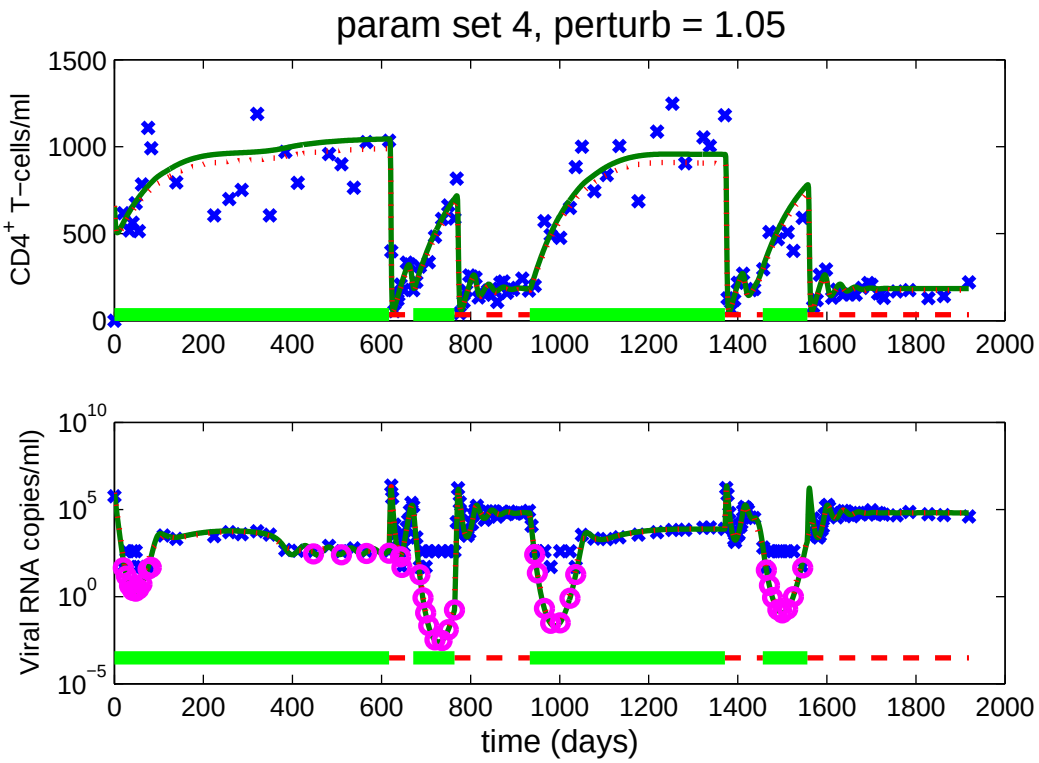


Figure 5.22: Sample fit to simulated censored data ($\bar{\sigma} = 0.2$) with censored data algorithm for fourth parameter set with initial iterate $1.05q^0$. Red dotted line indicates true model solution from which data were generated, solid green line is model at estimated parameters, 'x' denotes observed data, and 'o' the predicted value of a censored data point.

In Figure 5.27 we see similar plots for parameter set 3. The results for 0.9 and 0.99 perturbations are excellent, and while a bit more erratic, the 0.95 perturbation stabilizes. Perhaps most informative are the plots for perturbation 0.90 which demonstrate that the algorithm makes steady progress toward maximum likelihood when starting far away. Plots for other parameters are very similar: large perturbations result in essentially monotonic trajectories and some small perturbations yield slightly erratic behavior that eventually stabilizes.

We list in Table 5.5 parameter estimates and standard errors calculated for the case with initial iterate $0.9q^0$ for both $\bar{\sigma} = 0.2$ and 0.3 cases. These may be used to

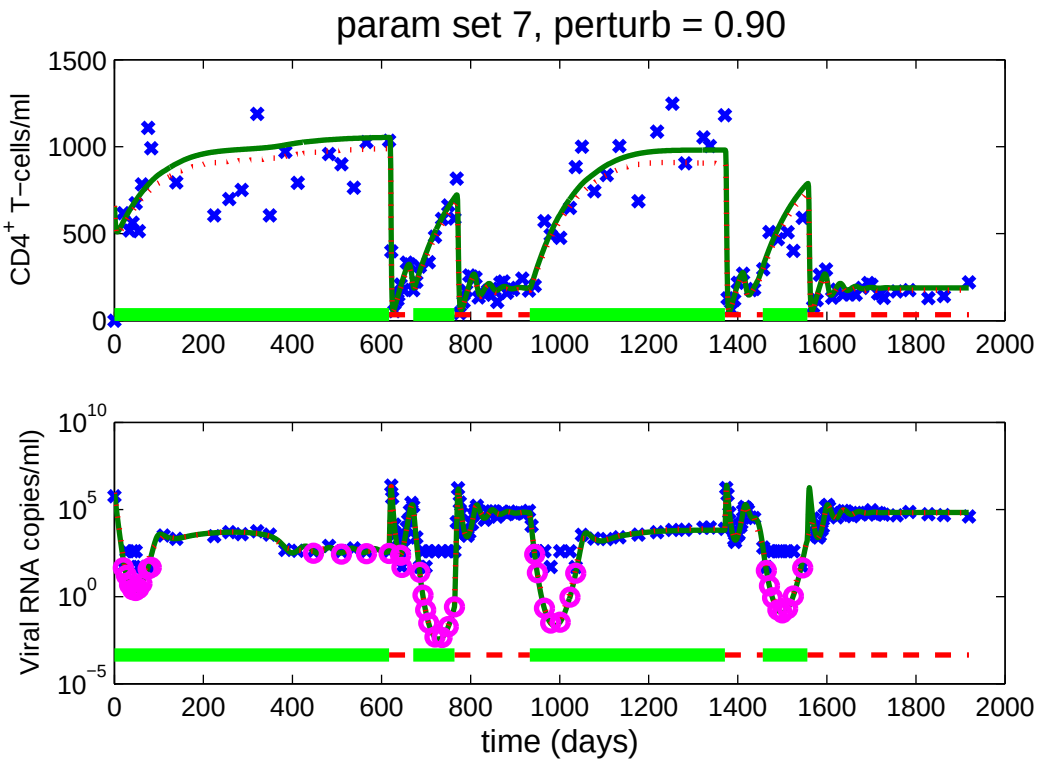


Figure 5.23: Sample fit to simulated censored data ($\bar{\sigma} = 0.2$) with censored data algorithm for seventh parameter set with initial iterate $0.90q^0$. Red dotted line indicates true model solution from which data were generated, solid green line is model at estimated parameters, 'x' denotes observed data, and 'o' the predicted value of a censored data point.

calculate confidence intervals at a desired confidence level for each parameter estimated. As expected, the parameters estimated from the data generated with larger coefficient of variation have less certainty. In the case of this simulated data even 99% confidence intervals would be fairly tight.

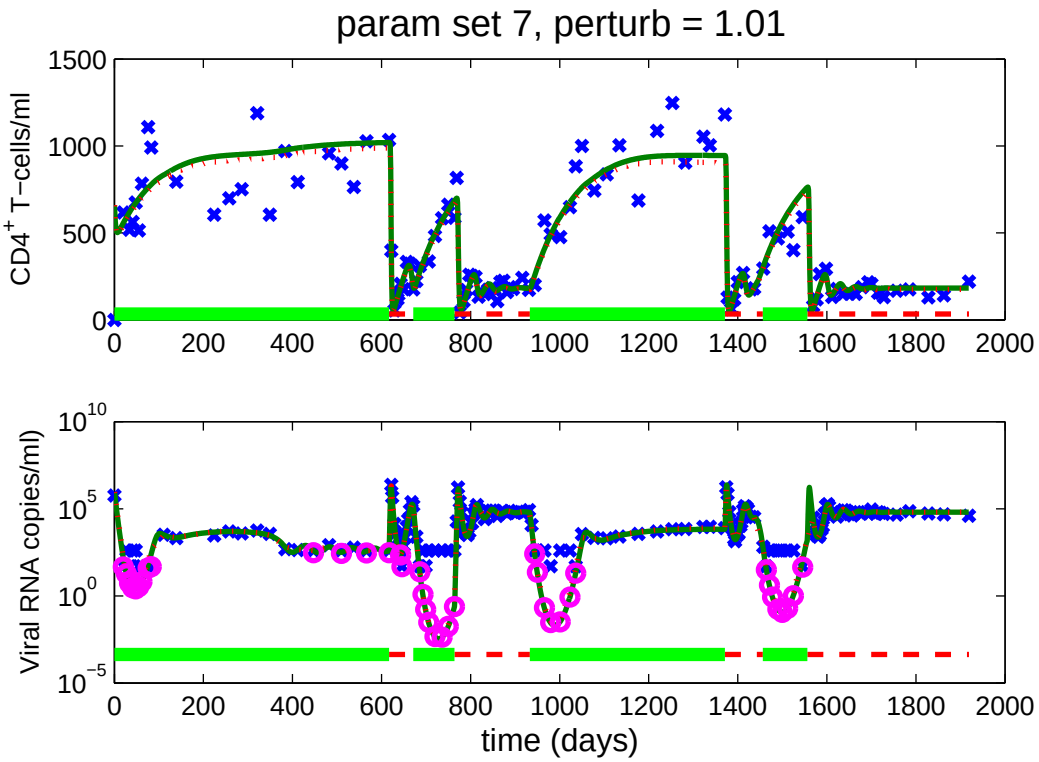


Figure 5.24: Sample fit to simulated censored data ($\bar{\sigma} = 0.2$) with censored data algorithm for seventh parameter set with initial iterate $1.01q^0$. Red dotted line indicates true model solution from which data were generated, solid green line is model at estimated parameters, 'x' denotes observed data, and 'o' the predicted value of a censored data point.

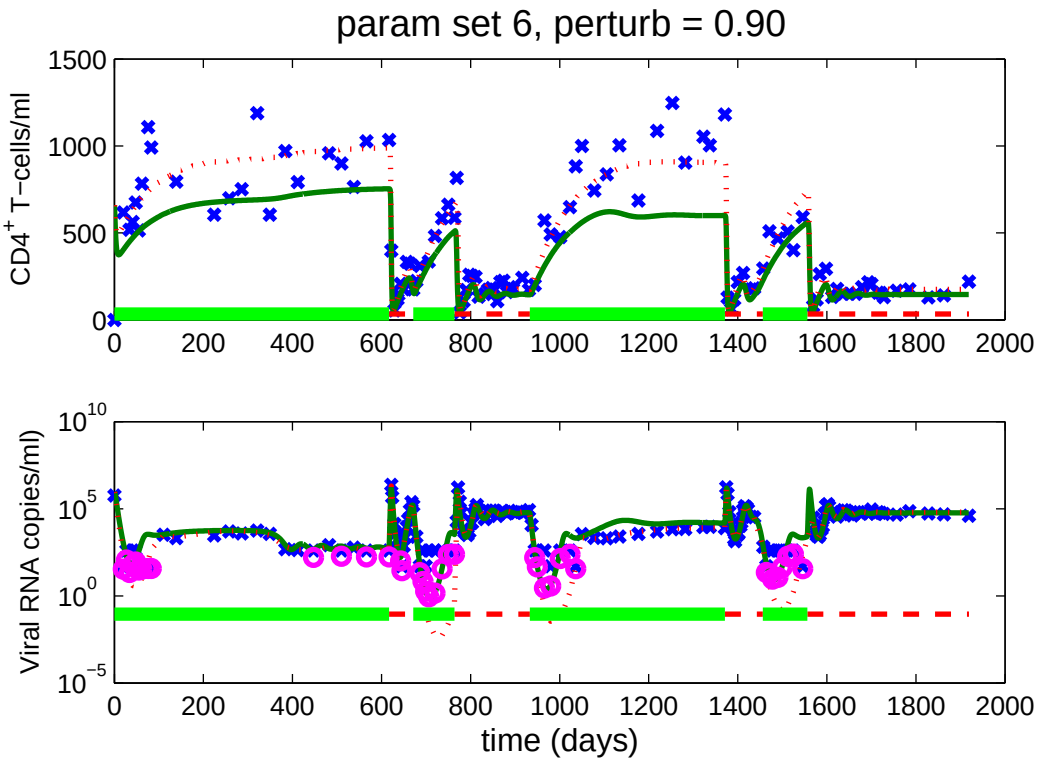


Figure 5.25: Sample fit to simulated censored data ($\bar{\sigma} = 0.2$) with censored data algorithm for sixth parameter set with initial iterate $0.90q^0$. Red dotted line indicates true model solution from which data were generated, solid green line is model at estimated parameters, 'x' denotes observed data, and 'o' the predicted value of a censored data point.

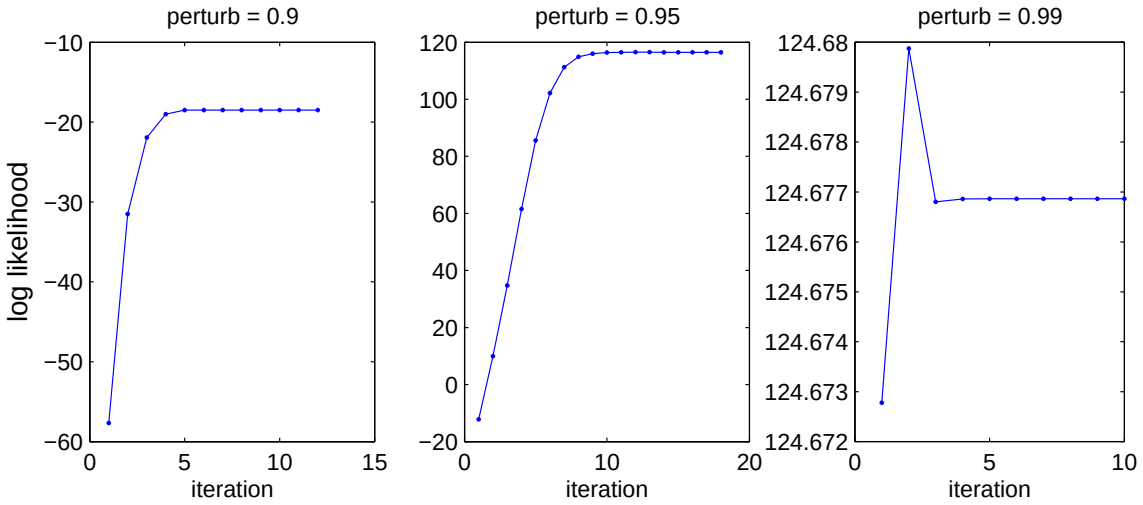


Figure 5.26: Simulated data censored at 400 and 50. Plot of log likelihood vs. iteration for various perturbations of true parameters considered, parameter set 2.

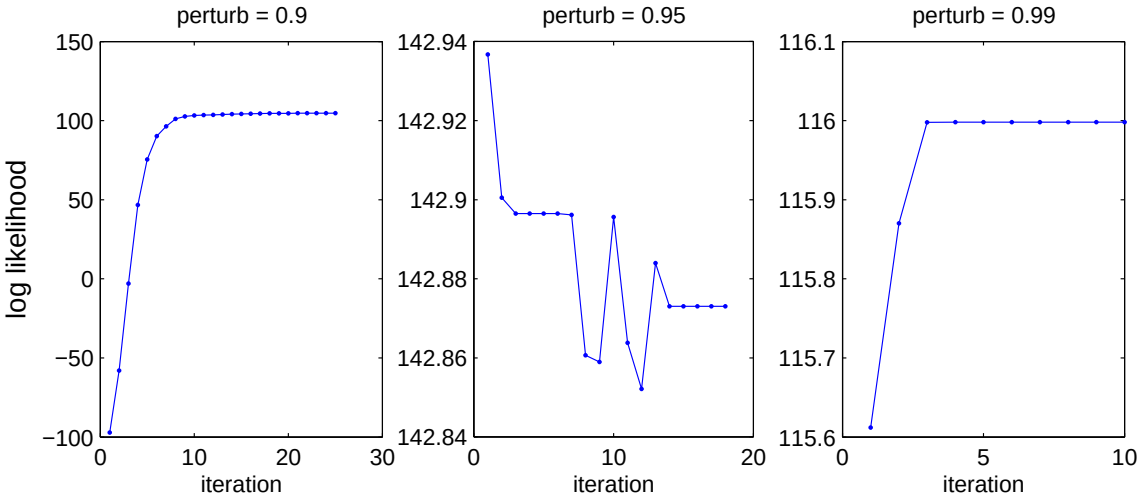


Figure 5.27: Simulated data censored at 400 and 50. Plot of log likelihood vs. iteration for various perturbations of true parameters considered, parameter set 3.

Table 5.5: Simulated censored data: sample parameter estimates and standard errors for perturbation $0.9q^0$ for $\bar{\sigma} = 0.2$ and 0.3 cases.

		$\bar{\sigma} = 0.2$		$\bar{\sigma} = 0.3$	
set	parameter	q^*	std. err.	q^*	std. err.
1	ϵ_1	6.9436e-01	1.3612e-03	6.9651e-01	1.4855e-03
	ϵ_2	4.0586e-01	2.8390e-05	4.1107e-01	3.0788e-05
	c	1.3189e+01	1.9757e-03	1.2976e+01	2.1429e-03
	b_E	2.9974e-01	8.7447e-04	2.9954e-01	9.5432e-04
2	ϵ_1.	6.8938e-01	2.2742e-03	6.9524e-01	2.5719e-03
	k_1	7.9385e-07	5.8771e-09	8.0319e-07	6.6437e-09
	ϵ_2	4.1819e-01	2.8514e-05	4.1521e-01	3.0843e-05
	c	1.3020e+01	1.8526e-03	1.2976e+01	2.0036e-03
3	ϵ_1.	6.9104e-01	2.2812e-03	6.9380e-01	2.4899e-03
	k_1	7.9747e-07	6.5685e-09	8.0464e-07	7.2000e-09
	ϵ_2	4.1755e-01	1.3725e-05	4.1572e-01	1.4867e-05
	b_E	3.0031e-01	9.1678e-04	3.0089e-01	9.9595e-04
4	λ_1	1.0460e+01	7.4973e-02	1.0184e+01	8.6784e-02
	ϵ_1	6.9822e-01	2.0118e-03	7.0359e-01	2.3248e-03
	c	1.3575e+01	9.0735e-04	1.3155e+01	1.0463e-03
	b_E	3.0449e-01	1.0498e-03	2.9982e-01	1.2124e-03
5	λ_1	1.0394e+01	1.2599e-01	1.0030e+01	2.0745e-01
	ϵ_1	7.0058e-01	2.1993e-03	7.2303e-01	4.9130e-03
	k_1	7.7989e-07	1.1874e-08	8.1755e-07	1.9072e-08
	c	1.3253e+01	5.8492e-04	1.2632e+01	1.3449e-03
6	λ_1	1.0561e+01	4.6609e-02	1.0171e+01	5.5700e-02
	ϵ_2	3.9166e-01	2.6023e-05	4.0444e-01	3.0515e-05
	c	1.3777e+01	1.8311e-03	1.3157e+01	2.1527e-03
	b_E	3.0544e-01	1.0349e-03	3.0069e-01	1.2370e-03
7	λ_1	1.0320e+01	1.2491e-01	1.0034e+01	1.8584e-01
	k_1	7.8575e-07	1.1156e-08	8.0967e-07	1.6595e-08
	ϵ_2	3.9955e-01	2.7317e-05	4.0522e-01	3.0625e-05
	c	1.3287e+01	1.7785e-03	1.3144e+01	1.9962e-03
8	λ_1	1.0615e+01	1.4547e-01	1.0076e+01	1.7155e-01
	k_1	7.8901e-07	1.1291e-08	8.2063e-07	1.3317e-08
	c	1.3526e+01	8.4411e-04	1.3386e+01	9.8614e-04
	b_E	3.0440e-01	1.0699e-03	3.0222e-01	1.2517e-03

Chapter 6

Model Fitting to Clinical Data

6.1 Fits to Individual Patients

In this section we review results for fitting the dynamic model to data from each of the 59 patients in PS59 (patients with at least ten viral load and ten CD4 measurements). As described in Section 4.4, we fit the model to each patient's data on the log scale using a least-squares cost criterion by (1) estimating all 20 model parameters and seven initial conditions with `DIRECT`, (2) refining those estimated with `lsqnonlin`, and (3) applying the iterative censored data algorithm to estimate a subset of eight model parameters $(\lambda_1, d_1, \epsilon_1, k_1, \epsilon_2, NT, c, b_E)$ and three initial conditions (T_1^0, T_1^{*0}, V_I^0). The estimates from step (2) serve as initial iterates for the censored data estimation process.

During both the initial least squares fitting and the censored data algorithm we employ bound constraints on all parameters as shown in Table 6.1. It is difficult to get reliable model predictions of CD4 data as it is erratic on a much faster time scale than the long-term model dynamics typically predict, but the estimation process is helped by further restricting the initial condition. If for patient j, $mCD4_j$ denotes the

Table 6.1: Bounds employed when estimating parameters from clinical data.

parameter	lower bd.	upper bd.	parameter	lower bd.	upper bd.
λ_1	0.1	100	m_2	1.0×10^{-8}	0.001
d_1	0.001	0.1	ϵ_2	0	0.99
ϵ_1	0	0.99	N_T	10	1000
k_1	1.0×10^{-9}	1.0×10^{-4}	c	1	100
λ_2	0.001	1	λ_E	1.0×10^{-5}	0.1
d_2	0.001	0.1	b_E	0.01	0.9886
f	0	0.99	K_b	0.001	10
k_2	1.0×10^{-7}	0.01	d_E	0.01	0.9886
δ	0.01	0	K_d	0.001	10
m_1	1.0×10^{-8}	0.001	δ_E	0.01	1
T_1^0 †	100	1.0×10^4	V_I^0	100	1.0×10^6
T_2^0	0.1	100	V_{NI}^0	1	1.0×10^4
T_1^{*0} †	0.01	100	E^0	1.0×10^{-4}	10
T_2^{*0}	1.0×10^{-5}	100			

minimum observed CD4 value over the time span and $MCD4_j$ the maximum, then we desire $0.9mCD4_j \leq \bar{T}_1^{0j} \leq 1.1MCD4_j$ and $\bar{T}_1^{*0j} \leq MCD4_j$ and when tighter, enforce these constraints over those in the table.

Appendix A contains sample model fits to data for all 59 patients. The model fit for Patient 5 is also shown in Figure 6.1. The top panel corresponds to $CD4^+$ T-cell data and includes crosses denoting clinical data and a solid line representing the model solution evaluated with the estimated parameters. In the bottom panel, crosses again denote data and solid line model solution. The added hollow circles represent censored data points estimated by the censored data algorithm. The line under each graph represents treatment, with a thicker solid line denoting on treatment and thin dashed line indicating off treatment. The model reasonably predicts the viral load behavior, including during therapy interruptions, and is capable of capturing the long-term off-treatment steady state exhibited by this patient. The model fits the overall trend in the $CD4^+$ T-cell data, which fluctuate substantially even in healthy

patients. Since long-term CD4 decline is an important clinical marker for transition to AIDS, model fits such as this may provide important interpretation of overall data dynamics. Figure 6.2 contains plots for all the model states.

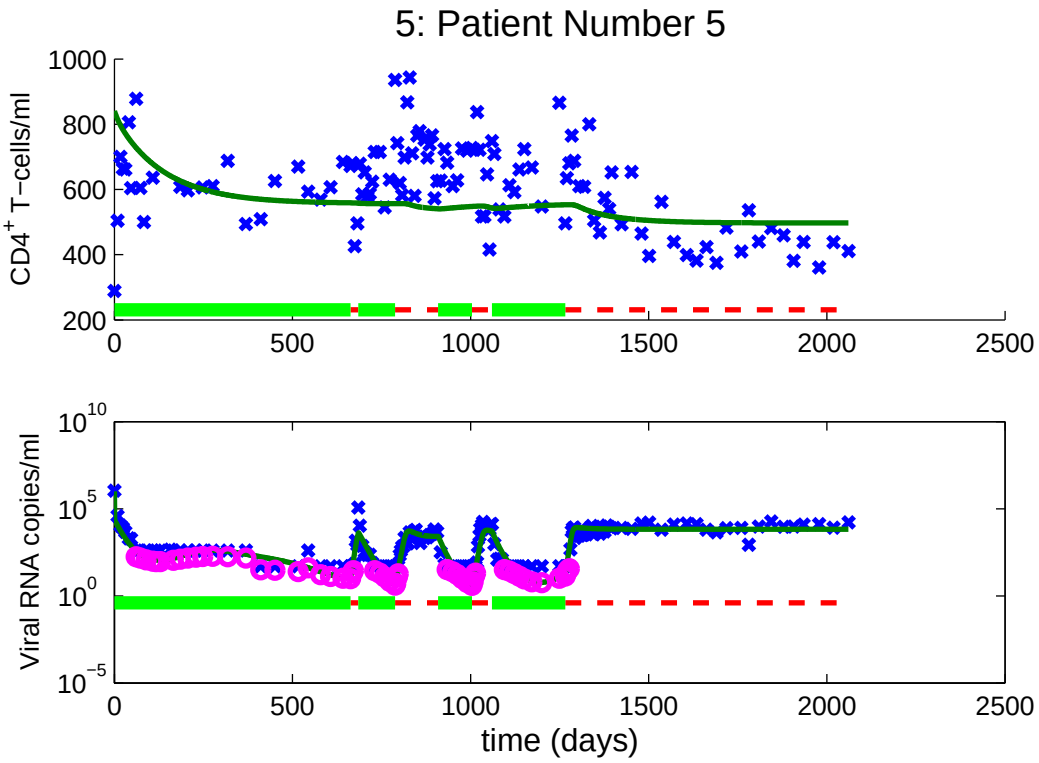


Figure 6.1: Model fit to $CD4^+$ T-cell data (top) and viral load data (bottom) for Patient 5. Crosses denote data and solid line, model evaluated with estimated parameters. Underlying thick solid line denotes on treatment period and thin dashed line, off.

Notice from Figure 6.3, for a patient on treatment for nearly five years, that the model is capable of predicting long-term suppression of virus at a non-zero level. Simpler models of HIV infection often predict extinction of virus during extended therapy periods. This patient also exhibits a slowly increasing CD4 count, possibly due to the success of therapy. Figure 6.4 contains plots for all the model states.

Tables 6.2 and 6.3 show estimated parameters and standard errors for patients with indices 1–30. Tables 6.4 and 6.5 show estimated parameters and standard errors

for patients with indices 31–59. Dashes indicate a reasonable estimate of standard error was not available due to the small number of data points. In these cases the unbiased estimator

$$\hat{\sigma}_s^2 = \frac{1}{N_s^j - p} \sum_{i=1}^{N_s^j} |z_s^{ij}(q^*) - y_s^{ij}|^2$$

is not well-defined, since $p = 11$ and for CD4 or RNA taken individually, N_s^j may be as small as 11 or 10, even though there are always at least 20 time series points in total for the two states. For most patients, the standard errors computed would lead to considerable confidence in the estimates of $\epsilon_2, c, T_1^0, V_I^0$ and less frequently, N_T. We have little confidence is most of the estimates of b_E and T_1^{*0}. For the first four parameters $\lambda_1, d_1, \epsilon_1,$ and k_1 we have substantial standard errors on the order of the parameter estimated. This may be due to identifiability issues since they all enter into the same dynamics equation for $\dot{T}_1$.

As a means for understanding the population variability, in Section 6.2 we present histograms with the distribution of the estimated parameters across the population.

Table 6.6 contains statistics for the censored data algorithm for all 59 patients. We present estimates of σ_1 and σ_2 on the log scale and number of iterations required for the censored data algorithm to converge. In general, the estimates of standard deviation in the viral RNA data tend to be considerably larger than the CD4 data. In most cases the algorithm converges in fewer than the specified 251 maximum iterations (including one preliminary least squares estimate).

Tables 6.7 and 6.8 show the predicted model equilibria, given the parameters estimated for each patient. We report the uninfected equilibrium and one or two equilibria with infection present. For many patients, estimated parameters dictate only a single stable equilibrium. All are presented for the case with no treatment applied to give an idea of what would happen to each patient if taken off therapy. In

Table 6.2: Clinical data: estimated parameters q^* and corresponding standard errors (se) for Patients 1–30, first six parameters estimated.

pat #		λ_1	d_1	ϵ_1	k_1	ϵ_2	N_T
1	q^*	1.01e+00	1.00e-03	5.75e-01	1.44e-08	7.43e-01	1.31e+02
	se	6.96e+00	7.40e-03	6.86e-02	6.09e-09	3.15e-04	1.35e+01
2	q^*	1.23e+00	1.73e-03	6.14e-01	4.35e-09	6.51e-01	1.26e+02
	se	6.23e+00	8.63e-03	1.33e-01	1.49e-08	2.35e-03	2.14e+01
3	q^*	1.39e+00	1.00e-03	2.71e-01	1.46e-07	4.65e-01	8.17e+01
	se	9.88e-01	1.17e-03	1.27e-01	1.61e-08	6.97e-04	9.72e+00
4	q^*	3.71e+00	6.06e-03	5.10e-01	1.65e-07	4.41e-01	9.41e+01
	se	2.81e+00	4.64e-03	6.73e-02	8.34e-09	3.97e-04	6.96e+00
5	q^*	4.06e+00	7.29e-03	7.01e-01	1.43e-07	6.06e-01	1.06e+02
	se	2.67e+00	4.87e-03	1.42e-01	5.93e-09	3.38e-04	6.48e+00
6	q^*	3.70e+01	4.27e-02	2.14e-01	4.65e-08	5.71e-01	1.01e+02
	se	1.16e+02	1.35e-01	9.05e-02	7.12e-09	4.57e-04	8.78e+00
7	q^*	5.46e+00	9.20e-03	6.56e-01	8.63e-07	7.16e-01	1.03e+02
	se	2.96e+00	5.18e-03	2.26e-01	7.62e-07	3.14e-03	4.09e+01
8	q^*	8.85e-01	1.13e-03	4.44e-01	1.43e-09	4.46e-01	9.18e+01
	se	7.54e-01	1.26e-03	3.89e+00	4.62e-07	1.14e-02	3.28e+02
9	q^*	8.90e+00	2.02e-02	7.97e-01	3.30e-07	5.71e-01	1.18e+02
	se	1.39e+01	3.29e-02	4.78e-01	2.10e-07	2.60e-03	7.22e+01
10	q^*	1.84e+00	2.37e-03	9.12e-02	3.51e-09	5.23e-01	8.72e+01
	se	4.75e+00	6.24e-03	2.87e-01	1.87e-07	1.84e-02	2.15e+01
11	q^*	1.11e+00	1.76e-03	1.19e-01	1.67e-09	6.92e-01	9.69e+01
	se	1.13e+00	2.28e-03	1.84e+00	1.47e-08	2.09e-03	1.46e+01
12	q^*	4.45e-01	1.13e-03	6.16e-01	2.00e-08	4.55e-01	8.53e+01
	se	1.32e+00	1.81e-03	8.76e-01	3.15e-08	1.55e-03	3.21e+01
13	q^*	2.79e+01	2.89e-02	4.70e-01	8.18e-08	5.05e-01	9.11e+01
	se	9.86e+01	1.03e-01	2.82e-01	6.72e-09	1.14e-03	2.19e+01
14	q^*	1.44e+00	2.20e-03	5.43e-01	3.12e-07	6.09e-01	7.60e+01
	se	6.42e+00	1.05e-02	1.24e-01	3.02e-07	5.25e-02	8.85e+01
15	q^*	1.04e+01	1.18e-02	6.73e-01	1.18e-09	6.78e-01	1.37e+02
	se	5.74e+02	6.52e-01	2.12e-01	1.47e-08	1.05e-03	3.16e+01
16	q^*	8.36e-01	2.61e-03	7.99e-01	2.86e-07	8.42e-01	2.18e+02
	se	5.87e-01	1.40e-03	5.77e-01	3.70e-07	9.31e-02	2.33e+02
17	q^*	3.12e+00	9.49e-03	3.72e-01	3.96e-07	4.83e-01	9.26e+01
	se	7.30e+00	1.80e-02	1.79e-01	2.41e-07	4.54e-02	2.63e+02
18	q^*	1.54e+00	1.16e-03	5.76e-01	3.64e-09	4.55e-01	7.89e+01
	se	4.78e+00	4.05e-03	3.39e-01	1.10e-07	3.92e-03	2.87e+01
19	q^*	4.95e+00	6.07e-03	6.00e-01	1.61e-07	4.69e-01	9.87e+01
	se	3.33e+00	4.51e-03	1.51e-01	1.07e-07	1.25e-03	1.74e+01
20	q^*	4.58e+00	1.14e-02	8.22e-01	1.34e-07	7.54e-01	1.73e+02
	se	4.93e+00	1.34e-02	8.90e-02	2.42e-08	9.72e-04	8.64e+01
21	q^*	4.28e+00	1.24e-02	6.39e-01	1.93e-07	6.00e-01	1.21e+02
	se	5.07e+00	1.60e-02	1.74e-01	3.14e-08	1.15e-03	3.83e+01
22	q^*	1.74e+00	4.41e-03	5.55e-01	1.18e-07	6.15e-01	1.17e+02
	se	2.27e+00	5.55e-03	5.27e-02	2.78e-08	1.78e-04	1.00e+01
23	q^*	7.70e+00	1.24e-02	6.15e-01	1.91e-07	5.19e-01	1.06e+02
	se	2.87e+00	4.92e-03	1.69e-01	2.13e-08	1.00e-03	3.17e+01
24	q^*	2.96e+01	3.58e-02	9.90e-01	1.55e-07	5.71e-04	7.01e+01
	se	3.12e+01	3.82e-02	1.15e-01	1.15e-07	2.76e-01	1.82e+01
25	q^*	3.07e+00	2.32e-03	7.73e-01	4.83e-07	7.21e-01	8.11e+01
	se	1.50e+00	1.27e-03	2.39e-02	1.11e-07	1.26e-03	1.37e+00
26	q^*	5.42e+00	6.56e-03	6.49e-01	4.38e-09	4.50e-01	9.42e+01
	se	7.38e+00	9.71e-03	6.81e+00	1.33e-06	2.10e-02	2.78e+02
27	q^*	7.59e+00	1.20e-02	6.31e-01	2.16e-09	6.24e-01	1.03e+02
	se	7.02e+00	1.13e-02	2.11e+00	2.20e-06	7.84e-03	1.15e+02
28	q^*	7.51e-01	1.98e-03	6.70e-01	3.53e-07	6.96e-01	5.87e+01
	se	1.28e+00	2.62e-03	5.06e-02	3.14e-07	1.20e-03	8.43e+00
29	q^*	9.18e+00	1.29e-02	7.69e-01	1.71e-07	4.82e-01	8.49e+01
	se	3.59e+01	5.18e-02	3.02e-01	1.99e-08	1.21e-03	2.55e+01
30	q^*	1.90e+00	5.10e-03	2.41e-01	2.87e-09	1.24e-01	3.76e+01
	se	7.11e+00	1.86e-02	6.55e+01	6.54e-07	1.96e+00	6.48e+02

Table 6.3: Clinical data: estimated parameters q^* and corresponding standard errors (se) for Patients 1–30, last five parameters estimated.

pat #		c	b_E	T_1^0	${T_1*}^0$	V_I^0
1	q^*	6.10e+00	1.01e-02	2.94e+00	-2.00e+00	3.79e+00
	se	1.27e-02	2.75e+06	6.96e-02	5.05e+00	1.12e-01
2	q^*	7.68e+00	1.09e-01	2.87e+00	-1.64e+00	5.45e+00
	se	1.49e-01	1.07e-03	3.69e-02	2.05e+03	3.56e-01
3	q^*	9.78e+00	1.30e-01	2.52e+00	1.69e+00	5.84e+00
	se	5.98e-03	2.84e-03	1.62e-01	3.01e-01	5.34e-01
4	q^*	9.95e+00	1.01e-01	3.00e+00	-1.81e+00	6.00e+00
	se	8.20e-03	1.64e-03	8.61e-02	2.98e+02	4.74e-01
5	q^*	8.70e+00	1.03e-01	2.92e+00	9.61e-01	6.00e+00
	se	8.38e-03	3.30e-04	8.08e-02	2.65e-01	1.78e-01
6	q^*	8.76e+00	1.36e-01	2.70e+00	2.15e-01	5.47e+00
	se	5.84e-03	9.43e-03	5.01e-01	4.99e+00	7.74e-01
7	q^*	8.47e+00	1.32e-01	2.54e+00	-1.10e+00	5.85e+00
	se	3.70e-01	4.56e-03	1.53e-01	8.03e+02	4.27e-01
8	q^*	9.97e+00	4.90e-02	2.62e+00	-1.42e+00	5.86e+00
	se	2.27e+00	4.42e+05	7.73e-02	1.20e+03	5.73e-01
9	q^*	8.03e+00	6.73e-02	2.63e+00	-8.95e-01	5.93e+00
	se	5.97e-02	2.63e+00	9.86e-02	1.66e+02	1.04e+00
10	q^*	1.02e+01	1.11e-01	2.86e+00	-7.36e-01	4.39e+00
	se	6.93e+00	2.39e-01	1.85e-02	6.62e+00	8.99e-01
11	q^*	4.67e+00	1.60e-01	2.52e+00	9.90e-01	4.77e+00
	se	1.30e-01	6.74e-03	5.95e-02	1.21e+00	2.16e-01
12	q^*	1.07e+01	2.59e-01	2.98e+00	8.39e-01	5.67e+00
	se	1.09e-02	3.80e-02	3.99e-02	3.38e-01	6.19e-01
13	q^*	1.03e+01	1.73e-01	2.90e+00	-9.78e-01	5.04e+00
	se	1.30e-02	8.88e-03	1.47e-01	7.85e+00	1.75e-01
14	q^*	1.20e+01	5.83e-02	2.71e+00	-1.89e-01	4.46e+00
	se	7.27e+00	8.93e+00	1.12e-01	2.34e+00	7.08e-01
15	q^*	7.46e+00	1.15e-01	3.09e+00	-4.76e-01	5.93e+00
	se	5.16e-03	2.00e-03	2.53e+00	4.40e+00	7.81e-01
16	q^*	3.76e+00	1.55e-01	2.95e+00	9.71e-01	5.86e+00
	se	3.12e+00	4.82e-03	6.83e-02	5.58e+00	2.92e-01
17	q^*	1.06e+01	1.02e-01	2.93e+00	6.62e-01	4.48e+00
	se	2.76e+01	1.03e-02	2.21e-01	2.58e+00	1.15e+00
18	q^*	1.00e+01	1.96e-01	3.00e+00	-7.07e-01	4.30e+00
	se	4.86e-01	2.71e-02	5.21e-02	1.49e+01	9.50e-01
19	q^*	9.36e+00	9.20e-02	2.71e+00	1.96e+00	5.87e+00
	se	1.05e-01	5.98e-03	1.66e-01	9.60e-01	4.47e-01
20	q^*	7.34e+00	5.18e-02	2.40e+00	-4.35e-01	3.90e+00
	se	8.14e-02	1.72e+02	1.09e-01	4.20e-01	7.83e-01
21	q^*	6.14e+00	1.12e-01	2.44e+00	-6.07e-01	5.83e+00
	se	8.00e-03	3.30e+03	1.84e-01	3.84e+01	2.97e-01
22	q^*	7.98e+00	1.50e-01	2.85e+00	2.00e+00	5.75e+00
	se	3.52e-02	5.96e-03	1.35e-01	2.95e-01	6.17e-01
23	q^*	9.25e+00	8.60e-02	2.99e+00	4.78e-01	5.70e+00
	se	4.15e-02	1.05e+01	1.45e-01	2.89e+00	5.27e-01
24	q^*	7.81e+00	7.28e-02	3.01e+00	-1.54e+00	5.54e+00
	se	1.82e+00	2.77e+03	9.42e-02	4.21e+01	3.33e-01
25	q^*	1.43e+01	1.65e-01	2.83e+00	1.31e+00	5.87e+00
	se	1.66e-01	1.84e-03	7.10e-02	7.04e-01	5.07e-01
26	q^*	1.06e+01	1.16e-01	2.71e+00	1.55e+00	4.88e+00
	se	1.72e+00	1.02e-01	1.28e-01	1.64e+00	9.51e-01
27	q^*	9.05e+00	2.23e-02	2.56e+00	-1.46e+00	4.99e+00
	se	1.94e-01	7.13e+02	3.26e-01	3.70e+03	5.65e-01
28	q^*	1.34e+01	2.42e-01	3.01e+00	5.10e-01	5.09e+00
	se	1.33e-01	5.44e-03	1.41e-01	2.96e+01	4.15e-01
29	q^*	1.03e+01	1.01e-01	2.77e+00	-2.28e-01	5.72e+00
	se	4.08e-03	1.04e-02	1.61e-01	6.17e+00	5.45e-01
30	q^*	1.09e+01	5.90e-01	2.65e+00	-1.03e+00	5.69e+00
	se	3.34e+01	8.95e-02	1.21e-01	2.25e+03	2.73e-01

Table 6.4: Clinical data: estimated parameters q^* and corresponding standard errors (se) for Patients 31–59, first six parameters estimated.

pat #		λ_1	d_1	ϵ_1	k_1	ϵ_2	N_T
31	q^*	2.68e+00	3.11e-03	8.40e-01	3.17e-07	3.01e-01	8.68e+01
	se	5.91e+00	7.74e-03	7.37e-01	1.33e-06	1.56e-02	3.89e+01
32	q^*	3.55e+00	6.94e-03	9.63e-01	4.84e-07	3.77e-02	2.16e+01
	se	7.07e+00	1.43e-02	2.35e+00	7.06e-06	1.36e-01	4.20e+01
33	q^*	4.04e+00	7.13e-03	8.08e-01	3.78e-07	4.30e-01	8.79e+01
	se	4.52e+00	9.31e-03	7.40e-01	1.53e-07	3.62e-03	6.12e+01
34	q^*	2.10e+01	2.69e-02	8.17e-01	4.33e-07	4.60e-01	8.12e+01
	se	8.07e+01	1.08e-01	7.05e-01	1.51e-07	6.55e-03	1.47e+02
35	q^*	3.13e+00	4.39e-03	7.21e-01	3.51e-08	7.25e-01	1.53e+02
	se	9.94e+00	1.32e-02	1.54e-01	1.50e-08	1.79e-03	7.49e+01
36	q^*	7.22e-01	1.00e-03	7.93e-01	1.02e-07	1.14e-01	8.51e+01
	se	2.61e+00	4.94e-03	1.75e+00	2.73e-07	1.16e-01	3.54e+02
37	q^*	4.06e+00	1.10e-02	6.91e-01	3.63e-07	6.00e-01	7.58e+01
	se	5.43e+00	1.35e-02	2.32e-01	6.15e-08	1.20e-03	1.61e+01
38	q^*	1.60e+00	1.09e-03	6.41e-01	8.27e-08	4.46e-01	8.98e+01
	se	3.04e+00	3.36e-03	3.44e-01	8.33e-09	1.03e-03	3.03e+01
39	q^*	6.19e+00	1.11e-02	3.71e-01	3.64e-07	5.14e-01	8.73e+01
	se	2.51e+00	5.61e-03	1.07e+00	7.50e-08	1.09e-02	1.94e+02
40	q^*	2.89e+01	3.11e-02	9.20e-01	2.21e-07	6.01e-01	9.19e+01
	se	1.13e+02	1.32e-01	3.47e+01	2.55e-07	1.99e-01	5.49e+03
41	q^*	5.88e+00	7.63e-03	7.34e-01	2.37e-07	6.18e-01	7.72e+01
	se	0.00e+00	0.00e+00	0.00e+00	0.00e+00	0.00e+00	0.00e+00
42	q^*	2.35e+00	7.45e-03	7.79e-01	1.12e-09	9.62e-02	9.98e+01
	se	6.23e+00	2.04e-02	1.12e+02	1.56e-07	5.76e-01	2.35e+03
43	q^*	3.14e+00	5.11e-03	8.78e-01	2.05e-07	2.79e-01	9.82e+01
	se	1.52e+01	2.53e-02	1.45e-02	1.92e-07	2.97e-03	2.40e+00
44	q^*	2.97e+00	7.21e-03	4.98e-01	3.34e-08	1.61e-01	9.93e+01
	se	4.14e+00	1.09e-02	1.05e+01	2.33e-07	5.58e-02	7.23e+02
45	q^*	4.99e+00	5.18e-03	7.90e-01	1.95e-07	6.43e-01	1.35e+02
	se	-	-	-	-	-	-
46	q^*	2.66e+01	3.44e-02	8.89e-01	1.43e-08	1.72e-01	8.71e+01
	se	5.49e+03	7.12e+00	2.53e+01	3.60e-06	1.19e-01	5.78e+02
47	q^*	3.01e+00	3.70e-03	4.54e-01	2.58e-07	5.08e-01	1.03e+02
	se	4.38e+00	6.30e-03	2.66e-01	1.56e-07	2.37e-03	1.03e+02
48	q^*	4.26e+00	9.94e-03	4.01e-01	3.47e-07	4.63e-01	8.85e+01
	se	5.61e+00	1.40e-02	6.29e-01	1.08e-07	4.34e-03	6.07e+01
49	q^*	5.15e+00	5.10e-03	4.32e-01	8.50e-08	4.15e-01	1.02e+02
	se	8.45e+00	1.12e-02	3.96e-01	1.00e-07	8.80e-03	4.09e+01
50	q^*	2.74e+00	3.27e-03	5.06e-01	1.29e-07	5.03e-01	1.18e+02
	se	1.60e+00	3.12e-03	2.00e-01	1.54e-07	8.15e-04	2.33e+01
51	q^*	4.96e+00	7.30e-03	9.86e-01	1.12e-06	3.86e-01	7.21e+01
	se	5.03e+00	1.05e-02	8.00e-01	1.46e-06	3.14e-02	3.62e+02
52	q^*	2.64e+00	4.92e-03	6.10e-01	1.38e-09	1.89e-01	9.33e+01
	se	3.15e+02	5.82e-01	4.44e+01	2.14e-07	1.31e-01	1.05e+03
53	q^*	3.22e-01	1.31e-03	2.58e-02	1.94e-09	1.25e-01	9.67e+01
	se	2.43e+01	3.18e-02	2.91e+01	1.83e-07	4.40e-02	3.95e+00
54	q^*	2.42e+00	2.36e-03	4.12e-01	1.25e-07	3.51e-01	7.79e+01
	se	6.56e+00	1.13e-02	3.06e+00	1.10e-06	3.06e-02	1.92e+02
55	q^*	2.83e+01	9.22e-02	5.73e-01	1.15e-06	5.11e-01	8.25e+01
	se	2.91e+01	1.33e-01	9.59e-01	1.79e-06	7.99e-03	9.86e+01
56	q^*	1.60e+00	2.26e-03	6.09e-01	6.76e-09	4.74e-01	9.75e+01
	se	-	-	-	-	-	-
57	q^*	3.63e+00	3.11e-03	8.37e-01	3.85e-07	3.80e-01	5.99e+01
	se	-	-	-	-	-	-
58	q^*	9.68e-01	1.00e-03	5.25e-01	1.06e-09	5.58e-01	8.87e+01
	se	1.20e+01	2.01e-02	8.18e+01	5.01e-06	1.15e+01	1.44e+04
59	q^*	2.79e+00	4.88e-03	7.26e-01	4.26e-08	6.47e-01	2.26e+02
	se	-	-	-	-	-	-

Table 6.5: Clinical data: estimated parameters q^* and corresponding standard errors (se) for Patients 31–59, last five parameters estimated.

pat #		c	b_E	T_1^0	$T_1{*}^0$	V_I^0
31	q^*	8.45e+00	1.84e-01	2.83e+00	-3.12e-01	6.00e+00
	se	4.34e-01	1.71e-02	1.63e-01	2.43e+02	5.28e-01
32	q^*	8.35e+00	5.04e-01	2.82e+00	1.91e+00	5.91e+00
	se	1.51e+00	1.12e-01	4.74e-01	3.38e+01	1.97e-01
33	q^*	1.13e+01	6.99e-02	2.83e+00	-1.84e-01	5.22e+00
	se	4.23e-02	4.04e+02	2.16e-01	9.53e+00	8.68e-01
34	q^*	9.63e+00	2.33e-01	2.78e+00	6.53e-02	5.47e+00
	se	2.01e-01	5.22e-02	1.46e-01	1.66e+00	5.48e-01
35	q^*	4.68e+00	1.22e-02	3.01e+00	-1.96e+00	2.09e+00
	se	8.47e-02	1.50e+05	2.12e-01	1.11e+01	4.42e+00
36	q^*	8.32e+00	2.70e-02	2.66e+00	-1.77e+00	3.75e+00
	se	6.89e-01	5.26e+05	1.03e-01	6.58e+03	6.11e-01
37	q^*	1.13e+01	1.30e-01	2.89e+00	3.48e-01	6.00e+00
	se	2.93e-02	1.21e+01	1.37e-01	8.59e+00	6.30e-01
38	q^*	9.97e+00	1.90e-01	2.73e+00	8.84e-01	5.12e+00
	se	1.31e-02	3.84e+02	2.20e-01	6.80e-01	5.27e-01
39	q^*	1.27e+01	1.04e-01	2.59e+00	-1.20e+00	4.44e+00
	se	1.92e+00	1.55e+00	8.77e-02	1.11e+02	4.79e-01
40	q^*	1.10e+01	4.17e-01	2.83e+00	-2.00e+00	2.02e+00
	se	5.12e+00	1.20e+00	2.23e-01	4.54e+02	1.78e+03
41	q^*	9.90e+00	9.28e-02	2.81e+00	7.80e-01	4.12e+00
	se	0.00e+00	0.00e+00	0.00e+00	0.00e+00	0.00e+00
42	q^*	8.83e+00	1.88e-01	2.38e+00	1.62e+00	4.95e+00
	se	6.60e+00	1.72e-02	2.16e-01	2.96e+00	7.53e-01
43	q^*	1.14e+01	4.92e-01	2.85e+00	6.54e-01	5.67e+00
	se	6.92e-02	2.39e-02	1.45e-01	9.08e+00	8.24e-02
44	q^*	1.04e+01	1.32e-01	2.39e+00	1.00e+00	5.16e+00
	se	7.31e-01	5.29e-02	1.29e-01	4.02e+00	1.01e+00
45	q^*	9.13e+00	1.41e-01	2.55e+00	1.85e+00	5.88e+00
	se	-	-	-	-	-
46	q^*	1.20e+01	2.05e-01	2.89e+00	5.87e-01	5.88e+00
	se	2.32e+00	1.97e-01	2.96e+00	6.01e+02	6.09e-01
47	q^*	1.01e+01	1.19e-01	2.70e+00	6.82e-01	5.73e+00
	se	5.12e+00	2.18e-02	8.25e-02	3.87e+00	2.46e-01
48	q^*	9.62e+00	1.51e-01	2.70e+00	6.63e-01	4.36e+00
	se	6.63e-02	9.66e-03	1.48e-01	1.35e+00	5.97e-01
49	q^*	8.58e+00	1.20e-01	2.56e+00	1.90e+00	5.79e+00
	se	6.19e-01	7.93e-03	2.71e-01	7.53e-01	5.69e-01
50	q^*	8.46e+00	1.22e-01	2.31e+00	2.00e+00	5.34e+00
	se	7.73e-02	2.26e-02	1.71e-01	5.08e-01	2.90e-01
51	q^*	8.93e+00	1.73e-02	2.50e+00	5.53e-01	5.92e+00
	se	9.11e-01	1.26e+10	4.36e-01	3.46e+01	1.87e+00
52	q^*	8.95e+00	1.02e-01	2.74e+00	1.50e+00	4.36e+00
	se	1.90e+00	6.84e-02	2.24e-01	2.52e+00	4.82e-01
53	q^*	9.98e+00	5.04e-01	2.99e+00	5.77e-01	5.17e+00
	se	3.20e-01	3.95e-02	3.44e-01	1.28e+03	1.00e+00
54	q^*	9.85e+00	1.11e-01	2.57e+00	7.52e-01	5.02e+00
	se	1.03e+00	1.84e-01	1.18e-01	2.31e+00	9.89e-01
55	q^*	9.40e+00	2.51e-02	2.41e+00	7.46e-01	5.86e+00
	se	2.05e-02	4.96e+02	1.10e+00	4.04e+01	7.46e-01
56	q^*	1.01e+01	4.99e-01	2.96e+00	-1.81e+00	2.81e+00
	se	-	-	-	-	-
57	q^*	1.39e+01	1.57e-01	2.97e+00	1.68e+00	5.89e+00
	se	-	-	-	-	-
58	q^*	8.41e+00	1.14e-02	2.69e+00	-1.55e+00	6.00e+00
	se	1.89e+03	1.58e+06	3.09e-01	1.31e+04	6.36e+01
59	q^*	2.18e+00	4.18e-01	2.81e+00	1.10e+00	5.58e+00
	se	-	-	-	-	-

Table 6.6: Summary statistics using clinical patient data to estimate eleven parameters with censored data algorithm. Columns include patient index number, estimates of σ_s (on the log scale), and number of iterations required for the censored data algorithm to converge. Entries with dashes represent cases where data are insufficient to provide a reasonable estimate.

pat #	σ_1	σ_2	num its	pat #	σ_1	σ_2	num its
1	0.08	0.48	180	31	0.09	0.53	46
2	0.08	0.68	12	32	0.16	0.20	20
3	0.19	0.54	20	33	0.17	0.87	45
4	0.11	0.53	14	34	0.11	0.60	32
5	0.11	0.44	12	35	0.11	0.51	188
6	0.13	0.75	29	36	0.08	0.61	251
7	0.18	0.43	181	37	0.18	0.63	61
8	0.09	0.57	225	38	0.11	0.54	12
9	0.11	1.04	12	39	0.11	0.38	19
10	0.05	0.70	16	40	0.22	0.63	20
11	0.10	0.54	90	41	0.00	0.00	11
12	0.08	0.62	17	42	0.09	0.75	224
13	0.10	0.47	22	43	0.12	0.08	47
14	0.08	0.62	25	44	0.12	1.01	35
15	0.07	0.97	24	45	0.00	-	17
16	0.21	0.36	251	46	0.11	0.61	40
17	0.33	1.21	14	47	0.10	0.25	36
18	0.13	0.83	52	48	0.12	0.59	31
19	0.15	0.44	16	49	0.31	0.57	26
20	0.10	0.53	25	50	0.12	0.29	41
21	0.26	0.63	20	51	0.23	1.86	126
22	0.13	0.62	38	52	0.24	0.48	6
23	0.11	0.52	18	53	0.29	1.00	6
24	0.12	0.35	35	54	0.13	0.99	21
25	0.05	0.50	16	55	0.51	0.74	18
26	0.15	0.95	18	56	0.00	-	89
27	0.08	0.55	210	57	-	-	23
28	0.17	0.41	178	58	0.10	0.48	251
29	0.09	0.53	2	59	-	0.67	31
30	0.15	0.27	2				

most cases the predicted T-cell count in equilibrium 1 is not substantially depleted, even in the absence of therapy. The viral loads vary substantially though, and for patients 20, 23, 24, 51, 55, and 59 a higher viral load is predicted. When the second predicted stable infected equilibrium exists, it includes an unrealistically high $CD4^+$ T-cell count and predicts a viral load controlled below the limit of detection. Due to the large number of degrees of freedom in the parameter estimation, it may be necessary to impose more stringent bounds on estimated parameters. Ideally one would have data available on all of the model states and thus obtain more reliable estimates of the parameters. While more work is necessary to tune parameters, these examples demonstrate the kinds of results one could obtain by fitting this model to patient data with the help of the censored data algorithm.

Table 6.7: Calculated model equilibria given each patient's estimated parameters, patients 1–30.

	uninfected	infected 1		infected 2	
pat #	T_1	$T_1 + T_1^*$	V_I	$T_1 + T_1^*$	V_I
1	1009.66	979.90	2121.29		
2	712.99	712.90	50.23		
3	1388.41	1079.39	1994.72		
4	611.91	611.58	20.49		
5	556.88	497.03	556.67	6634.65	19.78
6	866.37	865.36	866.36	1708.62	10.33
7	593.81	170.22	593.16	56063.89	11.83
8	784.77	784.27	510.01		
9	441.26	283.84	441.26	71135.60	0.20
10	777.32	776.93	380.75		
11	626.89	626.89	7.90		
12	394.04	392.56	215.52		
13	964.60	959.93	964.58	2356.46	7.91
14	654.63	474.50	3710.08		
15	880.47	880.47	10.82		
16	320.87	316.73	119.15		
17	329.10	260.82	328.70	7066.09	29.75
18	1324.79	1324.71	18.92		
19	814.75	577.62	814.24	16447.69	23.55
20	400.65	362.31	29323.25		
21	344.59	257.90	344.51	25336.22	16.58
22	394.36	393.73	59.54		
23	619.80	461.42	26140.23		
24	828.80	742.41	36452.56		
25	1324.69	1318.73	21.75		
26	827.14	827.12	37.21		
27	631.35	631.30	722.09		
28	379.52	378.01	22.60		
29	710.53	686.53	3411.80		
30	373.63	373.63	18.26		

Table 6.8: Calculated model equilibria given each patient's estimated parameters, patients 31–59.

	uninfected	infected 1		infected 2	
pat #	T_1	$T_1 + T_1^*$	V_I	$T_1 + T_1^*$	V_I
31	862.96	861.50	16.76		
32	511.87	511.03	23.53		
33	565.96	340.86	13694.61		
34	780.23	777.43	237.54		
35	712.74	653.18	11846.32		
36	721.22	605.91	1890.96		
37	369.66	369.49	13.88		
38	1464.69	1115.43	1464.56	4194.38	1.23
39	556.11	407.74	13601.73		
40	927.33	914.89	2102.68		
41	771.19	636.57	16137.49		
42	315.99	315.99	11.58		
43	615.13	614.61	21.18		
44	411.91	411.87	17.56		
45	963.52	961.50	55.67		
46	772.04	772.04	7.10		
47	811.97	810.81	20.65		
48	428.88	428.62	18.16		
49	1008.60	1008.32	16.95		
50	835.57	834.96	18.69		
51	680.01	152.00	33938.12		
52	537.47	537.46	21.90		
53	245.97	245.97	7.80		
54	1025.32	1024.41	16.79		
55	307.18	235.43	165464.89		
56	708.09	708.05	18.16		
57	1166.22	1165.68	3.72		
58	968.26	967.53	719.61		
59	570.95	404.02	51449.86		

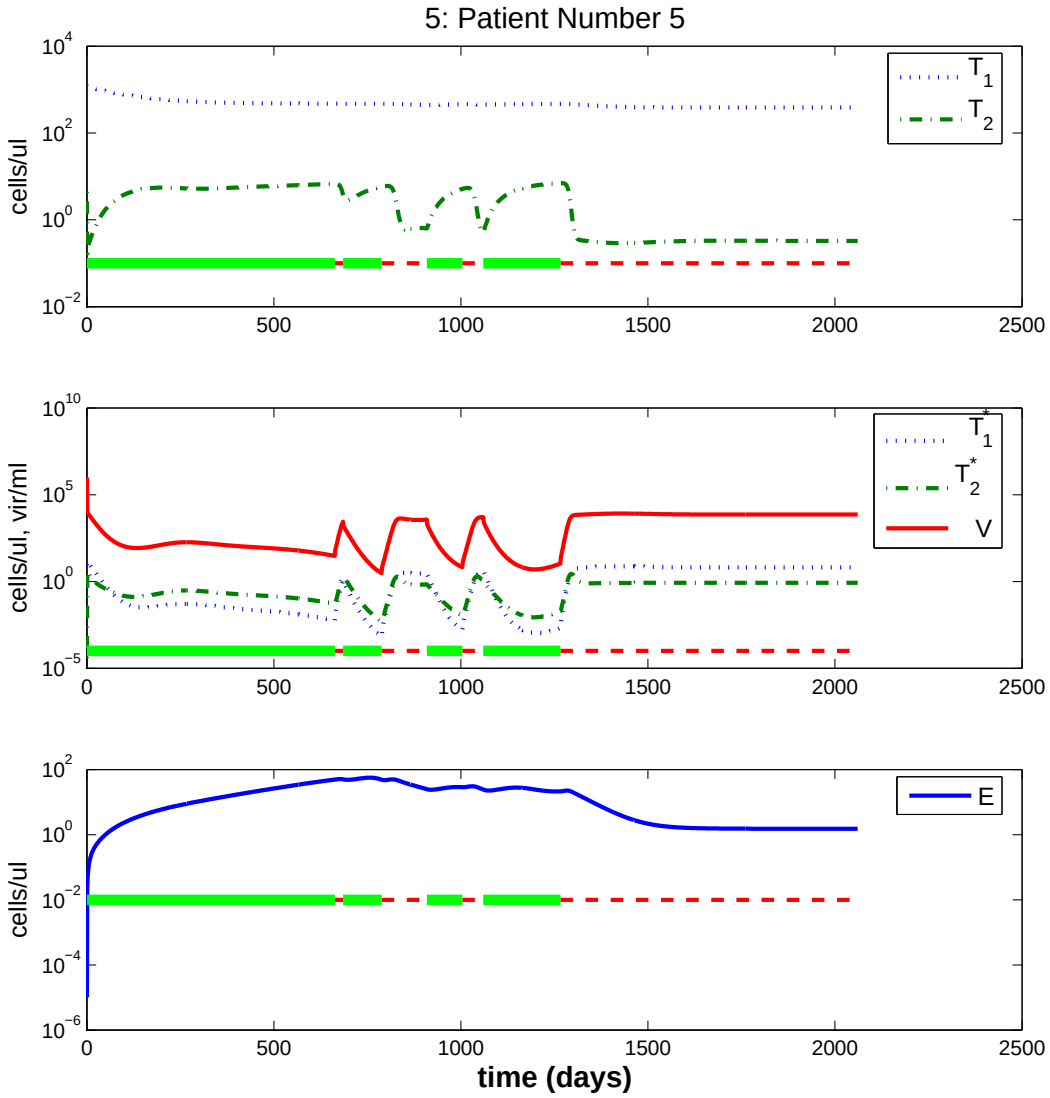


Figure 6.2: All model state dynamics corresponding to estimated parameters and initial conditions for Patient 5.

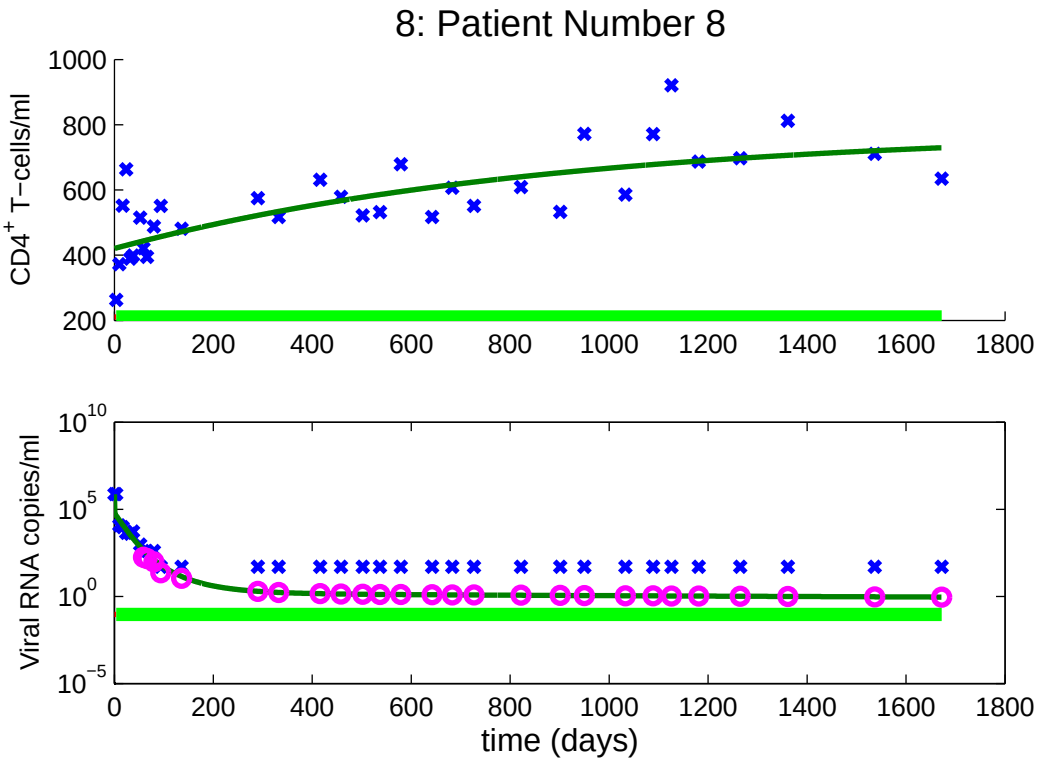


Figure 6.3: Model fit to $CD4^+$ T-cell data (top) and viral load data (bottom) for Patient 8. Crosses denote data and solid line, model evaluated with estimated parameters. Underlying thick solid line indicates constant treatment.

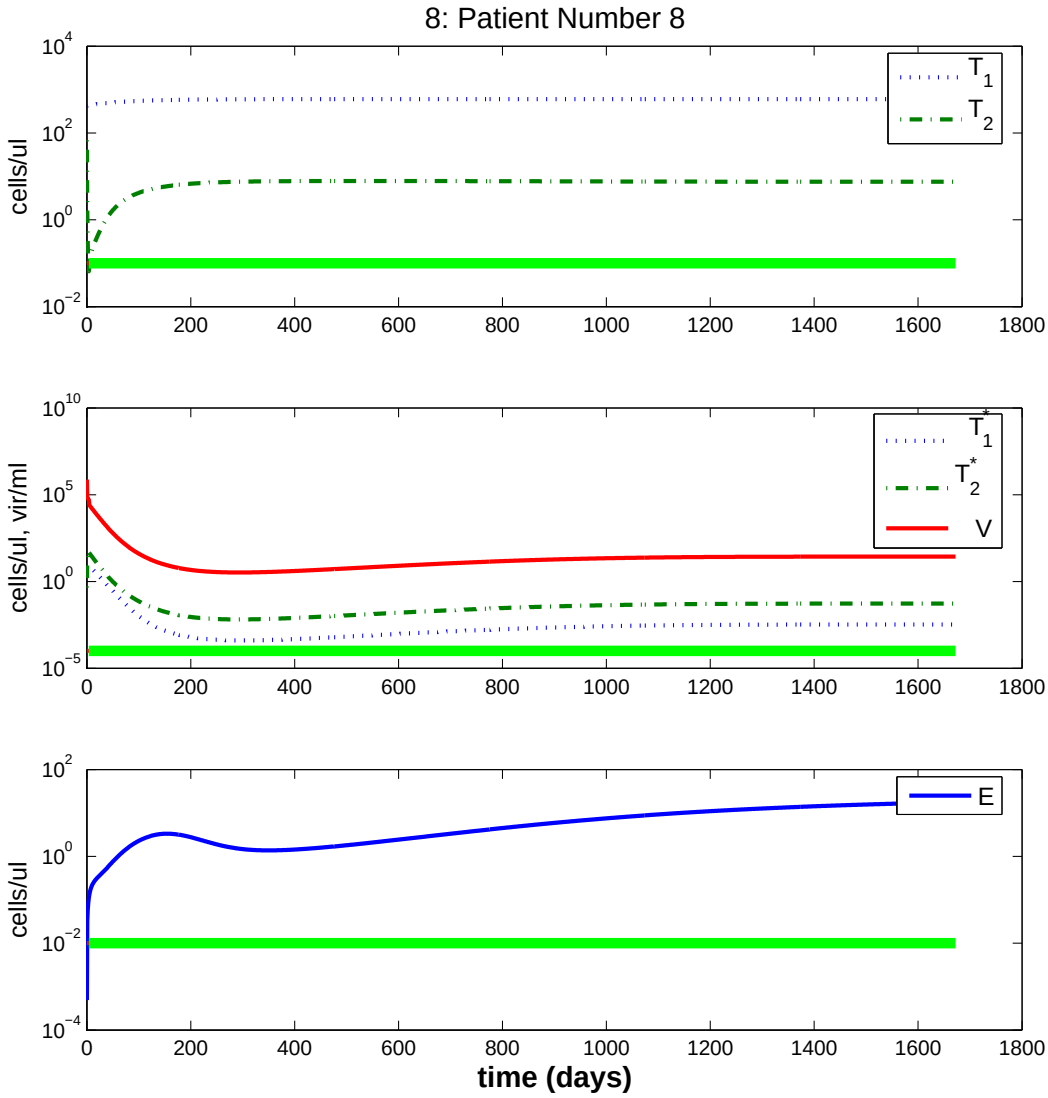


Figure 6.4: All model state dynamics corresponding to estimated parameters and initial conditions for Patient 8.

6.2 Estimation of Distributions

In this section we present results on the distribution of parameters across the population, comparing distributions of parameter estimates from the individual-patient method to direct estimates of distributions across the population with the density estimation method. To employ the latter method, we fix for each patient all but one parameter at estimated values, then assume the remaining parameter varies across the population and attempt to estimate its distribution. Instead of each patient's observed viral RNA data we use the best estimate of censored data points from the previous estimation process.

In Figures 6.5 and 6.6 we have results for each of the 11 estimated parameters. Histograms comprise the individual parameter estimates while solid lines denote the estimated probability densities. In some cases, e.g., parameters $\lambda_1, \epsilon_1, b_E, V_I^0$, there seems to be similarity between the predictions of the single- and multiple-patient methods. In most of the other cases, different distributions are suggested by the two methods. Given the small amount of data ($N_P = 59$) available here, it is not reasonable to conclude which of these estimates provide a better estimate of the true distribution across the population.

Intervals were initially chosen to roughly span the estimates from the individual patient method. Prompted by the number of cases in Figures 6.5 and 6.6 in which the ends of the splines are nonzero (and in some cases the nodal values at the interval boundaries are quite large), we test the effect of several factors on the estimation process. As indicated in the computational methods section, the optimization is robust to choice of initial iterate for the spline coefficients d_k. In addition to the uniform distribution, initial iterates represented schematically in Figure 6.7 (and chosen to integrate to one) all yield convergence to the same optimal parameter values. (This

includes initial iterates like that in the rightmost subplot with the mode at various locations in the interval.)

It is possible that the inverse problem is hindered by a lack of data and the requirement that the estimated function integrate to one is causing the minimum of the quadratic cost function to occur when mass is placed arbitrarily near the boundaries (perhaps where sensitivity is less). However, removing the requirement that the density integrate to one only slightly affects the estimates. In fact, they still integrate nearly to one (0.96–0.99) in all cases considered. Requiring that the ends of the splines be exactly zero also does not affect this arbitrary mass placement; it simply results in a bump in the spline estimate at the second node in from the end of the interval.

Finally, we experiment with moving the boundaries, since when working with experimental data it is not known where the interval of consideration should lie. Moving them by 10% yields similar results. The densities estimated in the interior of the interval remain similar, with boundary effects as discussed previously.

In Figures 6.8 and 6.9 we present the same parameter distribution results, but include nodal confidence intervals on the estimated probability density functions. The confidence intervals were calculated using the *ad hoc* estimates of $\hat{\sigma}_s$ from (4.14), and show considerable variability across parameters considered. With parameters N_T and b_E we have more confidence in the shape of the distribution, but in most cases cannot reasonably conclude the exact form. We see the potential effects of over-regularization in the plot for T_1^0, where the regularization has smoothed the estimate, but also likely biased it. However, since the inverse problem is better posed with regularization, the confidence intervals are tighter for this distribution.

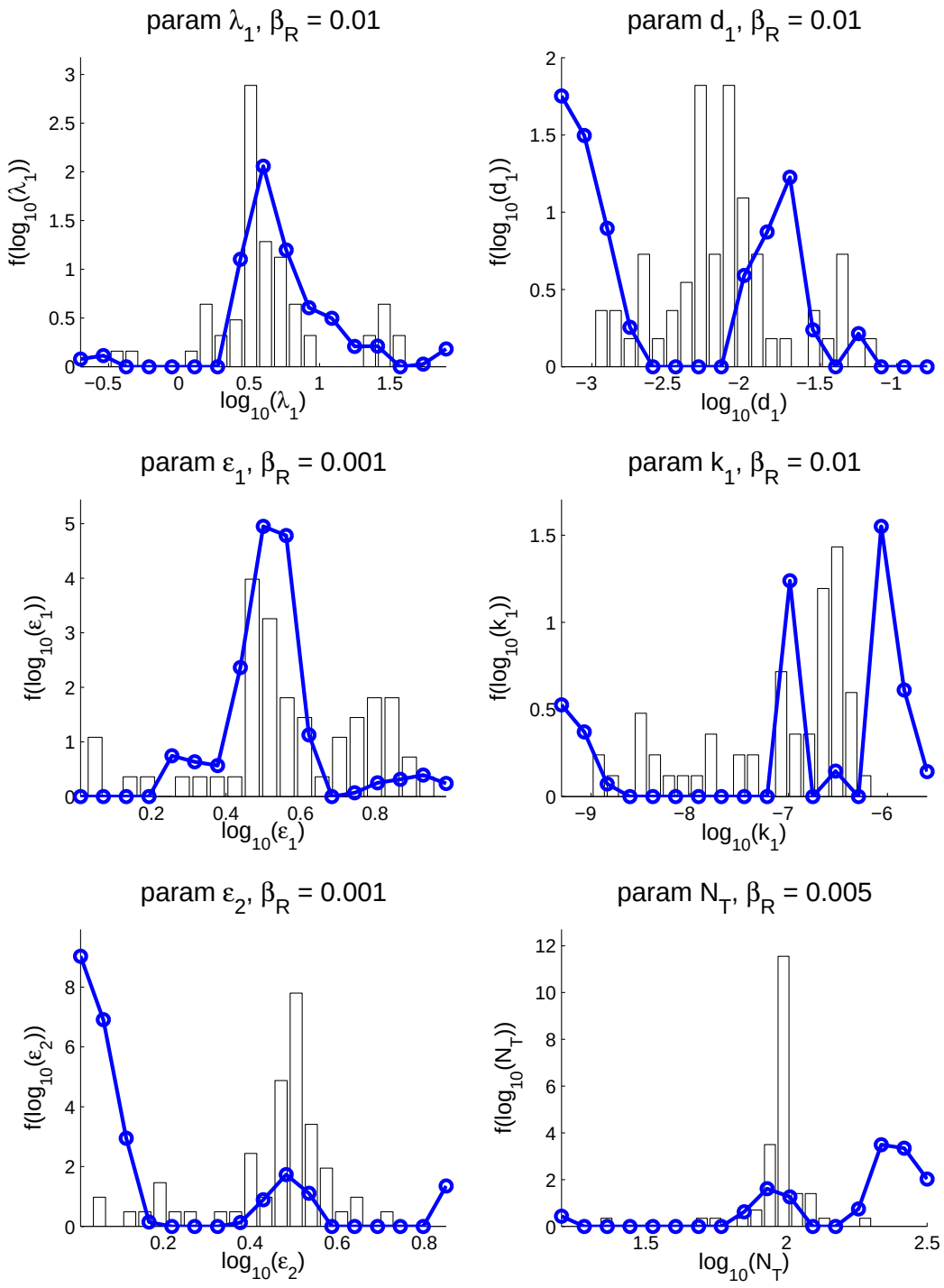


Figure 6.5: Histograms showing distribution of estimated values of various parameters across a population of patients from clinical data set PS59. Solid blue lines denote piecewise linear density estimates obtained from the multiple-patient inverse problem method.

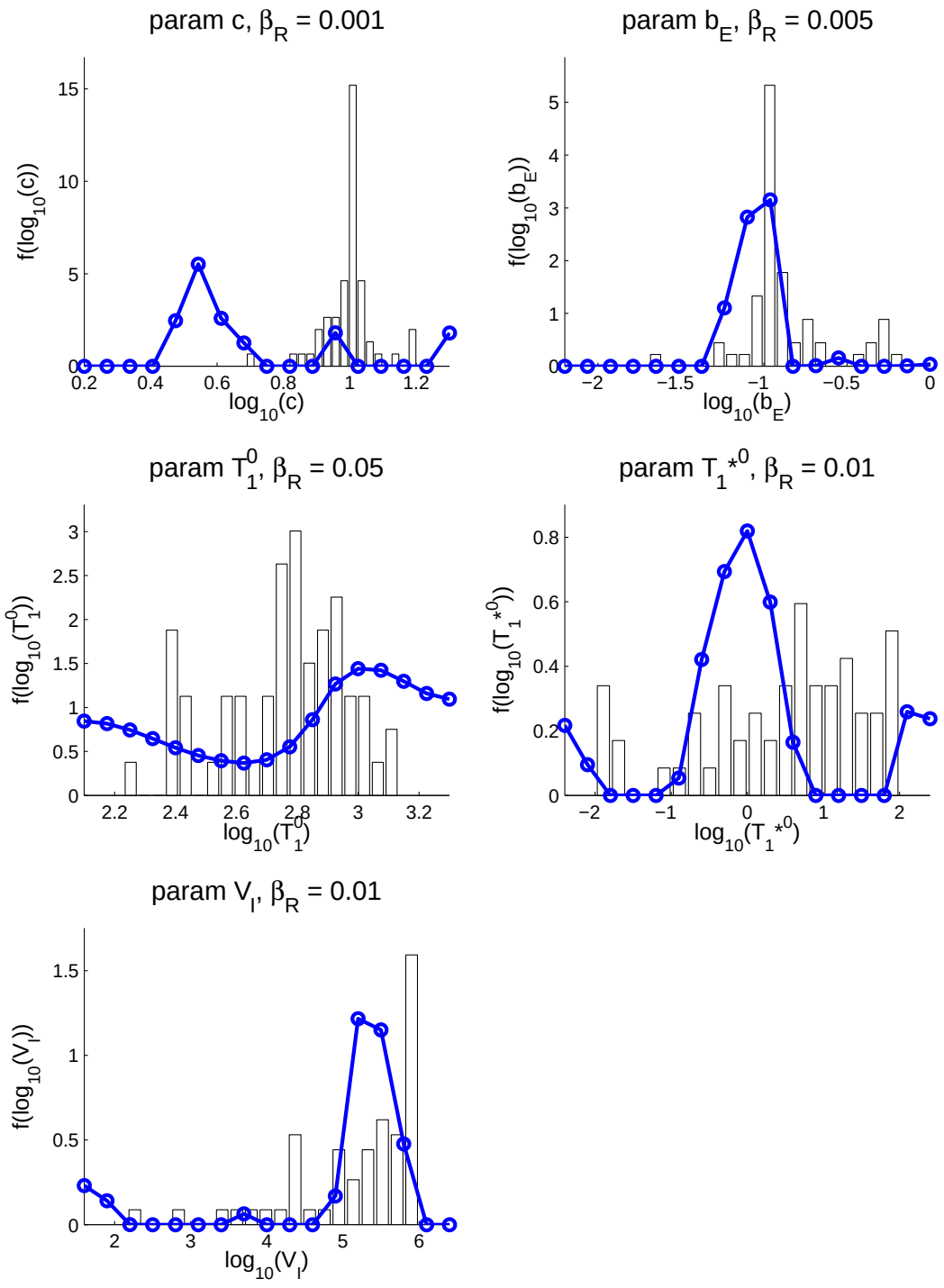


Figure 6.6: Histograms showing distribution of estimated values of various parameters across a population of patients from clinical data set PS59. Solid blue lines denote piecewise linear density estimates obtained from the multiple-patient inverse problem method.

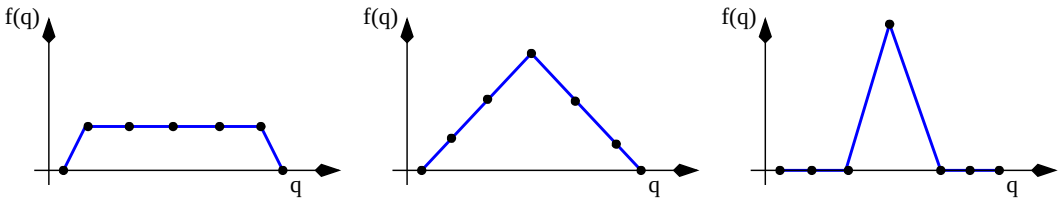


Figure 6.7: Schematics of initial iterates for density functions.

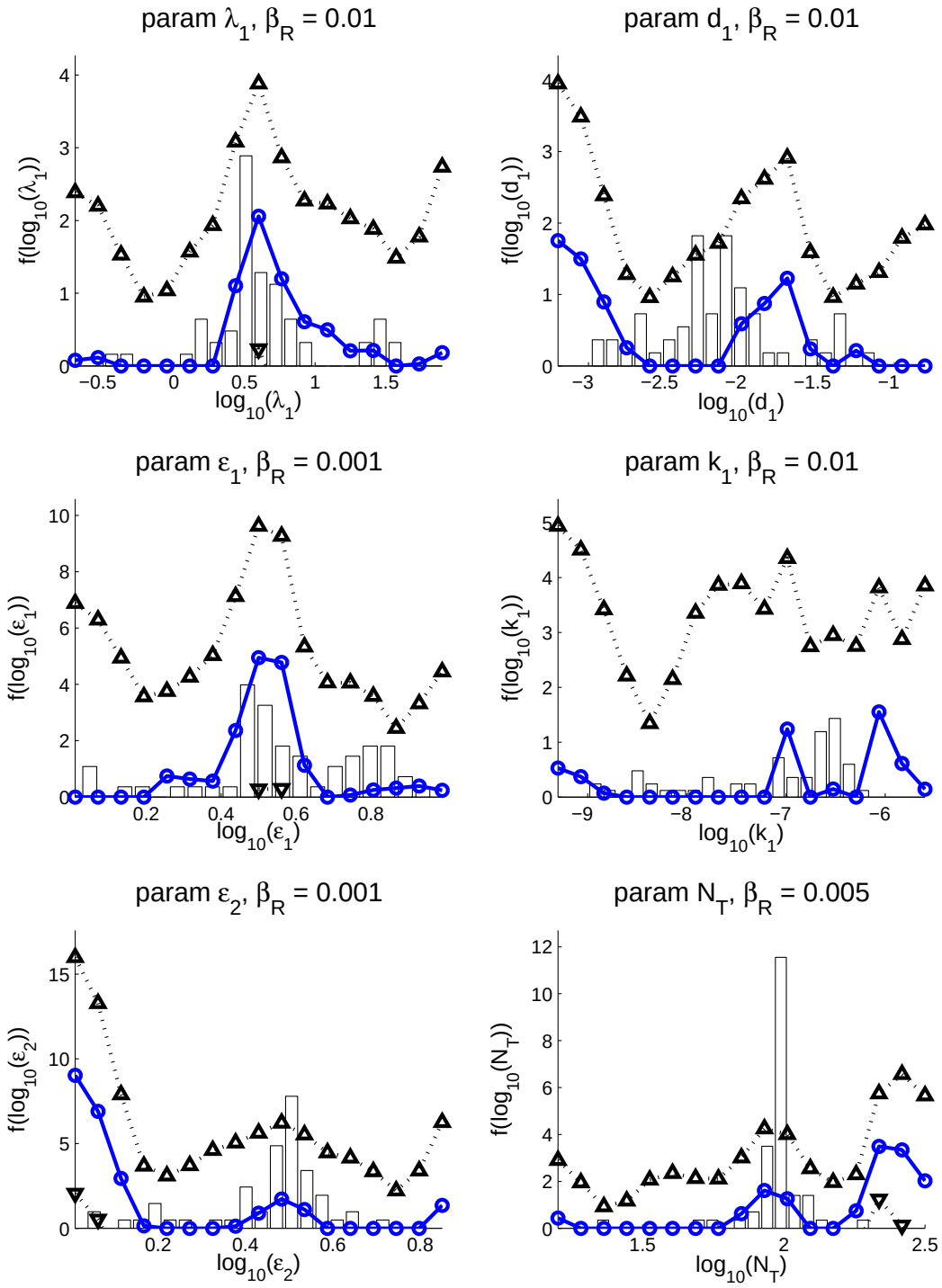


Figure 6.8: Histograms showing distribution of estimated values of various parameters across a population of patients from clinical data set PS59. Solid blue lines denote piecewise linear density estimates obtained from the multiple-patient inverse problem method.

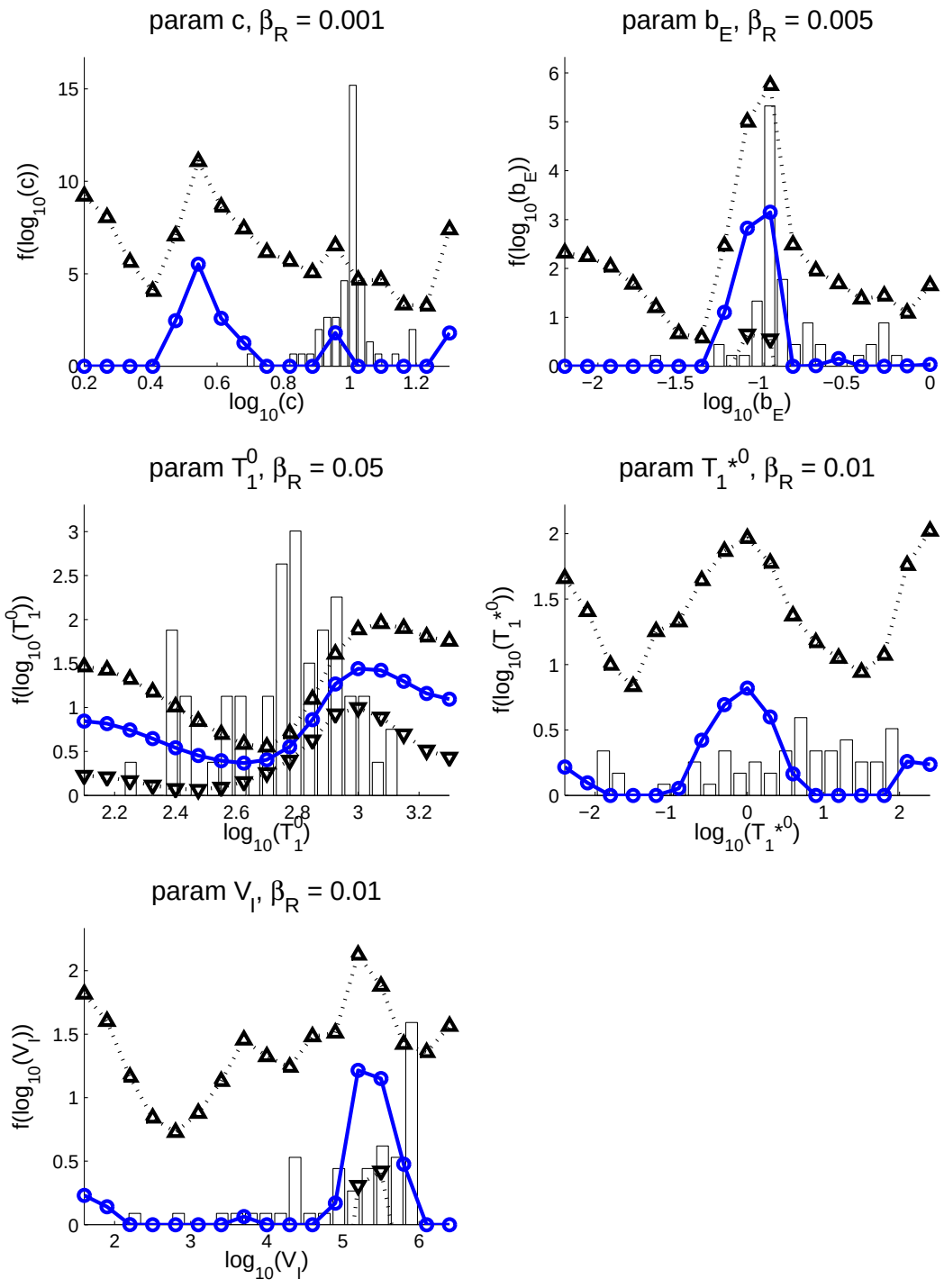


Figure 6.9: Histograms showing distribution of estimated values of various parameters across a population of patients from clinical data set PS59. Solid blue lines denote piecewise linear density estimates obtained from the multiple-patient inverse problem method.

Chapter 7

Conclusions and Future Directions

We developed a combined mathematical and statistical modeling approach for analyzing clinical data from an HIV acute infection study. Using an example system of differential equations to model in-host HIV infection dynamics we examined a method to estimate distributions of model parameters across a patient population, without prior presumption of the distributional form. After summarizing theory for the well-posedness of such inverse problems, we verified the method on simulated data. In the context of this model, we implemented the EM Algorithm for maximum likelihood estimation to handle patient measurements censored by assay resolution limits and again tested with simulated data. For each of these processes we developed and applied methods for estimating variability of the resulting parameter estimates by using sensitivity analysis to calculate confidence intervals. Finally we presented sample results for the application of these methods to actual clinical data.

Experiments with simulated data, generated to reflect clinical reality, reveal that the density estimation methods considered are capable of identifying underlying distributions of parameters from noisy observations of the dynamic model's behavior. Even with a modest number of virtual patients the methods yield insight into the

form of the distribution. In certain cases, these methods offer computational benefits over traditional individual-based inverse problem approaches, since one directly estimates a population level distribution without the need for fitting each individual and then aggregating and analyzing the results.

In addition to enabling fitting observed patient data, the censored data method produces important estimates of patient behavior when measurements are below the limit of detection. Most HIV patients who successfully suppress virus while on drug therapy immediately have a viral rebound when on drug holiday and the model can predict this. The censored data method allows us to calibrate the dynamic model to gain insight into the possible dynamics of the virus during periods of suppression. This is one way to explain patient behavior in this extremely data-limited setting. It is possible that due to the small quantities of virus believed present during suppression, models including stochastic effects may be more representative. Regardless of the modeling approach chosen, more data is needed to soundly validate any theory as to the true dynamics in this regime.

As implemented, the censored data method produces high estimates of variability in patient data, especially for viral load measurements. Since we observe this for simulated data as well, further testing may be warranted to validate the method.

We demonstrated the construction of piecewise linear confidence bands (based on nodal standard errors) on the probability density function estimates. While computationally one might be satisfied with the nodal confidence intervals presented, as they closely relate to the estimated quantities (spline coefficients), a more sophisticated theory would involve true $L^2(Q)$ functional confidence intervals. Such functional confidence intervals would be constructed to contain the true density function with confidence (almost) everywhere, rather than just at the nodes. (See [28] and [52] for some examples of methods developed to construct functional confidence intervals

of a similar sort in the context of smoothing splines.) A future goal is to develop functional analytic theory for the abstract case where the inverse problem operator is considered as a map from a smooth $L^2(Q)$ density function to the data and try to develop confidence intervals formulations based on its pseudoinverse map. Then one could verify if the sort of piecewise linear band approximations considered here converge to the smooth confidence intervals.

While we estimated parameters from patient data, it is not clear that the viral and T-cell dynamics suggested by the calibrated model are realistic in all cases. Further work is necessary to refine the parameter estimates obtained. Ideally, this would be done in the context of richer data. Some immune response data are available, but they typically summarize responses to specific HIV epitopes, so work must be done to correlate them to the E immune response compartment in the model. Even more useful would be the development of more detailed models of the mechanisms of HIV-specific immune responses, including modeling of activation and differentiation and the role of helper cells in the establishment of memory pools of immune responders. The model also includes a secondary cell population for which we have no observational data. Since in the model these secondary compartments are responsible for the persistent low level virus property, acquiring additional data to validate them should be a high priority.

References

[1] B.M. Adams, H.T. Banks, J.E. Banks, and J.D. Stark. Population dynamics models in plant–insect herbivore–pesticide interactions. Technical Report CRSC-TR03-12, NC State Univ., March 2003. *Math. Biosci.*, to appear.

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

[3] S. Armstrong, C. Fontaine, and A. Wilson. *2004 Report on the Global AIDS Epidemic*. UNAIDS/Joint United Nations Programme on HIV/AIDS, Geneva, Switzerland, 2004. http://www.unaids.org.

[4] H.T. Banks and K.L. Bihari. Modeling and estimating uncertainty in parameter estimation. *Inverse Problems*, 17:95–111, 2001.

[5] H.T. Banks and V.A. Bokil. Parameter identification for dispersive dielectrics using pulsed microwave interrogating signals and acoustic wave induced reflections in two and three dimensions. Technical Report CRSC-TR04-27, NC State Univ., July 2004.

[6] H.T. Banks and D.M. Bortz. A parameter sensitivity methodology in the context of HIV delay equation models. *J. Math. Biol.*, 2005. to appear.

[7] H.T. Banks, D.M. Bortz, and S.E. Holte. Incorporation of uncertainty in mathematical modeling of HIV infection dynamics. *Math. Biosci.*, 183:63–91, 2003.

[8] H.T. Banks, D.M. Bortz, G.A. Pinter, and L.K. Potter. Modeling and imaging techniques with potential for application in bioterrorism. Technical Report CRSC-TR03-02, NC State Univ., January 2003. Chapter 6 in *Bioterrorism: Mathematical Modeling Applications in Homeland Security,* (H.T. Banks and C. Castillo-Chavez, eds.), Frontiers in Applied Math, **FR28**, SIAM, Philadelphia, PA, 2003, pp. 129–154.

[9] H.T. Banks, S.L. Grove, S. Hu, and Y. Ma. A hierarchical bayesian approach for parameter estimation in HIV models. Technical Report CRSC-TR05-19, NC State Univ., April 2005.

[10] H.T. Banks and K. Kunisch. *Estimation Techniques for Distributed Parameter Systems*. Birkhauser, Boston, 1989.

[11] H.T. Banks, H.D. Kwon, J.A. Toivanen, and H.T. Tran. An SDRE based estimator approach for HIV feedback control. Technical Report CRSC-TR05-20, NC State Univ., April 2005.

[12] H.T. Banks, Y. Ma, and L.K. Potter. A simulation-based comparison between parametric and nonparametric estimation methods in PBPK models. Technical Report CRSC-TR04-25, NC State Univ., June 2004.

[13] H.T. Banks and G.A. Pinter. A probabilistic multiscale approach to hysteresis in shear wave propagation in biotissue. *Multiscale Modeling and Simulation*, 3(2):395–412, 2005.

[14] P. Billingsley. *Convergence of Probability Measures*. Wiley, New York, 1968.

[15] S. Bonhoeffer, J.M. Coffin, and M.A. Nowak. Human immunodeficiency virus drug therapy and virus load. *J. Virol.*, 97:3275–3278, 1997.

[16] D.M. Bortz. *Modeling, analysis, and estimation of an* in vitro *HIV infection using functional differential equations*. PhD thesis, NC State Univ., 2002.

[17] D. S. Callaway and A. S. Perelson. HIV-1 infection and low steady state viral loads. *Bull. Math. Biol.*, 64(1):29–64, January 2002.

[18] R.J. Carroll and D. Ruppert. *Transformation and Weighting in Regression*. Chapman and Hall, New York, 1988.

[19] T.F. Coleman and Y. Li. An interior, trust region approach for nonlinear minimization subject to bounds. *SIAM Journal on Optimization*, 6:418–445, 1996.

[20] M. Davidian and D.M. Giltinan. *Nonlinear Models for Repeated Measurement Data*. Chapman and Hall, London, 1995.

[21] A.P. Dempster, N.M. Laird, and D.B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Stat. Soc., Ser. B*, 39(1):1–38, 1977.

[22] N. Dunford and J.T. Schwartz. *Linear Operators*. Interscience Publishers, New York, 1958.

[23] M. Eslami. *Theory of Sensitivity in Dynamic Systems*. Springer-Verlag, Heidelberg, 1994.

[24] D.E. Finkel. *Global optimization with the DIRECT algorithm*. PhD thesis, NC State Univ., 2005. http://www4.ncsu.edu/~definkel/research/Direct.m.

[25] P.M. Frank. *Introduction to System Sensitivity Theory*. Academic Press, New York, 1978.

[26] H.T. Banks N.L. Gibson and W.P. Winfree. Electromagnetic crack detection inverse problems using terahertz interrogating signals. Technical Report CRSC-TR03-40, NC State Univ., September 2003.

[27] P.E. Gill, W. Murray, and M.H. Wright. *Practical Optimization*. Academic Press, London, UK, 1981.

[28] C. Gu and G. Wahba. Smoothing spline (ANOVA) with component-wise Bayesian "confidence intervals". *J. Computational and Graphical Statistics*, 2:97–117, 1993.

[29] A.C. Hindmarsh. *Scientific Computing*, chapter ODEPACK, A Systematized Collection of ODE Solvers, pages 55–64. North-Holland, 1983. http://www.llnl.gov/CASC/odepack/.

[30] D.D. Ho, A.U. Neumann, A.S. Perelson, et al. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature*, 373:123–126, 1995.

[31] P.R. Johnston and R.M. Gulrajani. An analysis of the zero-crossing method for choosing regularization parameters. *SIAM J. Sci. Comput.*, 24(2):428–442, 2002.

[32] S. Kassutto, K. Maghsoudi, M. Johnston, et al. Longitudinal analysis of clinical markers following antiretroviral therapy initiated during acute or early HIV-1 infection. 2005. Submitted.

[33] C.T. Kelley. *Iterative Methods for Optimization*. SIAM, Philadelphia, 1999. Frontiers in Applied Mathematics FR18.

[34] T.B. Kepler and A.S. Perelson. Drug concentration heterogeneity facilitates the evolution of drug resistance. *Proc. Natl. Acad. Sci. USA*, 95:11514–11519, 1998.

[35] F. Lori and J. Lisziewicz. Structured treatment interruptions for the management of HIV infection. *J. American. Med. Assoc.*, 4286(23):2981–2987, 2001.

[36] K. Luzuriaga et al. Dynamics of human immunodeficiency virus type 1 replication in vertically infected infants. *J. Virol.*, 73:362–367, 1999.

[37] G.J. McLachlan and T. Krishnan. *The EM algorithm and extensions*. John Wiley, New York, 1997.

[38] A.J. Melvin et al. HIV-1 dynamics in children. *AIDS*, 20:468–473, 1999.

[39] J.E. Mittler, B. Sulzer, A.U. Neumann, and A.S. Perelson. Influence of delayed viral production on viral dymanics in HIV-1 infected patients. *Mathematical Biosciences*, 152:143–163, 1998.

[40] A. Neumaier. Solving ill-conditioned and singular linear systems: A tutorial on regularization. *SIAM Rev.*, 40(3):636–666, September 1998.

[41] A.U. Neumann, R. Tubiana, V. Calvez, C. Robert, T.S. Li, H. Agut, and B. Autran. HIV-1 rebound during interruption of highly active anti-retroviral therapy has no deleterious effect on re-initiated treatment. *AIDS*, 13:677–683, 1999.

[42] D.W. Notermans, J. Goudsmit, and S.A. Danner. Rate of HIV-1 decline following anti-retroviral therapy is related to viral load at baseline and drug regimen. *AIDS*, 12:1483–1490, 1998.

[43] M.A. Nowak and C.R.M. Bangham. Population dynamics of immune responses to persistent viruses. *Science*, 272:74–79, April 1996.

[44] A.S. Perelson, A.U. Essunger, P.Y. Cao, et al. Decay characteristics of HIV-1 infected compartments during combination therapy. *Nature*, 387:188–191, 1997.

[45] A.S. Perelson and P.W. Nelson. Mathematical analysis of HIV-1 dynamics in vivo. *SIAM Review*, 41(1):3–44, 1999.

[46] A.S. Perelson, A.U. Neumann, M. Markowitz, et al. HIV-1 dynamics *in vivo*: virion clearance rate, infected cell life-span, and viral generation time. *Science*, 271:1582–1586, 1996.

[47] C. Robinson. *Dynamical Systems: Stability, Symbolic Dynamics, and Chaos*. CRC Press, Boca Raton, FL, 1995.

[48] E.S. Rosenberg, M. Altfield, S.H. Poon, et al. Immune control of HIV-1 after early treatment of acute infection. *Nature*, 407:523–526, September 2000.

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

[50] L.F. Shampine and M.W. Reichelt. The MATLAB ODE suite. *SIAM Journal on Scientific Computing*, 18:1–22, 1997. http://www.mathworks.com.

[51] A.A. Tsiatis, V. DeGruttola, and M.S. Wulfsohn. Modeling the relationship of survival to longitudinal data measured with error: Applications to survival and CD4 counts in patients with AIDS. *J. American Stat. Assn.*, 90:27–37, 1995.

[52] G. Wahba. Bayesian "confidence intervals" for the cross-validated smoothing spline. *J.R. Statist. Soc. B*, 45(1):133–150, 1983.

[53] X. Wei, S.K. Ghosh, M.E. Taylor, et al. Viral dynamics of HIV-1 infection. *Nature*, 373:117–122, 1995.

[54] D. Wodarz and M.A. Nowak. Specific therapy regimes could lead to long-term immunological control of HIV. *Proceedings of the National Academy of Sciences*, 96(25):14464–14469, December 1999.

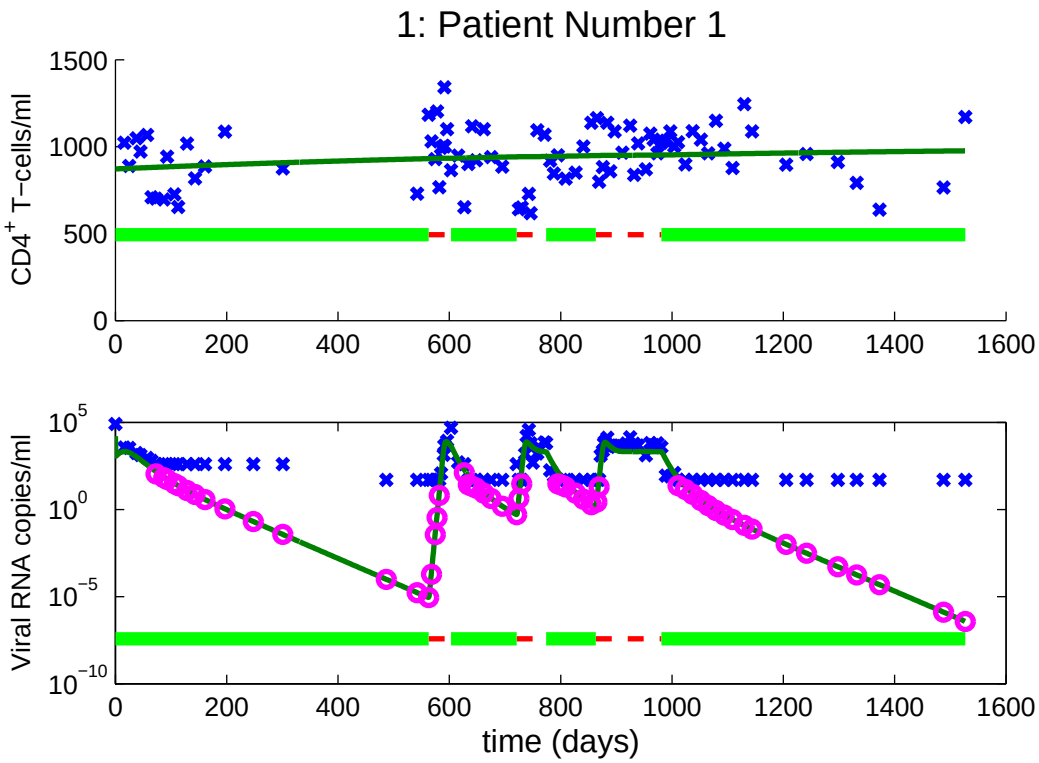
[55] H. Wu and A.A. Ding. Population HIV-1 dynamics *in vivo*: Applicable models and inferential tools for virological data from AIDS clinical trials. *Biometrics*, 55:410–418, 1999.

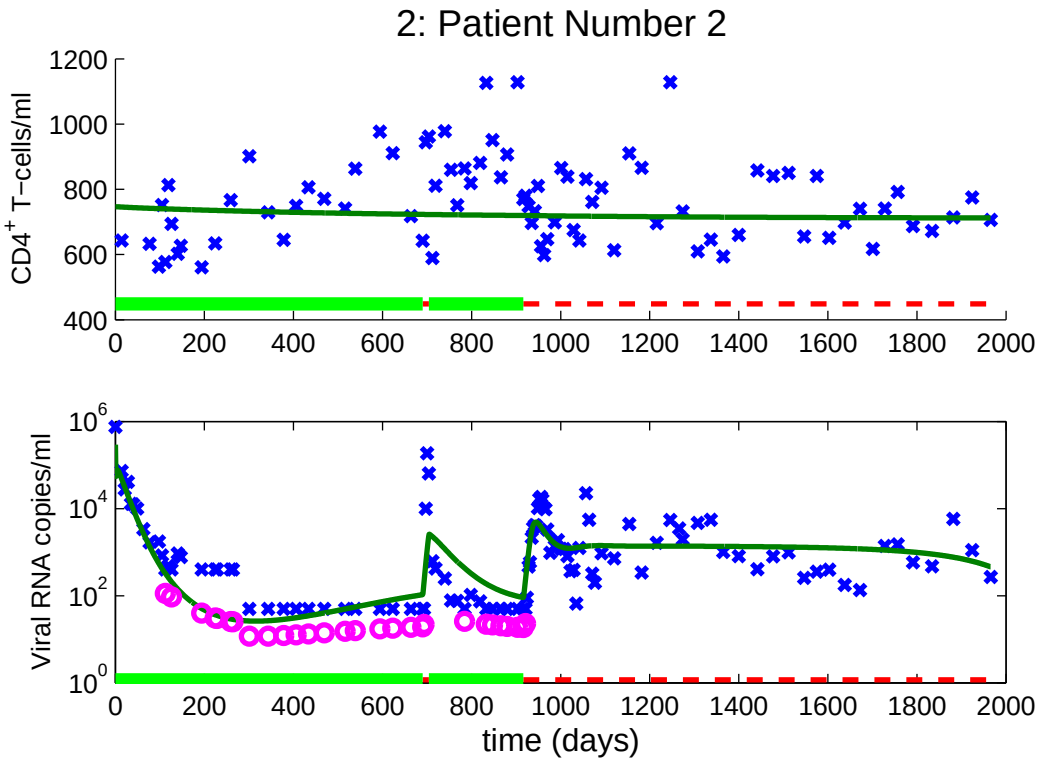
[56] H. Wu et al. Characterization of viral dynamics in human immunodeficiency virus type 1-infected patients treated with combination anti-retroviral therapy: Relationships to host factors, cellular restoration, and virologic end points. *J. Inf. Dis.*, 179:799–807, 1999.

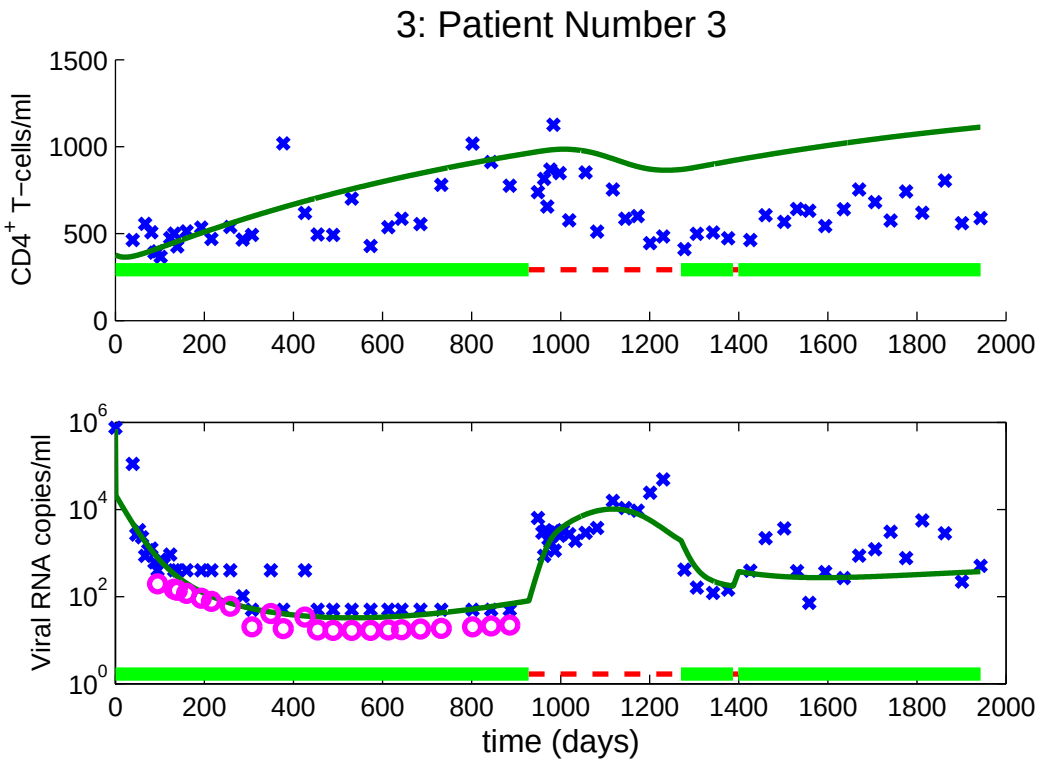
Appendix A

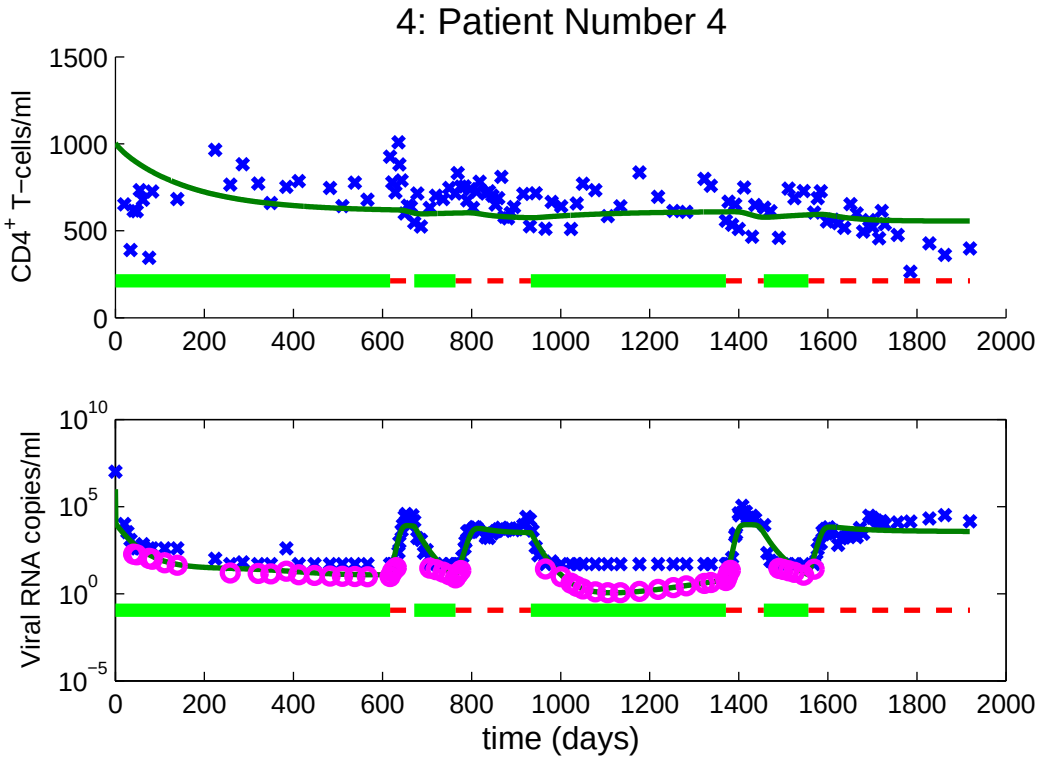
Model Fits to Clinical Data

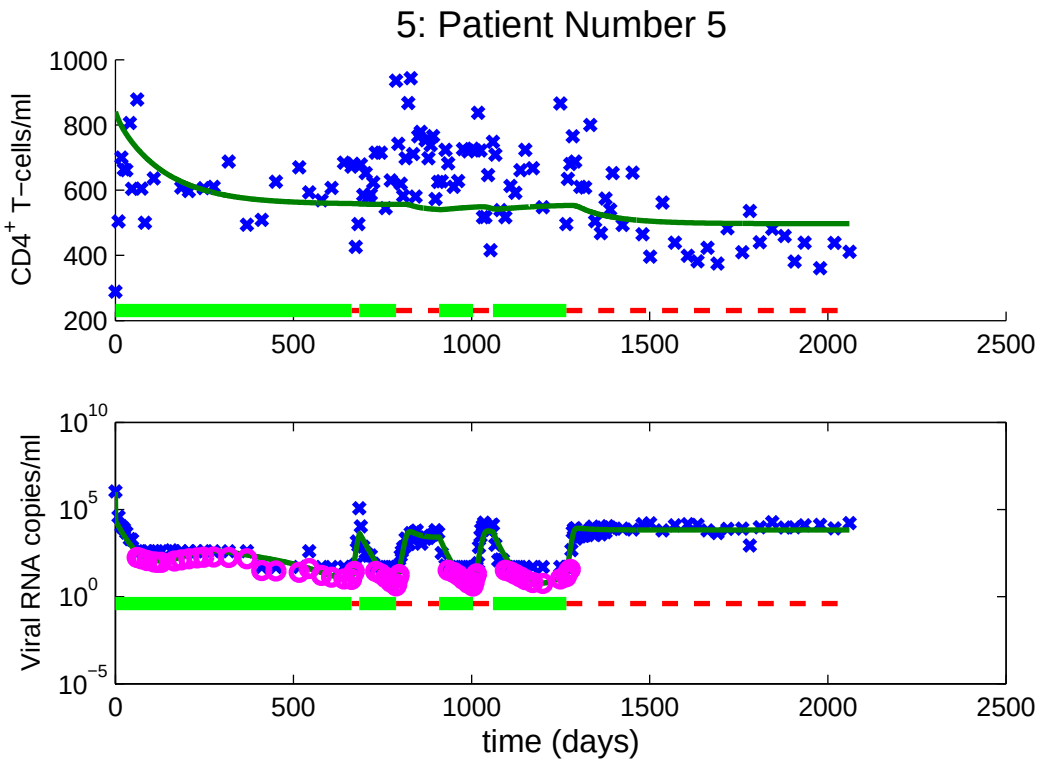
The following plots contain the model fits to data obtained for each of the 59 patients in PS59. For each patient, the top graph shows $CD4^+$ T-cell data and the bottom, viral load data. Model solutions evaluated at optimal parameters are plotted with solid lines, blue crosses denote observed data, and magenta circles, the estimated values for censored data points. On-treatment periods are indicated with a thick solid green line and off periods by a thin red dashed line under the data. We achieve reasonable fits to patient data in most cases.

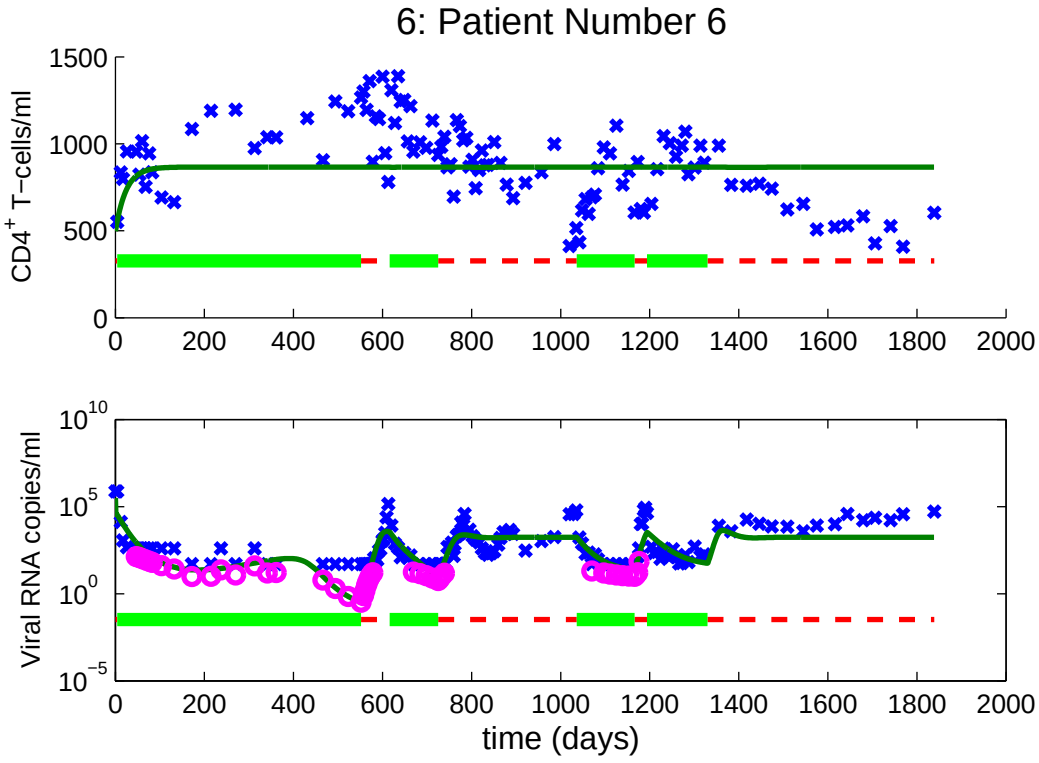

1: Patient Number 1
CD4+ T−cells/ml
Viral RNA copies/ml
time (days)

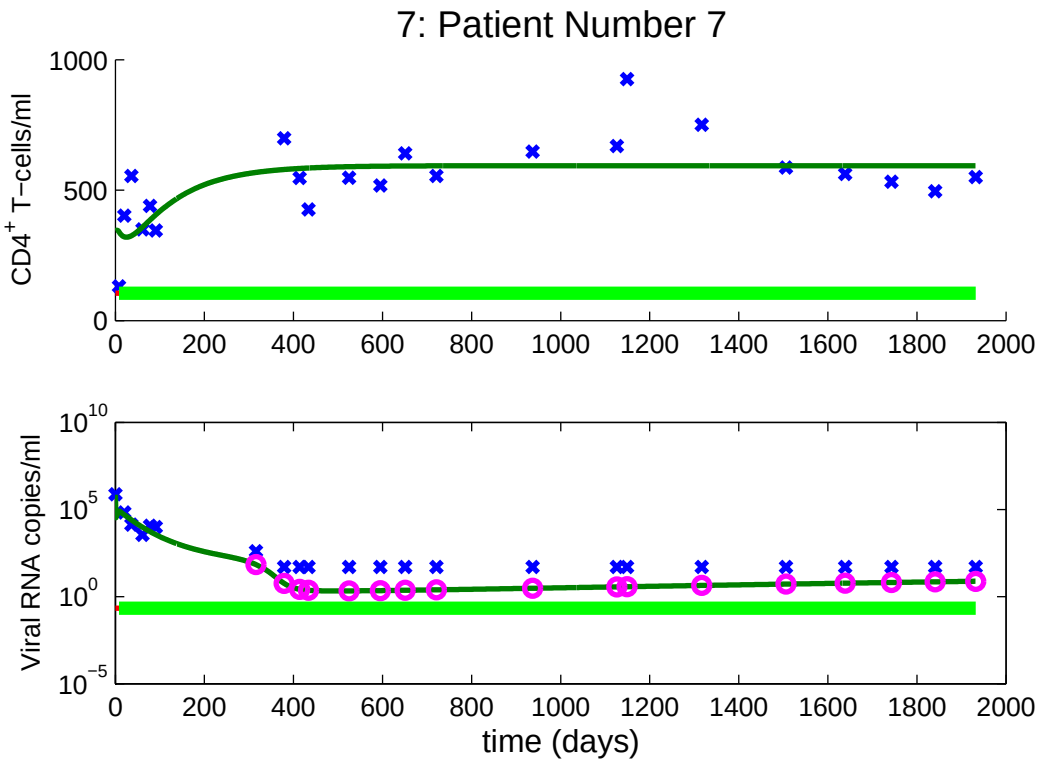

2: Patient Number 2
CD4+ T−cells/ml
Viral RNA copies/ml
time (days)

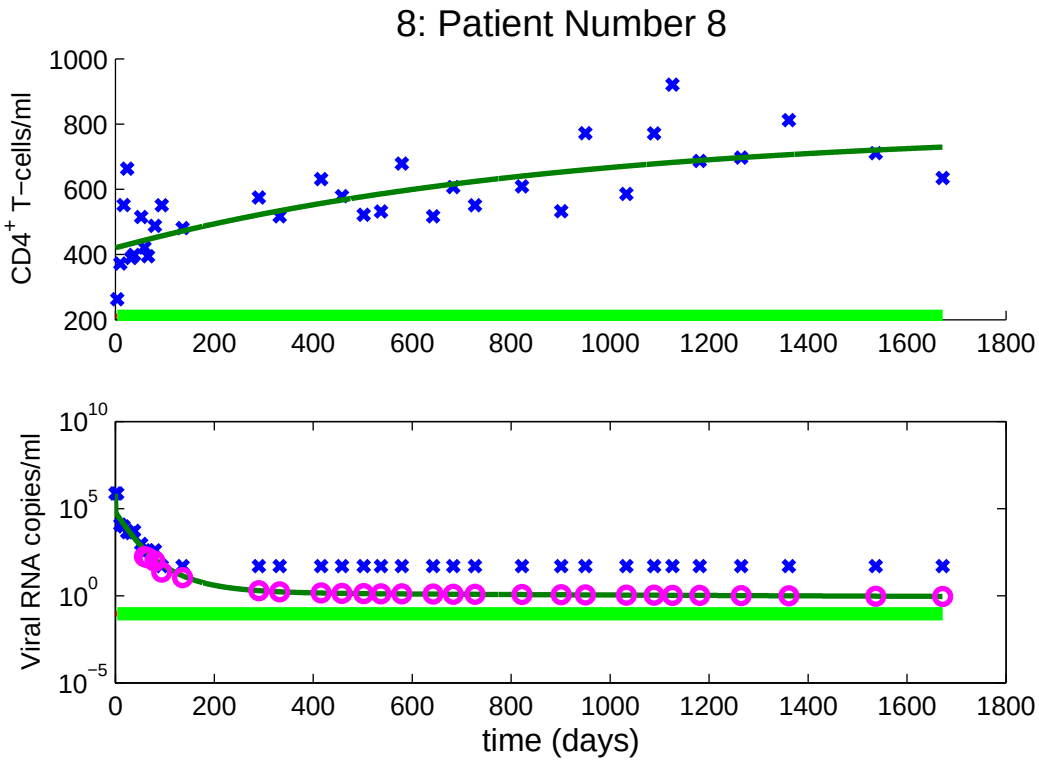

3: Patient Number 3
CD4+ T−cells/ml
Viral RNA copies/ml
time (days)

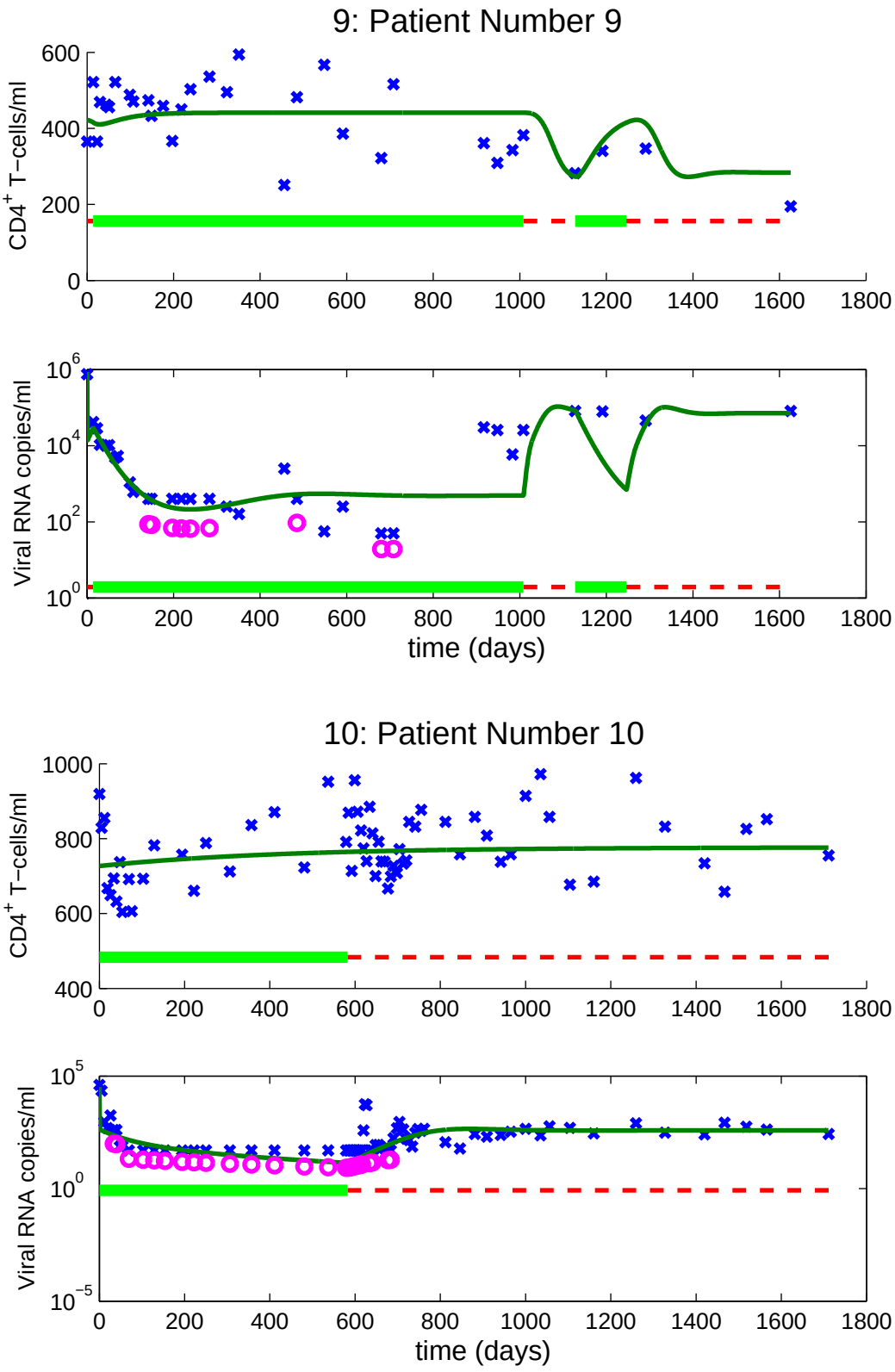

4: Patient Number 4
CD4+ T−cells/ml
Viral RNA copies/ml
time (days)

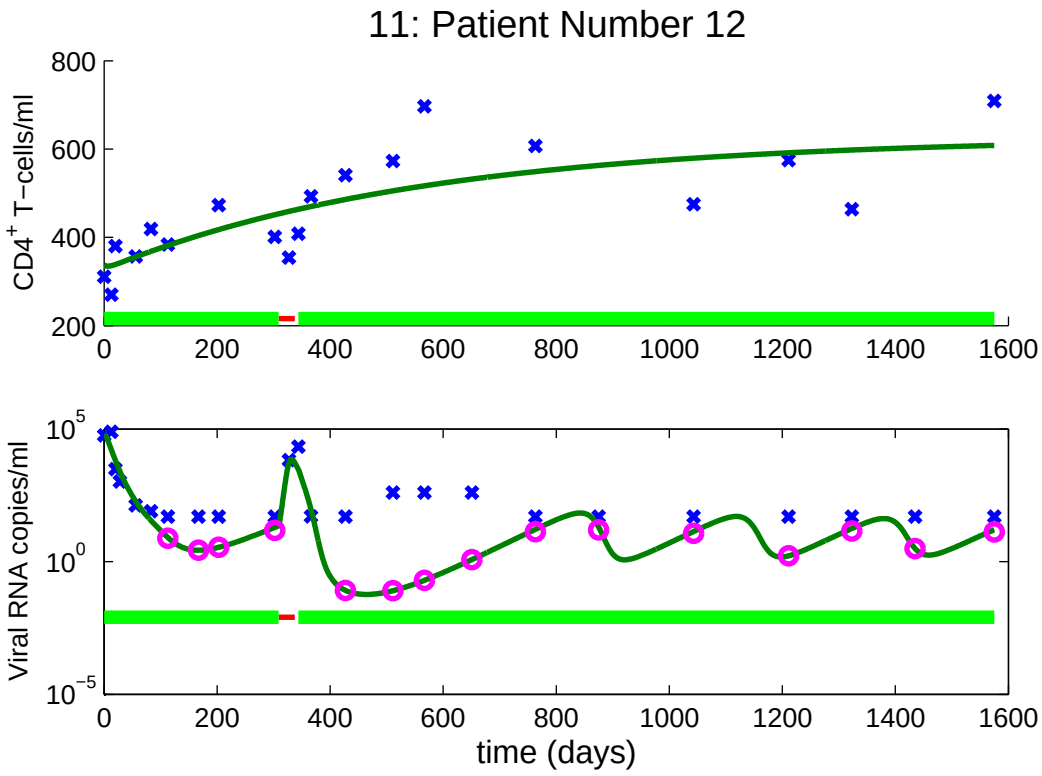

5: Patient Number 5
CD4+ T−cells/ml
1000
800
600
400
200
0
500
1000
1500
2000
2500
Viral RNA copies/ml
10^10
10^5
10^0
10^−5
time (days)

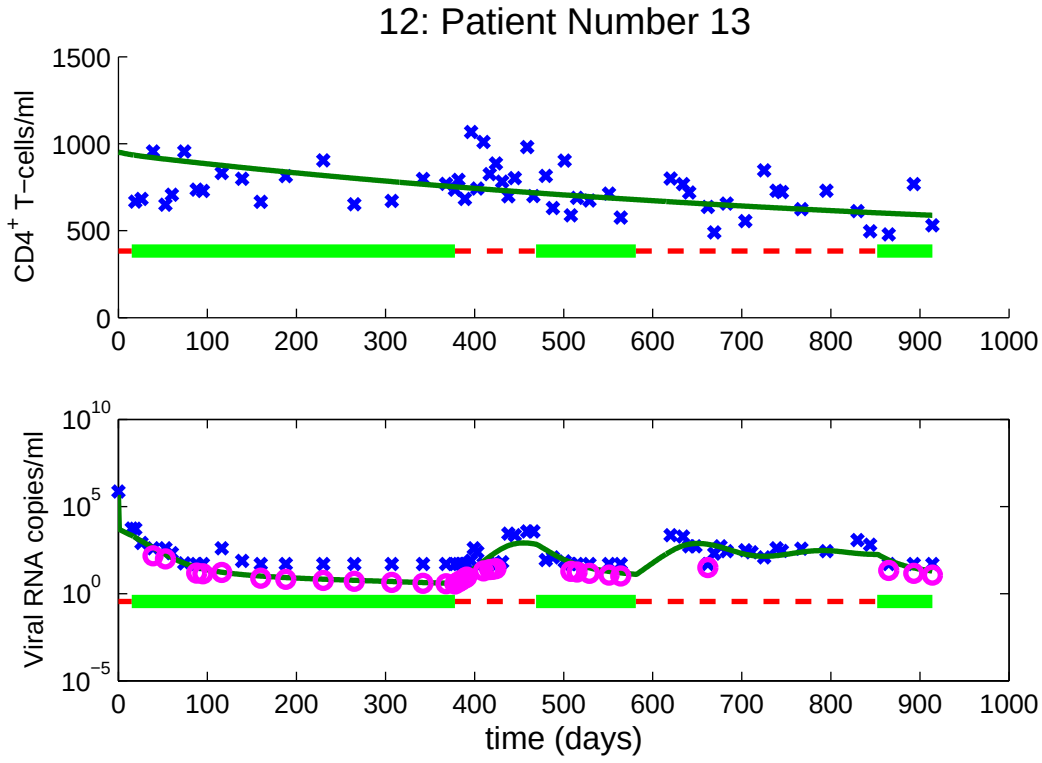

6: Patient Number 6
CD4+ T−cells/ml
1500
1000
500
0
200
400
600
800
1000
1200
1400
1600
1800
2000
Viral RNA copies/ml
10^10
10^5
10^0
10^−5
time (days)

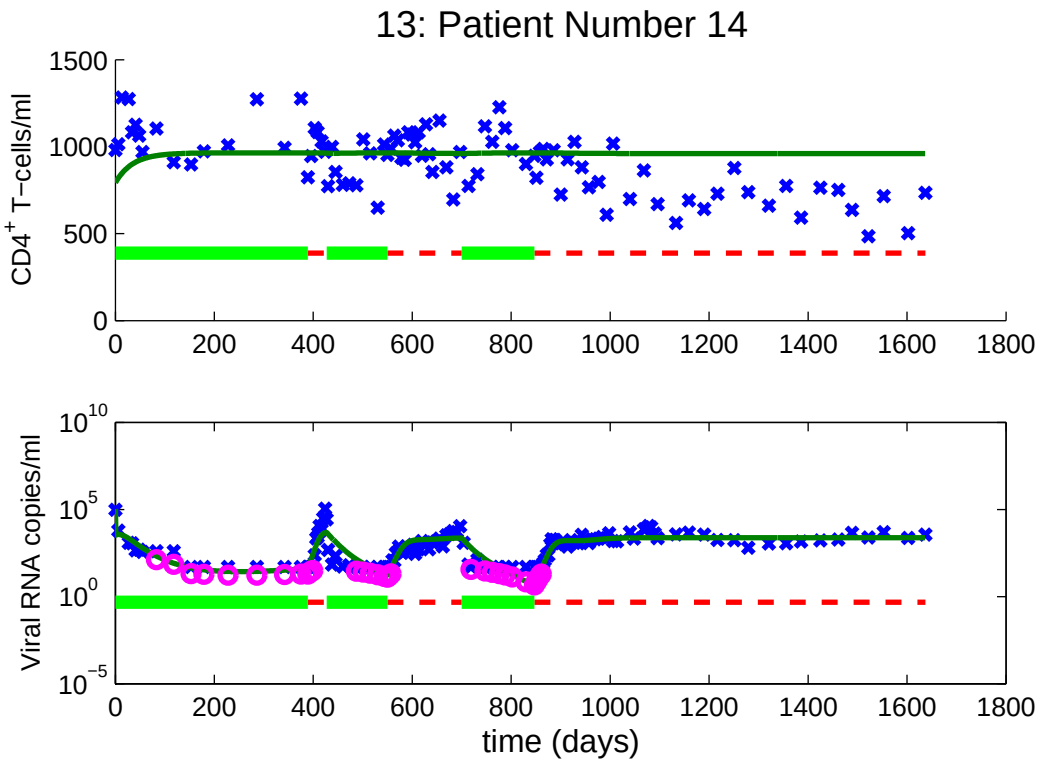

7: Patient Number 7
CD4+ T−cells/ml
1000
500
0
0 200 400 600 800 1000 1200 1400 1600 1800 2000
Viral RNA copies/ml
10^10
10^5
10^0
10^−5
0 200 400 600 800 1000 1200 1400 1600 1800 2000
time (days)

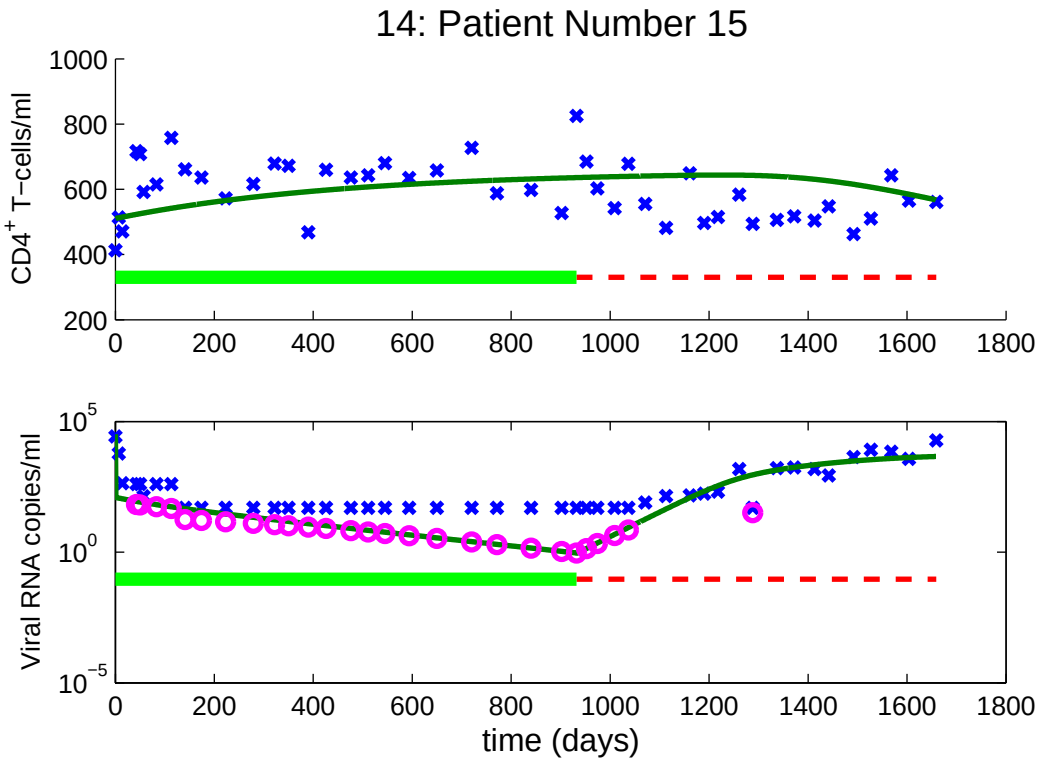

8: Patient Number 8
CD4+ T−cells/ml
1000
800
600
400
200
0 200 400 600 800 1000 1200 1400 1600 1800
Viral RNA copies/ml
10^10
10^5
10^0
10^−5
0 200 400 600 800 1000 1200 1400 1600 1800
time (days)

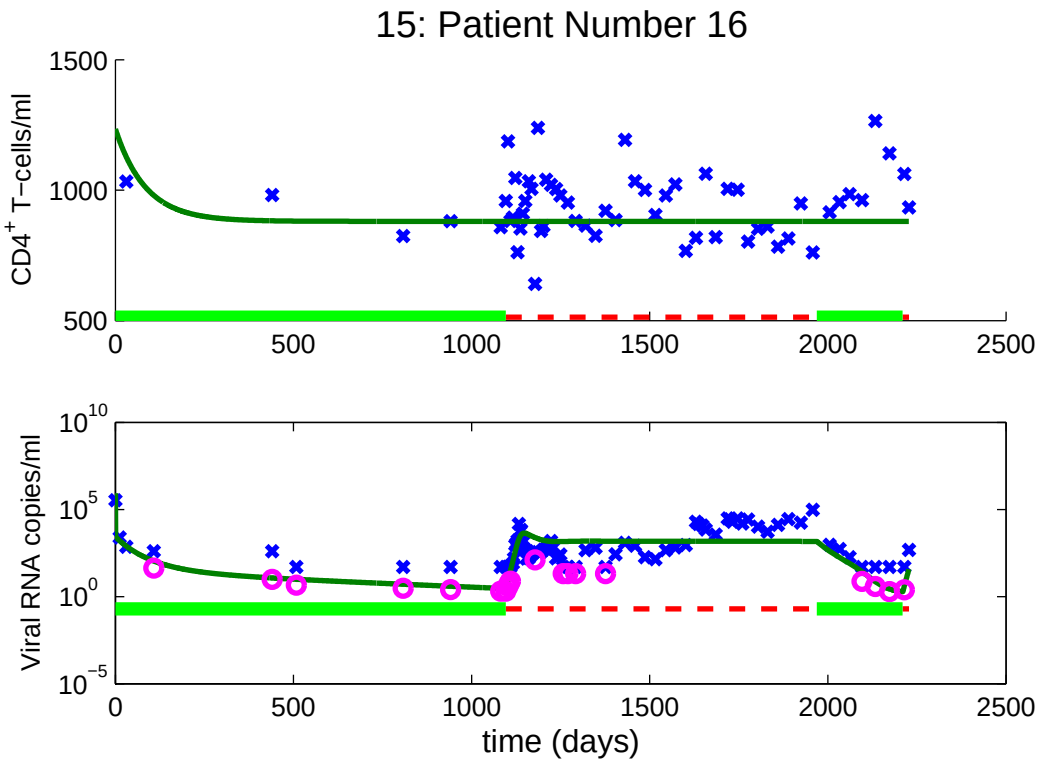

9: Patient Number 9
CD4+ T−cells/ml
Viral RNA copies/ml
time (days)
10: Patient Number 10
CD4+ T−cells/ml
Viral RNA copies/ml
time (days)

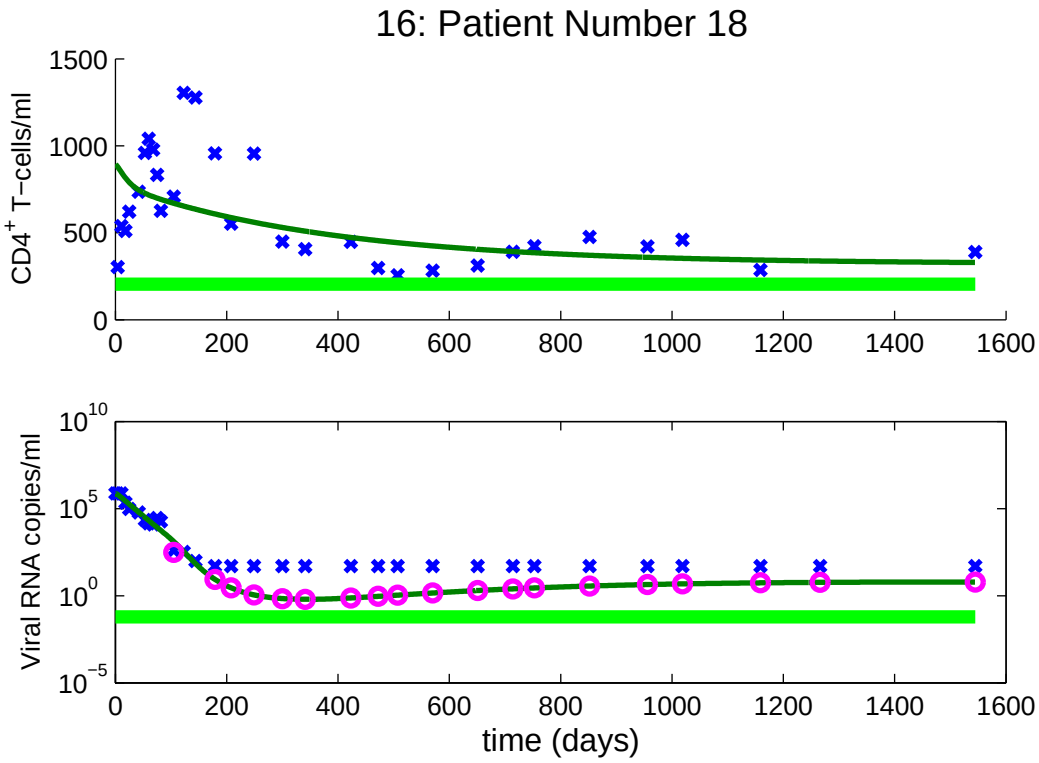

11: Patient Number 12
CD4+ T−cells/ml
800
600
400
200
0
200
400
600
800
1000
1200
1400
1600
Viral RNA copies/ml
10^5
10^0
10^−5
time (days)

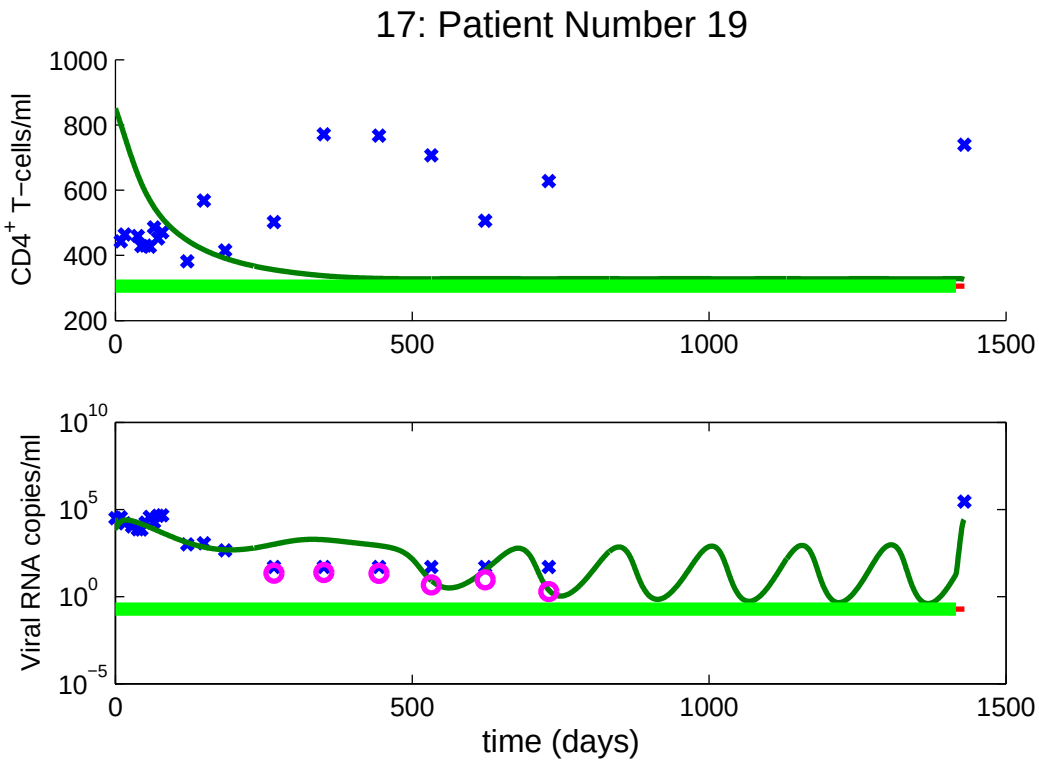

12: Patient Number 13
CD4+ T−cells/ml
1500
1000
500
0
100
200
300
400
500
600
700
800
900
1000
Viral RNA copies/ml
10^10
10^5
10^0
10^−5
time (days)

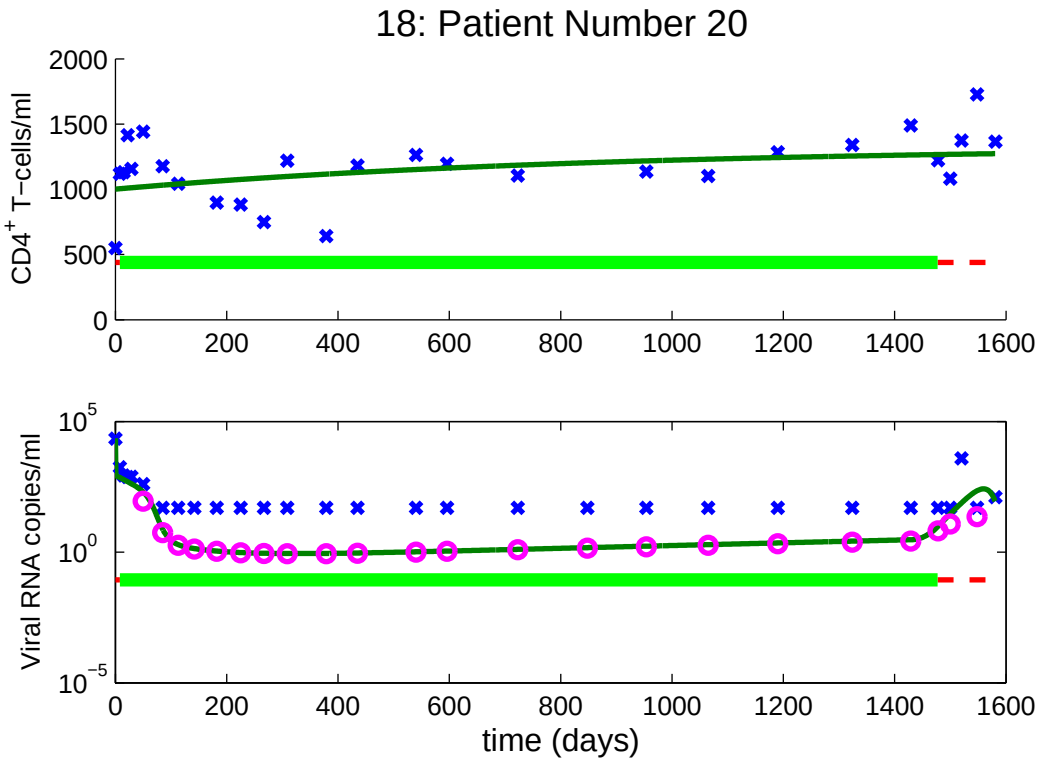

13: Patient Number 14
CD4+ T−cells/ml
Viral RNA copies/ml
time (days)

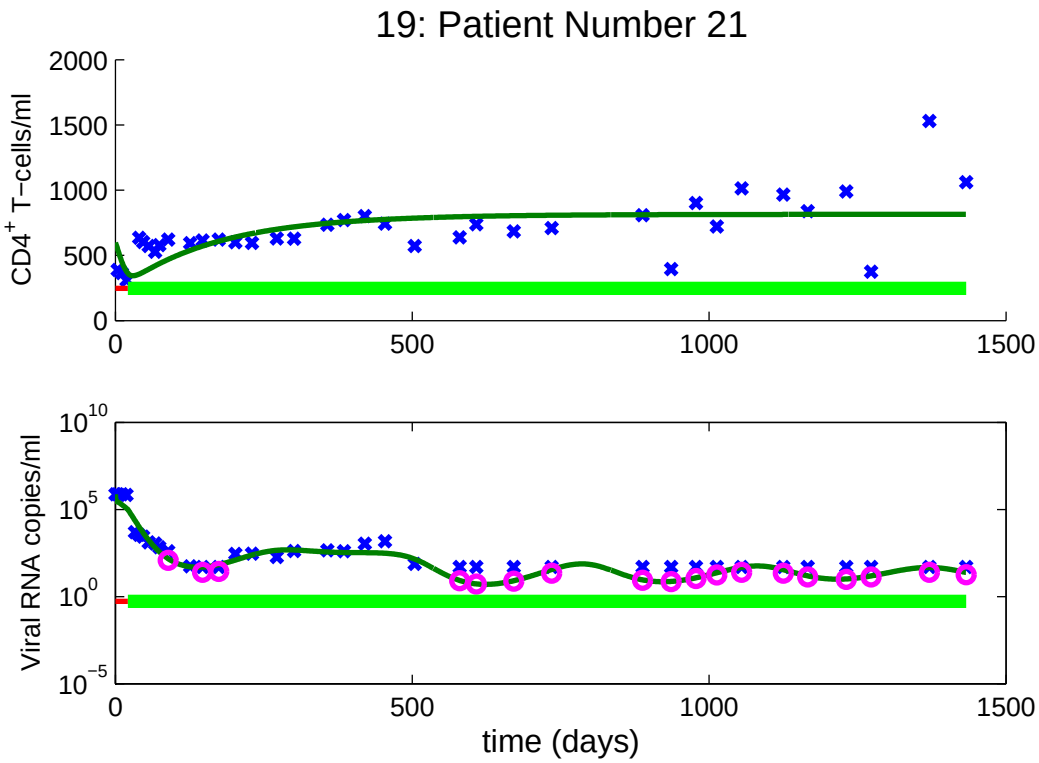

14: Patient Number 15
CD4+ T−cells/ml
Viral RNA copies/ml
time (days)

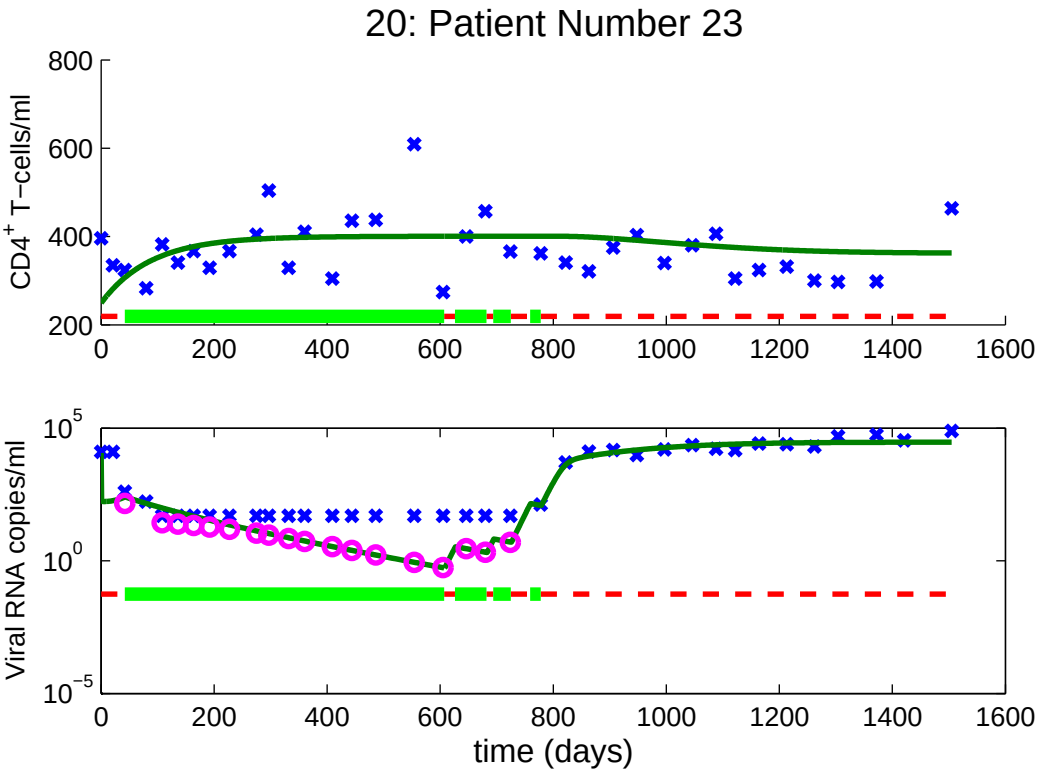

15: Patient Number 16
CD4+ T−cells/ml
1500
1000
500
0
500
1000
1500
2000
2500
Viral RNA copies/ml
10^10
10^5
10^0
10^−5
time (days)

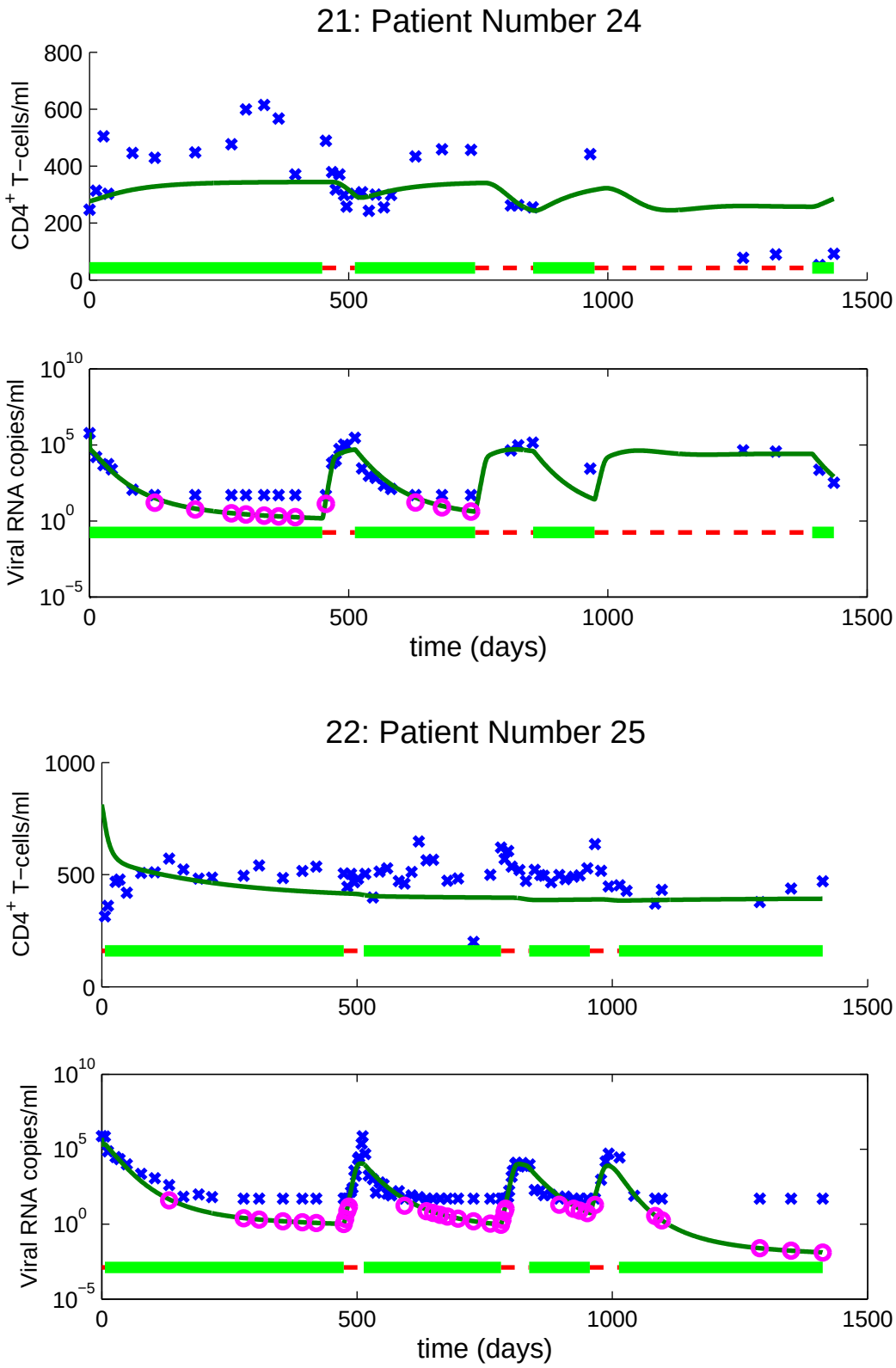

16: Patient Number 18
CD4+ T−cells/ml
1500
1000
500
0
200
400
600
800
1000
1200
1400
1600
Viral RNA copies/ml
10^10
10^5
10^0
10^−5
time (days)

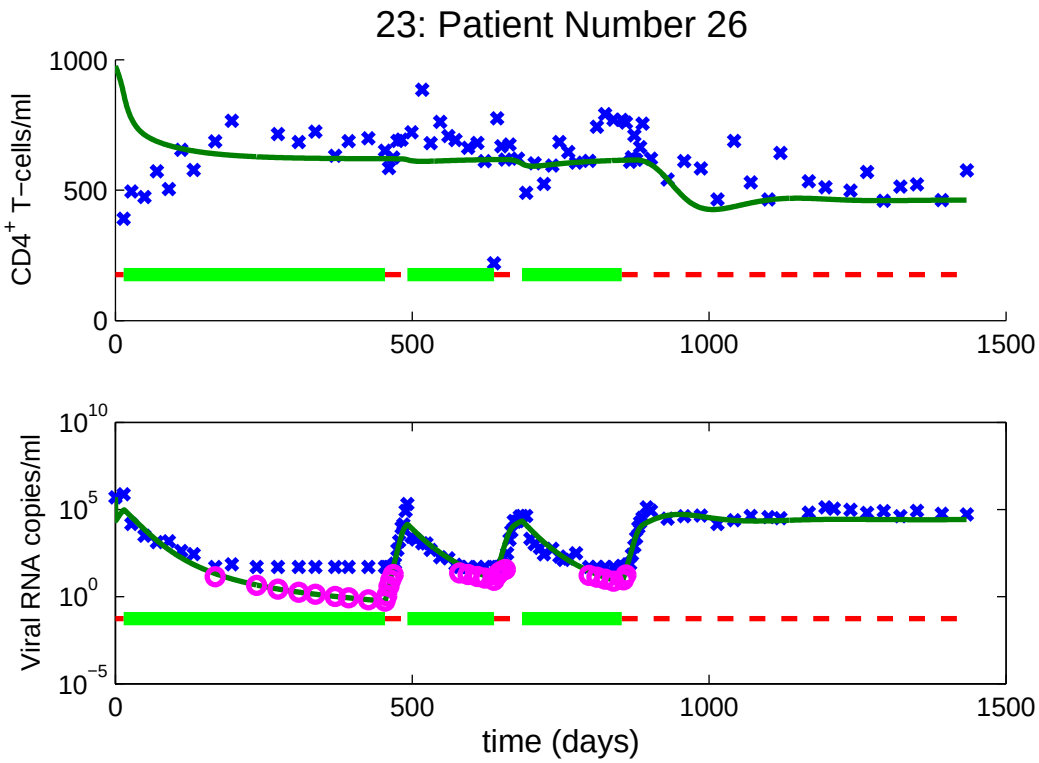

17: Patient Number 19
CD4+ T−cells/ml
1000
800
600
400
200
0
500
1000
1500
Viral RNA copies/ml
10^10
10^5
10^0
10^−5
0
500
1000
1500
time (days)

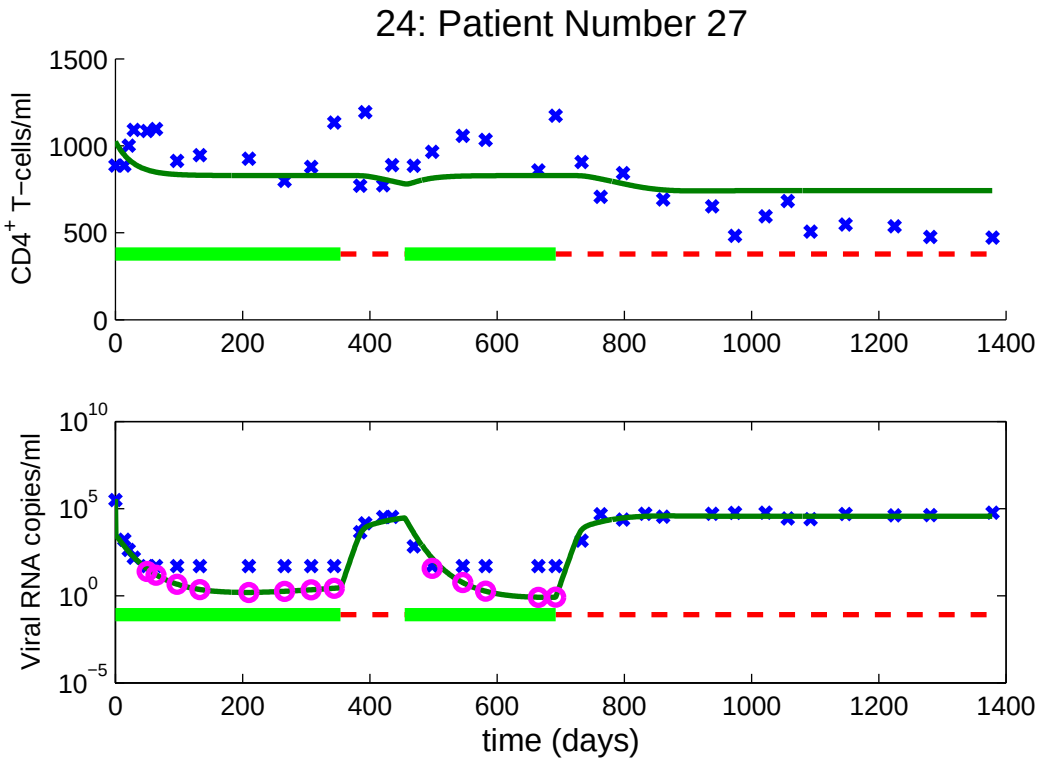

18: Patient Number 20
CD4+ T−cells/ml
2000
1500
1000
500
0
0
200
400
600
800
1000
1200
1400
1600
Viral RNA copies/ml
10^5
10^0
10^−5
0
200
400
600
800
1000
1200
1400
1600
time (days)

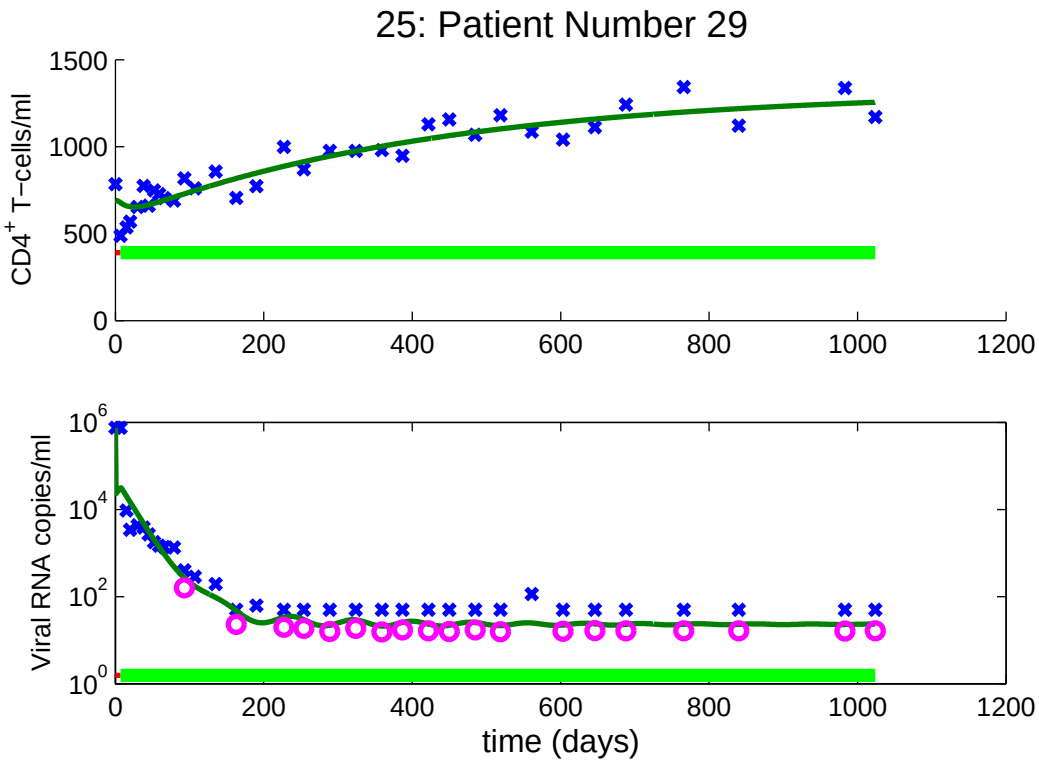

19: Patient Number 21
CD4+ T−cells/ml
2000
1500
1000
500
0
0
500
1000
1500
Viral RNA copies/ml
10^10
10^5
10^0
10^−5
0
500
1000
1500
time (days)

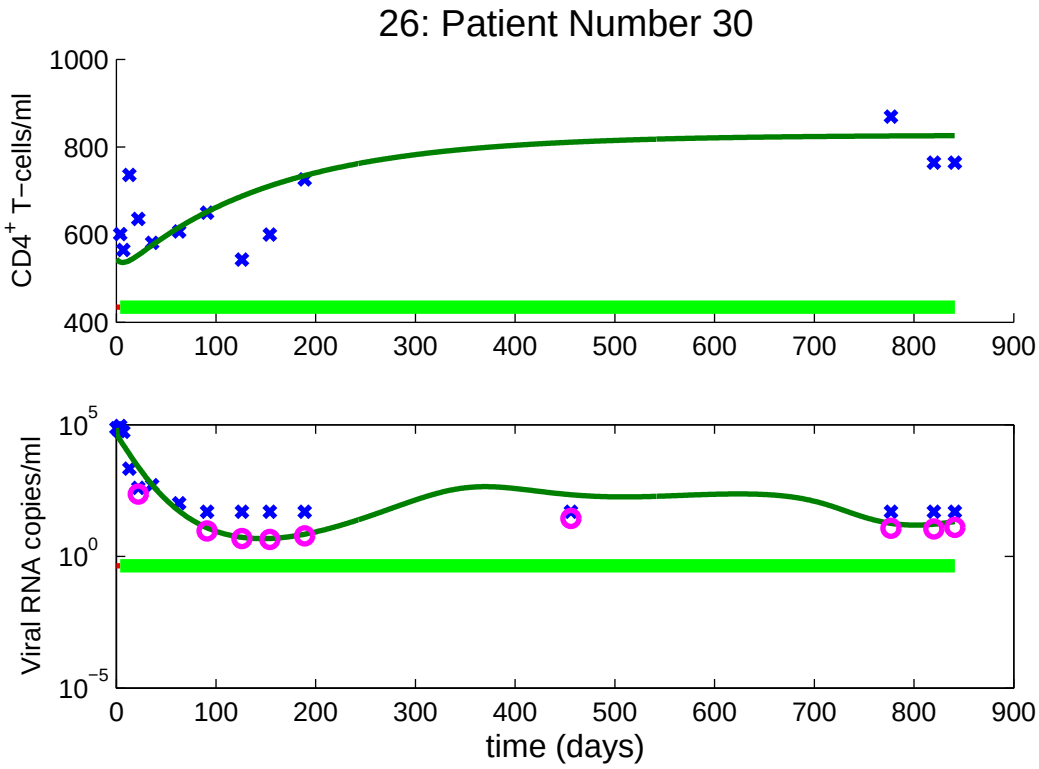

20: Patient Number 23
CD4+ T−cells/ml
800
600
400
200
0
200
400
600
800
1000
1200
1400
1600
Viral RNA copies/ml
10^5
10^0
10^−5
0
200
400
600
800
1000
1200
1400
1600
time (days)

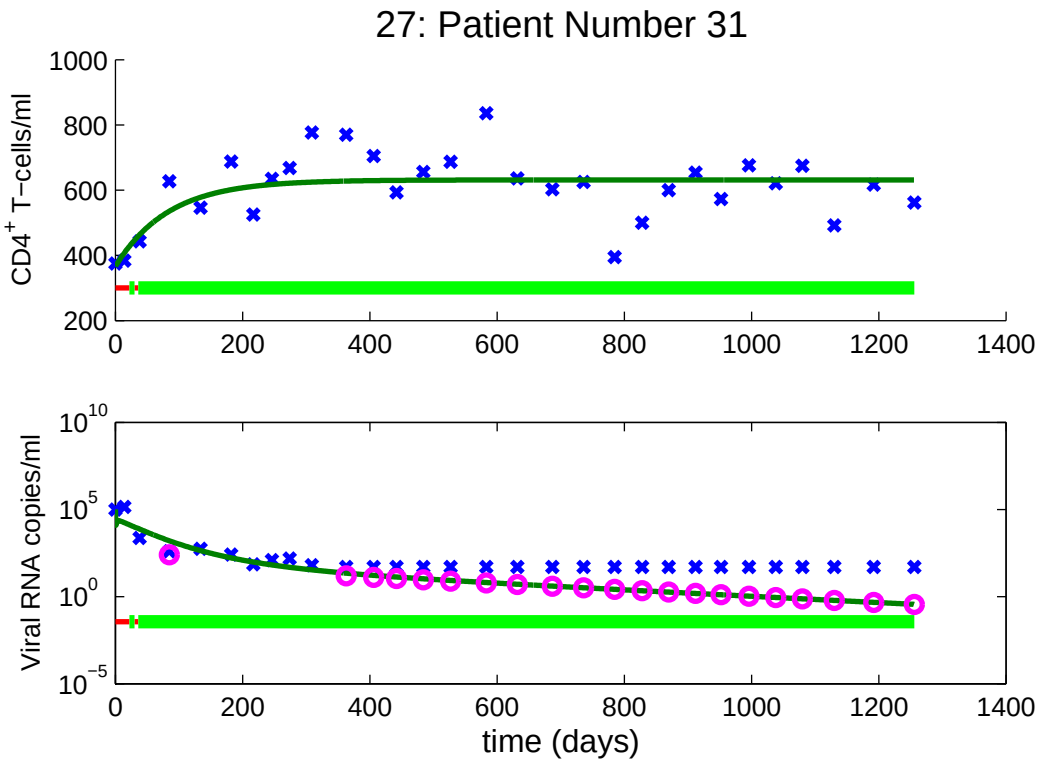

21: Patient Number 24
CD4+ T−cells/ml
800
600
400
200
0
0
500
1000
1500
Viral RNA copies/ml
10^10
10^5
10^0
10^−5
0
500
1000
1500
time (days)
22: Patient Number 25
CD4+ T−cells/ml
1000
500
0
0
500
1000
1500
Viral RNA copies/ml
10^10
10^5
10^0
10^−5
0
500
1000
1500
time (days)

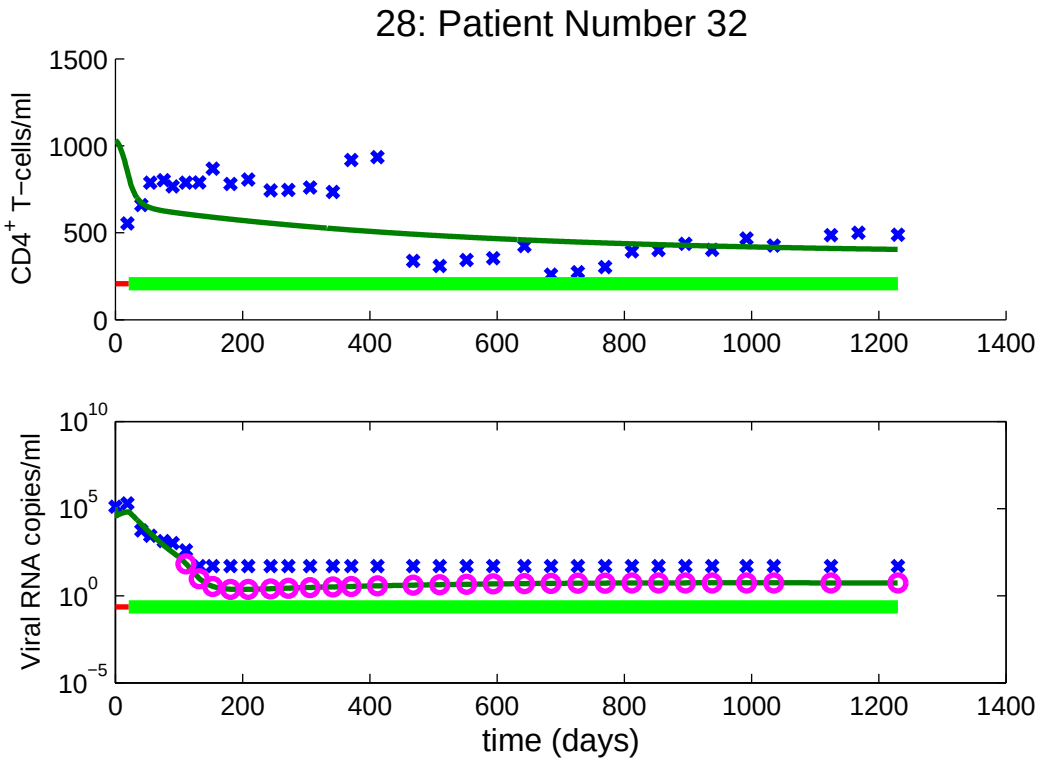

23: Patient Number 26
CD4+ T−cells/ml
1000
500
0
0
500
1000
1500
Viral RNA copies/ml
10^10
10^5
10^0
10^−5
0
500
1000
1500
time (days)

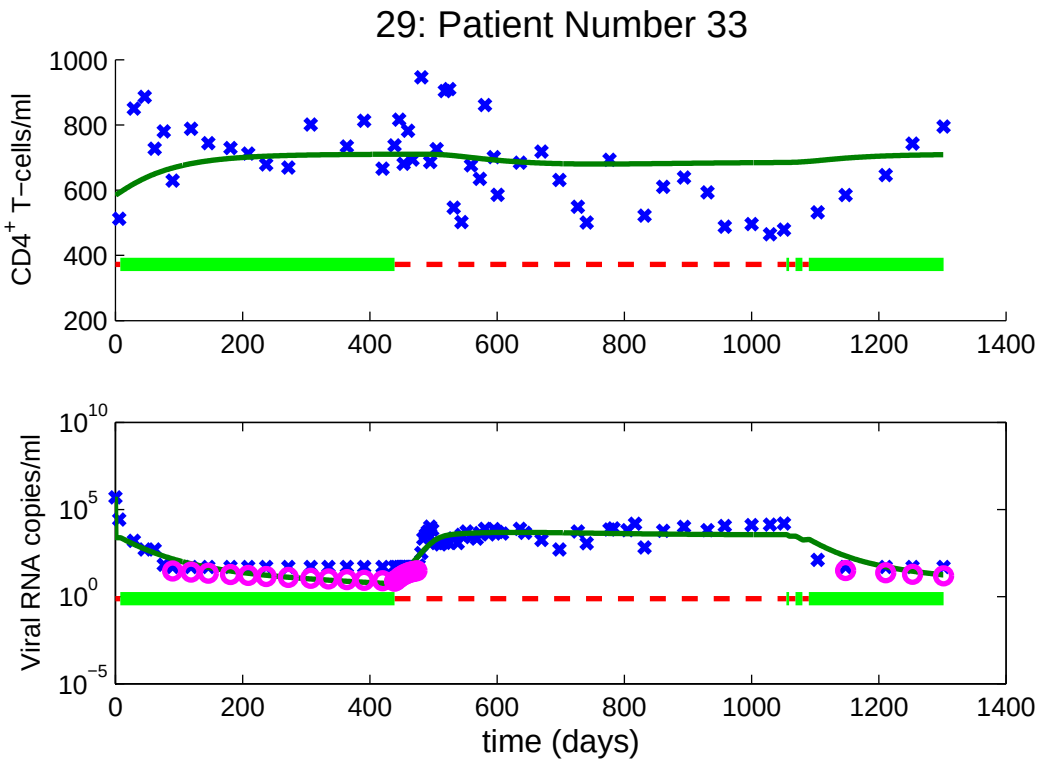

24: Patient Number 27
CD4+ T−cells/ml
1500
1000
500
0
0
200
400
600
800
1000
1200
1400
Viral RNA copies/ml
10^10
10^5
10^0
10^−5
0
200
400
600
800
1000
1200
1400
time (days)

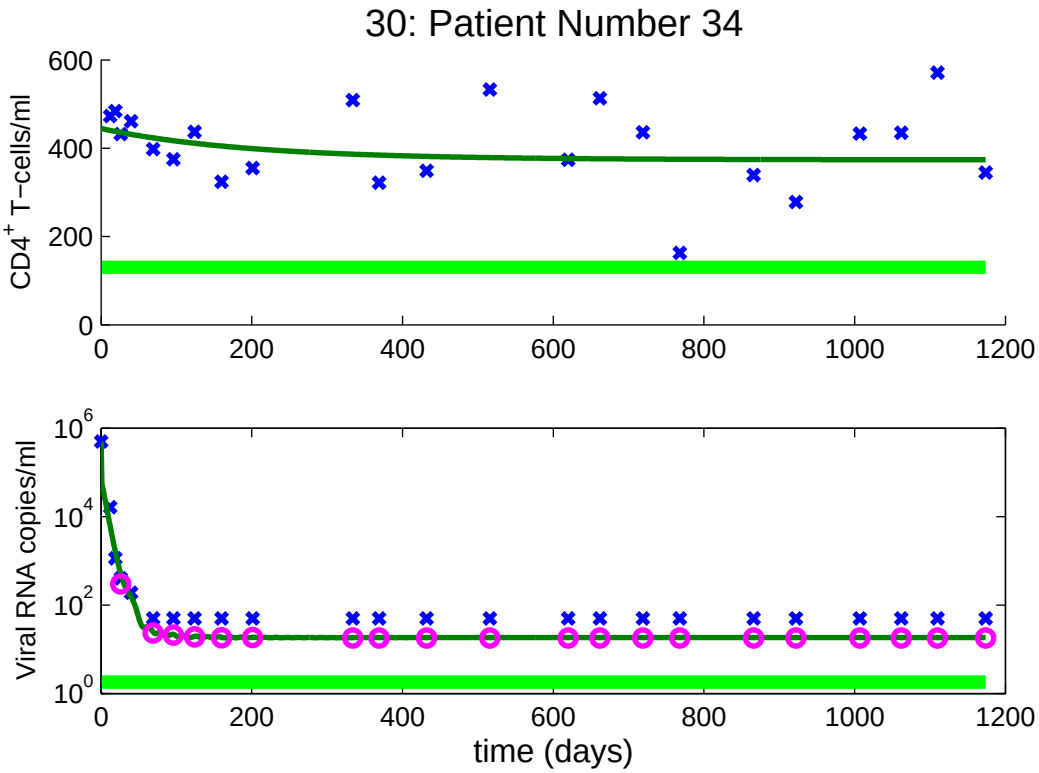

25: Patient Number 29
CD4+ T−cells/ml
1500
1000
500
0
0
200
400
600
800
1000
1200
Viral RNA copies/ml
10^6
10^4
10^2
10^0
time (days)

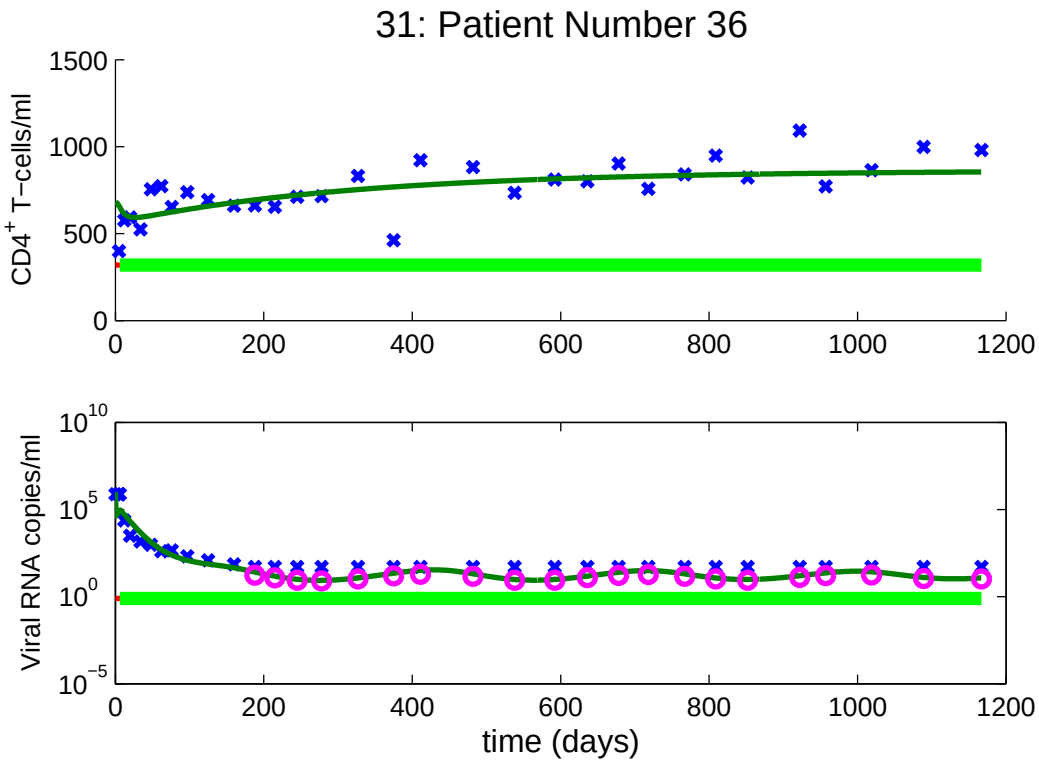

26: Patient Number 30
CD4+ T−cells/ml
1000
800
600
400
0
100
200
300
400
500
600
700
800
900
Viral RNA copies/ml
10^5
10^0
10^−5
time (days)

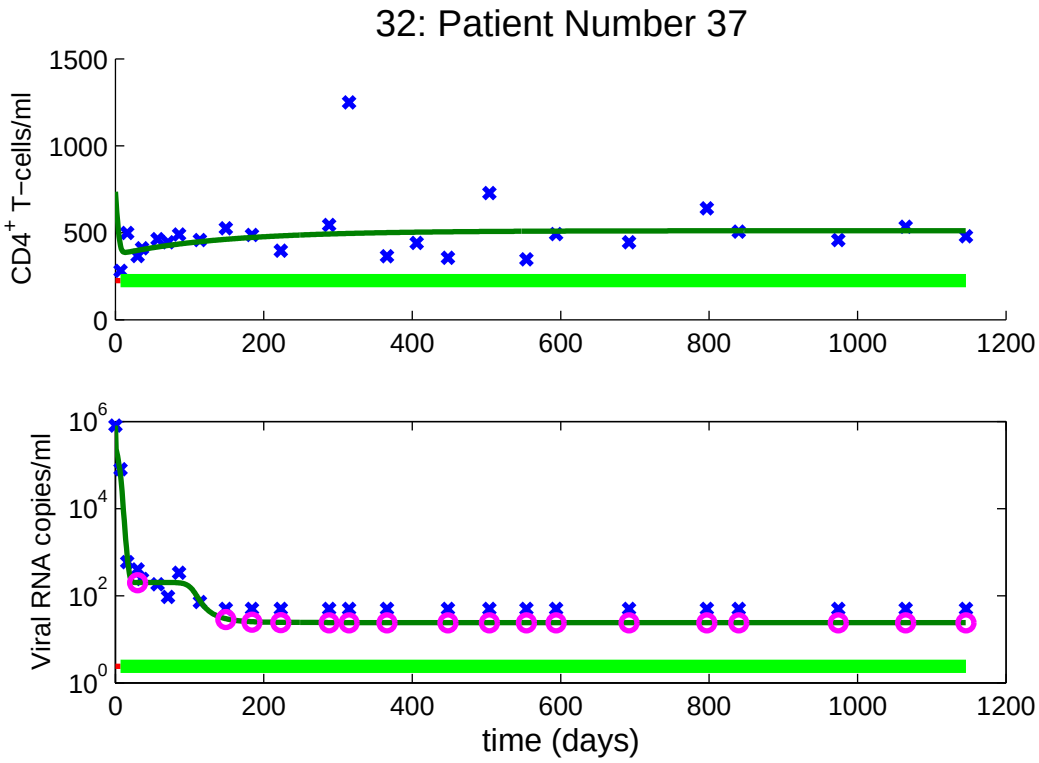

27: Patient Number 31
CD4+ T−cells/ml
1000
800
600
400
200
0
200
400
600
800
1000
1200
1400
Viral RNA copies/ml
10^10
10^5
10^0
10^−5
time (days)

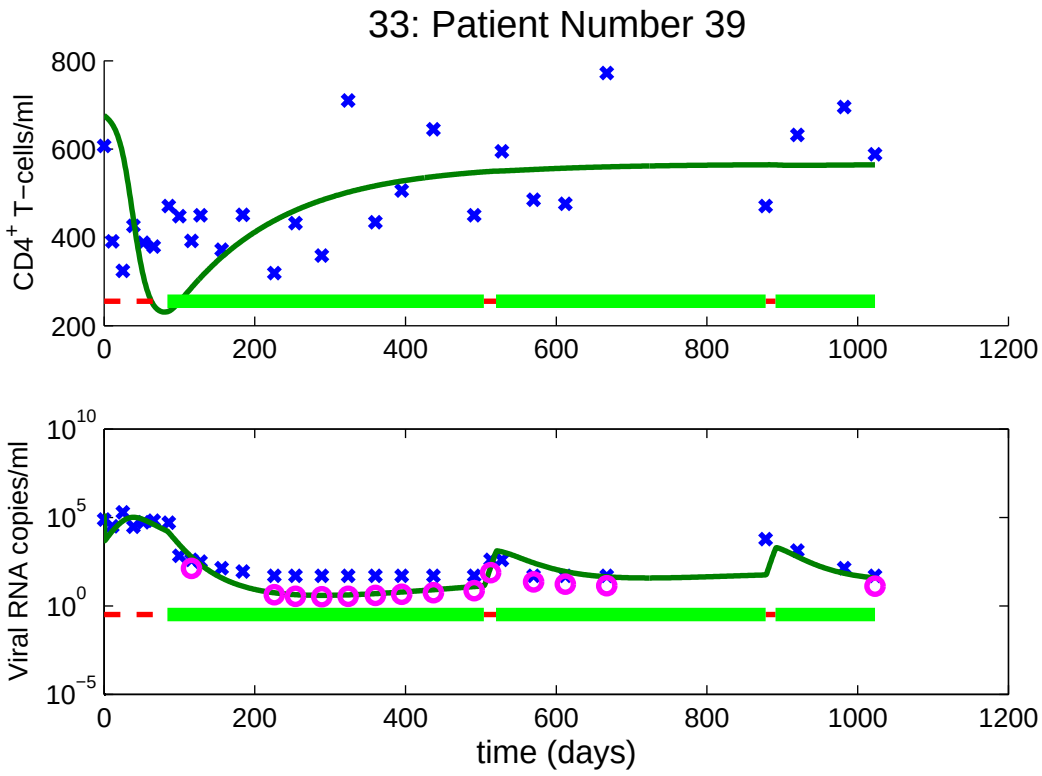

28: Patient Number 32
CD4+ T−cells/ml
1500
1000
500
0
200
400
600
800
1000
1200
1400
Viral RNA copies/ml
10^10
10^5
10^0
10^−5
time (days)

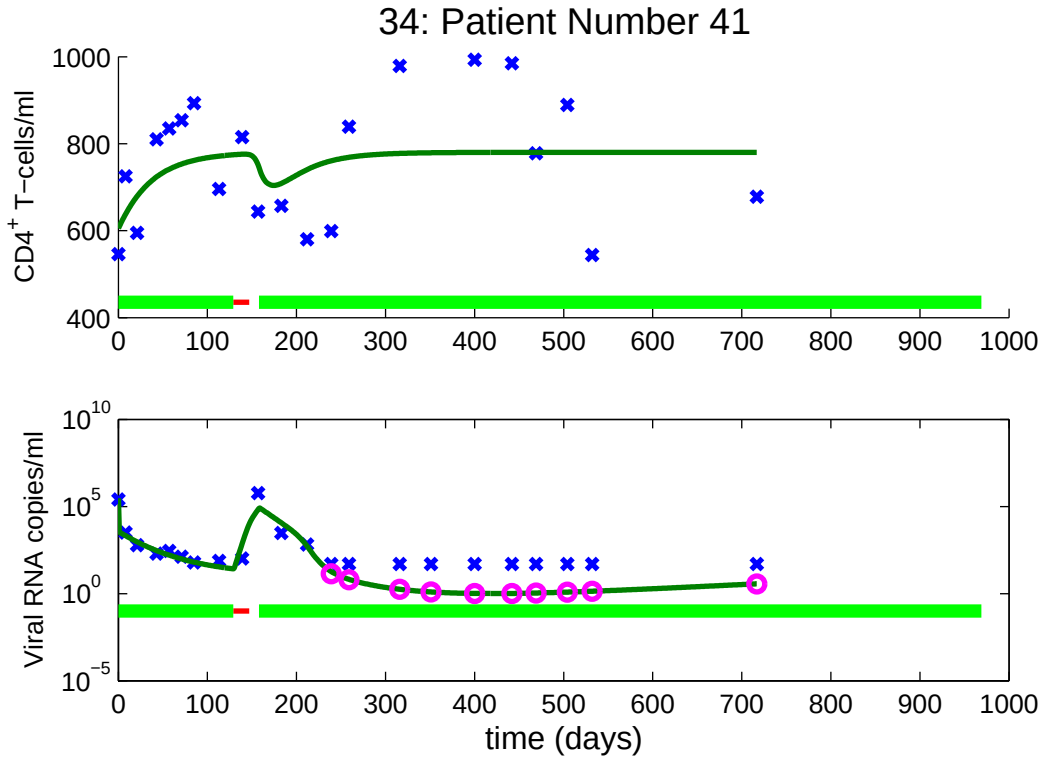

29: Patient Number 33
CD4+ T−cells/ml
1000
800
600
400
200
0
200
400
600
800
1000
1200
1400
Viral RNA copies/ml
10^10
10^5
10^0
10^−5
time (days)

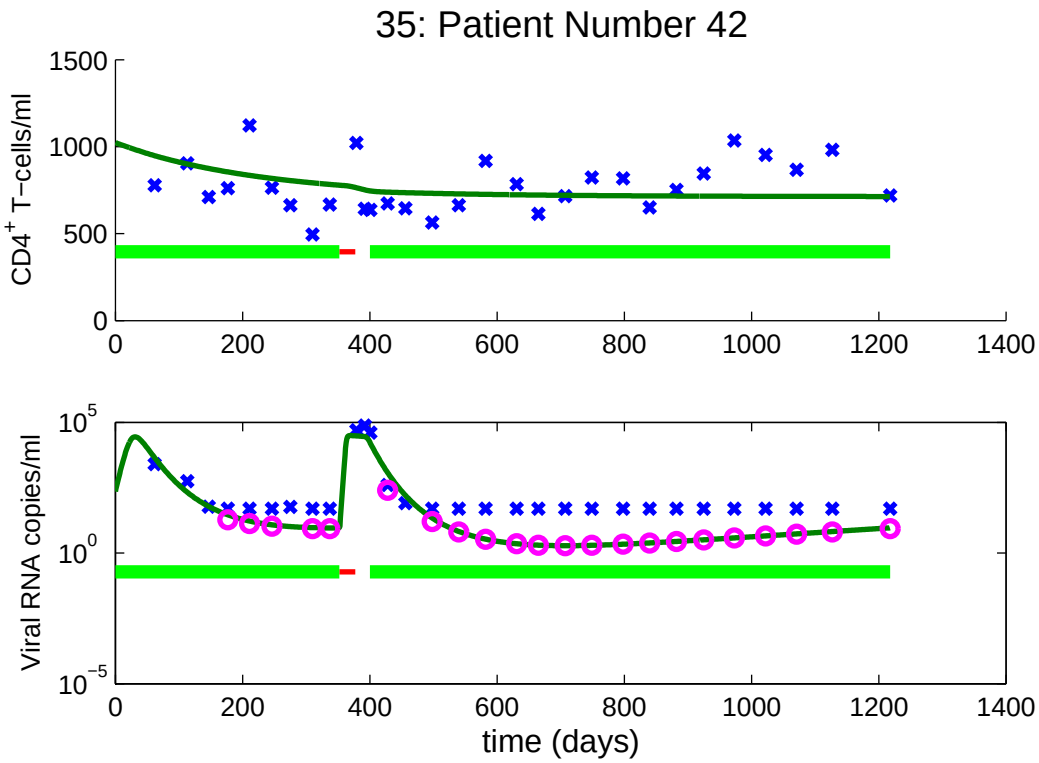

30: Patient Number 34
CD4+ T−cells/ml
600
400
200
0
200
400
600
800
1000
1200
Viral RNA copies/ml
10^6
10^4
10^2
10^0
time (days)

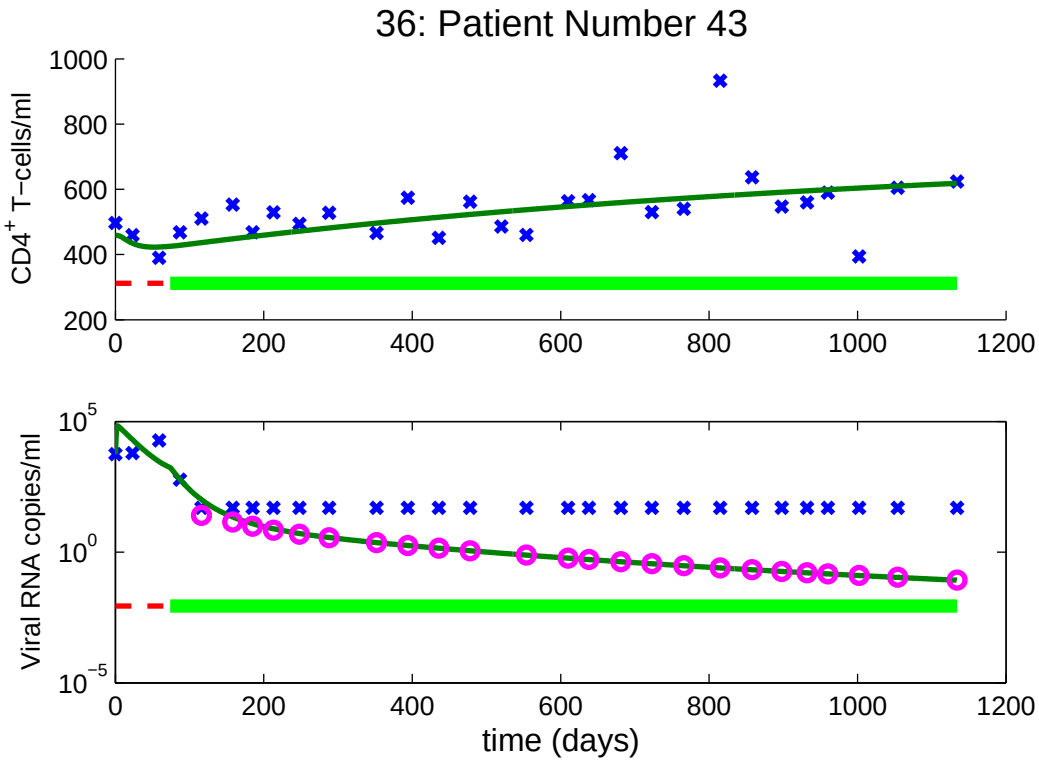

31: Patient Number 36
CD4+ T−cells/ml
1500
1000
500
0
0
200
400
600
800
1000
1200
Viral RNA copies/ml
10^10
10^5
10^0
10^−5
time (days)

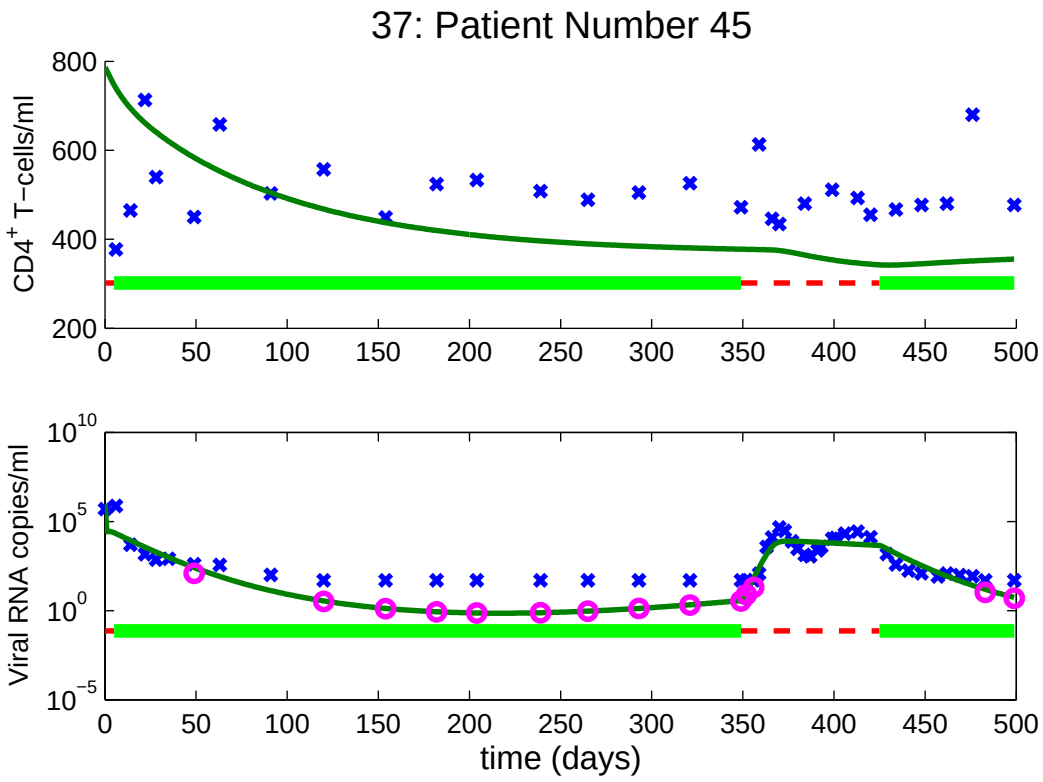

32: Patient Number 37
CD4+ T−cells/ml
1500
1000
500
0
0
200
400
600
800
1000
1200
Viral RNA copies/ml
10^6
10^4
10^2
10^0
time (days)

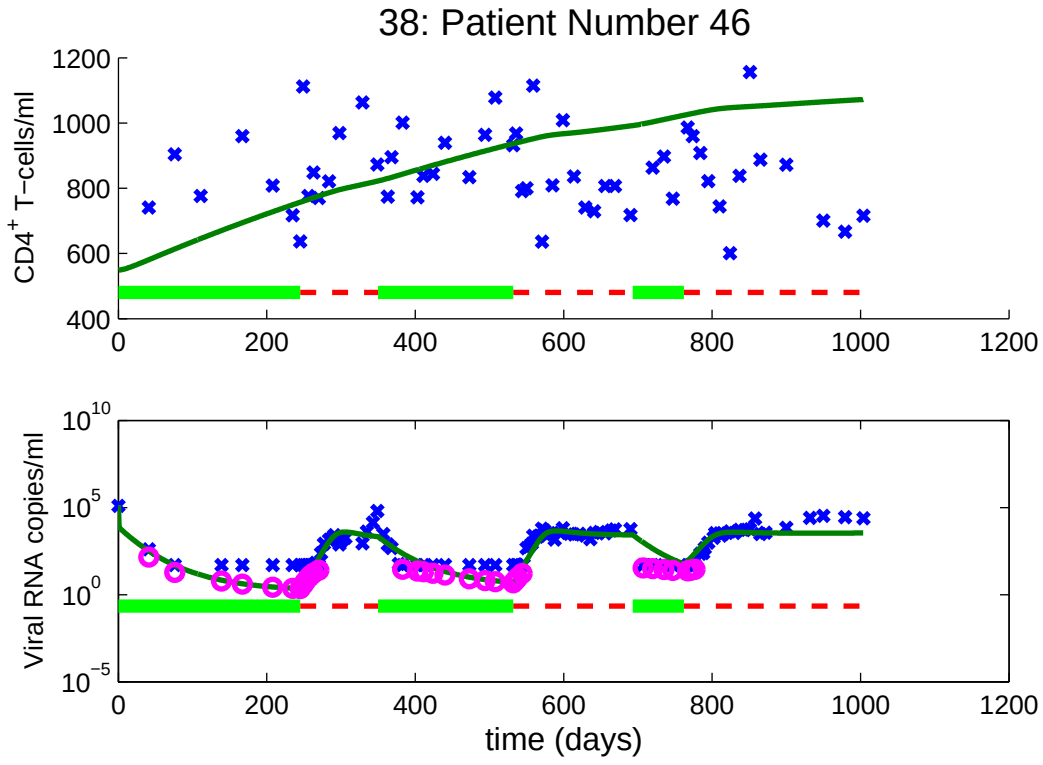

33: Patient Number 39
CD4+ T−cells/ml
800
600
400
200
0
200
400
600
800
1000
1200
Viral RNA copies/ml
10^10
10^5
10^0
10^−5
time (days)

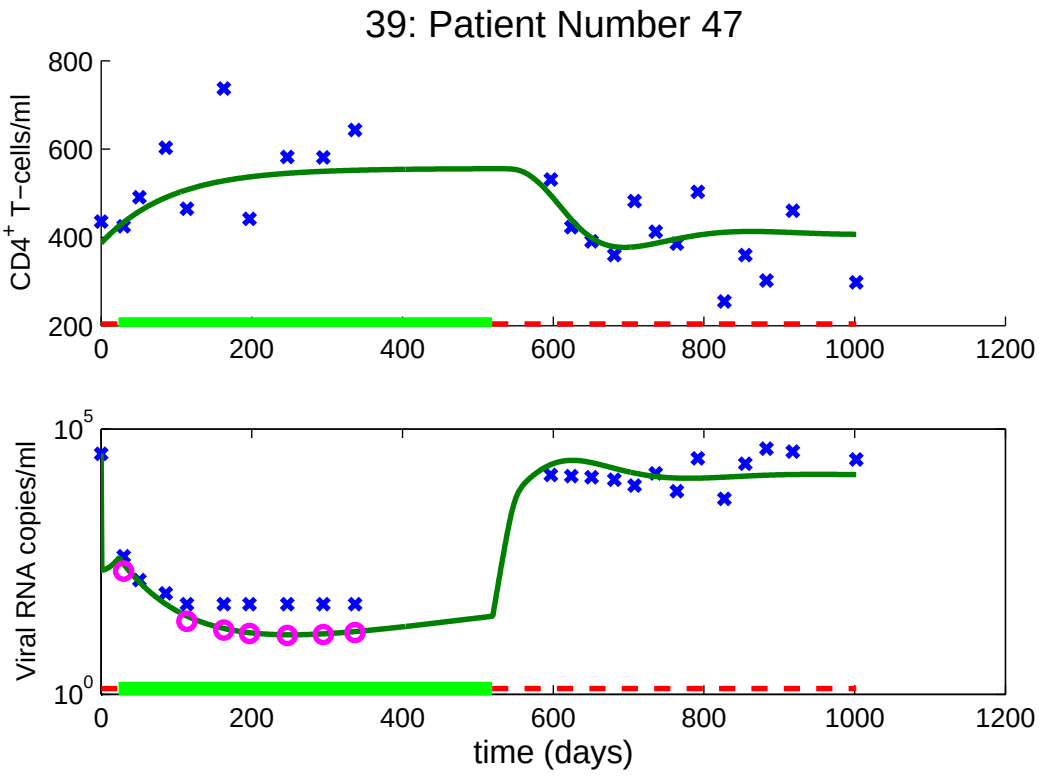

34: Patient Number 41
CD4+ T−cells/ml
1000
800
600
400
0
100
200
300
400
500
600
700
800
900
1000
Viral RNA copies/ml
10^10
10^5
10^0
10^−5
time (days)

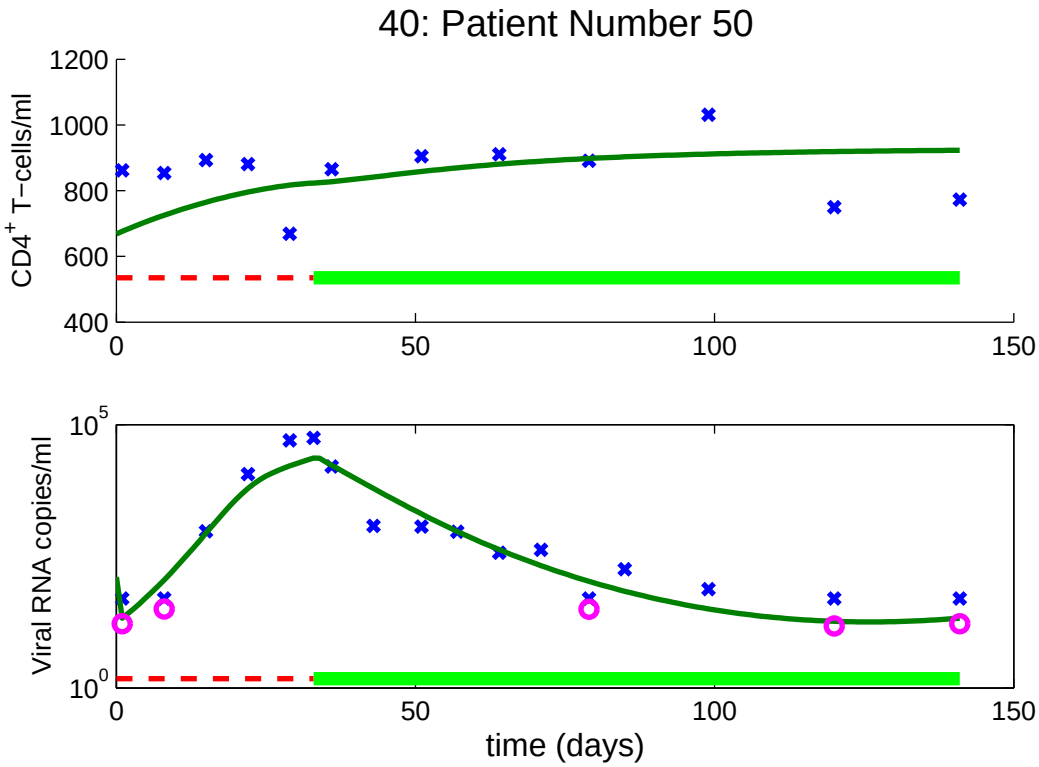

35: Patient Number 42
CD4+ T−cells/ml
Viral RNA copies/ml
time (days)

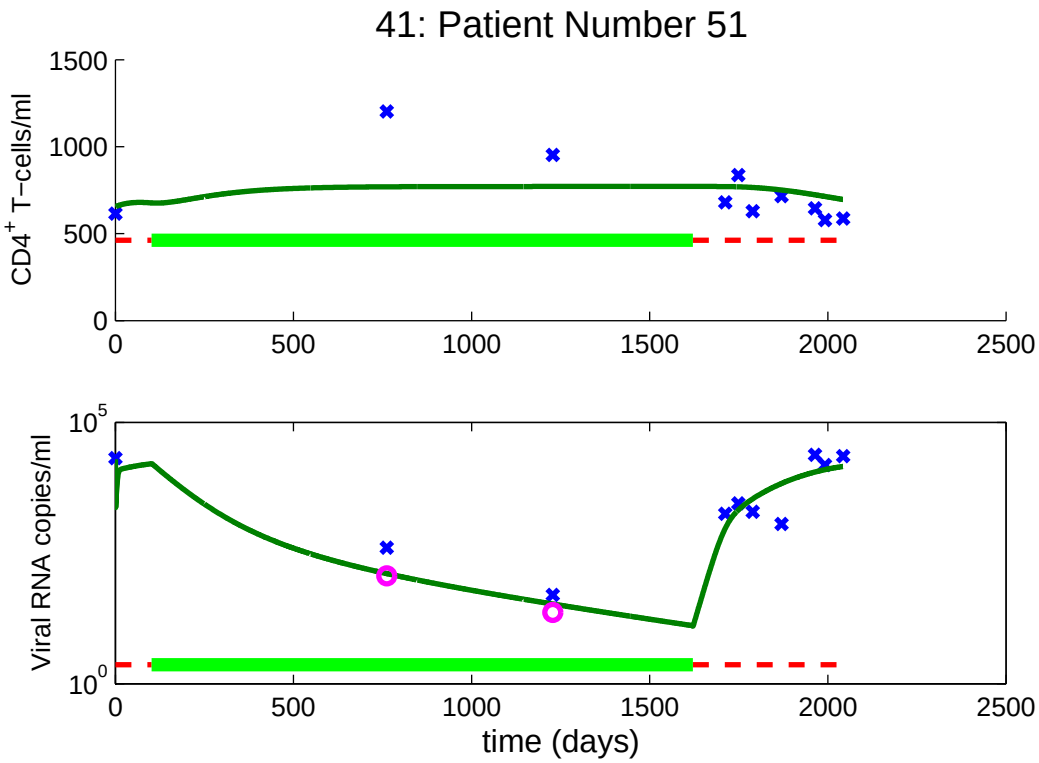

36: Patient Number 43
CD4+ T−cells/ml
Viral RNA copies/ml
time (days)

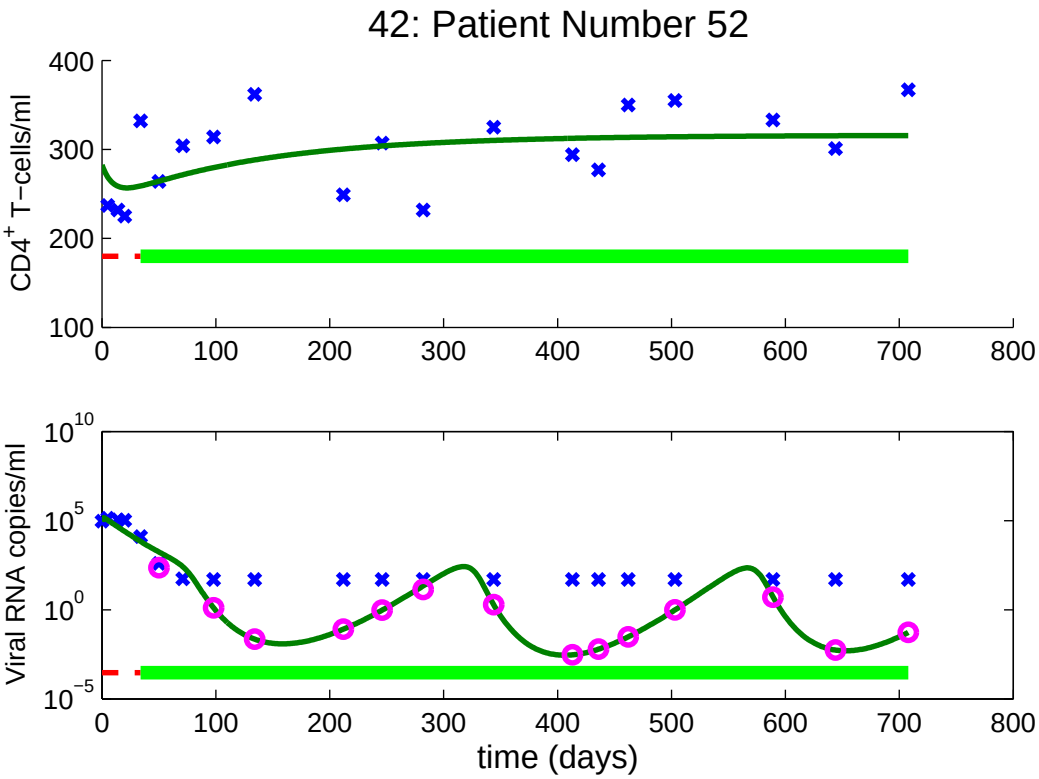

37: Patient Number 45
CD4+ T−cells/ml
800
600
400
200
0
50
100
150
200
250
300
350
400
450
500
Viral RNA copies/ml
10^10
10^5
10^0
10^−5
time (days)

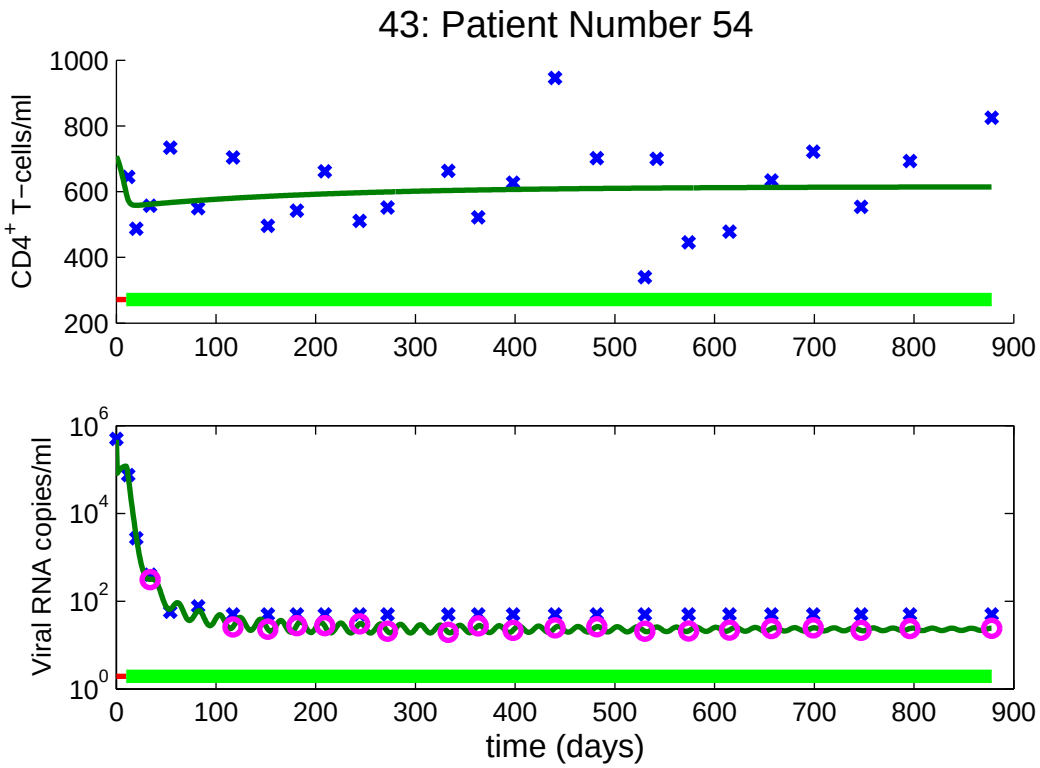

38: Patient Number 46
CD4+ T−cells/ml
1200
1000
800
600
400
0
200
400
600
800
1000
1200
Viral RNA copies/ml
10^10
10^5
10^0
10^−5
time (days)

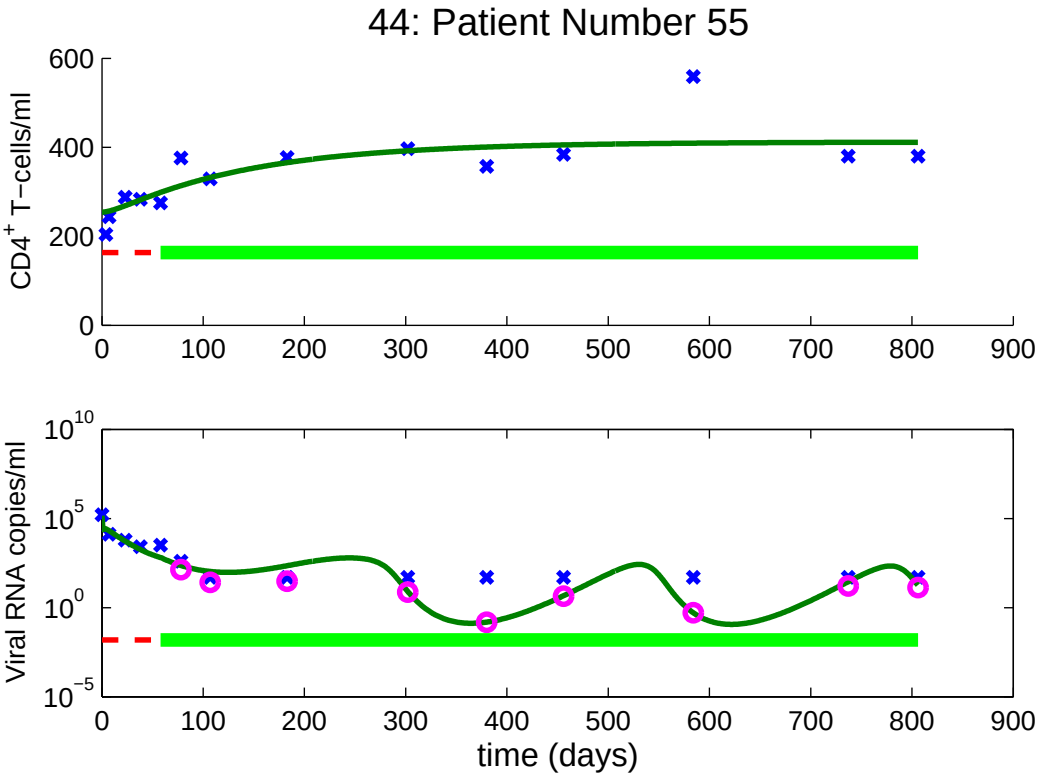

39: Patient Number 47
CD4+ T−cells/ml
800
600
400
200
0
200
400
600
800
1000
1200
Viral RNA copies/ml
10^5
10^0
time (days)

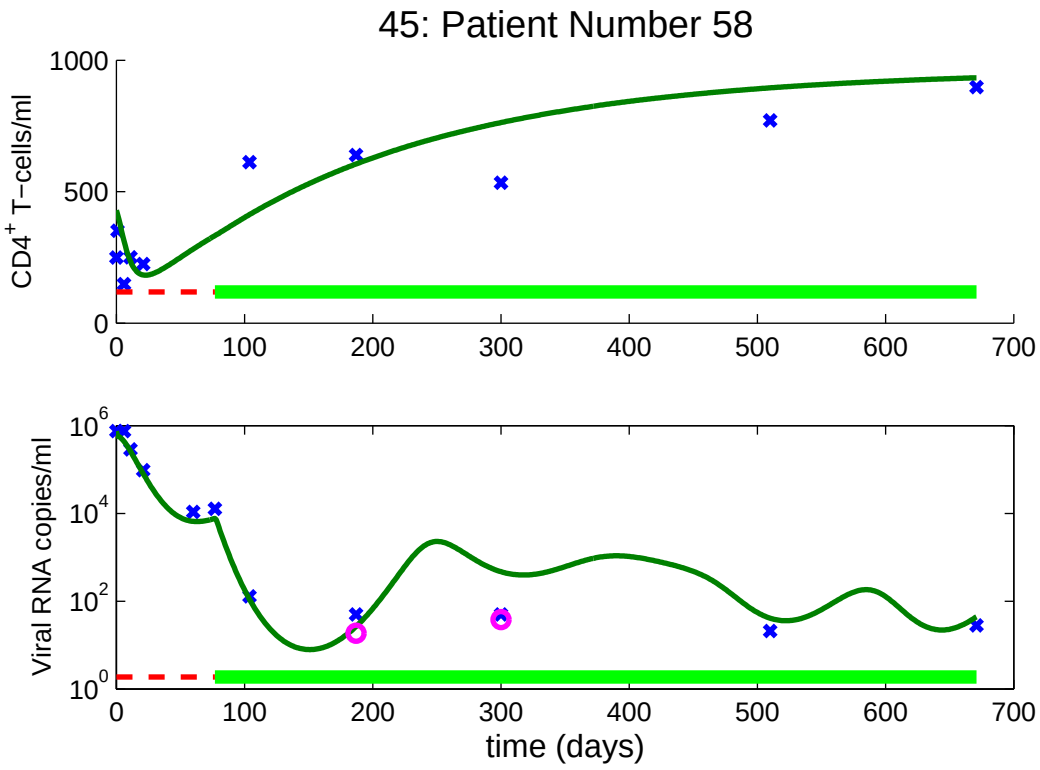

40: Patient Number 50
CD4+ T−cells/ml
1200
1000
800
600
400
0
50
100
150
Viral RNA copies/ml
10^5
10^0
time (days)

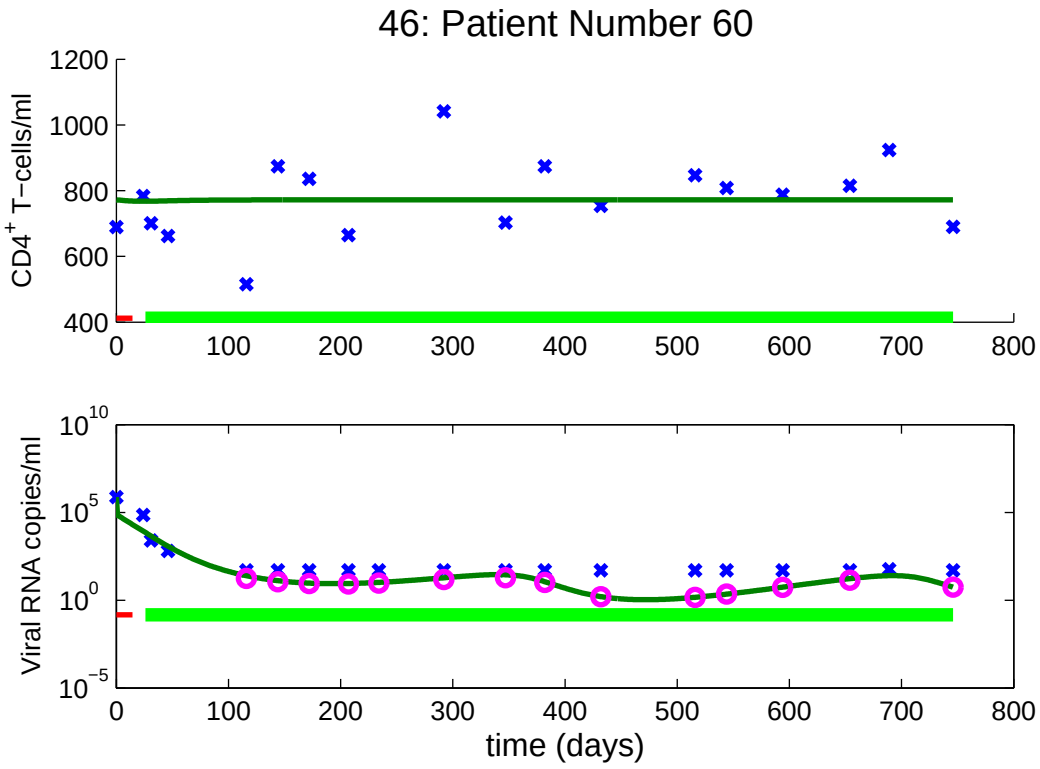

41: Patient Number 51
CD4+ T−cells/ml
1500
1000
500
0
0
500
1000
1500
2000
2500
Viral RNA copies/ml
10^5
10^0
time (days)

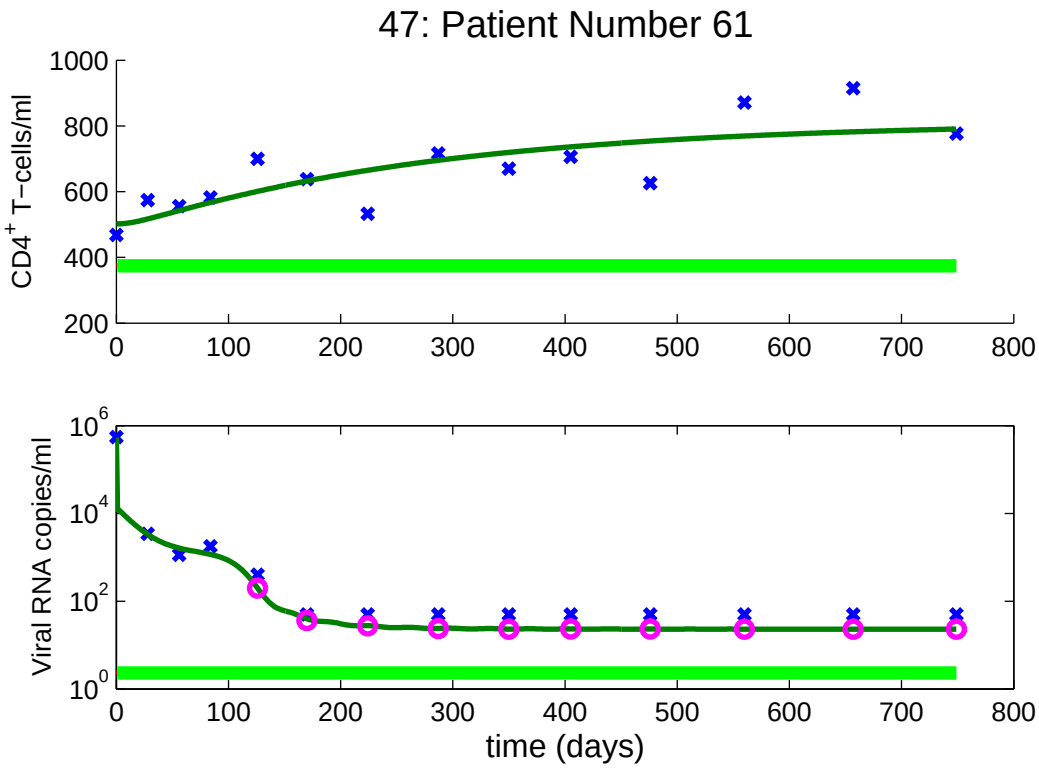

42: Patient Number 52
CD4+ T−cells/ml
400
300
200
100
0
100
200
300
400
500
600
700
800
Viral RNA copies/ml
10^10
10^5
10^0
10^−5
time (days)

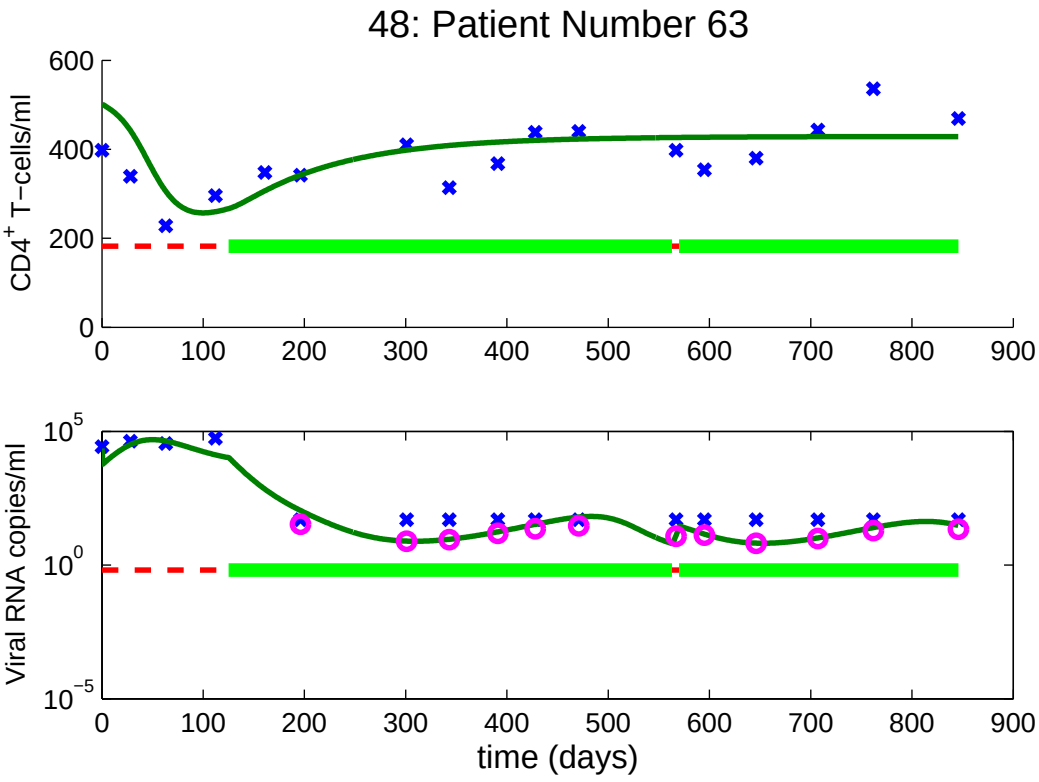

43: Patient Number 54
CD4+ T−cells/ml
1000
800
600
400
200
0
100
200
300
400
500
600
700
800
900
Viral RNA copies/ml
10^6
10^4
10^2
10^0
time (days)

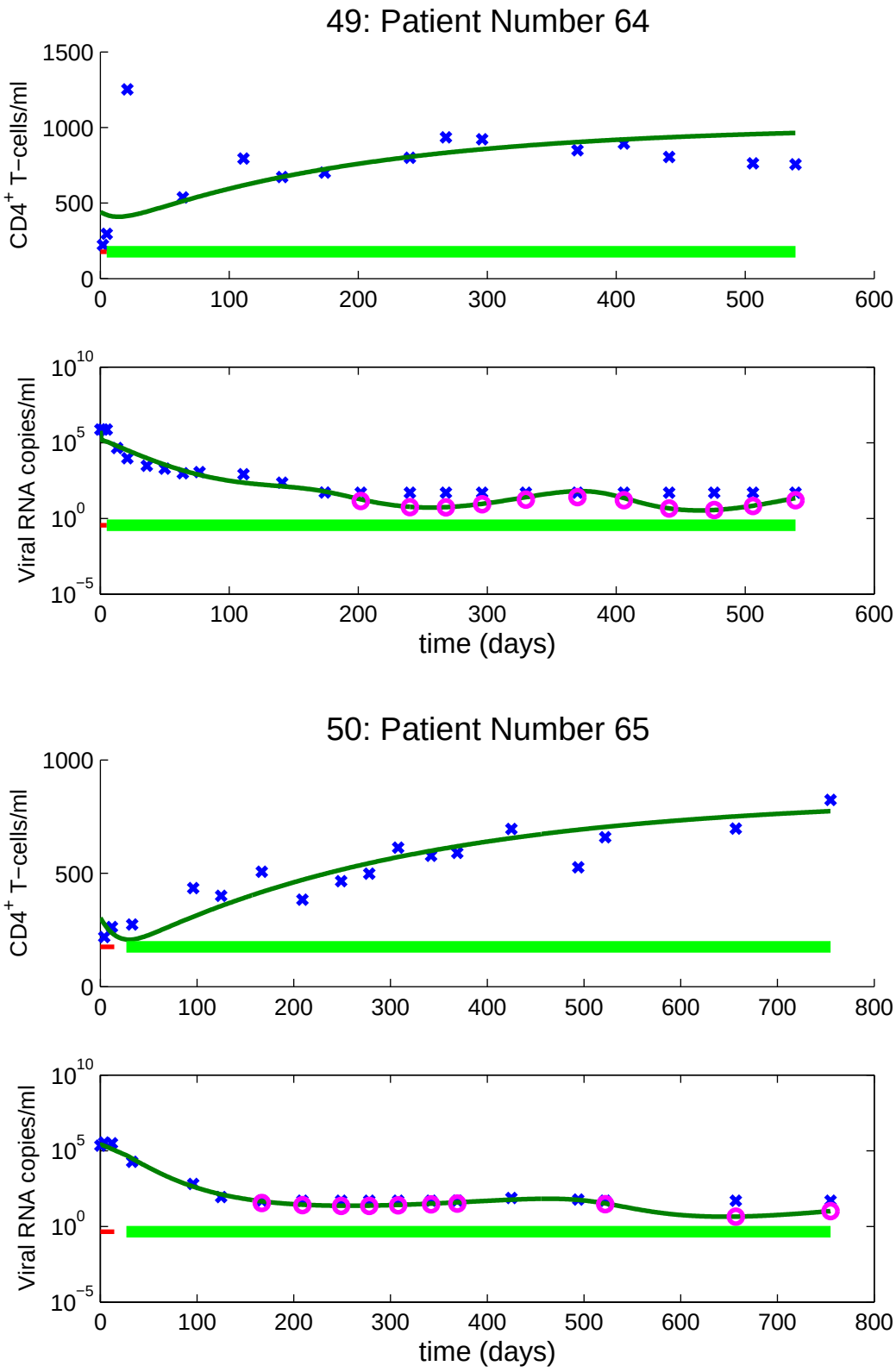

44: Patient Number 55
CD4+ T−cells/ml
600
400
200
0
100
200
300
400
500
600
700
800
900
Viral RNA copies/ml
10^10
10^5
10^0
10^−5
time (days)

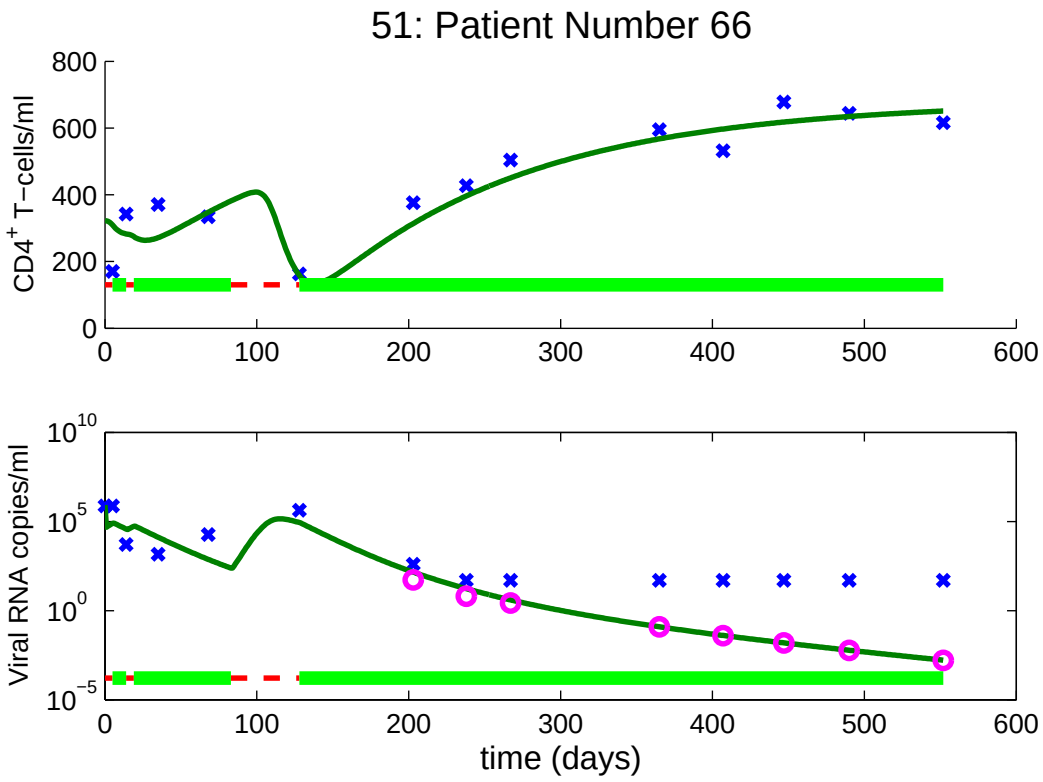

45: Patient Number 58
CD4+ T−cells/ml
Viral RNA copies/ml
time (days)

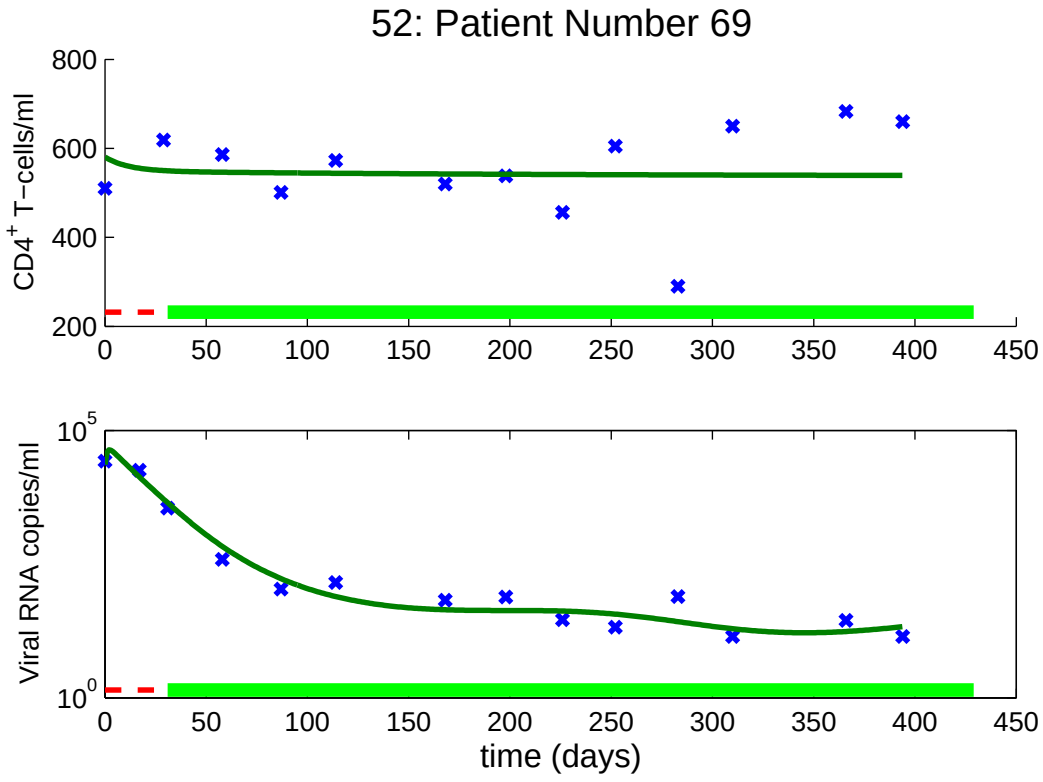

46: Patient Number 60
CD4+ T−cells/ml
Viral RNA copies/ml
time (days)

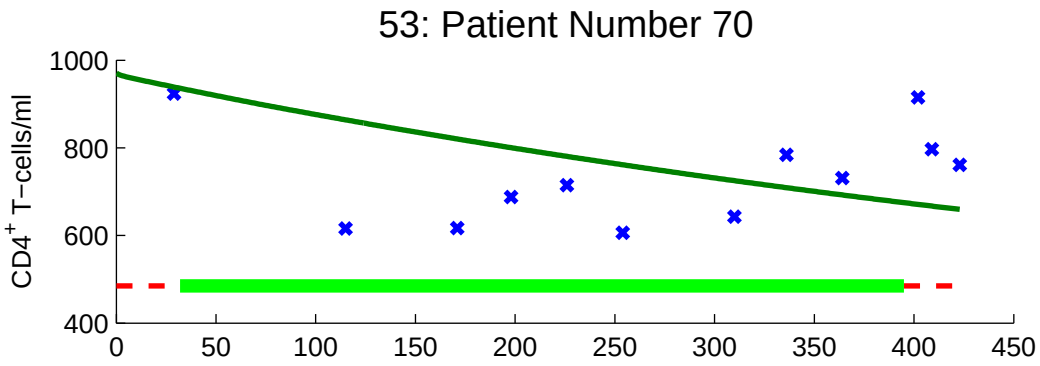

47: Patient Number 61
CD4+ T−cells/ml
1000
800
600
400
200
0
100
200
300
400
500
600
700
800
Viral RNA copies/ml
10^6
10^4
10^2
10^0
time (days)

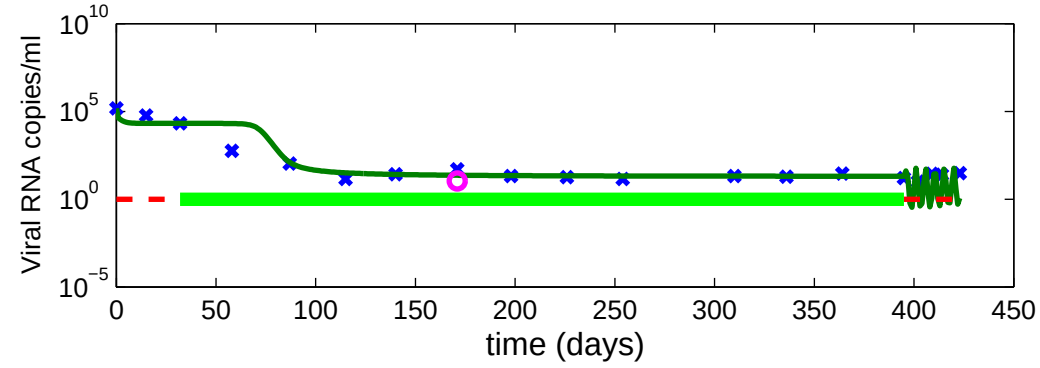

48: Patient Number 63
CD4+ T−cells/ml
600
400
200
0
100
200
300
400
500
600
700
800
900
Viral RNA copies/ml
10^5
10^0
10^−5
time (days)

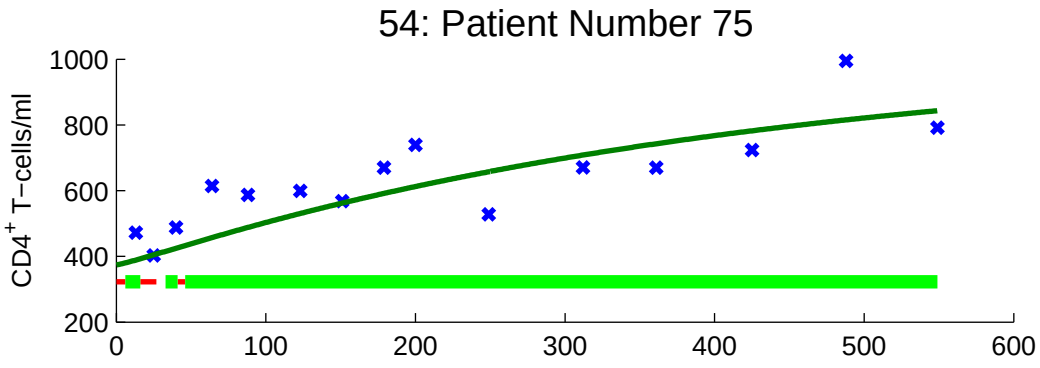

49: Patient Number 64
CD4+ T−cells/ml
Viral RNA copies/ml
time (days)
50: Patient Number 65
CD4+ T−cells/ml
Viral RNA copies/ml
time (days)

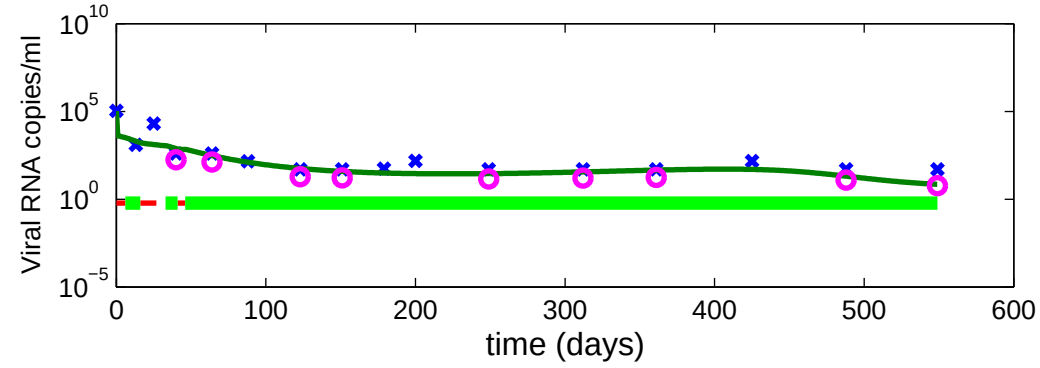

51: Patient Number 66
CD4+ T−cells/ml
800
600
400
200
0
0
100
200
300
400
500
600
Viral RNA copies/ml
10^10
10^5
10^0
10^−5
0
100
200
300
400
500
600
time (days)

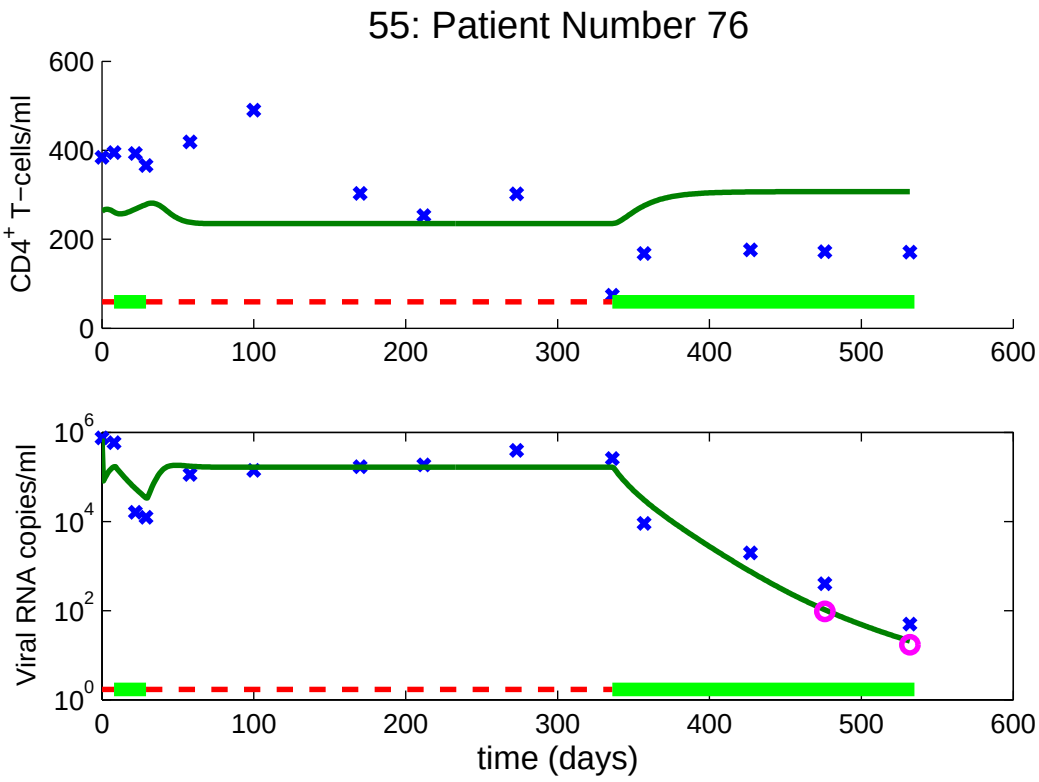

52: Patient Number 69
CD4+ T−cells/ml
800
600
400
200
0
50
100
150
200
250
300
350
400
450
Viral RNA copies/ml
10^5
10^0
0
50
100
150
200
250
300
350
400
450
time (days)

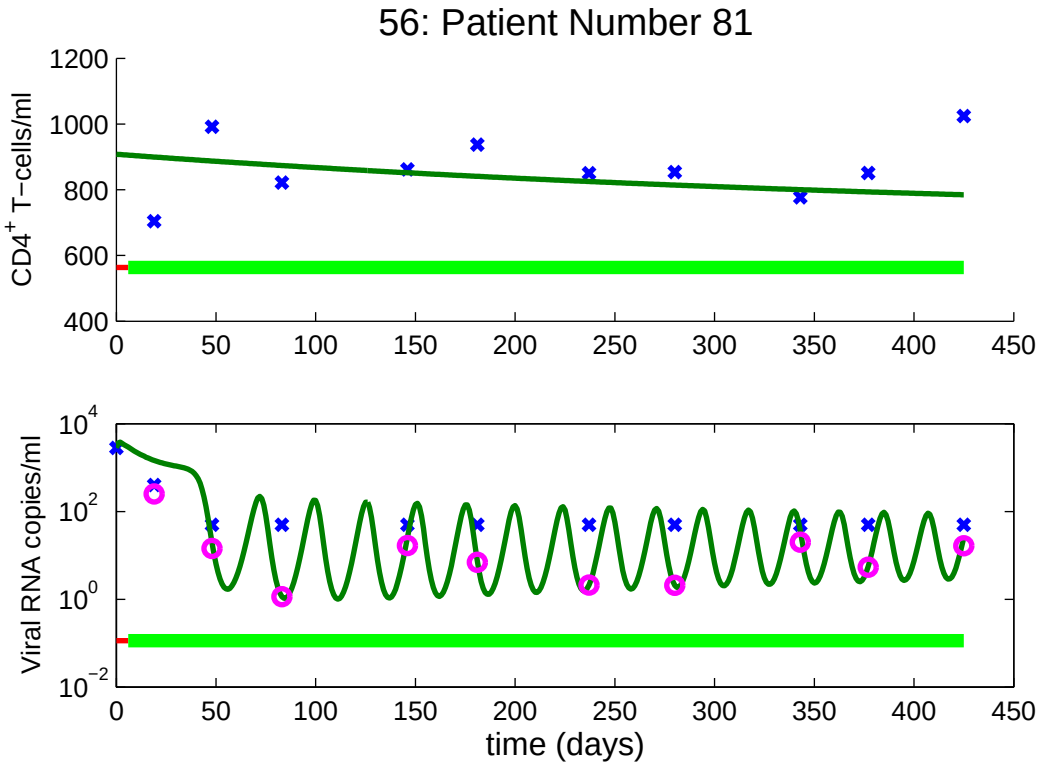

53: Patient Number 70
CD4+ T−cells/ml
1000
800
600
400
0
50
100
150
200
250
300
350
400
450

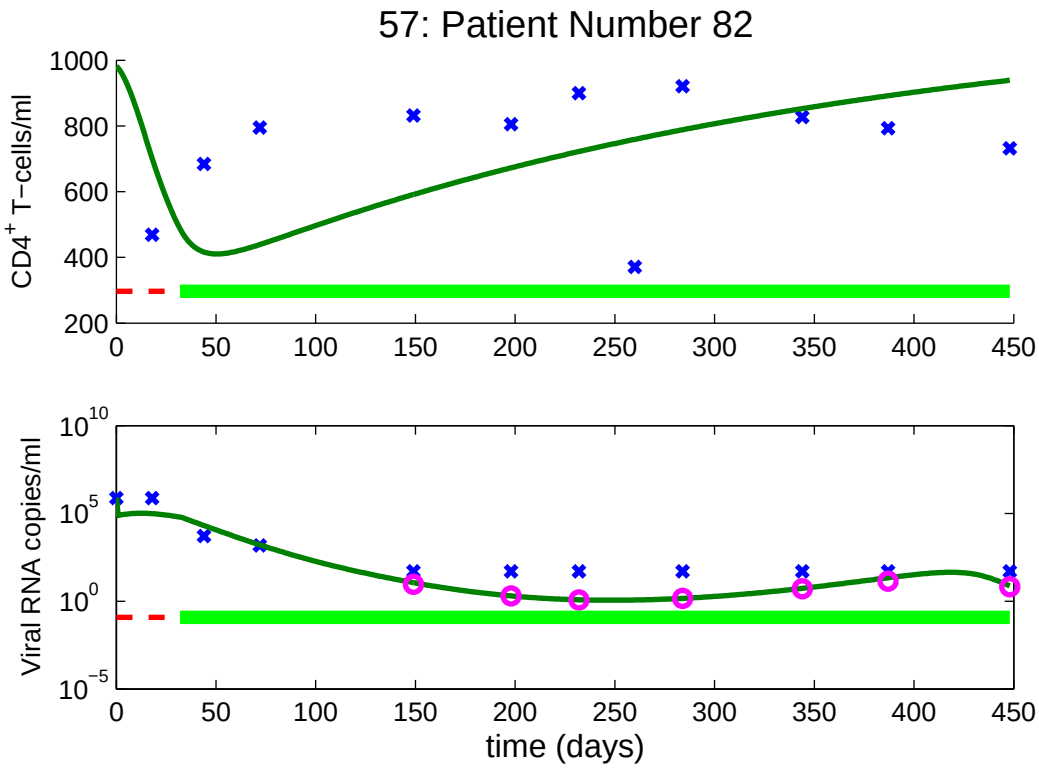

Viral RNA copies/ml
10^10
10^5
10^0
10^−5
0
50
100
150
200
250
300
350
400
450
time (days)

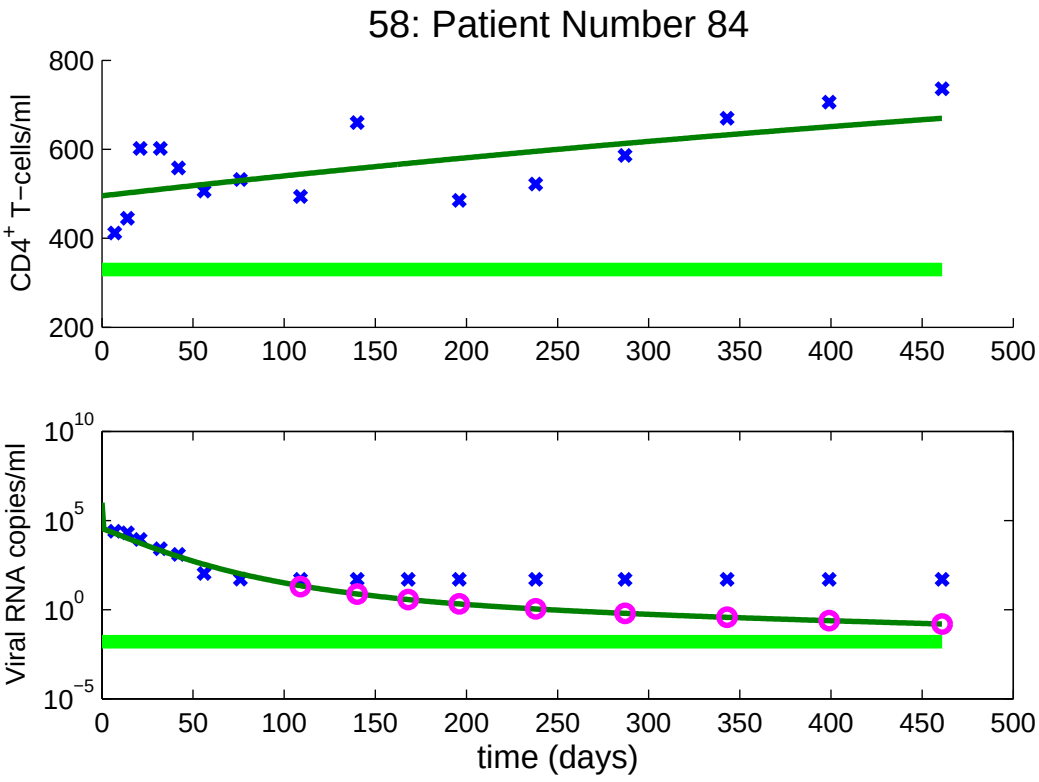

54: Patient Number 75
CD4+ T−cells/ml
1000
800
600
400
200
0
100
200
300
400
500
600

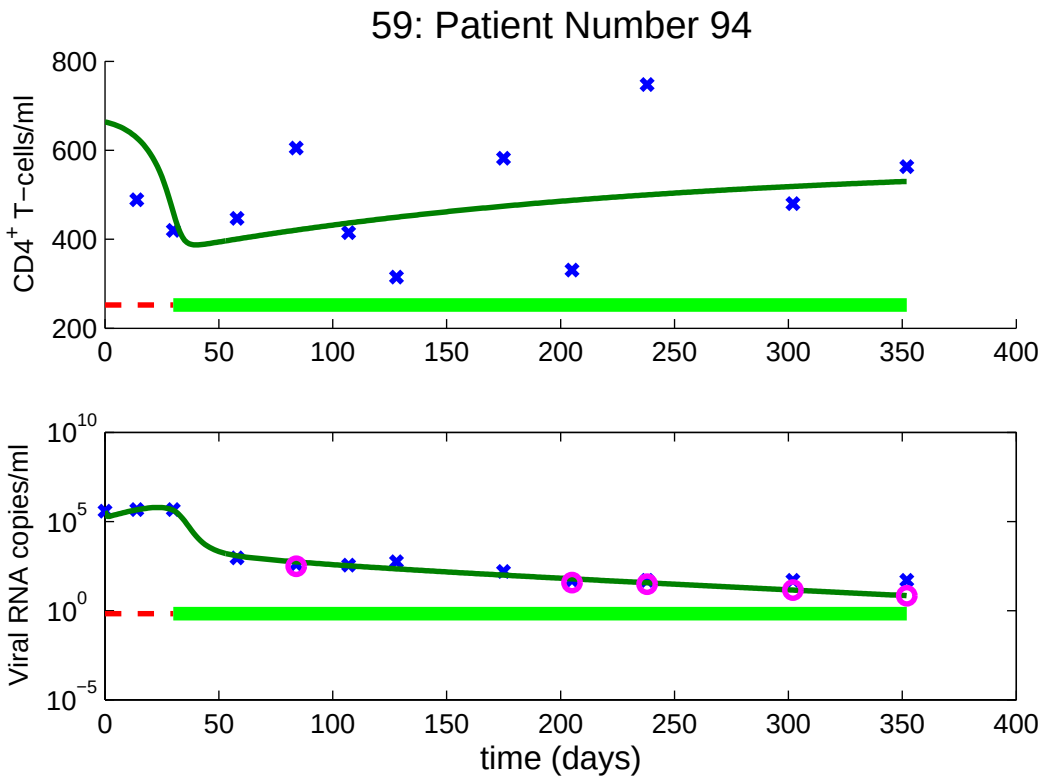

Viral RNA copies/ml
10^10
10^5
10^0
10^−5
0
100
200
300
400
500
600
time (days)


55: Patient Number 76
CD4+ T−cells/ml
Viral RNA copies/ml
time (days)


56: Patient Number 81
CD4+ T−cells/ml
Viral RNA copies/ml
time (days)


57: Patient Number 82
CD4+ T−cells/ml
1000
800
600
400
200
0
50
100
150
200
250
300
350
400
450
Viral RNA copies/ml
10^10
10^5
10^0
10^−5
time (days)


58: Patient Number 84
CD4+ T−cells/ml
800
600
400
200
0
50
100
150
200
250
300
350
400
450
500
Viral RNA copies/ml
10^10
10^5
10^0
10^−5
time (days)


59: Patient Number 94
CD4+ T−cells/ml
800
600
400
200
0
50
100
150
200
250
300
350
400
Viral RNA copies/ml
10^10
10^5
10^0
10^−5
time (days)