DYNAMICS OF NAÏVE AND MEMORY CD4 T-CELLS IN CHRONICALLY INFECTED HIV PATIENTS POST-INJECTION OF CCR5-DOWN-MODULATED MEMORY CD4 T-CELLS

ANGIE RAAD

A THESIS SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS
FOR THE DEGREE OF

MASTERS OF SCIENCE

GRADUATE PROGRAM IN DEPARTMENT OF MATHEMATICS AND STATISTICS
YORK UNIVERSITY
TORONTO, ONTARIO
AUGUST 2016

Abstract

HIV/AIDS, a sexually transmitted diseases continues to affect the lives of millions of individuals worldwide. This retrovirus targets CD4 T-cell populations, the main driver of the immune system by using the chemokine co-receptor 5 (CCR5). Despite the success of the highly active antiretroviral therapy in reconstituting the immune system, HIV infected individuals still suffer from low CD4 T-cell counts. Recently, researchers were able to highlight the success of immunotherapy in restoring the CD4 T-cell count. To further, investigate such importance, our collaborators at case Western University injected CCR5-down-modulated memory CD4 T-cells into 9 chronically infected HIV patients. Using a linear transitions from the naive to the effector memory state, a non linear ordinary differential equation model was used to model the experiment. Various data fitting techniques in Matlab Stan and Monolix software were used to estimate the model parameters (proliferation, death, transition and birth rates) before and after the initiation of the treatment to study the change of the cell dynamics. Our fittings have indicated an increase in the memory stem

and naïve cell lifespan post-clinical trial. Using sensitivity analysis, we showed that the naïve cell birth rate from the thymus λ, the memory stem cell proliferation rate p_{ST} and the central memory cell death rate d_C played an important role in restoring the CD4 T-cell count. A stochastic model for the CD4 T-cells population was developed to examine if fluctuations from the stochastic simulation were able to capture the experimental data measurements. The findings of this study indicates the importance of looking further into how modified CD4 T-cells are able to restore the T-cell counts which thereby decrease the HIV virus pool and help HIV patients to maintain a low level of the virus and most importantly a high level of T-cell count.

Dedication

To my family; Hanaa, Ornella and Tony

Acknowledgements

I would like to express my great appreciation to everyone that has helped to make this work possible.

First, I would like to express my very great appreciation to my supervisor Dr. Jane Heffernan. I am grateful for her patience, motivation, support and her willingness to spend the needed time throughout the past two years of my studies to make this work possible. I was very fortunate to be one of her graduate student and be able to learn a great amount of knowledge in disease modeling and to learn the essential skills to become a better researcher and writer in the mathematical modeling field. The opportunities she gave me through attending conferences and summer schools was a valuable experience in helping me to get introduced to various research. Last but not the least, I would also like to thank her for creating a healthy environment to complete this work.

I would also like to thank my co-supervisor Dr. Georges Monette for his patience in

teaching me data fitting, a topic that was totally new to me when I first started this project. His willingness to spend the needed amount of time to accomplish the data fitting task was a great blessing.

I would also like to extend my thanks to all of my committee members Dr. Xin gao, Dr. Seyed Moghadas and Dr Yi Sheng for the time they made to read my thesis, attend my talks and provide me with their valuable comments.

My sincere thanks also goes to our collaborators Dr. Rafick Sekaly and his Lab at Case Western University for sharing with us the data and allowing us to present it in this thesis.

Finally, I would like to thank my family for their great support and encouragement throughout the last two years.

Table of Contents

List of Tables

List of Figures

Abbreviations

ODE	Ordinary Differential Equations
NLMEM	Non Linear Mixed effect Model
AIC	Akiake Information Criterion
BIC	Bayesian Information Criterion
-2LL	Log-likelihood
HIV	Human Immunodeficiency Virus
S.E	Standard Error
CPU	Central processing unit
AIDS	Acquired Immunodeficiency Syndrome
HAART	Highly active antiretroviral therapy
CCR5	C-C chemokine receptor type 5
N	Naive CD4 T-cell
MN	Modified Naive CD4 T-cell
CM	Central Memory CD4 T-cell

MCM	Modified Central Memory CD4 T-cell
TM	Transitional Memory CD4 T-cell
MTM	Modified Transitional Memory CD4 T-cell
EM	Effector Memory CD4 T-cell
MEM	Modified Effector Memory CD4 T-cell
SD	Standard deviation
SEM	Standard Error Over Mean
MCMC	Markov Chain Monte Carlo
STD	Sexually Transmitted Disease
PRCC	Partial Rank Correlation Coefficient
LHS	Latin Hypercube Sampling

1 Introduction

Mathematical Modeling of infectious disease has become a growing area in the past two centuries. This is because infectious diseases are the number one cause of human morbidity in the world [9]. This discipline has helped to gain insights about the dynamics of infectious diseases on both population and in-host levels. Some of the most important advantages of using mathematical models, its ability to analyze and unfold unobserved dynamics experimentally.

Mathematical immunology studies the spread of the disease in the host by analyzing the interaction between the immune cells and the pathogens. This field stemmed in the early 1980's with the emergence of the Human Immunodeficiency Virus(HIV) pandemic [15]. One of the primary reasons for the success in this area is the ability for researchers to collect experimental data and use it in the mathematical models to draw important conclusions about the dynamics of the immune system, the virus and the effect of vaccines and other drug treatments. Many great mathematicians have elaborated on this area such as Perleson, Nowak and May and others [16,17,18].

These models were mainly used to study the dynamics of the HIV and the immune system in the presence of the virus in the body. Some of the most remarkable findings that corrected our understanding about the HIV dynamics in the body of chronically infected individuals were achieved by Georges et al [17] and Perleson et al in 1995 [30]. Using experimental data measurements, they were able to show that in the chronic infection stage, the CD4 T-cell count is constant as a result of an interesting immune system dynamics where millions of the CD4 T-cells are being killed but replaced daily by the thymus. However, a great amount of mathematicians have and are applying these models to other diseases but not limited to: measles [19], hepatitis B and C [21], flu [22] and HPV [20]. Besides studying disease dynamics, mathematical immunology was used to study B and T-cells dynamics[23]. Most of mathematical models used in epidemiology and immunology consisted of coupled ordinary differential equations (ODEs). Despite the success of using ODE models, many researchers have started using stochastic models due to its ability to account for stochasticity in the disease dynamic unlike the deterministic models [14]. As mentioned above, mathematical models have shed light on various processes in HIV infection, including effective drug therapy regimens [25], activation of the immune system [28], and latently infected cells [27]. Some researchers have focused on studying the dynamic of memory CD4 T-cells in HIV infected individuals [24,26]. A recent clinical study by Sekaly et al at Case Western has determined that

introducing a dose of CCR5-down- modulated memory CD4 T-cells is able to induce the activity of the immune cell in chronically infected HIV patients. Hence the work of this thesis will focus on modeling this experimental trial in order to quantify the resulting immune system dynamics after the introduction of such perturbation to the CD4 T-cell population.

In the sections below, I will elaborate more about the immune system, disease, its structure, transmission and progression to AIDS. In addition, I will present the previous models used to study the HIV-immune system dynamics. Deterministic and stochastic models are both considered. Later, I will describe the clinical trial performed by our collaborators (Sekaly et al., 2013). Lastly, I will describe the longitudinal data of the three cohorts that was obtained from the study [52].

1.1 Basic Facts About the Immune System

The immune system is among one of the most important systems in the human bodies that sustains survival [5]. The understanding of its function began in the 19th century and up to date some aspects of the immune system remains unclear [5]. The immune system consists of a variety of cells, organs and tissues that work collectively in order to fight any foreign organisms such as bacteria, viruses and fungus that poses a threat to the host. The immune system distinct any foreign organisms by identifying two patterns called Danger or Pathogen -Associated- molecular patterns,

also known as DAMPs and PAMPs [5]. It is composed of two major responses; the Innate and Adaptive response. These responses are distinguished based on the components, activity, response time, duration and the ability to build memory against a specific pathogen. As the Adaptive immune system is our main focus in this paper, we will not be discussing the details about the Innate immune response. Briefly, the innate immune response is mainly composed of macrophages, white blood cells and natural killer cells. The innate immune responses is fast and unspecific as these cells act directly on any pathogen once it is recognized. It is activated within hours of the foreign pathogen's discovery [1,5].

In many cases the innate immune response is enough to contain an infection, however in some cases it gets overwhelmed by the rapid replication of some viruses as in the case of HIV infection. In this case the acquired immune system gets activated. It is known as acquired because this immunity is built from previous exposure to pathogens (bacteria and viruses), vaccinations or maternal immunity. This response could take days or weeks to be activated and it is known to be specific, where the presence of Antigen Presenting cells (APCs) is required for activation. There exist two types of adaptive immune systems responses; The humoral and cell-mediated responses. Both are carried out by two classes of white blood cells also known as lymphocytes [1]. The mediated immune system is composed of T-cells, which are produced in the bone marrow and is responsible for killing infected cells. However

these cells mature in the Thymus either to contribute to lymphocyte maturation or kill infected cells [1]. There exist two types of T-cells; CD4 or helper T-cell and CD8 or cytotoxic lymphocyte T-cells, depending on the receptor present on the cell surface. Once activated by the antigen presented by the MHCII on the macrophage cell surface, the helper T-cells are activated and produce cytokines. The cytokines activate the B-cells and the production of CD8 T-cells that are responsible in killing infected cells. The main focus of this thesis are the naïve and memory CD4 T-cell population.

B-cells are produced and mature in the bone marrow. These cells compose the humoral immune response where they bind to antigen to produce the right antibodies. Later, these antibodies are used to kill free viruses and bacteria. The activation process of the B and T-cells is illustrated in Figure 1.1.

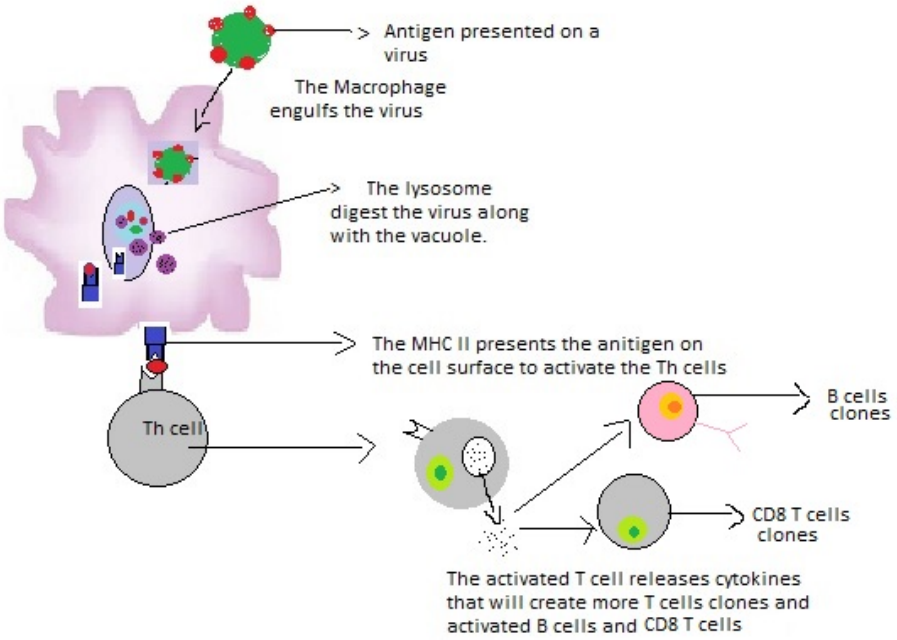

Figure 1.1: Activation process of B and T- cells.

In this figure the activation process of the T and B cells upon infection is illustrated. First the macrophage engulfs the virus where the MHCII presents the antigen on the macrophage cell surface. The naïve helper T-cells receptors binds to the MHCII-Antigen complex. This step activates the helper T-cells and cytokines are secreted. The secretion of the cytokines activates the B cells and produces CD8 or cytotoxic T-cell. The immune system is then activated.

1.1.1 CD4 T-cells

One of the most important feature of the immune system is the ability to develop memory against any pathogen. They are shown to have an imperative role in the adaptive immune response to infectious disease [7]. These memory cells have been the main focus of researchers when designing vaccines [3]. A study conducted by (STEVEN M. SCHNITTMAN et al., 1990) showed the HIV-1 virus preferentially infect memory CD4 T-cells subsets.

CD4 T-cells are divided into two kind of cells: memory and naïve [28]. After being exposed to antigen, the naïve T-cells proliferate and differentiate to memory CD4 T-cells if able to survive the contraction phase [7]. Memory CD4 T-cells are composed of 4 subsets, where the naïve CD4 T-cells transition to the memory effector state [6]. It is still not clear how the transition from the naïve to the memory effector state occurs in the body. However, in this paper will follow the model proposed by (Mahnke et al., 2013) and our experimental collaborator (Sekaly et al.) where a linear transition from the naïve state to effector memory state is considered as illustrated in Figure 1.2 below.

- Naïve cells: These cells are produced by the thymus. No Antigen and marker of cellular, activation is expressed on the cell surface [5,6].

- Memory stem cells: It is the first subsets of memory T-cells that expresses stem like property with having a naïve phenotype. These cells cannot be made by any other memory T cell subset [3]. It was found that these cells have a higher rate of survival and proliferation compared to the naïve and other subsets of memory T-cells [7].

- Central memory cells: These cells are produced by the stem memory cells. They circulate primarily between the blood and lymphs.

- Transitional memory cells: It was shown to be located in the peripheral blood. These cells have a higher proliferation rate than TCM cells [6].

- Effector memory cells: It was shown that these cells are short lived[8]. They express effector phenotype such as cytokines secretion faster than the other CD4 memory subsets. They move from the blood to peripheral tissues [8].

Figure 1.2: The Transition of CD4 T-cells from naïve to effector memory state.

In this figure the transition from naïve to effector memory cells is illustrated. Where the expression of CCR5 increases gradually moving from the naïve to the effector state [7].

1.2 Human Immunodeficiency Virus-HIV

1.2.1 Pandemic

HIV, is a sexually transmitted disease. The virus is classified as a Retrovirus and belongs to the family of lentivirinae that is characterized by its long incubation period. There exist two types of HIV : HIV-1 and HIV-2. Even though both of these types will progress to the Acquired immunodeficiency syndrome(AIDS) stage, each has its own origin, biological and molecular characterizations [4]. The first HIV condition reported to the public in 1980,in New York and San Francisco in the homosexual community where rare cancer cases such as Kaposi's sacroma were observed among young individuals. Few months later, the disease was reported in the intravenous drug users communities as well as in hemophiliacs and heterosexuals partners[4]. Soon HIV became a serious epidemic. Since the discovery of the first case, at the end of 2014 approximately 36.9 million individuals are currently living with HIV [2]. According to the WHO, since its discovery, AIDS-related disease have caused the death of over 34 million people up to date[2]. The majority of the affected individuals are from low income countries, where Sub-Saharan Africa accounts for more than 70 percent of the cases as illustrated in Figure 1.3 [2]. Since the discovery of the disease, a tremendous amount of research have focused on better understanding the emergence, spread of the disease and most importantly

its unique pathogencity with the intention of developing a potential vaccine or a cure.

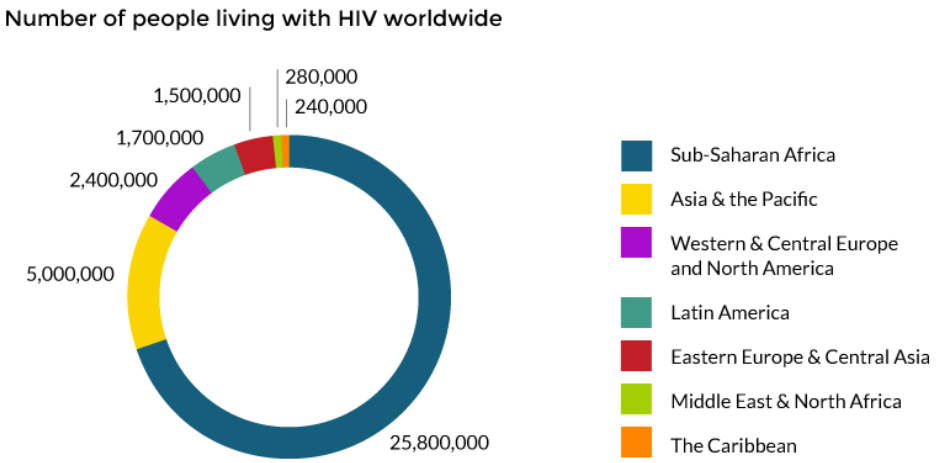

Figure 1.3: Number of people Living with HIV by Region in 2014

This figure illustrates the proportion of reported HIV positive cases worldwide per region. It is evident that Sub Saharan Africa account for more than 70 % of the cases in the world. This figure was adapted from avert 2014 statistics, [32].

1.2.2 Transmission and Progression

HIV transmission is sensitive to the amount of free virions and the degree of contact upon exposure. It has been shown that the virus has a high concentration of free virons in the blood stream and the genital fluids. Therefore, HIV can only be transmitted by the exchange of blood or body fluids such as semen and vaginal fluids from an infected individual. Sexual activities account for almost 75 percent of all new HIV cases [4]. In addition, HIV could be passed on by the use of unsterile needles in cases of blood transfusion or drug users. A mother could pass HIV to her fetus/baby during or after delivery as well as during breast feeding. However the degree of transmission depends on the stage of the infection. Early HIV phase yields a higher concentration of free virons in the breast milk which leads to a higher chance of transmission [4]. As the saliva contains glycoproteins and fibronectins it is thought to inhibit the cell to cell transfer of virus, saliva, tears sweat and faeces have low virions level and cannot be a mode of HIV transmission [4].

1.2.3 Structure

HIV consists of an outer membrane which consists of two layers of lipids proteins; gp120 and gp41, these lipid proteins are uniformly arranged into 72 knobs [4]. The

Glycoprotein 120 is located on the outer membrane of the virus and gp41 is embedded in the lipid matrix. As HIV is a retrovirus, it stores its two copies of ribonucleic acid, also known as RNA, in the inner core used to encode the necessary viral proteins for maturation. Along with the RNA, it contains three enzymes; reverse transcriptase, integrase and protease which plays a primary role in the viral replication and maturation process. In addition, it contains the protein of the last host cell that has infected previously[4]. Figure 1.4 illustrates in more details the structure of the virus.

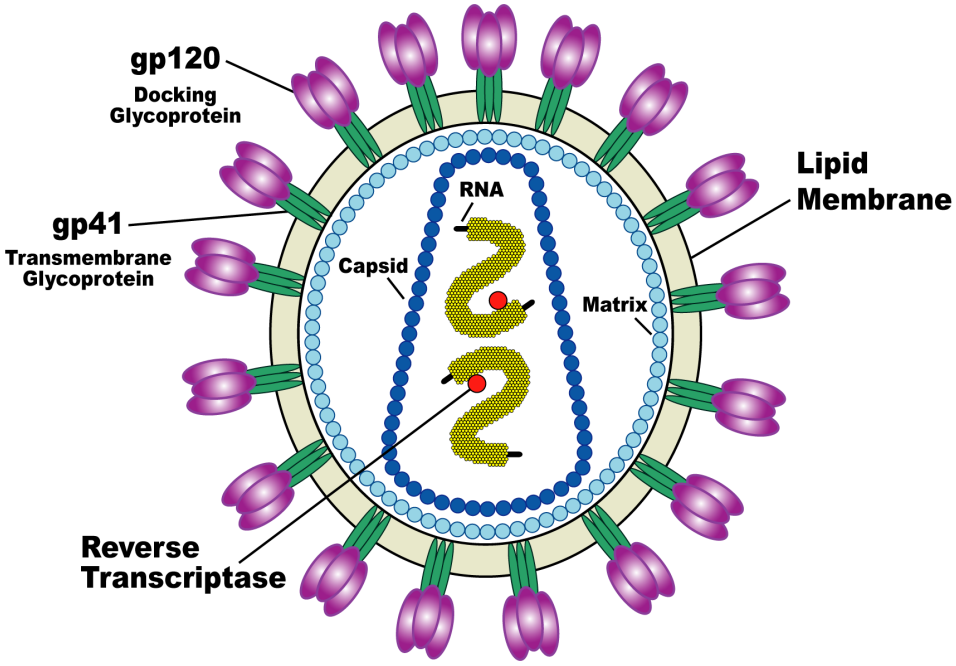

Figure 1.4: Structure of the HIV.

HIV is a retrovirus where it contains 2 copies of RNA and 3 essential enzymes. An outer membrane consisting of gp120 and gp 41. This virus uses the host cells such as CD4 T-cell to replicate. This Figure was adapted from [34].

1.2.4 Life-Cycle

Like any other virus, HIV needs a host cell in order to replicate. When HIV succeed to enter the human body, its main target is to replicate by infecting one kind of the immune system cells that has CD4 receptors on the surface. Figure 1.5 illustrates the 3 major steps of the replication process.

(1) Binding/ Fusion with the CD4-T cells

The first stage of the process is when the HIV binds itself to the CD4 T cells. This is accomplished by the binding of the gp120 glycoprotein to the CD4 receptors. Later, the transmembrane gp41 binds to the co-receptor CCR5 or CCRX4 present on the CD4-T cells. This binding causes a conformational change which allows the fusion of the cell-virus membranes. Once it is fused, the viral nucleocapside enters the cells and releases the two copies of the RNA along with three enzymes essential for the viral replication.

(2) Transcription and Translation

Once inside the cell, the reverse transcriptase begins the transcription of the single stranded viral RNA into a double helix DNA. Later, the viral DNA enters the host cell's nucleus where it is integrated to its DNA by the integrase enzyme. Now the virus genetic material are embedded in the CD4 T- cells DNA. Two possible events could occur. First, if the infected CD4 T-cell is activated, proviral DNA will be

transcribed into messenger RNA. Then the mRNA moves to the cytoplasm where it is translated into essential viral proteins. Otherwise, if not activated, the infected cells remains latent.

(3) Maturation and Budding

Once the mRNA is fully translated into viral proteins, the protease enzymes cleaves the long strains of proteins. This is an essential step for the HIV maturation as some of these proteins becomes enzymes and others turn into structural elements. Once the viral elements are assembled, it buds off the cell and a new virus is created and ready to infect other CD4 T-cells.

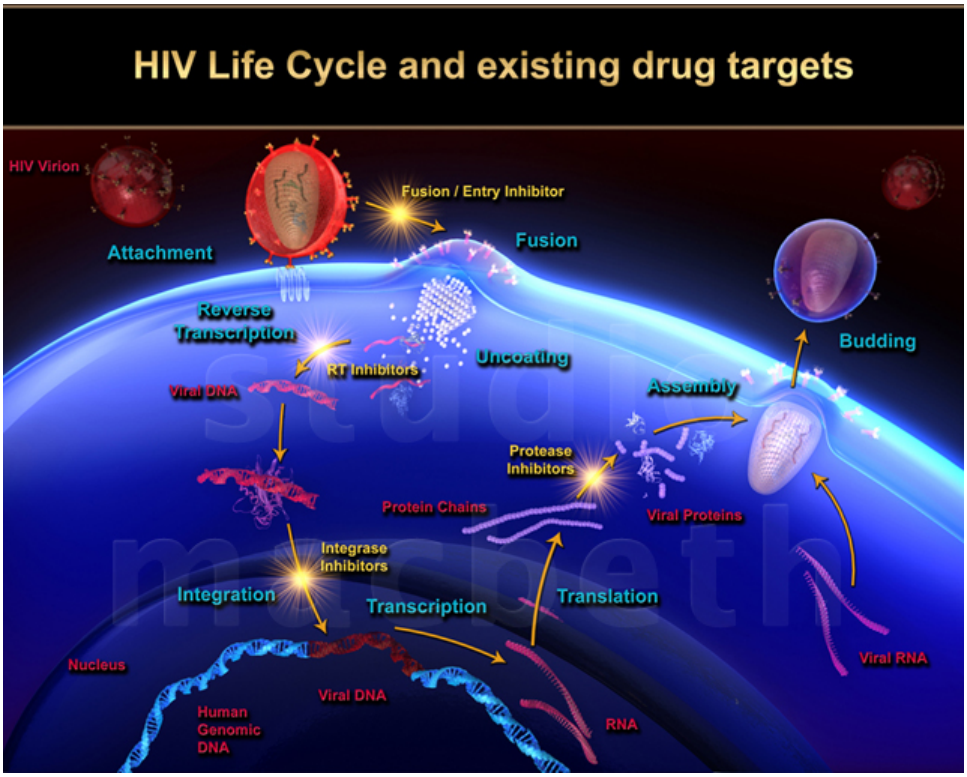

Figure 1.5: HIV Viral Life-cycle.

HIV viral Life-cycle is divided into three major steps: First the binding and fusing of the HIV and CD4 T-cell occurs. Once inside the virus begins its transcription and translation process to embed a copy of the viral DNA into the host DNA. Once the mRNA is translated, the virus buds and a new mature virus is released.This figure was adapted from [57].

1.2.5 Treatment

Since the emergence of HIV, various drugs have been introduced to target several important steps in the virus infection life-cycle. These drugs are used to slow down the progression to AIDS by making the viral load low enough to be considered undetectable. As shown in Figure 1.6, the number of deaths associated with HIV has decreased upon the introduction of the drugs in the early 1990's. The decrease in the number of death is associated with the increase of the number of people living with HIV. However, these drugs are not equally available to all individuals especially those infected in third world countries such as Africa. Below is the list of different classes of antiretroviral drugs:

- Protease Inhibitor, also known as PI's. This class of drugs inhibits the function of the protease enzyme. This inhibition prevents the budding of the newly made virions.
- Integrase Inhibitor, inhibits the function of the integrase enzyme so the viral DNA is not integrated into the host cell's DNA.
- Fusion inhibitor, these medication inhibits the the fusion between the virus and the CD4 T-cells. However, this class of drugs have proven to be not very effective in blocking the CD4 T-cell and preventing fusion.

- Reverse transcriptase Inhibitor, prevents the viral RNA from being transcribed into DNA. This will later stop the lifecycle as the virus does not have the DNA to incorporate in the host cell and begin the replication process.

HAART (Highly Active Anti-Retroviral Therapy), is a cocktail of 3 or more of the drug classes mentioned above. The main reason behind this cocktail is to reduce the emergence of mutated HIV types that are resistant to the drugs. Despite the effort of the drugs, mutations are highly likely to occur in the infection lifetime of a patient. However, we will not be considering the possible mutations throughout this work.

As this is a lifetime therapy, these drugs have various health effects on the human body, which leaves the patients in poor health conditions at times. Hence in this work we are trying to understand the effectiveness of using the immunotherapy approach in order to reconstitute the total CD4 T-cell count.

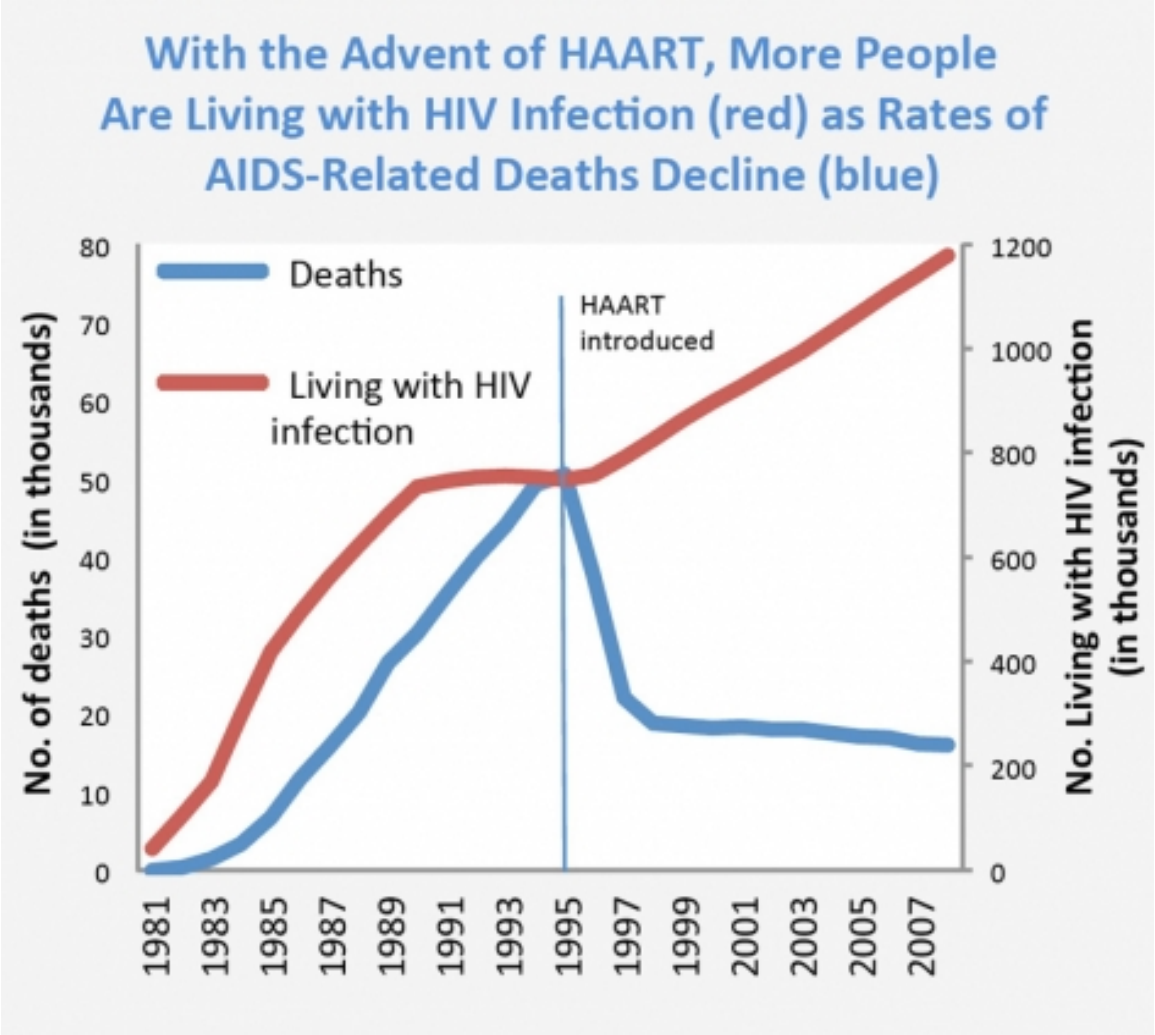

Figure 1.6: Number of Death and HIV positive individual in US from the period before and After the introduction of HAART.

The number of death and poeple living with HIV in US from 1981-2007. In this graph it is shows that when the HAART was introduced the number of death decreased and the number of people living with HIV increased as a result. This figure was adapted from [35].

1.3 Modeling HIV and Immune system

Since the discovery of the disease, a tremendous amount of research have focused on better understanding the emergence, spread of the disease and most importantly its unique pathogencity with the intention of developing a potential vaccine or a cure. Over the past two decades, thousands of studies were published about HIV/AIDS[29]. Many of those studies used statistical, deterministic and stochastic mathematical models to study the disease dynamics [29]. These models have made great advancements in better understanding epidemiological and immunological aspects of the virus. Some of the most important early works were done by A. Perelson and D. Ho. They used clinical data from HIV positive individuals to study the dynamics of HIV and T-cells in the absence and presence of antiviral drug therapies [30,31].

These studies have mainly focused on using ordinary differential equations to model such dynamics. These compartmental models are used to examine the interactions between several classes of populations to study and predict the changes in each of the population size over a certain time period. This is accomplished by representing the rate of change in these populations using ODEs. Any of the population size could be estimated by solving these differential equations at a specific time point. The basic ODE model for describing the HIV-immune system interaction is composed

of three compartments; the uninfected T-cell, the infected T-cells and the free HIV virions particles. This model was used in several papers such as [30,31,25]. Using the basic model, researchers were able to draw some important conclusions about the viral clearance rate, the lifespan of the infected T-cells and the viral production rate from each infected cell [38,39]. The Basic model is illustrated in the system of ODEs Eq's 1.1, where λ is the birth rate of uninfected cell, d and a are the death rates of the infected and uninfected immune cells respectively, k is the virus bud rate and u is the virus clearing rate.

$$x' = \lambda - dx - \beta xy \tag{1.1}$$

$$y' = \beta xy - ay$$

$$v' = ky - uv$$

Still the simplicity of this basic model was not enough to study various important dynamics such as the HIV mutation strains, the effect of the antiretroviral drug and so on, so extensions to the basic model are formulated. These models were able to better understand the role of the antiretroviral drug in blocking and slowing down the HIV infection process [17,18]. This was achieved by adding extra parameters ϵ_1 and ϵ to the basic model to illustrate the inhibition of both the reverse transcriptase and protease inhibitors correspondingly. A factor of (1-ϵ_1) and (1-ϵ) were multiplied

to each of the β and k parameters respectively.

When a patient is on the antiretroviral cocktail therapy, the HIV viral load is undetectable which could be ignored in the basic model. This assumption will allow us to incorporate the dynamics of various T-cell subsets such as naïve and memory as shown in [24,26].

1.3.1 Stochastic Modeling

Despite the simplicity and elegance of the ODE models, we cannot ignore the fact that the human cell dynamics are not deterministic and are subject to random fluctuations due to natural variability. Hence various mathematical modelers have switched gears and started using stochastic models to describe the HIV-immune system dynamics. One of the earliest stochastic models was developed by Merill's et al, where he was able to model aspects of the immune system response in the presence of HIV using a branching process [40]. In addition, Perelson et al [39] were able to estimate the probability that one virus could infect on average one CD4 T-cell. However, there aren't many studies that modeled the dynamics of naïve and memory CD4 T-cells subsets of chronically infected HIV patients. In this thesis, we will be developing a stochastic model for the naïve and memory CD4 T-cells so we are able to observe the variability in this model that can be a result from the natural variability, experimental error measurements and uncertainty in the

parameter estimations of the T-cell dynamics.

1.4 Experiment

Highly active antiretroviral therapy (HAART), does a fairly good job in keeping a low HIV viral load in the body of infected individuals. Despite the great effort from scientists to help in restoring the immune system in HIV positive individuals, patients remain to have a low count of T-cells in the body.

As we have seen in the previous section, CCR5 is a major co-receptor for the entry of the HIV virus into the CD4 T-cells. Our Collaborator, (Sekaly et al., 2013) at Case Western, have designed an experiment to study the effects of the introduction of CCR5-down-modulated memory CD4 T-cells into HIV positive patients. The main intention of this experiment is to examine if the perturbation of the system will result in any augmentation in the CD4 memory T-cell count and the activity of the immune system.

Information about the 3 cohorts

In this study 9 chronically infected HIV patients participated in this study. Where all of these patients were receiving a HAART drug therapy. Table 1 summarize the infection state, age , ethnicity and drug dose infused. The mean age of the patients was 49 ± 6.49 years and the mean period of the infection diagnosis is 20.77 ± 6.47 years.

Injection of CCR5-down-modulated memory CD4 T-cells and its quantification.

As illustrated in Table 1.1, the nine HIV patients were separated into three groups based on the injected dose of zinc finger nuclease (ZFN) driven CCR5-disrupted CD4 T-cells (SB-728-T). In order to measure the cells count for the 5 T-cell populations, samples of Peripheral blood mononuclear cells (PBMCs) were collected from each of the 9 patients at several time intervals. These PBMC were transduced by an adenovirus encompassing a Zn Finger endonucleae that targets the CCR5 gene. To identify each of the memory T-cell subsets, PBMCs were stained by different cytometry panels(Sekaly et al., 2013). In this work, we defined the total CD4 T-cells as the sum of the the five T-cell subsets unlike how it is quantified in the experiment. The distribution of the injected CCR5-down-modulated memory CD4 T-cells is illustrated in Figure 1.7 below.

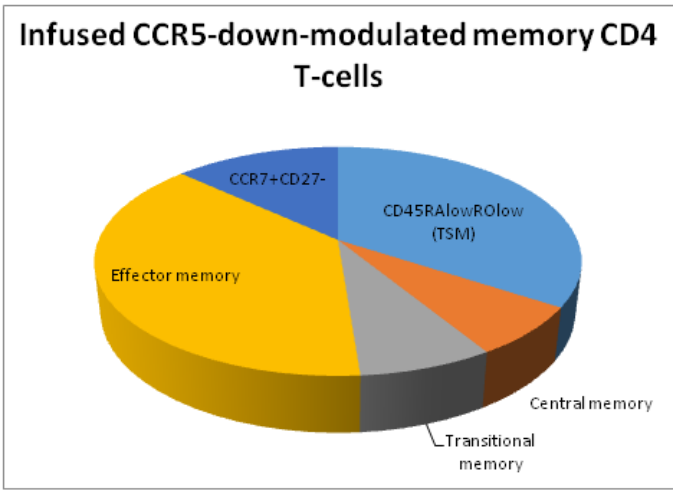

Figure 1.7: Ratio of infused CCR5-down-modulated- CD4 T-cell subsets.

In here the CD45RAlowROlow are considered as memory stem cells. These ratios were adapted from the experimental study report by (Sekaly et al., 2013).

ID	Period of Inf.(years)	Dose infused	CD4 T-cell count at M0	CD4 T-cell count at M36
103	20	$1x10^{10}$	188 -134.4	315-262.1
104	21	$1x10^{10}$	261-205.4	455-378.7
102	25	$1x10^{10}$	439-349	518-428.1
203	21	$2x10^{10}$	294-227.7	617-4628
302	19	$2x10^{10}$	413-279.6	606-468.7
201	30	$2x10^{10}$	525-354	848-651.4
304	13	$3x10^{10}$	306-215.6	757-444
303	28	$3x10^{10}$	330-211.8	525-391.5
305	10	$3x10^{10}$	480-347	340-271.8

Table 1.1: Information about the three cohorts.

This table illustrates the dose infused, infection period and the CD4 T-cell count at baseline and after three years. The highlighted numbers are the T-cell count defined by our collaborators Sekaly et al. and the non highlighted ones are simply defined as the sum of the naïve and memory CD4 T-cells, which is what is used throughout this work.

1.4.1 Data Description

The analysis in this thesis uses an unbalanced longitudinal data obtained from the clinical described above. Longitudinal data is defined as observations usually taken from several individuals at various time points. At time t=0, the CCR5-down-modulated memory CD4 T-cells were injected into the three cohorts with different dosages as illustrated in Table 1.1. After the injection, the nine patients were followed over three years and measurements of the naïve and memory CD4 T-cells and their CCR5-down-modulated versions were taken at different time intervals for each of the patients. Looking at the spaghetti plots in Figures 1.8 and 1.9, we can make the following observations about the collected data :

- In total there are approximately 450 observations, where some of the missing points could be due to patient inability to respect the follow up appointment, or experimental error measurements.
- The down-modulated CCR5 CD4 T-cells have a very low count.
- The modified naïve CD4 T-cell sub-population data is sparse.
- Memory stem cell populations measurement are missing as these were not determined from the total T-cell count in the laboratory until much later in the experiment.

- The naïve and central memory cell populations seem to increase in count in most of the nine patients from its baseline value.
- The effector and transitional memory cells seem to increase or stabilize at a value close to its baseline in most of the nine participants.
- The total number of the CD4 T-cell in here is not defined as the sum of all the CD4 T-cell sub-populations due to the strict definition used in the experiment protocol used in the laboratory.

When working with experimental data, it is important to draw some general observations before starting any analyses. These observations are usually important in understanding what fitting routines are more suitable and why some would fail in quantifying some parameters.

Figure 1.8 below presents the raw experimental data for each of the three cohorts: Low (103, 104, 102), medium (203,201,302). Figure 1.9 illustrates the high dose cohort(304,305,303). In all of the three cohorts, the different cell counts are represented by different colors where we have the the modified and its corresponding natural cell subset on the same graph. The figure legend indicates each cell population and the color where:

- N= naïve cell
- MN= modified naïve cell

- ST = memory stem cell
- MST= Modified memory stem cell
- C= central memory cell
- MC= modified central memory cell
- T= Transitional memory cell
- MT= modified transitional memory cell
- E= effector memory cell
- ME= modified effectot memory cell

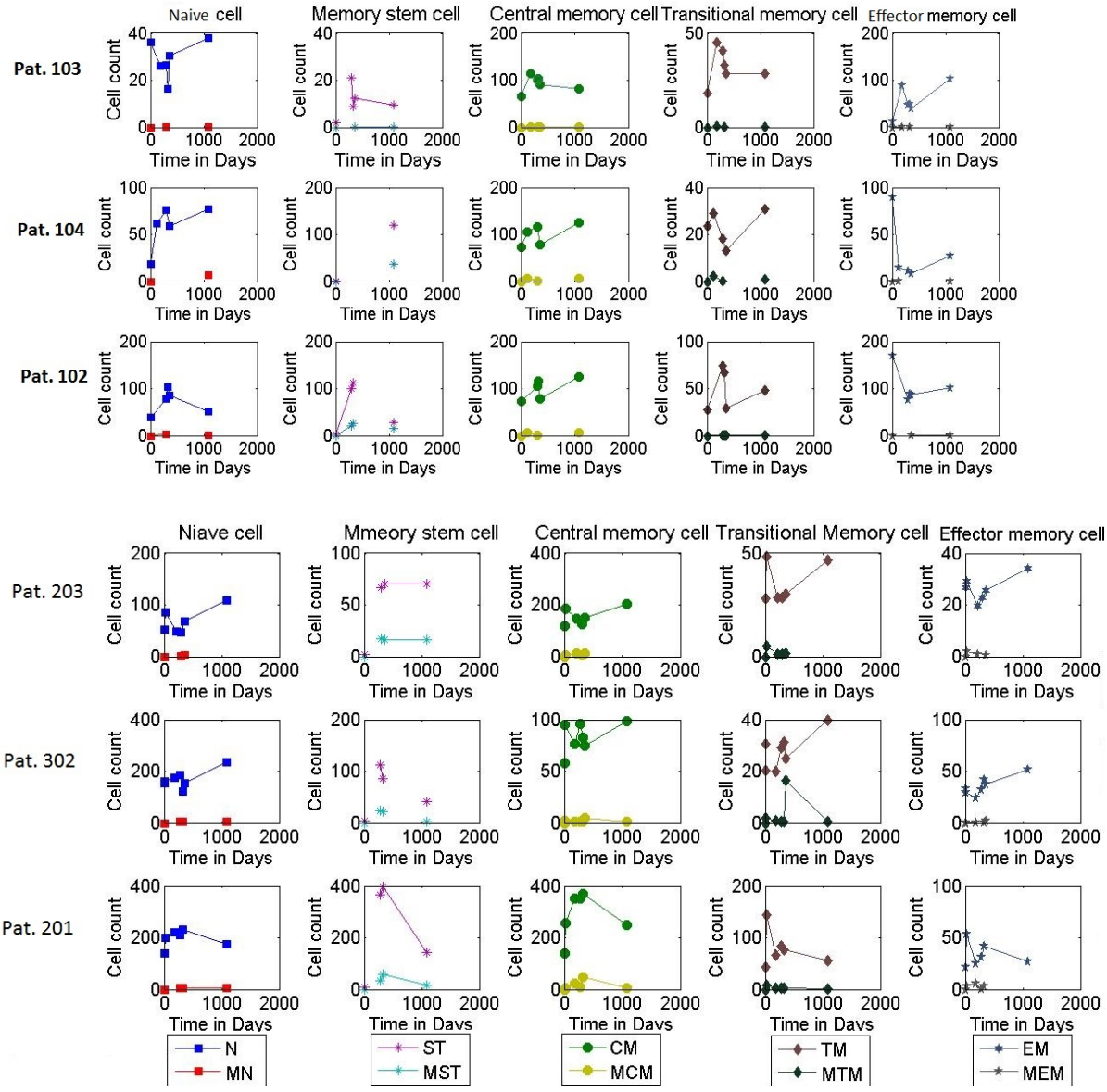

Figure 1.8: Naïve and Memory CD4 T-cell count for the low and medium dose cohorts.

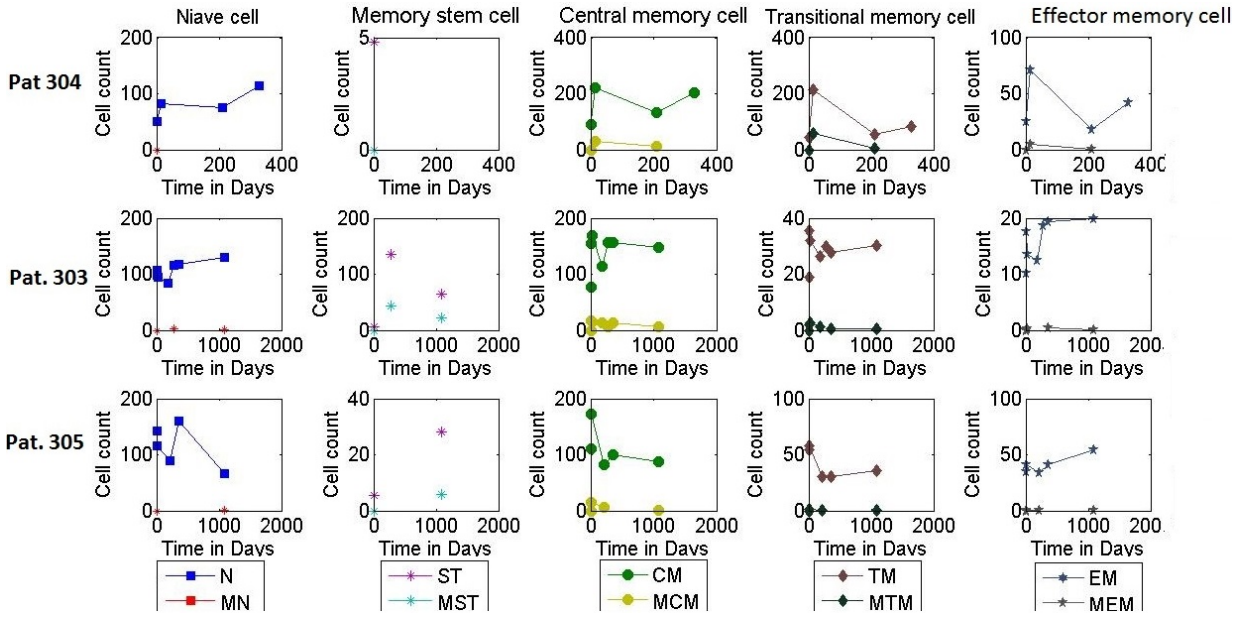

Figure 1.9: Naïve and Memory CD4 T-cell count for the high dose cohort.

1.5 Scope of Thesis

Several studies have shown that memory CD4 T-cells are a major contributor to the control of the HIV infection in the host [24]. In (Ostrowski et al, 1999), it was shown that naïve and memory CD4 T-cells were infected by the HIV-1 virus in vivo. Some researchers focused on studying the dynamics of the memory CD4 T-cell population in HIV infected individuals [24,28]. Others have focused on studying the importance of CD4 memory T-cells in the viral reservoir, as these memory cells are latently infected which when activated can infect hundreds of cells. The role of

these cells in relation to the progression to the AIDS condition is studied as well [26,36].

Since CCR5 is a primary co-receptor for the entry of the HIV virus into the host cells, and since the memory CD4 T-cells are primary HIV reservoir in the host [58], our work will evaluate the effect of introducing CCR5-down-modulated memory CD4 T-cells into HIV positive individuals, on the CD4 T-cell dynamic. The goal of this thesis is to determine what immune system naïve or memory cells proliferation, death, transition or birth rates are important in augmenting the memory T-cell population, and if there exists a significant augmentation at all.

We will address this question by introducing deterministic, stochastic and statistical models. My thesis is structured in following manner:

Chapter 2, I will present the deterministic and stochastic models used in this work. The non-linear ordinary differential equation model was developed by considering a linear transition model for the naïve and memory CD4 T-cell as suggested by our immunologist[52]. Later, a stochastic model is developed to account for the T-cell natural variability using the Gillespie Algorithm.

Chapter 3, I will present the fitting results obtained using various fitting procedures. Fitting routines were carried in Monolix software, Matlab programming language and and Stan. We will present baseline fit (using only data points at time

0) and fits carried after the injection of the CCR5-down-modulated memory CD4 T-cells.

Chapter 4, using an uncertainty sensitivity analysis technique, we will use the Latin Hypercubic sampling, coded in Matlab, to study the importance of each of the 15 model parameters, on each of the 5 subsets of CD4 T-cells using the Partial Rank Correlation Coefficients. This correlation is evaluated continuously after the treatment initiation over 200 days. Studying this correlation is important in order to understand what cell functions (death, proliferation or transition) and population size was affected by the experimental treatment.

Chapter 5 will be devoted to present the results and discussion about the stochastic model that we developed. Stochastic simulations are used to study variability in the cell dynamics by estimating the variance in the 5 CD4 memory T-cells populations. This will allow us to further determine whether the observed increase in the memory T-cell population in the laboratory is not lost in the variance of the cell dynamics from the model.

Lastly, in Chapter 6, I will present the conclusion obtained from this work where we determine which of the memory CD4 T-cell population functions was important in augmenting the CD4 T-cell count. As well, I will present some new directions for future research in this area.

2 The Model

2.1 Introduction

The goal of this thesis is to model a new HIV treatment, and determine whether natural fluctuations in CD4 T-cell count can include observed increases in CD4 T-cell count in patients. Before we introduce any analysis, the first step needed is to develop a mathematical model that describes the dynamics of the naïve and memory CD4 T-cells. In this chapter, first we will describe the deterministic model used in this work. This model will be used in chapter 3 to estimate the model parameters and in chapter 4 to perform an uncertainty and sensitivity analysis to study the relative significance of each of the model parameters with respect to each of the naïve and memory CD4 T-cell subsets. Second, a stochastic model is derived based on the Gillespie Algorithm, which will be used in chapter 5 to study variability in the naïve and memory CD4 T-cell populations. This is important to understand the occurrence of random variation in the CD4 T-cell count in the body.

2.2 Naïve and Memory CD4 T-cell Dynamical Model

In order to perform our parameter estimation, first we need to specify a mathematical model that describes the dynamics of the naïve and memory CD4 T-cell subsets. Our collaborators in Case Western University [52], proposed that memory CD4 T-cell subsets have a linear transition. Using this assumption, we constructed a system of ODE's to describe the transition process from the naïve to the effector memory state. The naïve and memory CD4 T-cell population is divided into 5 subsets, naïve (N), stem (ST), central (C), transitional (T), and effector (E), where each of these populations have death, proliferation and transition rates. In addition, the naïve cells have a birth rate as these cells are produced by the thymus. The effector memory cells do not transition to any other subset, as this is a terminal state. The following are some model assumptions made for the natural and CCR5 down-modulated memory CD4 T-cells.

- No backward transitions occur between the naïve an memory CD4 T-cell subsets.
- The CCR5-down-modulated memory CD4 T-cells follow the same dynamics presented for the natural CD4 T-cells.
- There is no interaction between the natural and injected modified memory CD4 T-cells.

- For model simplicity, and since the injected modified memory CD4 T-cells have a very small population size, we will assume that they have the same proliferation, death and transition rates as the natural ones.

- The modified naïve T-cells will not have a birth rate as these cells could not be produced naturally by the thymus.

- We assume that the blood is a well mixed homogeneous environment.

Figures 2.1 and 2.2 illustrate a flow diagram that describes the dynamics of the natural and CCR5-down-modulated memory CD4 T-cells respectively.
Parameters and variables used in both system of ODEs (Eq's 2.1 and 2.2) and the flow diagrams are described in Table 2.1 and 2.2 below.

Both natural and modified CD4 T-cells follow a linear transition from the naïve to the effector terminal state. The model is described as follows:

- The thymus is producing λ natural naïve cells per day.

- The naïve cells have a death rate d_N, a proliferation rate p_N, and transition to become memory stem cells with a rate ϕ_N.

- Memory stem cells have a death rate d_{ST}, a proliferation rate p_{ST} and transition to become a central memory cell with a rate ϕ_{ST}.

Variables	**Definition**
N	CD4 naïve T-cell
ST	CD4 memory stem T-cell
C	CD4 central memory T-cell
T	CD4 transitional memory T-cell
E	CD4 effector memory T-cell
MN	Down-modulated CCR5 CD4 naïve T-cell
MST	Down-modulated CCR5 CD4 memory stem T-cell
MC	Down-modulated CCR5 CD4 central memory T-cell
MT	Down-modulated CCR5CD4 transitional memory T-cell
ME	Down-modulated CCR5 CD4 effector memory T-cell

Table 2.1: Variables used in systems of ODEs 2.1 and 2.2.

Param.	**Definition and units**
λ	Number of naïve cells produced by the thymus per day
p_N	proliferation rate of natural and down-modulated naïve CD4 T-cells per $day^{-}1$
ϕ_N	Transition rate of natural and down-modulated naïve CD4 T-cells per $day^{-}1$
d_N	Death rate of natural and down-modulated naïve CD4 T-cells per $day^{-}1$
p_{ST}	Proliferation rate of natural and down-modulated stem memory CD4 T-cells per $day^{-}1$
ϕ_{ST}	Transition rate of natural and down-modulated stem memory CD4 T-cells per $day^{-}1$
d_{ST}	Death rate of natural and down-modulated stem memory CD4 T-cells per $day^{-}1$
p_C	Proliferation rate of natural and down-modulated central memory CD4 T-cells per $day^{-}1$
ϕ_C	Transition rate of natural and down-modulated central memory CD4 T-cells per $day^{-}1$
d_C	Death rate of natural and down-modulated central memory CD4 T-cells per $day^{-}1$
p_T	Proliferation rate of natural and down-modulated transitional memory CD4 T-cells per $day^{-}1$
ϕ_T	Transition rate of natural and down-modulated transitional memory CD4 T-cells per $day^{-}1$
d_T	Death rate of natural and down-modulated transitional memory CD4 T-cells per $day^{-}1$
p_E	Proliferation rate of natural and down-modulated effector memory CD4 T-cells per $day^{-}1$
d_E	Death rate of natural and down-modulated effector memory CD4 T-cells per $day^{-}1$
pterm	Limiting factor for the proliferation rate

Table 2.2: Variables used in the systems of ODEs 2.1 and 2.2.

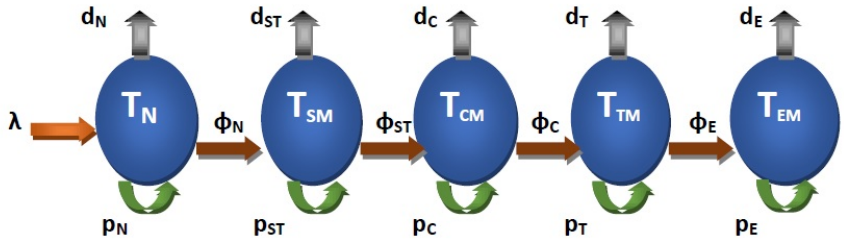

Figure 2.1: Flow Diagram presenting the dynamics of the naïve and memory CD4 T-cell

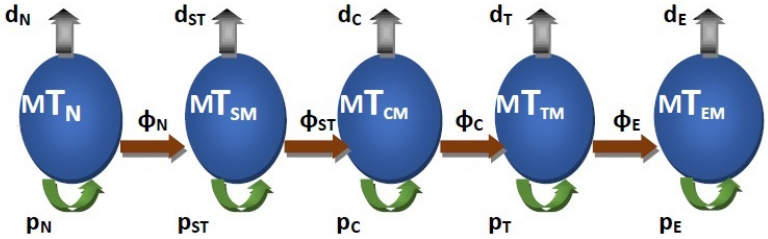

Figure 2.2: Flow Diagram presenting the dynamics of the down-modulated CCR5 naïve and memory CD4 T-cell

- Central memory cells have a death rate d_C, a proliferation rate p_C and transition to become a transitional memory cell with a rate ϕ_C.

- Transitional memory cells have a death rate d_T, a proliferation rate p_T and transition to become an effector memory cell with a rate ϕ_T.

- Effector memory cells have a death rate d_E, a proliferation rate p_E. This is a terminal state.

Modified CD4 T-cell follow the same dynamic described above but naïve cells are not produced by the thymus, so we do not have the λ term as shown in Figure 2.2.

The system of ODEs, Eq's 2.1 and 2.2, represent the dynamics of the natural naïve and memory CCR5 CD4 T-cell and down modulated CCR5 memory CD4 T-cell populations. The pterm in these equations represents a limiting factor as the T-cells cannot proliferate infinitely. At maximum there is a 1000 T-cells in 1 μL of plasma in the body of any healthy individual. So the proliferation rate of the CD4 T-cells is described by a logistic growth[54], So any subset of the CD4 T-cell can

only generate with a chance 1-$\frac{N+MN+ST+MST+C+MC+T+MT+E+ME}{1000}$

$$
\begin{aligned}
N' &= \lambda - d_N N - \phi_N N + p_N ptermN \\
ST' &= \phi_N N - d_{ST} ST - \phi_{ST} ST + p_{ST} ptermST \\
C' &= \phi_{ST} ST - d_C C - \phi_C C + p_C ptermC \\
T' &= \phi_C C - d_T T - \phi_T T + p_T ptermT \\
E' &= \phi_T T - d_E E + p_E ptermE
\end{aligned} \tag{2.1}
$$

$$
\begin{aligned}
MN' &= -d_N MN - \phi_N MN + p_N ptermMN \\
MST' &= \phi_N MN - d_{ST} MST - \phi_{ST} MST + p_{ST} ptermMST \\
MC' &= \phi_{ST} MST - d_C MC - \phi_C MC + p_C ptermC \\
MT' &= \phi_C MC - d_T MT - \phi_T T + p_T ptermMT \\
ME' &= \phi_T MT - d_E ME + p_E ptermME
\end{aligned} \tag{2.2}
$$

2.3 Stochastic Model

2.3.1 Introduction

As mentioned in chapter 1, in mathematical modeling there exist two types of models that are widely used to study the dynamics of infectious disease at the cellular and molecular scales; deterministic and stochastic models. Despite the advantages of using deterministic models to describe biological phenomenons, in a realistic world we know that human cells thrive in a Brownian world. The cell motion in the body can be described as partly discreet and partly random. Hence, the appropriate mathematical tools describing such motion are stochastic models that can capture variability in the Brownian world [49].

2.3.2 Continuous Time Markov Chain Model

Using the deterministic model presented by Eq's. 2.1 and 2.2, we develop a stochastic model. Here, time, represented as t, is considered to be continuous where $t \in [1, \infty)$ and variables are discrete.

Let N(t), ST(t), C(t), T(t), E(t) be random variables representing the number of naïve, stem, central, transitional, and effector memory CD4 T-cells at time t, respectively. These random variables are defined to be discrete and non-negative. And, let the transition probabilities that represent the change in state of the 5 CD4

T cells subsets be:

$$prob\Delta N(t) = i, \Delta ST(t) = k, \Delta C(t) = w, \Delta T(t) = v, \Delta E(t) = l \qquad (2.3)$$

Here we assume that the time step Δt is sufficiently small that only one event can occur at time t [49] i.e. each of the variables i, k, w,v and l can take only three values +1, -1 , 0 to describe what event is taking place during the time interval Δt. Table 2.3 lists all the fifteen possible outcomes that could occur in one time step Δt, their transition events and their corresponding transition probabilities. Note that this stochastic model is based on Eq's 2.1, as the modified memory CD4 T-cell population size is very small and is not the main focus in this work.

Reaction	**Events**	**Transition**	**Rate at Which Each Event Occur**	**Transition Probability In Time Interval [t,t+Δt]**
1	Birth of naïve Cells	N → N+1	λ	λ Δt
2	Death of N	N → N-1	d_N	d_N Δt
3	N proliferation	N → N+1	p_N pterm	p_N pterm Δt
4	Transition from N to STM	N → N-1 and STM → STM+1	ϕ_N	ϕ_N Δt
5	death of STM	STM → STM-1	d_{ST}	d_{ST} Δt
6	proliferation of STM	STM → STM+1	p_{ST}pterm	p_{ST}pterm
7	transition of STM to CM	STM→STM-1, CM→CM+1	ϕ_{ST}	ϕ_{ST} Δt
8	death of CM	CM →CM-1	d_{CM}	d_{CM} Δ t
9	CM proliferation	CM →CM+1	p_{CM}pterm	p_{CM}pterm Δt
10	Transition from CM to TM	CM→CM-1, TM→TM+1	ϕ_{CM}	ϕ_{CM} Δt
11	Death of TM	TM → TM-1	d_{TM}	d_{TM} Δt
12	Proliferation of TM	TM → TM+1	p_{TM}pterm	cell3
13	Transition from TM to EM	TM→TM-1, EM→EM+1	ϕ_{TM}	ϕ_{TM} Δt
14	Death of EM	EM → EM-1	d_{EM}	d_{EM} Δt
15	Proliferation of EM	EM → EM+1	P_{EM}pterm	P_{EM}pterm Δt

Table 2.3: Transition Events and their corresponding probabilities for the fifteen possible outcomes for the 5 CD4 T-cell subsets.

The transition probabilities of the fifteen possible different state for the five cell populations are presented below, where high order term are assumed to be 0 ($lim_{\Delta t \to 0} \frac{o(\Delta t)}{\Delta t}$=0). Equation 2.4 below, describes the Markov jump process, where each event takes place at a particular rate given the current state of the system.

$$prob\Delta N(t) = i, \Delta ST(t) = k, \Delta C(t) = w, \Delta T(t) = v, \Delta E(t) = l| \qquad (2.4)$$

$$N(t), ST(t), C(t), T(t), E(t) =$$

$$
\begin{cases}
\lambda\Delta t + o(\Delta t), (i,k,w,v,l) = (1,0,0,0,0) \\
d_N\Delta t + o(\Delta t), (i,k,w,v,l) = (-1,0,0,0,0) \\
p_N pterm\Delta t + o(\Delta t), (i,k,w,v,l) = (1,0,0,0,0) \\
\phi_N\Delta t + o(\Delta t), (i,k,w,v,l) = (-1,1,0,0,0) \\
d_S T\Delta t + o(\Delta t), (i,k,w,v,l) = (0,-1,0,0,0) \\
p_S Tpterm\Delta t + o(\Delta t), (i,k,w,v,l) = (0,1,0,0,0) \\
\phi_S T\Delta t + o(\Delta t), (i,k,w,v,l) = (0,-1,1,0,0) \\
d_C\Delta t + o(\Delta t), (i,k,w,v,l) = (0,0,-1,0,0) \\
p_C pterm\Delta t + o(\Delta t), (i,k,w,v,l) = (0,0,1,0,0) \\
\phi_C\Delta t + o(\Delta t), (i,k,w,v,l) = (0,0,-1,1,0) \\
d_T\Delta t + o(\Delta t), (i,k,w,v,l) = (0,0,0,-1,0) \\
p_T pterm\Delta t + o(\Delta t), (i,k,w,v,l) = (0,0,0,1,0) \\
\phi_T\Delta t + o(\Delta t), (i,k,w,v,l) = (0,0,0,-1,1) \\
d_E\Delta t + o(\Delta t), (i,k,w,v,l) = (0,0,0,0,-1) \\
p_E pterm\Delta t + o(\Delta t), (i,k,w,v,l) = (0,0,0,0,1) \\
1 - (\lambda + (d_N + p_N pterm + \phi_N)N(t) + (d_{ST} + p_{ST} pterm + \phi_{ST})ST(t) + \\
(d_C + p_C pterm + \phi_C)C(t) + (d_T + p_T pterm + \phi_T)T(t) + (d_E + p_E pterm)E(t)), \\
(i,k,w,v,l) = (0,0,0,0,0) \\
o(\Delta t), Otherwise.
\end{cases}
$$

2.3.3 Implementation of the MCMC simulations using the Gillespie's Algorithm

We implemented the stochastic simulations framework in matlab using the Gillespie's Direct algorithm [55].

In this method two uniformly distributed random variables defined in the interval (0,1) are used in each iteration. The first random variable is used to simulate the time step and the second is to select the event. Below is a description of the Algorithm used.

- Label all possible events of birth, proliferation, death and transition for all the five cell populations as E_1, ..., E_n.
- Define the rate at which event occurs as R_1 ... R_n.
- Generate two uniformly distributed random variables: $RAND_1$ and $RAND_2$
- the rate at which any event could occur is expressed as:

$$R_{total} = \sum_{i=1}^{n} R_i$$

- The time step between events is :

  $\Delta t = \frac{-log(RAND_1)}{R_{total}}$

- Set P = $RAND_2 R_{tot}$

- Event P occurs if

$$\sum_{i=1}^{p-1} R_i < P < \sum_{i=1}^{p-1} R_i$$

- Time t is now update as t t+ δt and the event P occurred.

- Return to step 2

In this algorithm, the transition rates are converted into probabilities where one random event is selected at each time step. Once selected, the time step and the number in each of the events is updated accordingly following the algorithm above [49].

3 Data Fitting Results

3.1 Introduction

The most important step in this work is the parameter estimation task. Using the non-linear ODE model described by Eq's 2.1 and 2.2, we will explore several fitting routines using different software to estimate the model parameters. This chapter is organized as follows:

- In section 3.2, we will introduce the use of the non linear mixed effect models in fitting dynamical biological data using several software.
- In section 3.3, we will present the baseline fit obtained using Matlab, where only data measurement at time 0 (before the infusion of the modified CD4 T-cells)is used.
- In section 3.4, we will present the data fitting obtained using Stan software where a hierarchical technique is used to fit all the data measurements obtained after the dose for the ten populations of cells.

- In section 3.5, Monolix fitting results is presented, where a non linear mixed effect approach is used. Model diagnostics for the three cohorts are achieved using several statistical tests and outputs to compare the different fits obtained in Monolix. Here data measurements after the dose are used.

- In the last section 3.6, parameter identifiablity and over fitting issues that occurred in the fits are discussed.

3.2 Non Linear Mixed-effects models and software

We proposed a system of ODEs for this study. However, many of the ODE parameters are unknown. Several statistical methods have been developed to estimate model parameters from experimental data such as Bayesian estimation and non linear mixed effect models [42]. Using non-linear ODE models made the fitting task harder as many of them do not have an analytic solution [42].

Bayesian hierarchical framework estimation uses a prior distribution obtained from the subject's data to estimate the individual parameters. This method is able to capture the within and between individual's variability [43,42].

Non-linear mixed effect model(NLME) uses a hierarchical framework which take into account the inner and intra individual variability. This method allow to perform statistical analysis on unbalanced longitudinal data. It is widely used to fit clinical

trials data in sociology, biology, psychology and many more [43]. Mixed effect models were originally introduced in the pharmacokinetics studies [43]. For more information the reader could refer to Phinero and Bates (2000).

Many software packages have been developed in the area of fitting experimental data to biological models. In this section we will introduce some software such as NONMEM, STAN, Monolix and Matlab.

- NONMEM developed by S.L. Beal and L.B Sheiner in 1980 is a very famous and widely used statistical software package in the pharmacokinetics community. It performs a population parameter estimation where the maximum likelihood is estimated using several approximation techniques such as the Laplacian and first order methods [45]. This software has various advantages. One of the most important advantage, is the ability to use compartmental and ODE based models, where several dosing methods could be considered. However, this software has several pitfalls. For instance, the graphical are sub-optional and the user needs to pair it with other software such as R and excel to output some graphics. In addition, fits are sensitive to the initial prior guess [46]. For further reference the reader could refer to Bates (2000).

- Stan is a statistical software package that estimates parameters using a Bayesian approach. The user will have to write a Stan program in order

to estimate the posterior distribution of the parameters as an initial estimates. A hierarchical framework could be adopted in this case. Like any statistical software, Stan has some limitations. For instance, the convergence time might be very long. In addition, Stan does not allow the user to make inferences about discrete parameters such as the case in mixture models [44].

- Monolix is a new statistical software package developed by Marc Lavielle and implemented in Matlab, that models non linear mixed effect models [51]. It is based on the Stochastic Approximation of the Expectation-Maximization,SEAM, using a Monte Carlo Markov Chain (MCMC)iterative algorithm. This algorithm is used to estimate the maximum likelihood estimator of the model population parameters, where a simulated annealing version improves the convergence of the model to a global maximum. The MCMC iterative method uses the Metropolis- Hastings approach. One of the drawbacks of this software package is its sensitivity to the initial guess provided when performing the fit [50,51].

  A more detailed explanation of the algorithm is presented in Appendix A.

- Matlab is a very famous mathematical software package that has a statistical toolbox that consists of many statistical packages. Depending on the kind of data, one may choose some linear and non linear least square methods

functions such as lsqln, lsqnonlin and fminsearch. These functions works by minimizing the distance between the ODE curve and the observations. The second fitting option available in Matlab is to use a non linear mixed effect method functions such as nlmefit and nlmefista. nlmefit uses the Likelehood Maximization Expectation (LME) or Laplacian first order to estimate the parameters. However, nlmefista use the SEAM algorithm to approximate the model parameters [47].

Parameter estimations in this thesis are carried out using three different fitting software; Monolix version 4.4, Matlab and Stan. In the next sections, we will present the fitting results obtained from these three software. Model diagnostics and goodness of fit were studied using the Akiake Information criterion (AIC), standard error values (s.e) and other figure output such as individual fits and prediction vs observations graphs.

3.3 Data Fitting

We first determine model parameter values at baseline, and then fit the model to the data after the dose. This methodology will us to determine which parameters are most affected by the dose of CCR5 down-modulated cells.

3.3.1 Baseline Fit- Matlab

In order to better understand the effect of the injected CCR5-down-modulated memory CD4 T-cell in each of the nine patients, we first performed a baseline fit in Matlab using a hierarchical mixed model using a log-norm distribution for the parameter's mean. This approach allows one to fit all the nine patients to their baseline count (data point at time 0) while allowing for inter patient variability. We used a non-linear least square data fitting method. Table 3.1 summarizes the results obtained for all the patients where we grouped them by the dose of injected CCR5-down-modulated memory CD4 T-cells (high(yellow), medium (white), low (gray)). The standard deviation and mean for the parameters is illustrated.

Pat	λ	d_N	d_{ST}	d_C	d_T	d_E	ϕ_N	ϕ_{ST}	ϕ_C	ϕ_T	p_N	p_{ST}	p_C	p_T	p_E
303	8.4457	.0096	0.0112	0.0033	0.0231	0.5151	0.0823	0.4648	0.1079	0.2479	0.0011	0.0110	0.0109	0.0215	0.0332
304	9.0864	0.0102	0.0110	0.0033	0.0231	0.3887	0.1709	0.4648	0.1075	0.2426	0.0011	0.0110	0.0109	0.0216	0.0332
305	10.5165	0.0114	0.0110	0.0033	0.0230	0.3215	0.0693	0.4648	0.1059	0.2333	0.0011	0.0110	0.0109	0.02179	0.0332
201	10.9220	0.0118	0.0110	0.0033	0.0232	0.4467	0.0746	0.4648	0.1022	0.2444	0.0011	0.0110	0.0110	0.0216	0.03320
203	9.1081	0.0102	0.0111	0.0033	0.0234	0.3750	0.1644	0.4648	0.1055	0.2493	0.0011	0.0110	0.01089	0.0215	0.0332
302	8.7649	0.0100	0.0111	0.0033	0.0232	0.2889	0.0497	0.4651	0.1107	0.2477	0.0011	0.0109	0.0108	0.0216	0.0332
102	9.1713	0.0104	0.0111	0.0033	0.0234	0.0909	0.2153	0.4636	0.1053	0.2478	0.00115	0.0110	0.0109	0.0216	0.0335
103	6.1309	0.0073	0.0111	0.0033	0.0232	0.4530	0.1558	0.4643	0.1071	0.2452	0.0011	0.0110	0.0109	0.0215	0.0331
104	7.6965	0.0089	0.0111	0.0033	0.0232	0.1297	0.2673	0.4673	0.1085	0.2495	0.0011	0.0109	0.01090	0.0216	0.0336
mean	8.9189	0.0100	0.0111	0.0033	0.0231	0.3388	0.1391	0.4648	0.1069	0.2453	0.0011	0.0110	0.0109	0.0216	0.0332
std	1.3756	0.00126	0.00085	0.00029	0.0017	0.1364	0.0702	0.0121	0.00405	0.0117	8.6456e-05	0.00074	0.00071	0.0015	0.0024

Table 3.1: Baseline fit for the three cohorts obtained using Matlab.

3.3.2 STAN Fitting results

Dr. Georges Monette performed a parameter estimation in Stan. In the fit, he used a hierarchical approach where all the patients are fit using an iterative MCMC approach model using sampling and population variance. A more detailed explanation about this fit is presented in Appendix B.

Table 3.2 shows the fit obtained for the nine patients.

Pat	λ	d_N	d_{ST}	d_C	d_T	d_E	ϕ_N	ϕ_{ST}	ϕ_C	ϕ_T	p_N	p_{ST}	p_C	p_T	p_E
303	10.2921	0.00936	0.0006	0.0804	0.0431	0.1	0.0986	0.1419	0.0196	0.0575	0.0078	0.0425	0	0.0007	0
304	8.3817	0.0092	0.0004	0.0467	0.0534	0.1	0.135	0.2234	0.0533	0.0466	0.00277	0.1237	0	0	0
305	11.2318	0.0511	0.0002	0.0658	0.0073	0.1	0.0492	0.2358	0.0342	0.1082	0.0003	0.1360	0	0.0155	0
201	19.3183	0	0.0058	0.0724	0.0571	0.1	0.1576	0.0984	0.0276	0.0429	0.05761	0.0042	0	0	0
203	6.9429	0.0080	0	0.0775	0.0249	0.1	0.1014	0.2243	0.02255	0.0780	0.0094	0.01243	0	0.0030	0
302	16.1964	0.0483	0.0083	0.0686	0.0015	0.1	0.0517	0.1024	0.0314	0.1341	0	0.0107	0	0.0356	0
102	7.4924	0.0028	0	0.07919	0	0.1	0.1171	02375	0.0281	0.188	0.02	0.1375	0	0.0880	0
103	2.7117	0.0136	0	0.0627	0	0.1	0.0973	0.3590	0.0373	0.1886	00109	0.2590	0	0.0887	0
104	6.47731	0.0012	0.0054	0.0744	0.0007	0.1	0.1401	0.1158	0.0256	0.1584	0.0413	0.0213	0	0.0590	0

Table 3.2: Hierarchical Fit obtained from Stan software. 2000 iterations are used.

3.3.3 Monolix Fitting Results

A non-linear mixed effect approach was employed to estimate the model parameters. Using Monolix software, we have used 3 fitting approaches. In the first two approaches, we used the data points provided for all ten cell populations. In the third approach, we used the data points for the non-modified naïve and memory CD4 T-cell only. The two main reasons behind ignoring the down-modulated CCR5 memory CD4 T-cell populations, is because of its small size and its short lifespan in the body, which varies between three to five years on average [56].

When performing the fit, Monolix requires the user to specify an initial guess for the fixed effect parameters. One of the good advantages this software provides, the ability to check the initial fixed effect. First, we have started with the baseline fit presented in Table 3.1 as an initial guess. Using the initial fixed effect tool, the model parameters were altered sometimes where we minimized the distance between the curve and the experimental data points by seeing the change of the curve instantly to get a good initial guess [50,51].

- In the First approach, we used a stochastic approximation of the Fisher Information Matrix. In here the Fisher Information Matrix is calculated using the exact model. The experimental data measurements provided for all the ten cell populations is used. Eq's 2.1 and 2.2 are used. This fitting routine

will be referred to as SNM throughout this work.

- In the second approach, we used a model linearization approach to estimate the Fisher Information Matrix where the model is linearized and is approximated by a Gaussian model. As well, the data points for all 10 cell populations is used. So we have used Eq's 2.1 and 2.2. This fitting routine will be denoted as LNM throughout this work.

- In the third approach, we have only used the data points of the non-modified naïve and memory CD4 T-cell. Eq's 2.1 are used. For this approach we used a model linearization to calculate the Fisher Information Matrix. This fitting routine will be referred to as LN throughout this work.

3.3.3.1 A non-linear mixed effect model for the T-cell dynamical model

As described earlier, our dynamical model consists of ten populations of naïve and memory CD4 T-cells for each of the nine patients. We present the concentration for each of the compartment at time, t. This work considers time to be the only dynamic variable in the model. In this fitting routine, we have used the same individual based model for the observations from the nine patients independently. The vector of observation at time t_j is presented as y_j where $1 < j < n$. The model is the distribution of the vector of observation y_i represented as $p_y(.; \phi, t)$.

A non-linear model is composed of fixed effect, random effect , residual error model and a covariate model can be added if needed. Where the fixed effect and random effect account for variability between and within individuals respectively. In this study, age, sex and ethnicity were not taken into account, hence there was no need to specify a covariate model. Having said that, we will have a 15 by 15 zero covariance model, as we have 15 unknown parameters to be estimated.

Due to a limitation in the Monolix software package where different initial conditions for each of the cell populations for each of the nine patients was not possible. Since patients are independent and share no information, we decided to consider an individual- based model where we fit the cell populations to patient separately. This is a drawback, since they will be related through distribution showing the inter-patient variation related over these parameter distributions. The Non-linear mixed effect model which represents our set of ODE equations is represented as:

$$Y_{ijl} = x_{il}(t_{ij}) + e_{ij} \tag{3.1}$$

Where we have the following:

- $i = 1....N$ = the ith individual.
- $j = 1...n_i$ = the time point.
- l= 5 or 10, which is the l'th number of compartment in the ODE model.

- β= the fixed effect, b = the random effect and e_{ij}= residual error model.

As many non-linear ODEs do not have an analytic solution, Monolix uses a numerical approximation technique to solve the ODEs. A non linear optimizing method is commonly adopted to fit these types of ODE's to experimental data. The user has to determine the initial conditions for each of the cell populations or else they are considered to be zero by the software. Hence, baseline measurements for each of the cell population are used as initial conditions.

In order to perform a fit in Monolix, the user has to determine the distribution for the parameters. Defined by the user, Monolix compares different paramater distributions and chooses the best model based on the Akiake Information Criterion (AIC), Bayesian Information Criterion (BIC) and -2 of the likelihoods (-2LL). For our model, it was shown that the a log-normal distribution resulted with the smaller AIC, BIC and -2LL values, where $log(\phi_i) = log(\theta) + \eta_i$. $log(\theta)$ represents the mean value of the fixed effect parameters and η_i is the random effect to account for inter-individual variability.

When performing data fitting, one cannot ignore the error that occurs when measuring experimental data. Hence it is important that for any given data set, we determine an error model that potentially estimates the distribution of the error in the data measurements. The residual error model defines the conditional probability distribution of the observations denoted by y_{ij}. Using the Akiake and Bayesian

Information Criterion (AIC and BIC) and the Likelihood (-2LL), Monolix proposed that a constant error model best describes our model and data data, where $e_{ij} = d_i\epsilon_{ij}$ and ϵ_{ij} is normally distributed $\sim$N(0,d^2).

3.3.3.2 Akiake Information Criterion

To evaluate the difference between the three fitting techniques, I have used the Akiake Information Criterion (AIC) and the standard errors to compare the goodness of fit.

The AIC takes into account the number of parameters, data points and the likelihood, with K and N being the number of parameters and data points respectively and is defined by the equation below [50]:

$$AIC = -2logl_m(y;\theta) + 2K + ((2K(K+1))/(N-K-1)) \tag{3.2}$$

In this section, the number of parameters(K=15), was the same for all the three fits. What changed is the number of observations and cell populations used in the model. Hence, having a smaller AIC indicates the most parsimonious model. In addition, to asses the best and most efficient model, the central processing unit time (CPU) and the standard error are used to asses the efficacy of the model [48]. The standard error is calculated by taking the ratio between the standard deviation of the estimated parameter and the estimated value for the parameter [48]. Besides

the above statistical tests, we have used the plot for the individual fits to asses the goodness of fit. Moreover, the observed vs prediction plots were used, where when the data points are more aligned to the 45^o degree line it indicates a better fit.

3.3.3.3 Parameter Estimation Results

In the subsections below, we will compare the three fitting routines for the three cohorts. A table showing the three different fits along with the standard error, AIC and CPU time are shown. In addition, I will present some of the plots obtained by Monolix such as the individual fits and the observed vs' predictions outputs to get a better insight about the quality of the fits. The ODE model described by Eq's 2.1 and 2.2 were solved over 1080 days in all the fits. In the case where we considered to fit the modified and non-modified memory and naïve CD4 T-cells, we have indicated that the modified-cell dose was injected into the patients after the measurement of their cell count at baseline at time 0. In all of the subsections below, results for non- modified CD4 naïve and memory T-cell populations will be presented, as these cell populations are the main interest of this work.

Patient 103- Low Dose

Table 3.3 below shows the parameter estimates for patient 103. When all the experimental data points for the ten cell populations of cell were used, adopting a stochastic approximation or model linearization to calculate the Fisher Information

matrix, we obtained the same values. However, using the model linearization approach, the software was able to calculate the s.e with the smaller CPU time. When comparing both the s.e values and the AIC, we can see that the fitting routine where only non-modified cell data points were used, resulted a better fit giving an AIC of 262, almost half that of the other models. Parameter values from all fitting exercises are somewhat similar in magnitude. A NaN value of the s.e means that either the coefficient is 0 or the software was not able to capture the parameter. Comparing the CPU time from all three fits, it is evident that when only the non-modified cell data is used, the fitting exercise is less computationally expensive. Looking at Figures 3.1, we can see that the model was not well fitted to the data as most of the observations were not close to the 45^opurple line. This further indicates that the model prediction is not perfectly close to the experimental data points. From Figure 3.2, we can make the following observation about the change in the memory and naïve CD4 T-cell populations from the best fit model:

- The naïve cell population decreased.
- The memory stem cell population had a very small increase.
- The central, transitional and effector memory cell populations have increased.

Param	**Lin-all**	**s.e**	**Stoch-all**	**s.e**	**Lin-non**	**s.e**
λ	5.83	9.5e+005	5.83	3.2	6.14	2.4e+007
p_N	0.000855	1.6e+005	0.000855	NaN	0.000201	NaN
p_{ST}	0.0064	NaN	0.0064	NaN	0.0064	NaN
p_C	0.09	5.7e+004	0.09	NaN	0.0821	1.6e+005
p_T	0.0976	5.6e+005	0.0976	NaN	0.0639	6.8e+005
p_E	0.483	8.2e+004	0.483	NaN	0.413	3.7e+006
ϕ_N	0.208	2.5e+004	0.208	0.066	0.217	2.7e+003
ϕ_{ST}	0.426	1.1e+004	0.426	NaN	0.461	5.8e+003
ϕ_C	0.104	2.3e+003	0.104	0.031	0.117	4.2e+003
ϕ_T	0.319	1.7e+004	0.319	0.092	0.34	5.1e+005
d_N	0.00565	4.9e+004	0.00565	0.084	0.006088	8.5e+005
d_{ST}	0.0105	5.6e+004	0.0105	NaN	0.00685	NaN
d_C	0.00249	3.2e+004	0.00249	0.025	0.00564	1.2e+005
d_T	0.0235	3e+005	0.0235	NaN	0.0327	NaN
d_E	0.437	5.3e+004	0.437	0.065	0.498	2.5e+006
AIC	479.55		479.55		262.33	
CPU (sec)	1.58e+003		7.37e+003		939	

Table 3.3: This table illustrates the fits obtained from the three fitting approaches with their AIC and standard error measurements for patient 103 .

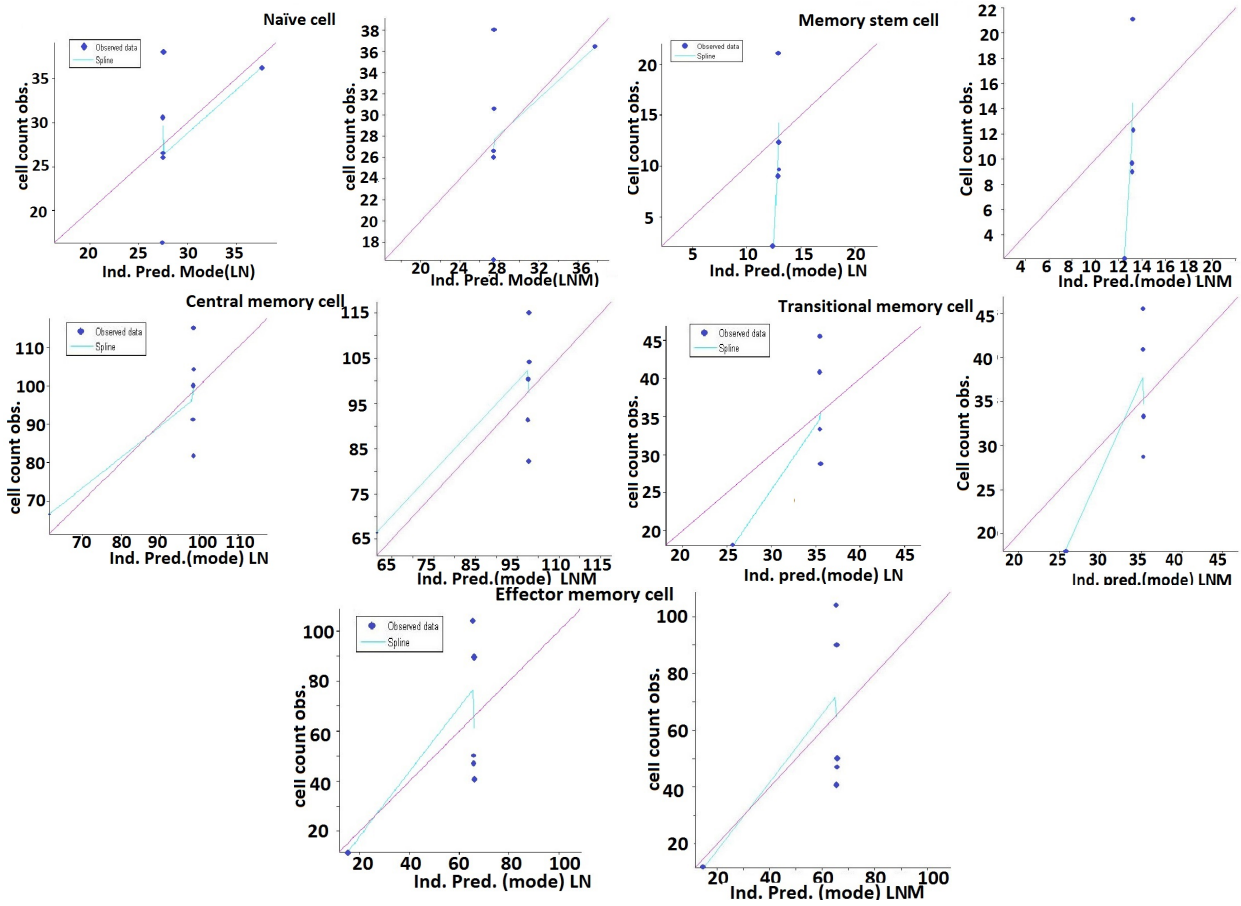

Figure 3.1: Pat 103 Prediction vs Observation.

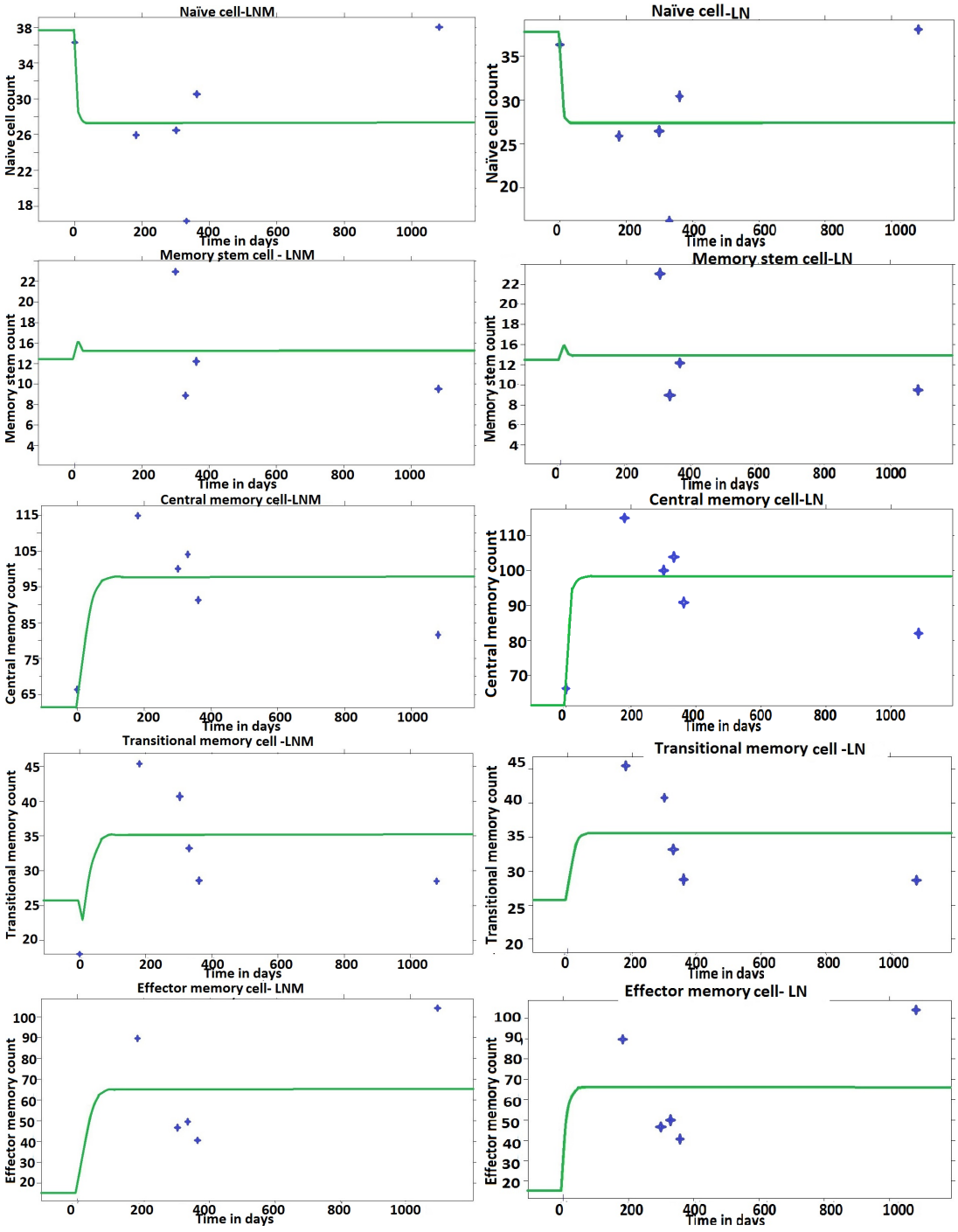

Figure 3.2: Individual fits of the naïve and memory CD4 T-cells for patient 103.

Patient 102-Low Dose

Table 3.4 shows the parameter estimates for patient 102. Similar to the previous patient, when all experimental data points for the ten cell populations are used, the stochastic approximation and the model linearization routines used to calculate the Fisher Information matrix resulted in similar parameter estimates. However, using the model linearization approach, the standard error was calculated in a shorter CPU time. Comparing both the s.e values and the AIC, we can see that the fitting routine where only non-modified cell data points were used gave a better fit where we had an AIC of 261. Parameter values from both fits are somewhat similar in magnitude except for the number of naïve cells produced by the thy thymus daily; λ. Looking at Figure 3.3, we can see that the model was not well fitted to the data as most of the experimental data points did not lie on the 45^opurple line, which indicates that the model did not succeed in replicating the clinical trial. However, the memory effector cell population seems to have a reasonable fit where the data measurement lie on the 45^o degree line.

From Figure 3.4, we can make the following observation about the change in the memory and naïve CD4 T-cell populations:

- The naïve, memory stem cell, central and transitional cell populations have increased from the baseline.

- The effector memory cells have decreased in count.

Param	**Lin-all**	**s.e**	**Stoch-all**	**s.e**	**Lin-non**	**s.e**
λ	18.2	3.8e+006	18.2	14	13	1.9e+007
p_N	0.0895	1e+006	0.0895	NaN	0.0871	4.4e+005
p_{ST}	0.159	1.3e+004	0.159	NaN	0.1	6.8e+005
p_C	0.0511	8e+005	0.0511	NaN	0.0318	2.9e+005
p_T	0.000167	NaN	0.000167	NaN	0.00181	NaN
p_E	0.268	3.7e+003	0.268	NaN	0.19	1.7e+006
ϕ_N	0.227	1.6e+003	0.227	0.18	0.19	1.7e+006
ϕ_{ST}	0.238	5e+004	0.238	0.2	0.235	3.7e+005
ϕ_C	0.0846	2.4e+005	0.0846	0.064	0.0963	3.4e+004
ϕ_T	0.319	1.7e+004	0.291	NaN	0.37	3.2e+005
d_N	0.00489	NaN	0.00489	NaN	0.00715	NaN
d_{ST}	0.00196	5.1e+004	0.00196	NaN	0.00836	NaN
d_C	0.00414	NaN	0.00414	2.5e+005	0.00474	NaN
d_T	0.0408	9.9e+005	0.0408	0.23	0.0209	4.4e+005
d_E	0.192	5.3e+004	0.192	0.032	0.316	9.6e+005
AIC	520.89		520.89		261.16	
CPU(sec)	2.71e+003		2.76e+004		3.45e+003	

Table 3.4: This table illustrates the fits obtained from the three fitting approaches with the AIC and standard error measurement values for patient 102.

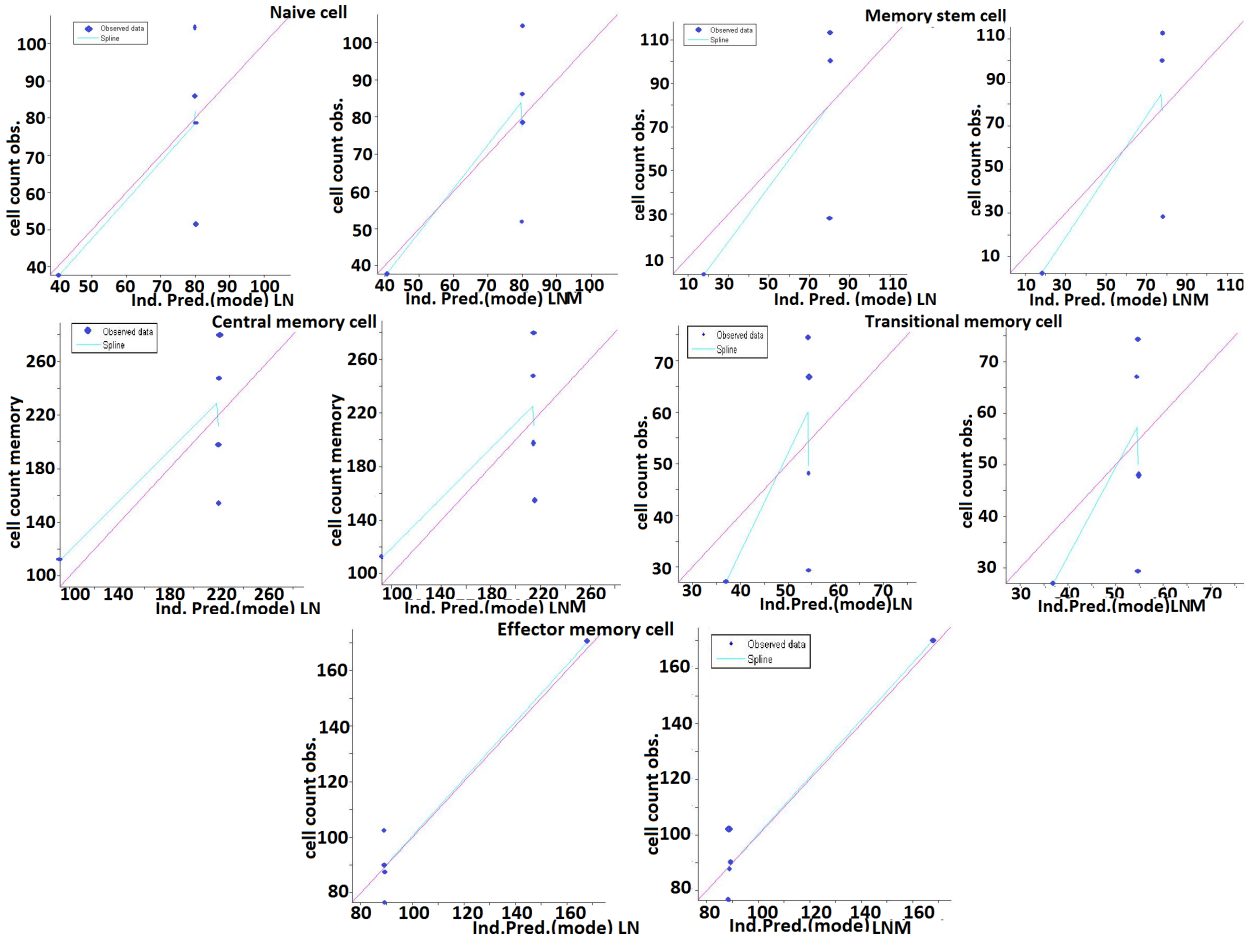

Figure 3.3: Observed vs prediction output for patient 102.

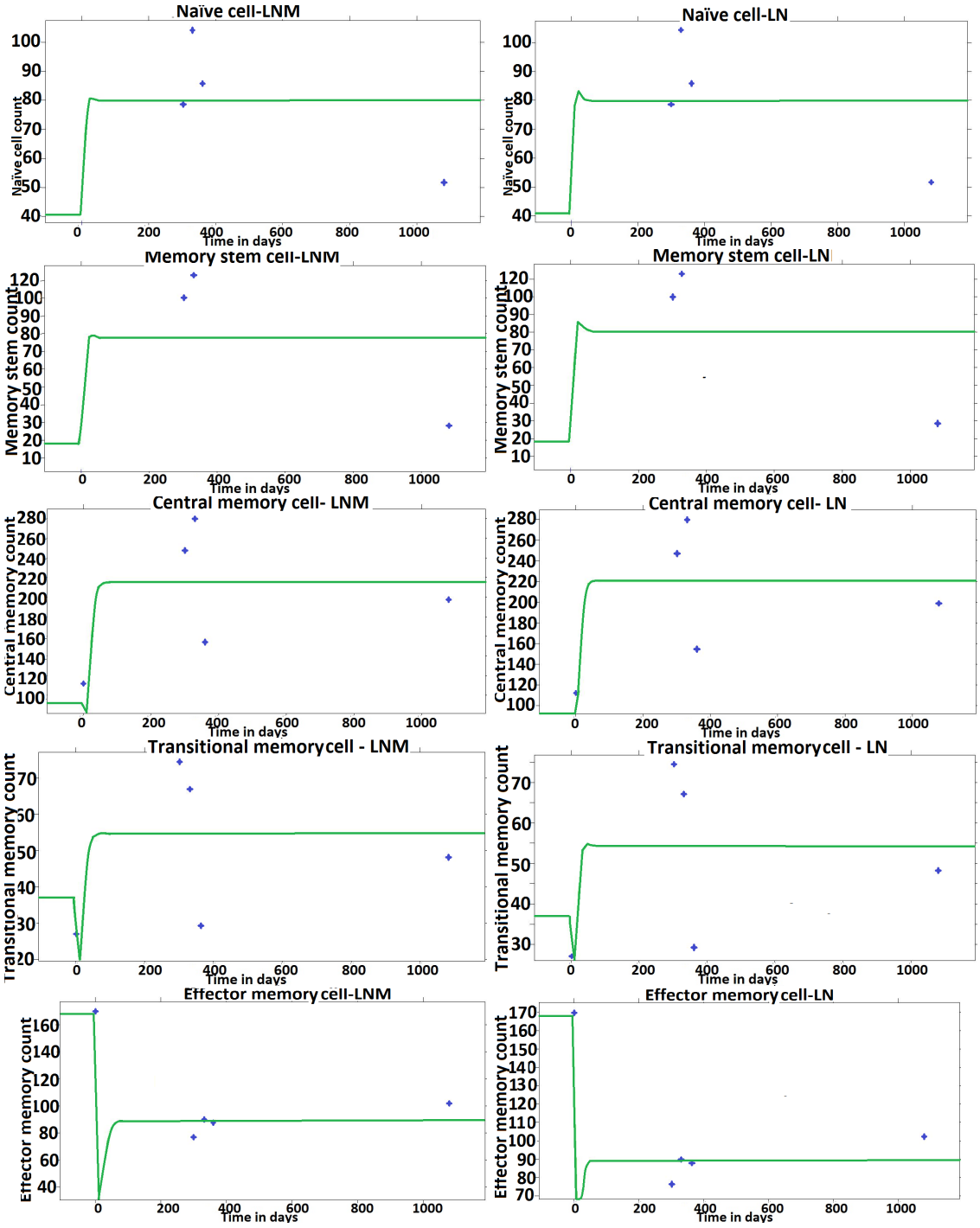

Figure 3.4: Individual fits for the naïve and memory CD4 T-cells for patient 102.

Patient 104 - Low Dose

Table 3.5 below shows the parameter estimates for patient 104. Again, when all the 10 cell populations data points are used, the stochastic approximation and model linearization approaches to calculate the Fisher Information matrix, result in similar parameter estimates. However, using the model linearization approach, the software was able to calculate the standard error. The fitting routine considering only non-modified cell population data gave a better fit, with an AIC of 210. Parameter values from all fits are similar in magnitude. Based on the CPU time, AIC and standard error values, the fitting routine considering only non-modified cell observations resulted in a better fit. Looking at Figure 3.5 , we can see that the model was not well fitted to the data for the naïve, central memory and transitional memory cells. However, observations for the stem and effector cells were well lined up with the 45^o degree purple line in both the fits. The results from the individual fits in Figure 3.6 shows a good fit for the central, effector and memory stem cell populations.
From Figure 3.6, we can make the following observation about the change in the memory and naïve CD4 T-cell populations:

- The naïve, central and memory stem cells have increased in count from the baseline.

- The effector and transitional memory cells have decreased in count from the

baseline.

Param	**Lin-all**	**s.e**	**Stoch-all**	**s.e**	**Lin-non**	**s.e**
λ	9.44	2.4e+007	9.44	9.5	8.73	2.6e+005
p_N	0.000226	7.2e+002	0.000226	NaN	1.97e-005	65
p_{ST}	0.00417	2.3e+003	0.00551	NaN	0.00551	6.3e+003
p_C	0.00365	1.9e+0053	0.0511	NaN	0.00365	NaN
p_T	0.00335	2.1e+004	0.00335	NaN	0.00181	NaN
p_E	0.00486	4.8e+003	0.00486	NaN	0.0139	6.9e+003
ϕ_N	0.135	3.7e+003	0.135	0.13	0.125	2.9e+003
ϕ_{ST}	0.0759	.8e+003	0.0759	0.048	0.0679	8.8e+002
ϕ_C	0.0827	4.2e+003	0.0827	0.095	0.0715	2.4e+003
ϕ_T	0.368	5.9e+0034	0.368	0.47	0.32	8.6e+003
d_N	0.0034	3.5e+005	0.0034	0.024	0.00261	1.1e+003
d_{ST}	0.0054	5e+003	0.0054	0.034	0.00836	NaN
d_C	0.0827	4.2e+003	0.00435	0.056	0.0715	2.4e+003
d_T	0.0187	3.8e+003	0.0188	NaN	0.0188	1.3e+004
d_E	0.527	9.4e+003	0.527	0.68	0.475	1.2e+004
AIC	383.08		385.08		210.41	
CPU (sec)	4.27e+003		1.78e+004		1.47e+003	

Table 3.5: This table illustrates the fits obtained from the three fitting approaches with their AIC and standard error measurements for patient 104.

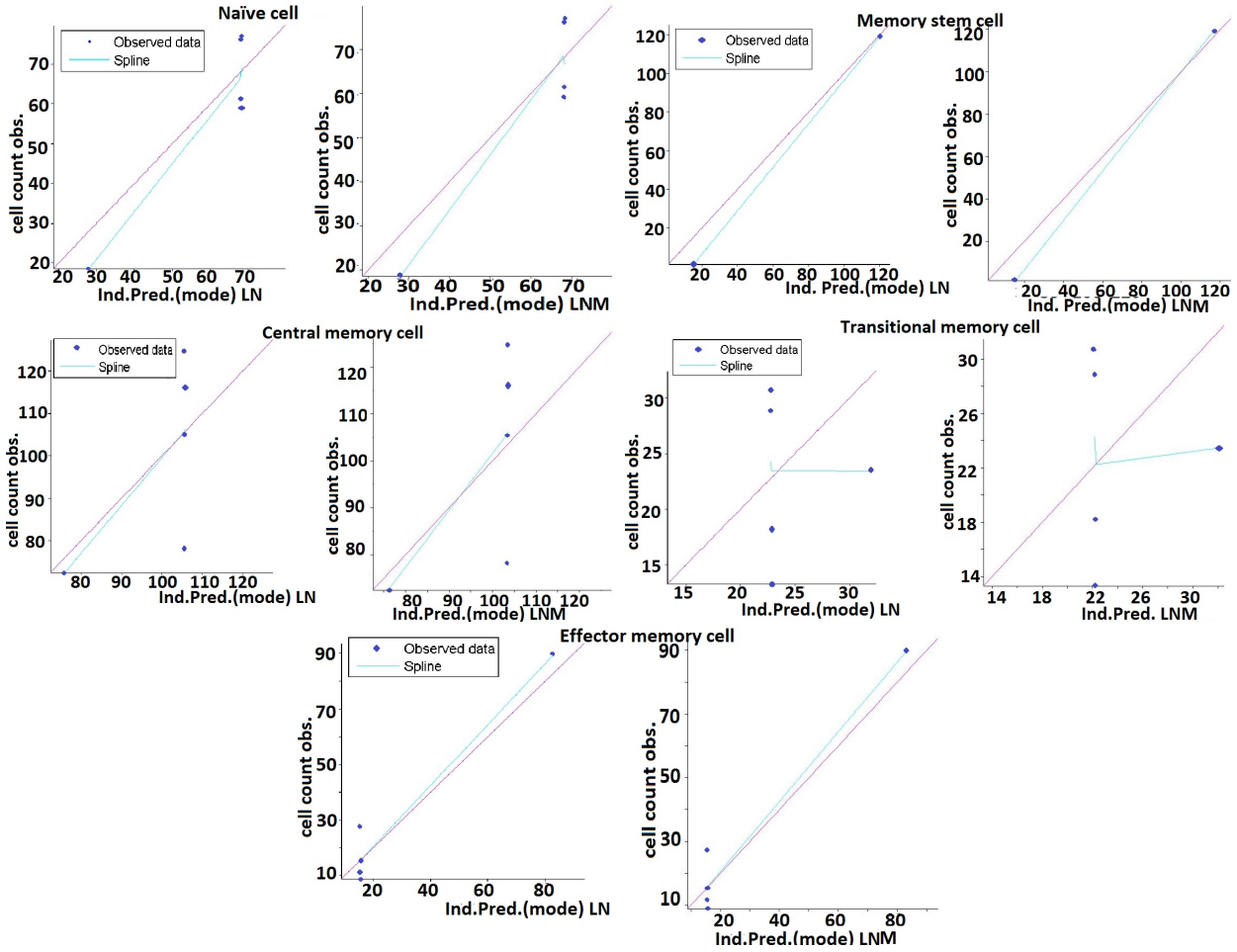

Figure 3.5: Observed vs prediction output for patient 104.

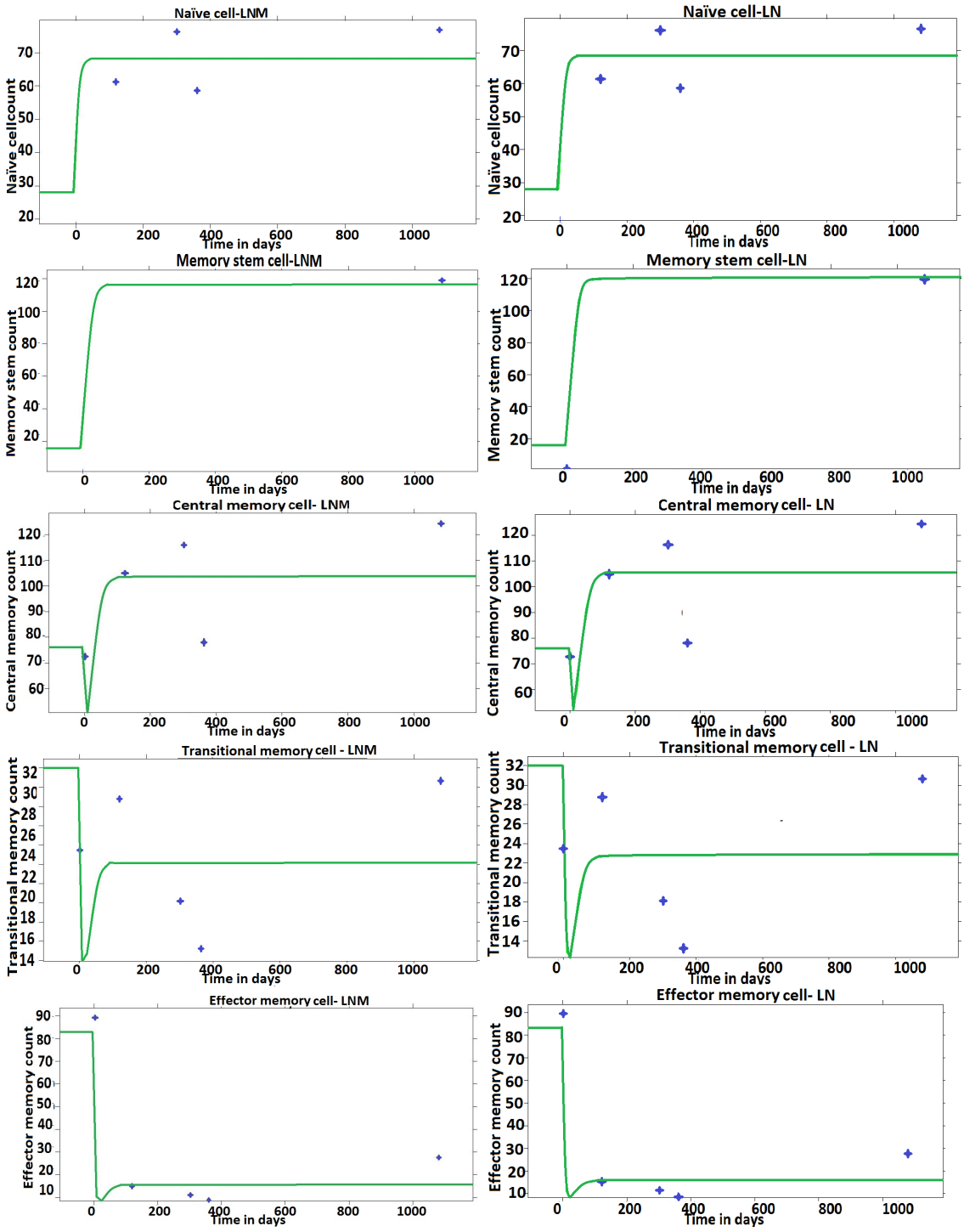

Figure 3.6: Individual fits for the naïve and memory CD4 T-cells for patient 104.

Patient 203- Medium Dose

Looking at Table 3.6, we can observe that when using all the 10 population of cells, Both the stochastic and model linearization approaches resulted in similar parameter estimates. However, the standard error values were obtainable using the linearization method with a faster CPU time. When the modified cell population observations are ignored, the AIC value of 268.61 is smaller compared to the AIC of the other two fits. Figure 3.7 we can see that the both of the fitting routines behaved similarly where the points are not aligned with the 45^o purple line except for the memory stem cell population. This indicates that the model does not appropriately fit the data very well. And this is evident in the individual fit Figure 3.8 where the model curve does not pass through all the data points like in the memory stem cell fit. The LN fitting routine, seems to pass through the data point better than the LNM method. The zero values for the proliferation rates indicates that these cells have increased to a level that cannot proliferate further.

Looking at Figure 3.8, we can draw the following observations about the dynamics of the five CD4 T-cell subsets.

- The naïve, memory stem cell and central memory cells increased in size.
- The transitional and effector cells exhibit no change.

Param	**Lin-all**	**s.e**	**Stoch-all**	**s.e**	**Lin-non**	**s.e**
λ	14.6	1.1e+004	14.6	2.8	10.6	4.1e+005
p_N	8.88e-007	31	8.88e-007	NaN	0.159	4.2e+004
p_{ST}	0.000107	2.4e+002	0.000107	0.00043	0.132	5.6e+005
p_C	0.000641	13	0.000641	0.0021	0.000514	9.3e+004
p_T	0.000305	53	0.000305	NaN	0.00117	1e+005
p_E	0.332	62	0.332	0.075	0.0221	7e+003
ϕ_N	0.209	35	0.209	0.047	0.251	7.1e+003
ϕ_{ST}	0.714	19	0.714	0.05	0.337	2.2e+005
ϕ_C	0.0862	15	0.0862	0.019	0.145	5.4e+003
ϕ_T	0.355	2.4e+002	0.355	0.09	0.638	3.2e+003
d_N	9.98e-007	1.1e+002	9.98e-007	0.0017	0.000576	2.4e+004
d_{ST}	5.86e-009	6.9	5.86e-009	NaN	1.12e-005	5.9e+005
d_C	0.00872	14	0.00872	0.018	5.98e-005	1.5e+005
d_T	0.0407	2.6e+002	0.0407	0.038	0.00224	4.4e+004
d_E	0.599	3.3e+002	0.599	0.085	0.9	1e+002
AIC	499.67		499.67		268.61	
CPU (sec)	1.35e+003		2.36e+004		890	

Table 3.6: This table illustrates the fits obtained from the three fitting approaches with their AIC and standard error measurements for patient 203.

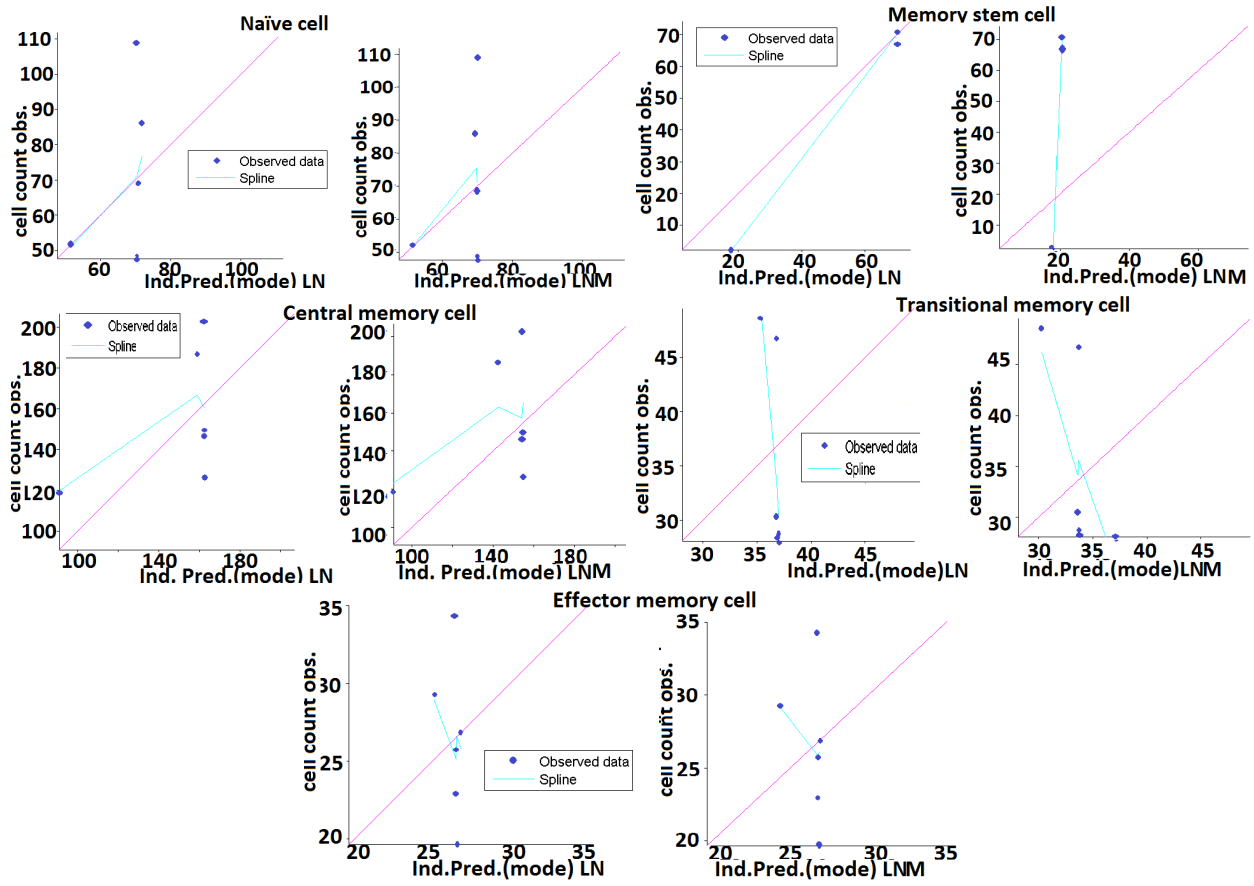

Figure 3.7: Observed vs Prediction output for patient 203.

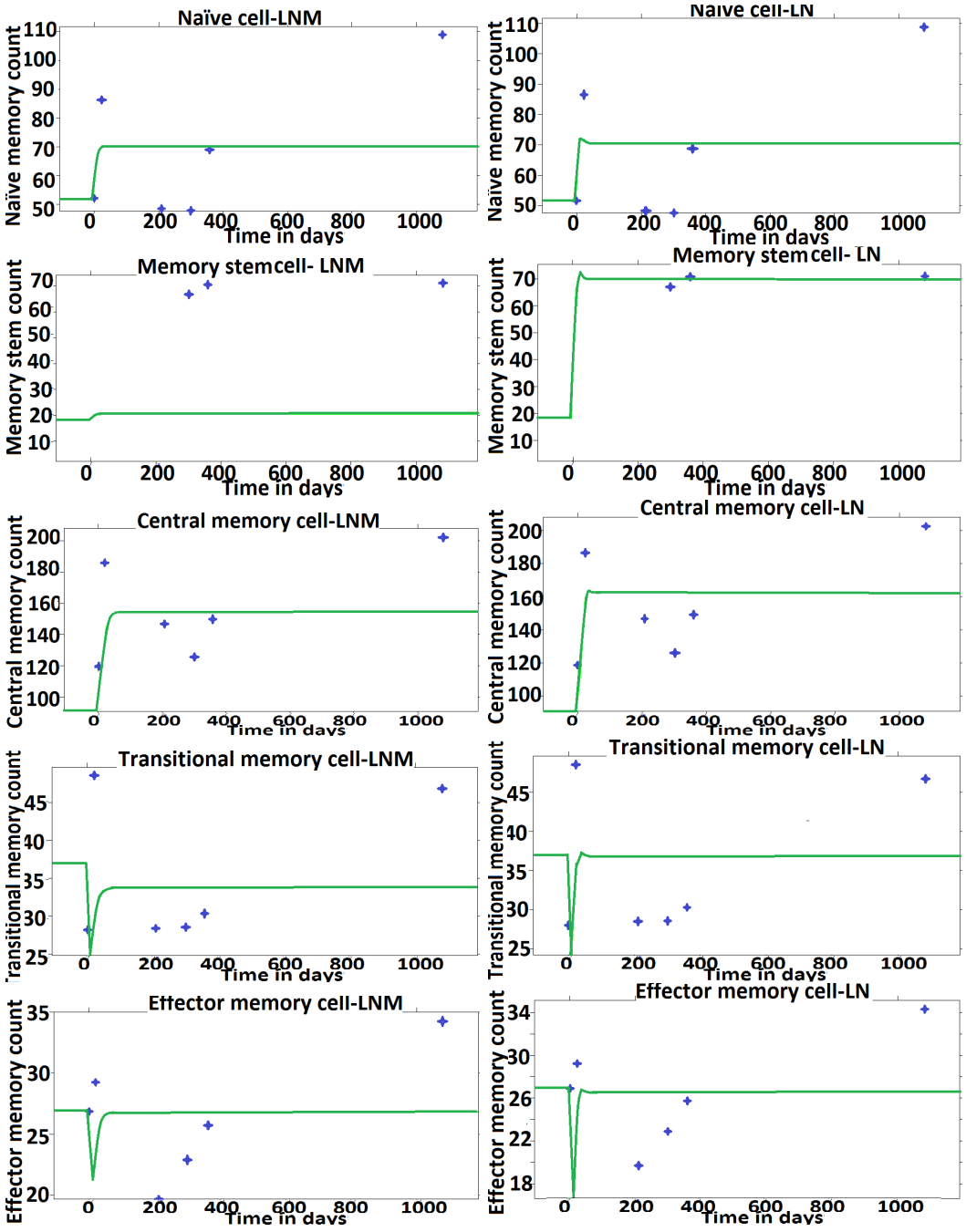

Figure 3.8: Individual fits for the naïve and memory CD4 T-cells for pat 203.

Patient 302 - Medium Dose

Table 3.7 shows the parameter estimates for patient 302. Again, when using all the 10 cell populations experimental data points are used, the stochastic approximation and model linearization approaches to calculate the Fisher Information matrix, result in similar parameter estimates. However, using the model linearization approach, the software was able to calculate the standard error with a shorter CPU time. The s.e values and the AIC value of 268 are smaller when using only non modified cell data points.

Figures 3.9, shows that the model is not well fitted to most of the cell populations,as the points and the spline line does not align with the 45^o purple line. However, the third approach, where only the non modified cells observations are used, the model was able to better fit the memory stem cell population which is evident as the observations and the spline line were closer to the purple line.

From Figure 3.10, we can make the following observation about the change in the memory and naïve CD4 T-cell populations:

- The naïve, memory stem and central have increased from the baseline.
- The effector and transitional memory cells have exhibit no change.

Param	**Lin-all**	**s.e**	**Stoch-all**	**s.e**	**Lin-non**	**s.e**
λ	13.8	1.8e+006	13.8	NaN	13.6	6.1e+005
p_N	0.173	3.3e+004	0.173	NaN	0.189	5.7e+003
p_{ST}	0.112	5.1e+004	0.112	NaN	0.0914	6.3e+004
p_C	0.000947	2.7e+004	0.000947	NaN	0.00649	1.2e+004
p_T	0.0472	5.3e+004	0.0472	0.4	0.00575	6.1e+003
p_E	0.846	1.7e+004	0.846	NaN	0.274	8.6e+004
ϕ_N	0.144	1.7e+003	0.144	0.01	0.191	3.8e+002
ϕ_{ST}	0.366	1.2e+004	0.366	0.081	0.459	3.7e+004
ϕ_C	0.277	3.1e+003	0.277	0.027	0.325	3.3e+002
ϕ_T	0.599	6.1e+002	0.599	0.16	0.832	1.3e+004
d_N	5.78e-015	0.0042	5.78e-015	2.8e-010	0.000541	9.6e+002
d_{ST}	0.000387	3.4e+004	0.000387	NaN	0.00057	7.4e+004
d_C	0.0449	3.8e+003	0.0449	NaN	0.106	2.8e+004
d_T	0.266	1e+004	0.266	NaN	0.19	1e+004
d_E	0.772	6.5e+003	0.772	0.13	0.796	4.1e+004
AIC	605.15		605.15		304.25	
CPU (sec)	1.43e+003		7.44e+003		950	

Table 3.7: This table illustrates the fits obtained from the three fitting approaches with their AIC and standard error measurements for patient 302.

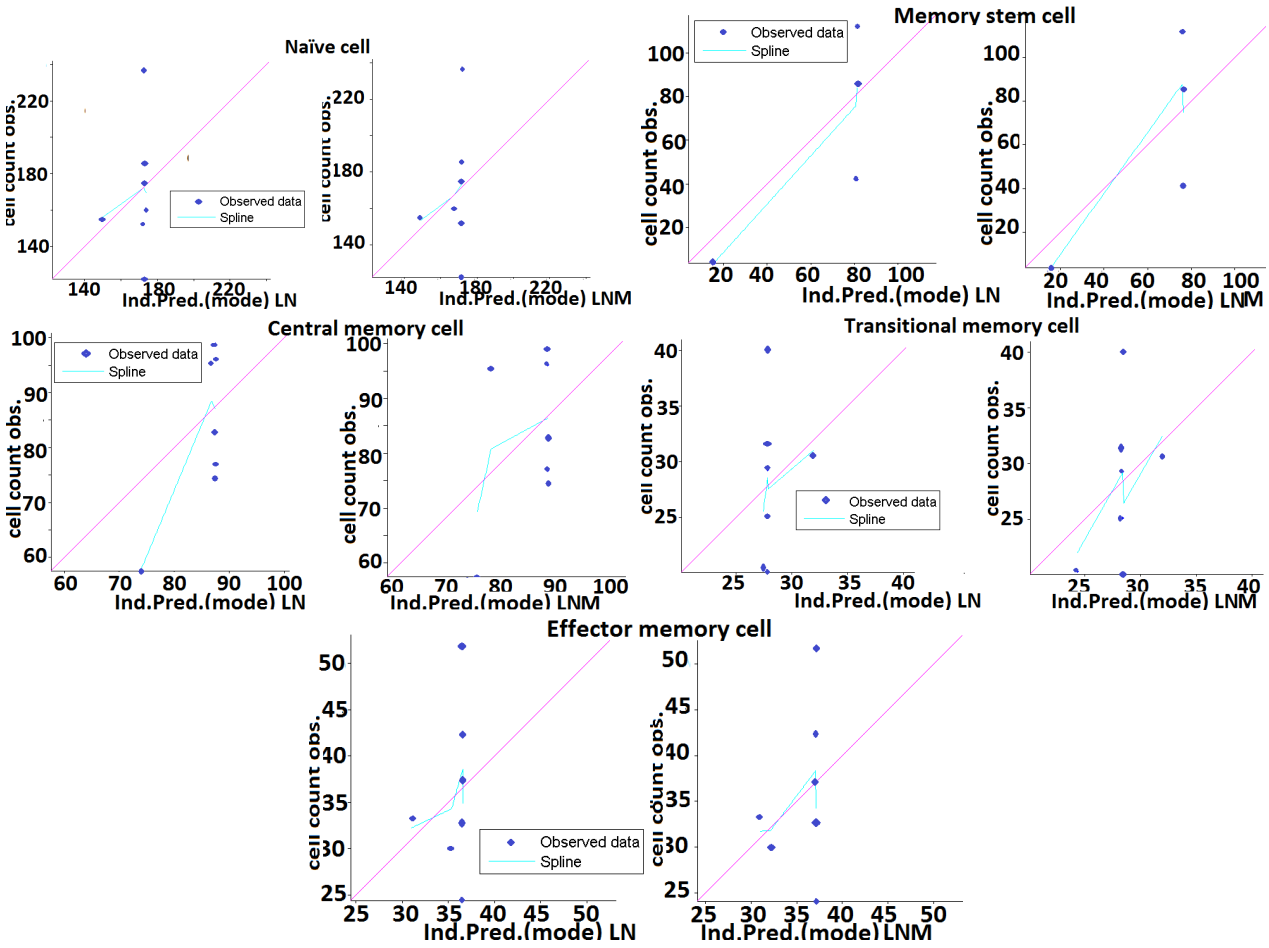

Figure 3.9: Observed vs Prediction output for patient 302.

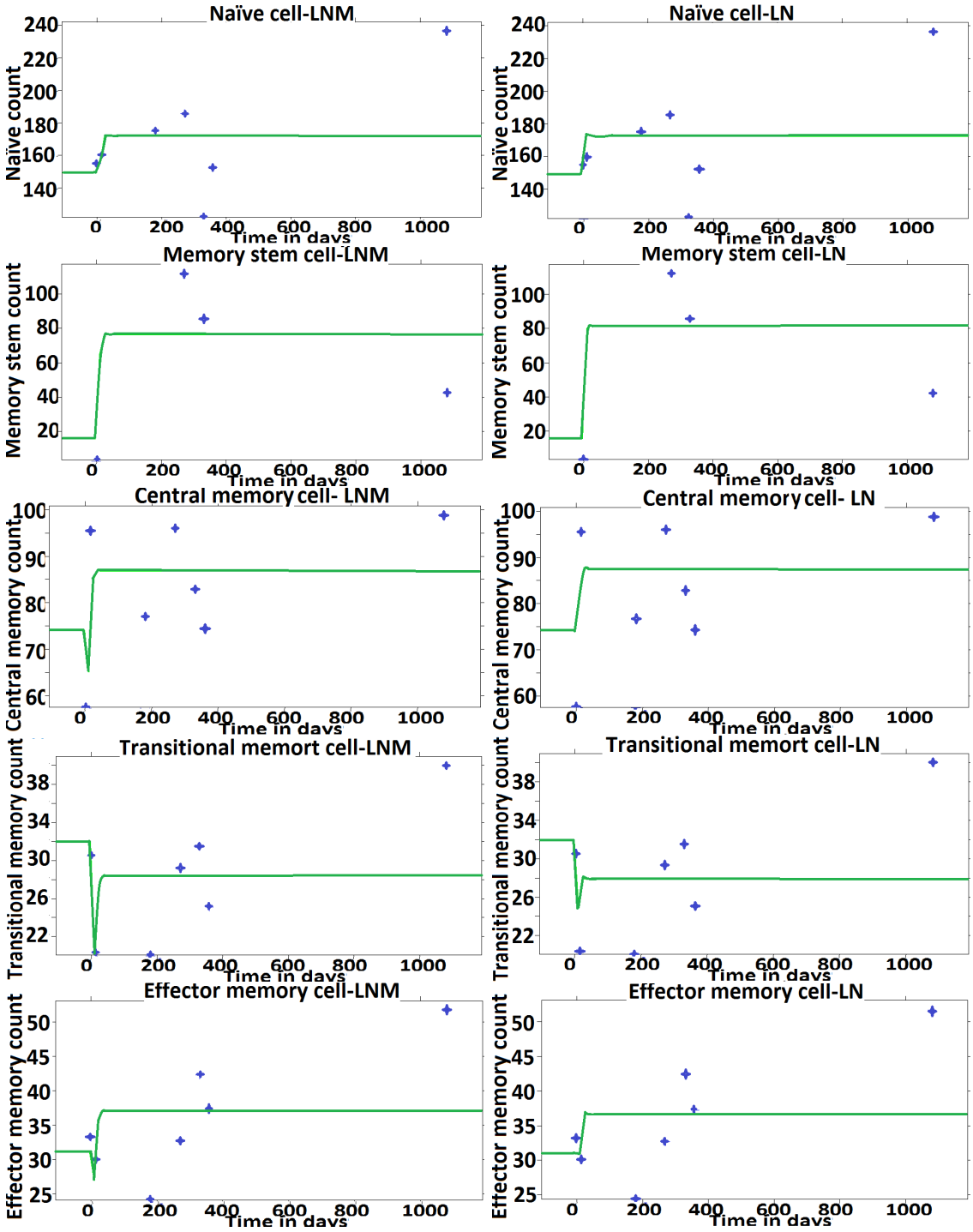

Figure 3.10: Individual fits for the naïve and memory CD4 T-cells for patient 302.

Patient 201 - Medium Dose

Table 3.8 lists the three parameter fit approaches for patient 201. When all the 10 cell populations data observations are used, the model linearization was able to capture a better fit than the stochastic approach with a sorter CPU time and a slightly smaller AIC value 627.80.Comparing both column 1 and column 5, the fit in column 5 had a smaller AIC value of 314, and smaller CPU time. The goodness of fit is evident when looking at the individual fits in Figure 3.11 as the ODE curve was closer to the data points.

Figure 3.12, shows that when using the non modified observations only, the spline and the observations were most of the time closer to the purple 45^o line.

From Figure 3.11, the naïve and memory CD4 T-cells subsets have increased from the baseline.

Param	**Lin-all**	**s.e**	**Stoch-all**	**s.e**	**Lin-non**	**s.e**
λ	24.9	7.3e+002	15.9	4.2	20.4	1.1e+002
p_N	1.03e-010	6	3.87e-012	NaN	1e-015	0.00096
p_{ST}	1.01e-010	9.5	1.75e-010	1.5e-007	0.528	1e+003
p_C	0.0444	36	0.11	NaN	2.77e-014	0.026
p_T	0.916	25	7.67e-007	NaN	5.96e-015	0.017
p_E	1.46e-010	7.8	0.903	NaN	2.88e-013	0.54
ϕ_N	0.119	3.8	0.0754	0.02	0.0942	87
ϕ_{ST}	0.567	16	0.285	0.12	0.0958	91
ϕ_C	0.0704	1.2	0.0231	0.0071	0.0912	5.3e+002
ϕ_T	0.125	6.4	0.101	0.029	0.262	1.1e+002
d_N	3.45e-005	2	2.39e-014	NaN	0.00297	86
d_{ST}	6.46e-012	0.78	1.76e-011	NaN	0.000293	35
d_C	5.34e-010	8.6	4.93e-013	8.1e-009	3.67e-005	6e+002
d_T	4.03e-011	9.2	6.54e-005	NaN	0.0711	1.9e+003
d_E	0.277	15	1.01e-005	NaN	0.617	2.5e+002
AIC	627.80		633.85		314.78	
CPU(sec)	1.8e+003		3.03e+004		1e+003	

Table 3.8: This table illustrates the fits obtained from the three fitting approaches with their AIC and standard error measurements for patient 201.

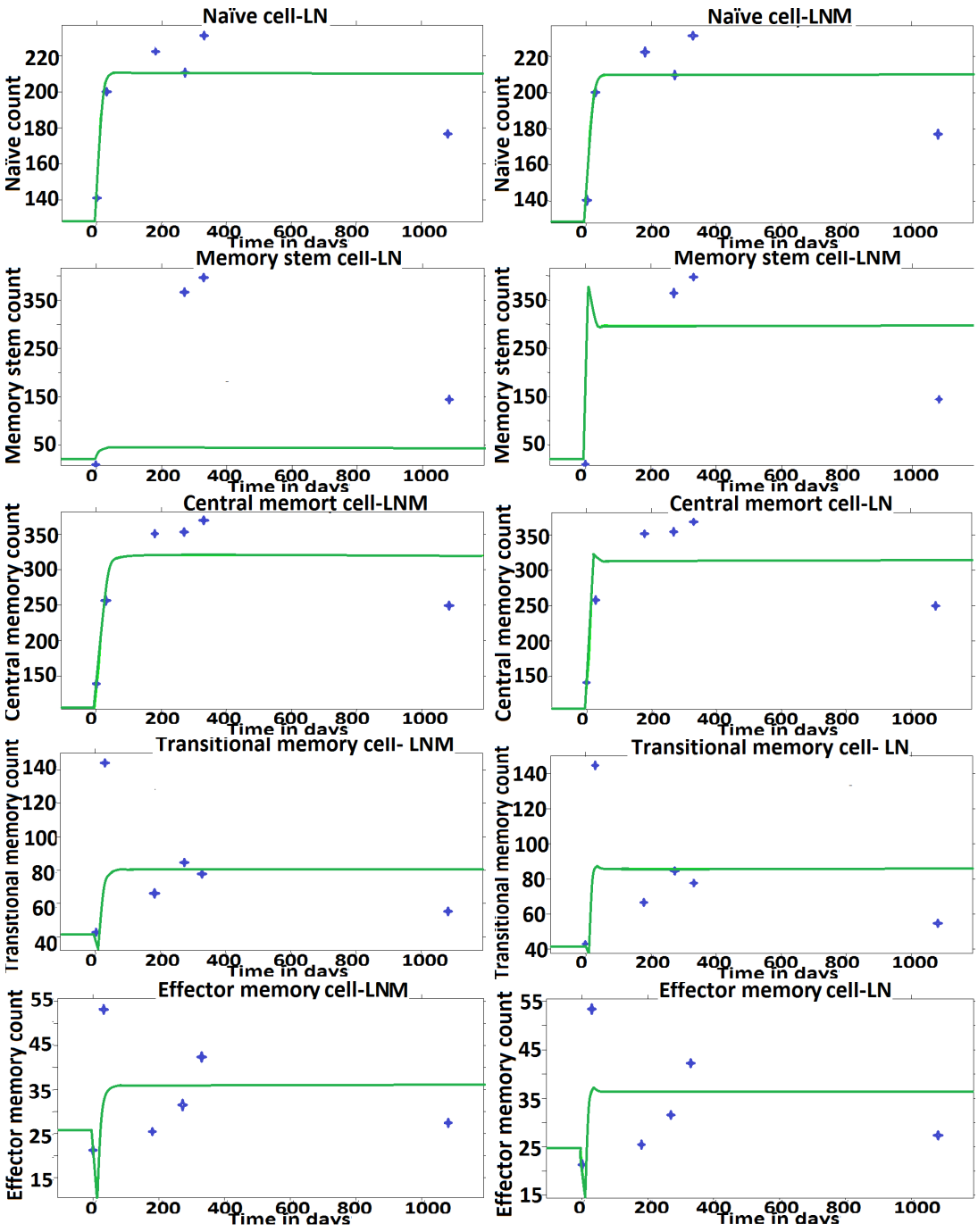

Figure 3.11: Individual fits for the 5 population of non modified CD4 T-cells for patient 201.

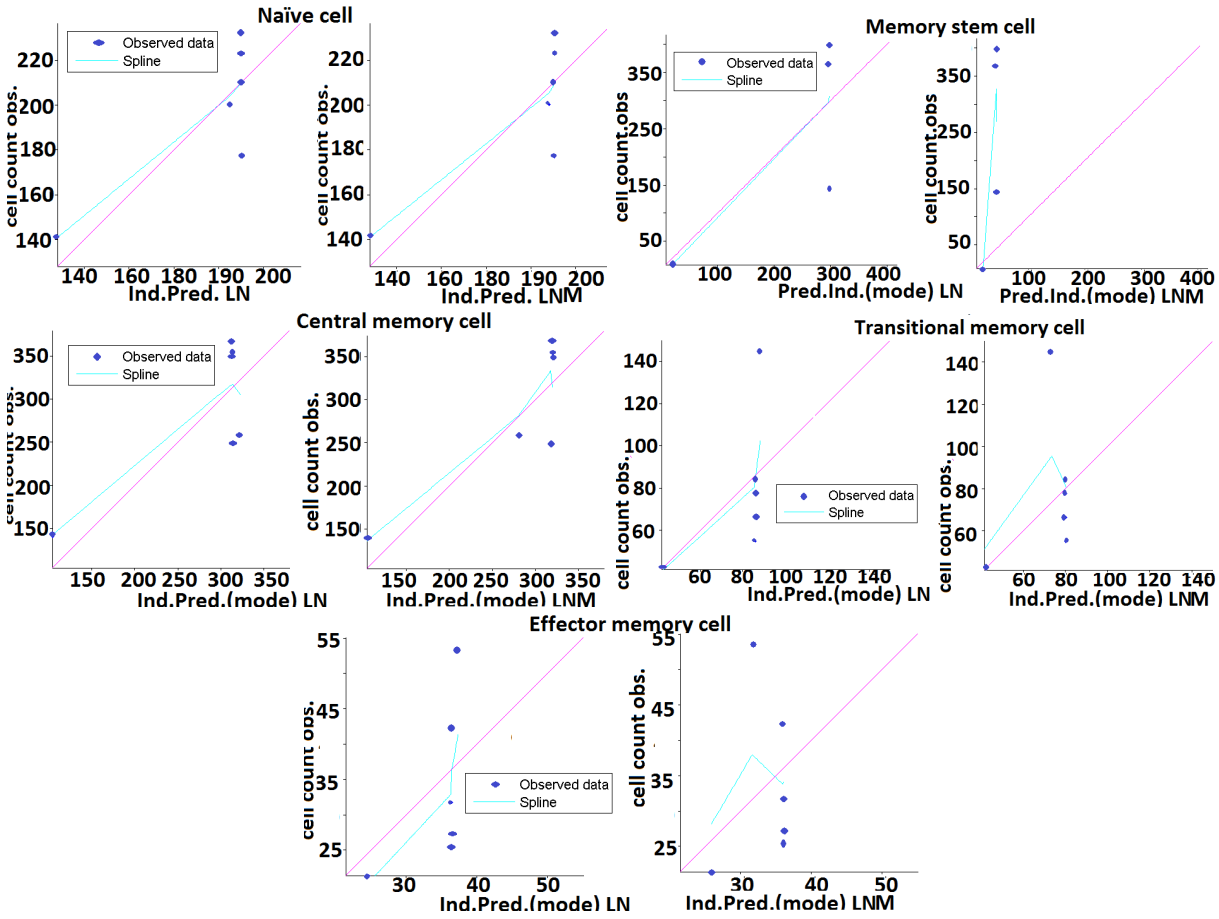

Figure 3.12: Observed vs prediction output for patient 201.

Patient 303- High Dose

Table 3.9 shows the parameter estimates for patient 303. Again, when all the 10 cell populations data points are used, both model linearization and stochastic approximation approaches to calculate the Fisher Information Matrix gave similar parameters estimates. However, the model linearization method was less computationally expensive and s.e were obtainable. The AIC value when using only non-modified cell population data points, the AIC value of 291.08 is half of that other two fits, indicating it is a better fit. This result is also evident in the individual fits Figure 3.13 as the ODE curve passes through most of the data points except for the central memory cell population.

Figure 3.14, Shows that when using the non modified observations only, the spline and the observations were laying on closer to the purple 45^o line. From Figure 3.13, we can see that all the cell populations have increased from the baseline except for the central memory cell which decreased in size according to the model predictions.

Param	**Lin-all**	**s.e**	**Stoch-all**	**s.e**	**Lin-non**	**s.e**
λ	32.5	2.6e+002	32.5	14	10.8	2.6e+002
p_N	8.67e-034	NaN	8.67e-034	NaN	0.0119	6.3
p_{ST}	0.991	15	0.991	NaN	0.0112	1.9e+002
p_C	2.17e-056	NaN	2.17e-056	NaN	0.00737	1.3e+003
p_T	0.00736	13	0.00736	NaN	0.27	1.1e+003
p_E	3.38e-012	0.075	3.38e-012	5.6e-008	0.0232	1.4e+003
ϕ_N	0.304	2.5	0.304	0.13	0.102	30
ϕ_{ST}	0.628	2.6	0.628	0.18	0.118	16
ϕ_C	0.0803	0.68	0.0803	0.032	0.096	2e+002
ϕ_T	0.397	0.93	0.397	0.16	0.247	9.9e+002
d_N	6.6e-053	NaN	6.6e-053	NaN	0.00559	27
d_{ST}	4.62e-077	NaN	4.62e-077	NaN	0.00538	1.2e+002
d_C	0.333	2.4	0.333	0.1	0.444	1.1e+003
d_T	1.85e-030	NaN	1.85e-030	NaN	0.018	49
d_E	0.705	1.7	0.705	0.29	0.499	2.9e+003
AIC	681.23		681.23		293.08	
CPU (sec)	4.74e+003		2.37e+004		2.18e+003	

Table 3.9: Illustrates the fits obtained from the three fitting approaches with their AIC and standard error measurements.

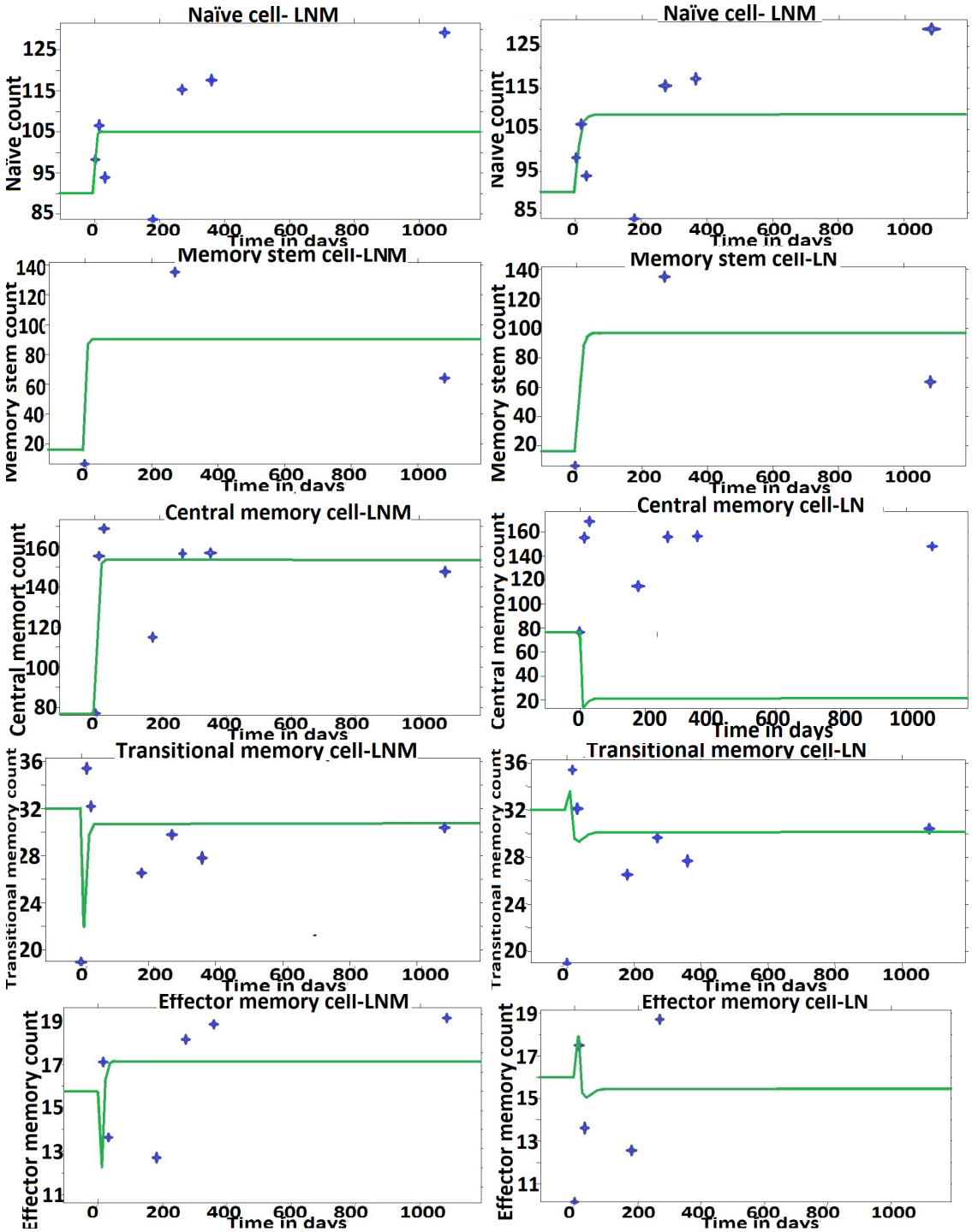

Figure 3.13: Individual fits for the naïve and memory CD4 T-cells for pat 303.

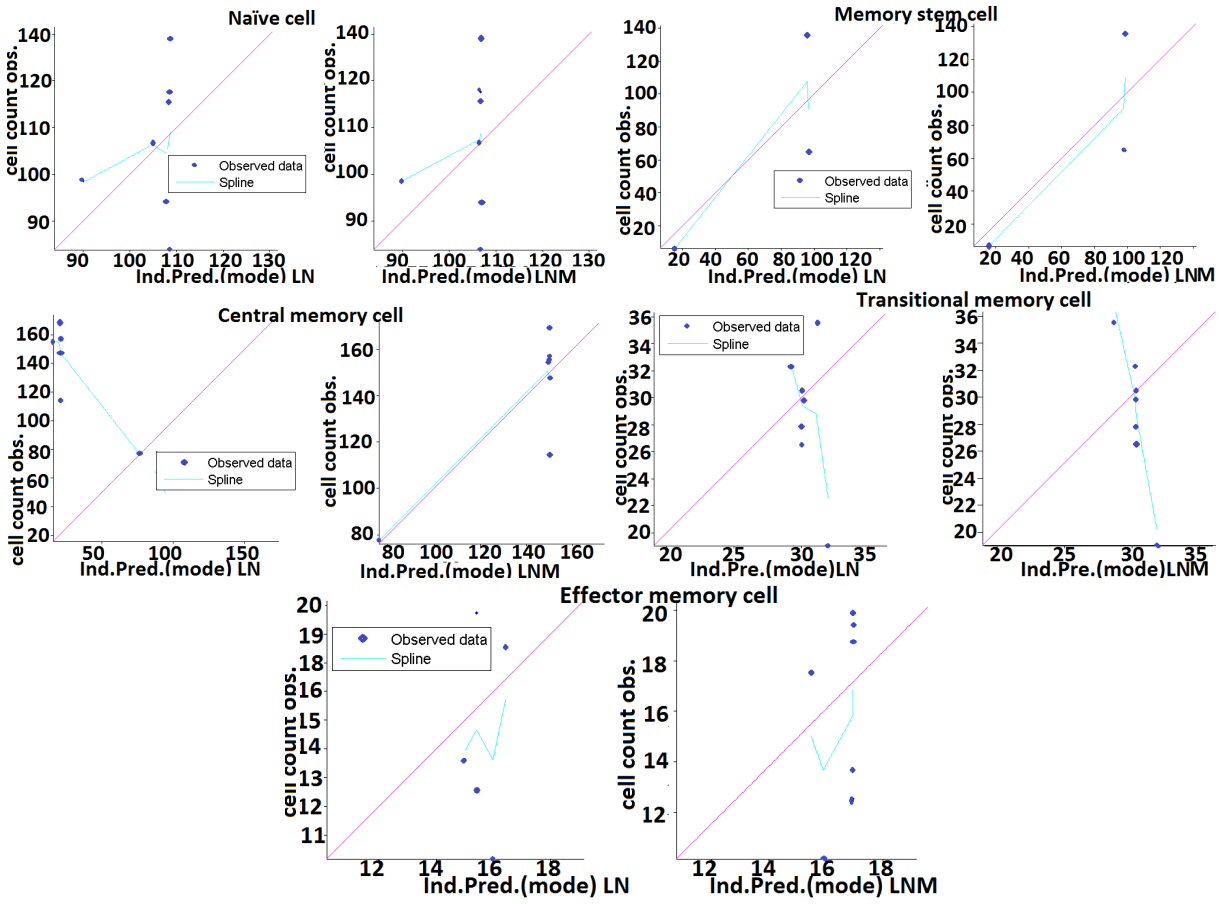

Figure 3.14: Observed vs prediction observation for pat 303.

Patient 304 - High Dose

Table 3.10 shows the parameter estimates for patient 304. Again, using all the 10 cell populations data points in the fit, both the model linearization and stochastic approximation approaches in calculating the Fisher Information Matrix, result in similar parameter estimates and AIC values with having a smaller CPU using the model linearization approach. Considering only the non-modified cell populations the AIC value of 208.82 indicates it is a better fit. However, when looking at Figure 3.15, we can observe that both fits(LN and LNM) result in similar individual fits. Figure 3.16, Shows that when using the non modified observations only, the spline and the observations were most of the time closer to the 45^o purple line, indicating that our model was able to replicate the clinical study. From Figure 2.15, it is evident that all the cell populations have increased from the baseline.

Param	**Lin-all**	**s.e**	**Stoch-all**	**s.e**	**Lin-non**	**s.e**
λ	48.2	1.9e+004	48.2	15	14.3	7.4e+003
p_N	6.57e-013	17	6.57e-013	2e-009	0.224	52
p_{ST}	4.84e-011	6.1e+002	4.84e-011	NaN	0.672	1.8e+002
p_C	8.4e-013	18	8.4e-013	NaN	3.74e-014	0.11
p_T	4.09e-011	2.3e+002	4.09e-011	NaN	1.87e-016	0.00031
p_E	3.38e-012	0.075	1.69e-010	4.4e-007	0.0232	1.4e+003
ϕ_N	0.304	2.5	0.532	0.18	0.266	52
ϕ_{ST}	0.628	2.6	0.981	0.022	0.618	1.4e+002
ϕ_C	0.0803	0.68	0.265	0.083	0.266	27
ϕ_T	0.491	1.4e+002	0.491	0.31	0.399	41
d_N	7.29e-012	32	7.29e-012	1.6e-008	9.99e-017	0.00024
d_{ST}	1.93e-012	69	1.93e-012	7.6e-009	2.08e-023	NaN
d_C	1.08e-010	1.3e+002	1.08e-010	NaN	0.444	1.1e+003
d_T	7.3e-006	2.8e+002	7.3e-006	NaN	1.94e-019	NaN
d_E	0.996	1.9e+002	0.996	0.0042	0.997	1e+002
AIC	364.49		364.49		208.82	
CPU(sec)	4.28e+003		1.56e+004		2.84e+003	

Table 3.10: Illustrates the fits obtained from the three fitting approaches with their AIC and standard error measurements for patient 304.

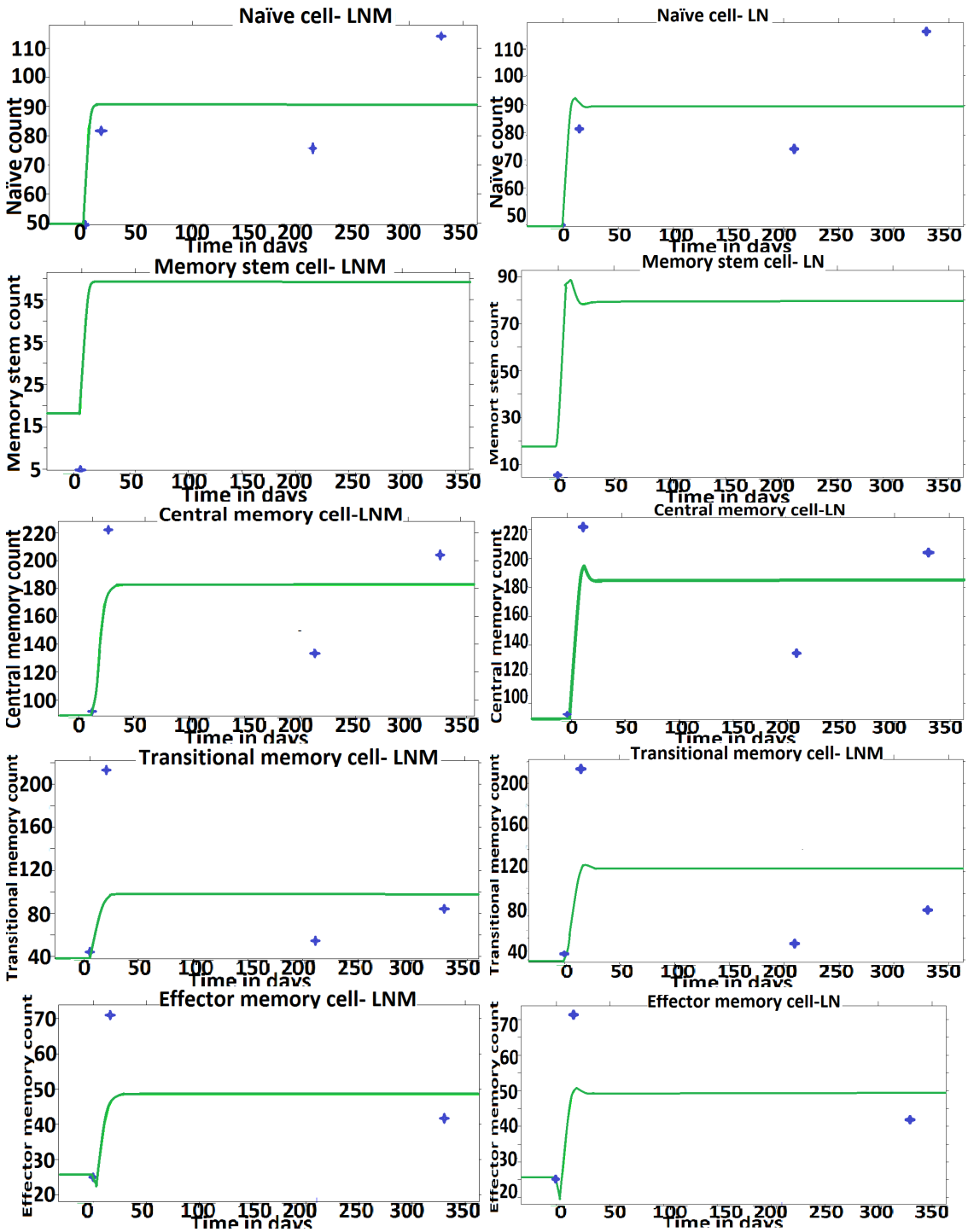

Figure 3.15: Individual fits for the naïve and memory CD4 T-cells for patient 304.

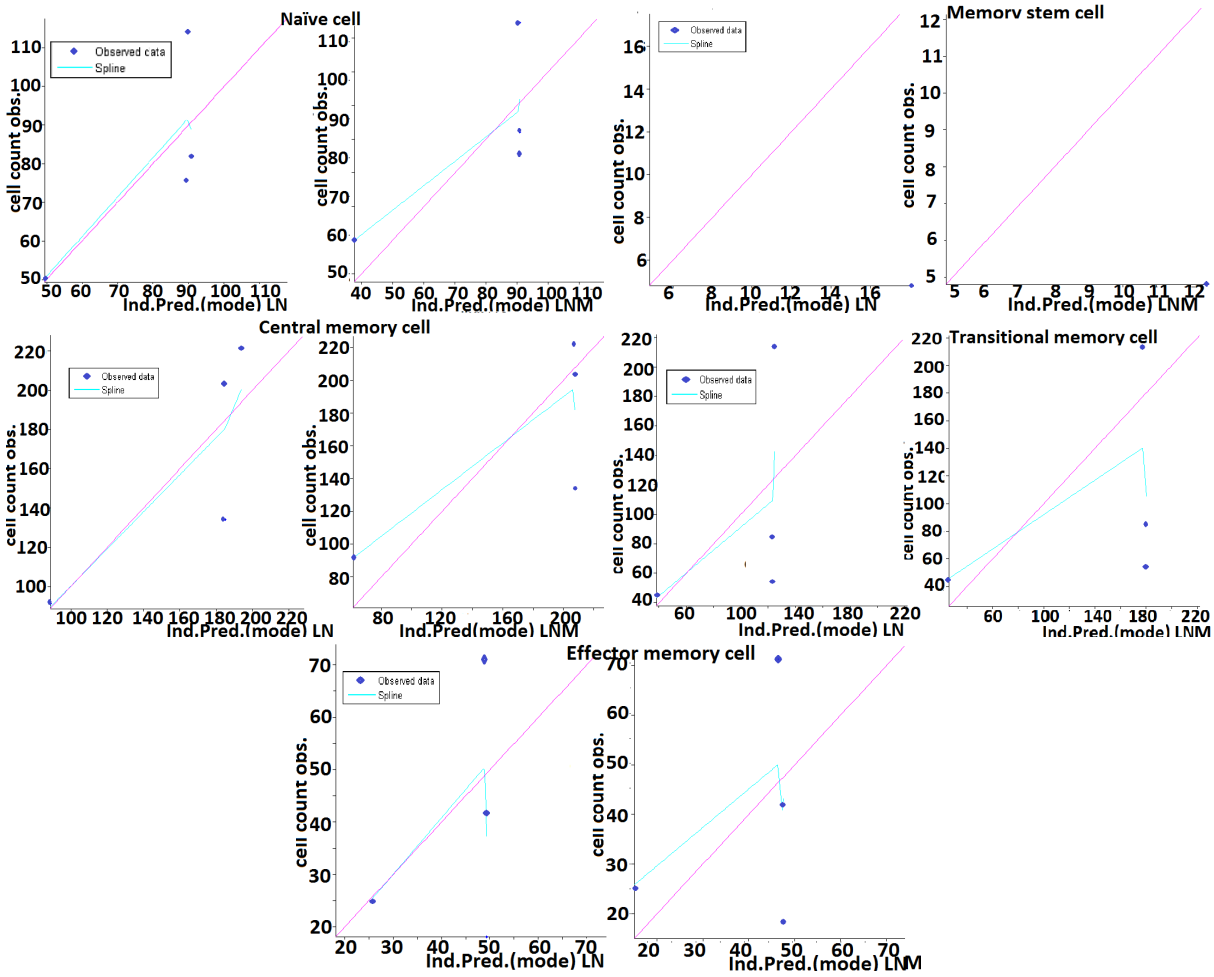

Figure 3.16: Observed vs prediction for patient 304.

Patient 305- High Dose

Table 3.11 shows the parameter estimates for patient 305. Similar to the patients above, fits obtained from using both a stochastic approximation or model linearization to calculate the Fisher Information Matrix had similar parameter values. Using When only non-modified cell populations the AIC value of 232.66 is the smallest compared to all the other fits.

The three fitting routines results in having a zero value for the memory stem cell death rate d_{ST}. This indicates that the death rate of the memory stem cell became negligible after the treatment. When the first two fitting routines failed to estimate the proliferation rates for most of the cell populations, the third one was able to do so. Looking at Figure 3.17, we can see that the naïve, central and effector cells had a good fit where the ODE curve passed through most of the data points. Figures 3.18 shows that the naïve and effector cell data points are closer to the 45^o purple line which means that the fit for these population was good. From Figure 3.17 we can conclude the following:

- The naïve and transitional memory cells decreased in cell count.
- The memory stem and effector cells increased in count.
- The transitional memory cell exhibit no change.

Param	**Lin-all**	**s.e**	**Stoch-all**	**s.e**	**Lin-non**	**s.e**
λ	12.56407	61	12.56407	8.5	8.96240	1.1e+006
p_N	0	0.064	0	7.5e-011	0.07369	6.2e+004
p_{ST}	0	0.29	0	2.2e-010	0.00396	7e+004
p_C	0	0.022	0	5.6e-011	0	0.039
p_T	0	0.21	0	6.6e-010	0.00141	5.7e+005
p_E	0.08672	4.1	0.08672	0.16	0.02939	1.3e+006
ϕ_N	0.13401	0.77	0.13401	0.075	0.13602	1.3e+003
ϕ_{ST}	0.51269	3.8	0.51269	0.58	0.48367	4.8e+004
ϕ_C	0.12176	1.5	0.12176	0.06	0.12000	2.3e+005
ϕ_T	0.31359	4.8	0.31359	0.31	0.32529	9.1e+005
d_N	0	0.21	0	2.1e-010	0.00918	3.2e+004
d_{ST}	0	0.8	0	1.9e-009	0	0.046
d_C	0.00082	2	0.8	0.00082	0.00148	2.3e+005
d_T	0.01461	8.4	0.01461	0.16	0.01267	1.4e+006
d_E	0.31986	6.1	0.31986	0.3	0.31500	1.7e+006
AIC	450.60		448.60		232.66	
CPU(sec)	4.13e+003		1.56e+004		2.48e+003	

Table 3.11: Illustrates the fits obtained from the three fitting approaches with their AIC and standard error measurements for patient 305.

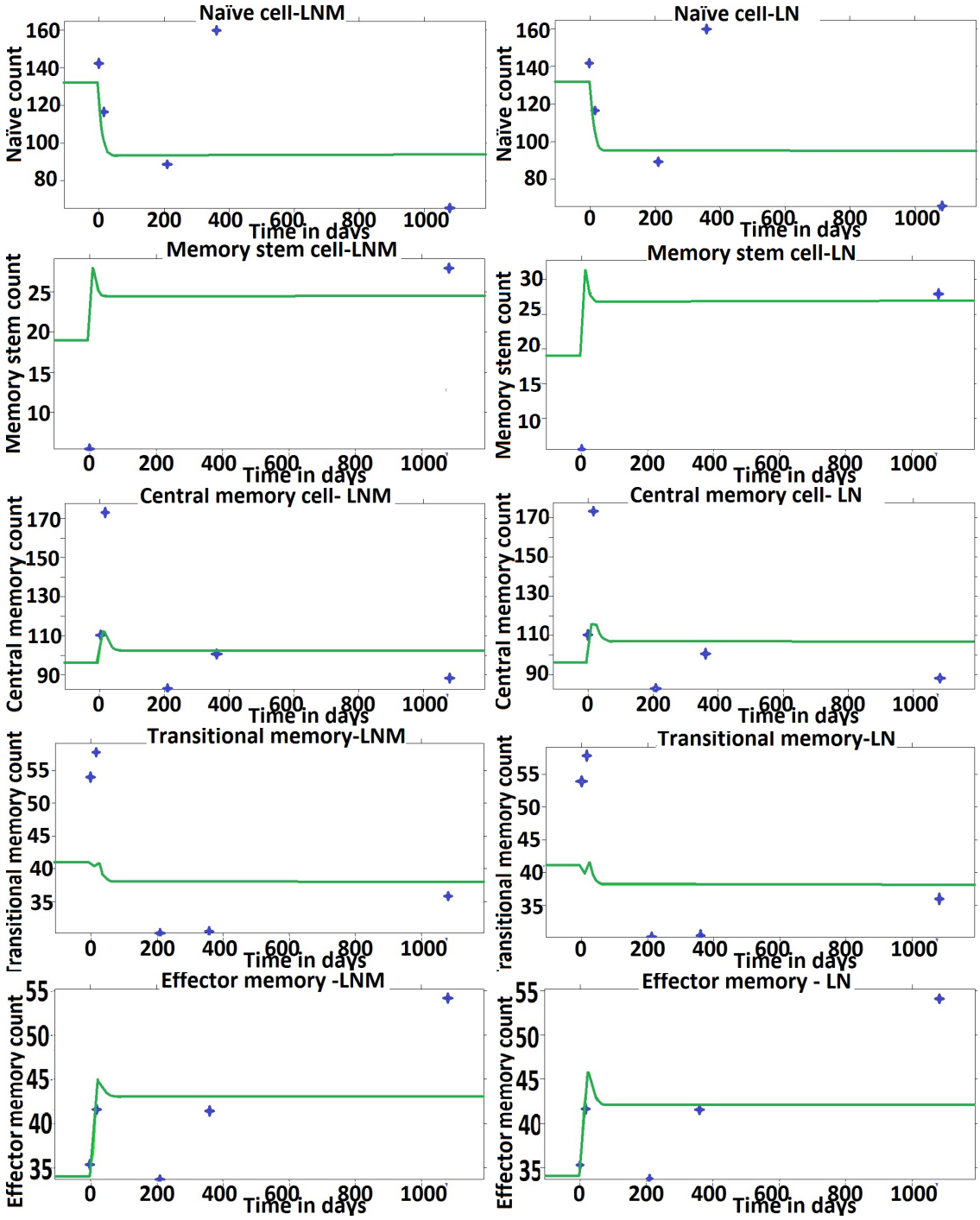

Figure 3.17: Individual fits for the naïve and memory CD4 T-cells for patient 305.

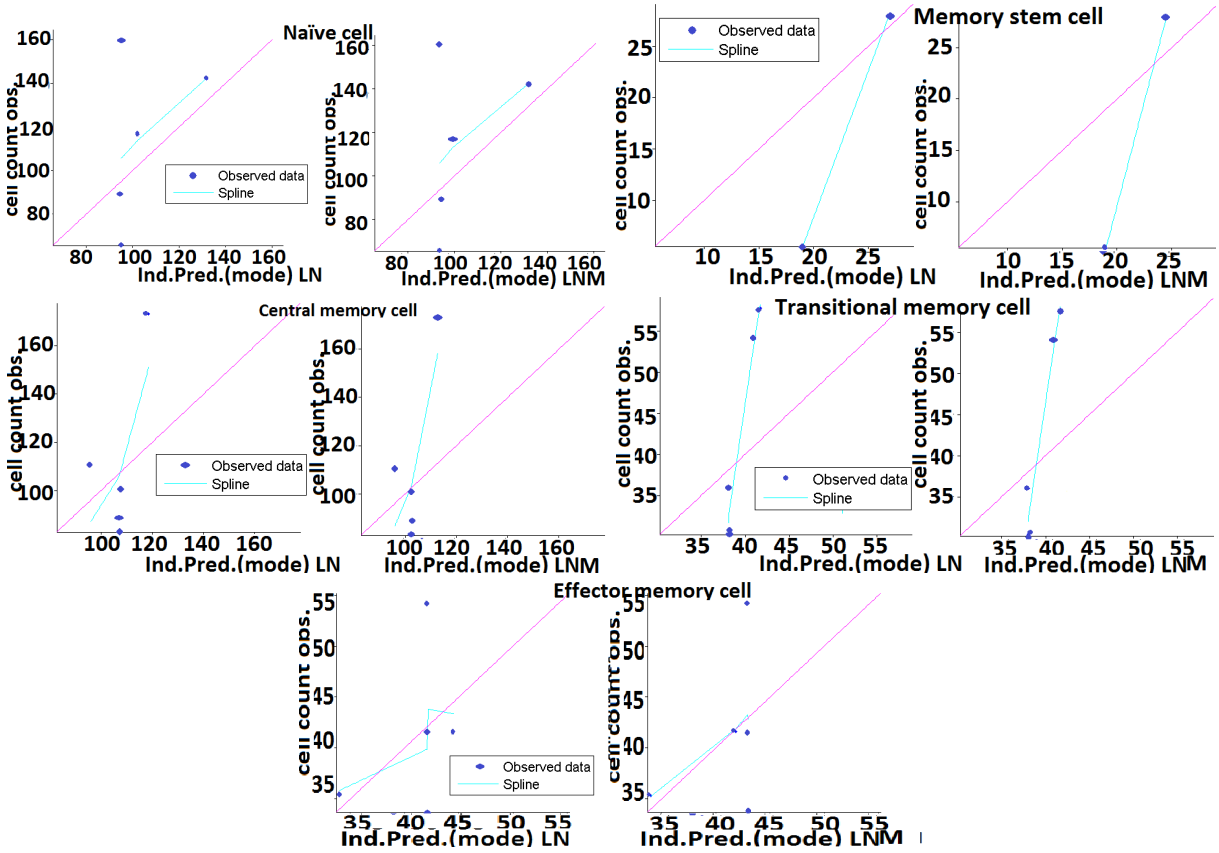

Figure 3.18: Observed vs prediction outputs for patient 305.

3.4 Parameter identifiablity, Over-fitting issues and Algorithm Convergence

One of the most common problems in complex biological models is that some parameters cannot be estimated by fitting the simplest proposed model. Using all the fitting routines in Monolix and Stan, we have faced several fitting problems where even if the parameters were identifiable they were not properly estimated and some parameters were not identifiable using the clinical data measurements.

When fitting the observations to our dynamical model, we had some parameters that were not identifiable such as the proliferation rates in some patients. This was evident in both the fits done in Monolix and Stan. These fitting issues could be a result of the assumptions we made when describing the transition from the naïve to the effector memory state as linear. In addition, in some cell populations for some patients, we had the over fitting issue where the number of unknown parameters exceeded the number of observations. The model specification was evident in the prediction vs observed Figures (3.3,3.5,3.7,3.9,3.12,3.14,3.16,3.17) where the spline and the data points were not always aligned with the 45^o line. This indicates that the model was not able to successfully replicate the clinical study.

An additional complication includes the fact that the fitting algorithm may get stuck in a local minimum or maximum rather than the global one. And, as we have

no prior knowledge about the parameters, it was challenging to give a closest initial guess in order to get a better fit.

In conclusion, we have seen that using the model linearization of the Fisher Information Matrix to calculate the standard error was the most suitable method. Restricting our model to the non-modified naïve and memory CD4 T-cell, the model was able to capture the data more successfully as it resulted in having smaller AIC values. As well, the observed vs prediction graphs showed that when ignoring the modified cell population, the data points were closer to the 45^o line, indicating a better fit.

4 Sensitivity and Uncertainty Analysis

4.1 Introduction

In chapter 3, we have presented several parameter estimation results using various fitting techniques. It is of interest to determine which parameters most affect the T-cell count at different observation points. We now use sensitivity and uncertainty analysis to determine such importance. In this chapter, using both Latin Hypercube Sampling (LHS) and Partial Rank Correlation Coefficient values (PRCC), we will explore the effect of each of the fifteen model parameters on the five naïve and memory CD4 T-cells along with the total CD4 T-cell counts over 200 days after the initiation of the treatment for each of the three cohorts. This chapter is organized as follow:

- In section 3.1 we will introduce the LHS and PRCC concepts.
- In section 3.2, we will present the observed changes in the proliferation, death and transition rates from baseline. Moreover, we will investigate the relative

significance of each of the model parameters in relation to each of the naïve and memory CD4 T-cell population size using the PRCC values.

4.2 Uncertainty Sensitivity Analysis: LHS and PRCC

LHS

Latin Hypercube Sampling, first proposed by Mckay in 1979, is a statistical sampling method that belongs to the Monte Carlo class of sampling method [41]. It is a stratified sampling without replacement techniques. Using a sample size N, the algorithm partitions the random parameter distributions into N equally probability intervals independently. A Latin Hypercube Sampling matrix is generated that consists of N rows and k columns, where k is the number of parameters and N is the number of run or simulations[41].

Partial Rank Correlation Coefficient- PRCC and Sensitivity Analysis

The PRCC is a correlation coefficient in statistics, known to measure the strength of the linear association between a given input x_j and an output y. It is defined by the ratio between the covariance of x_jy and sqrt of the product of the variance of x_j and y as illustrated in the equation below:

$$CC = \frac{COV(x_j y)}{\sqrt{Var(x_j)Var(y)}} \tag{4.1}$$

The PRCC value vary between -1 and +1, where -1 indicating a strong negative

correlation and the +1 a strong positive correlation. Any correlation less than +0.5 and greater than -0.5 is assumed to be a non significant correlation [41]. A positive correlation indicates that as we increase the model parameter the output will increase and similarly a negative indicates that as we increase the model parameter the output will decrease. Combining both of the PRCC and LHS, we are able to draw some conclusions about the significance of some unknown parameters on specific model outputs.

4.3 Sensitivity Analysis

In Chapter 3, results of the parameter estimations, have presented the fits obtained before the initiation of the treatment (at baseline), and the ones obtained after the initiation of treatment. The best two fits obtained from the Monolix and the fit obtained from the Stan software will be used in this section, in order to highlight any significant changes in the parameter values. This will better inform us how the perturbation by the CCR5-down-modulated memory CD4 T-cell affected any of the five naïve and memory CD4 T-cell sub-populations behaviors (birth, death, proliferation and transition). Later, using the PRCC values obtained by the sensitivity analysis, we analyze the correlation between the model parameters and each of the naïve, memory stem, central, transitional, effector and total memory CD4 T-cells using the deterministic model given by the Eq's 2.1 and 2.2. In this

work, we will explore the time dependent significance of the fifteen model parameters on the five naïve and memory CD4 T-cell subsets count over the course of 200 days. We have picked 200 days for these simulations as we have seen that the system populations reach a steady state approximately 80 days post initiation of the clinical trial. The PRCC variables are the fifteen model parameters and the output are the naïve, stem, central, transitional, effector and total memory CD4 T-cells. Throughout this work, we defined the total count for the naïve and memory CD4 T-cells as the sum of all the five sub-populations at time, t.

We used 10000 bins to run the LHS code in Matlab for all of the nine patients.In this work we will consider that any PRCC values smaller than 0.5 or bigger than -0.5 is said to be not significant[41]. A positive correlation indicates that the model variables and the output have a proportional relationship. This means that as we increase the parameter value the output will increase. A negative correlation indicates that the model parameter and the output have an inverse proportional relationship. This indicates that an increase in the parameter will yield to a decrease in the output. Throughout this chapter, we will be using the symbols to refer to each of the fifteen parameters as illustrated in Table 2.2.

4.3.1 Low Dose Cohort

The low dose cohort represents the three patients that received a single infusion of autologous CCR5-modified (SB-728-T) 1.0 x 10^{10} cells. All of the baseline measurements are measured seven days prior to the infusion of the CCR5 modified T-cells. The three patients had a mixed range of CD4 T-cell at baseline that varied between low, medium and high levels.

4.3.1.1 Patient 103

At baseline, this patient had the lowest count of 188 CD4 T-cell per μL. Looking at table 4.1 we can note the following observations about each of the five sub-population behaviors:

- The number of naïve CD4 T-cells λ produced by the thymus decreased.
- The proliferation rates for the central memory p_C and transitional memory p_T increased. The proliferation rate of the effector memory cells p_E has increased by one order of magnitude.
- The transition rates for all of the cells remained the same except for the transition rate of the transitional memory cell ϕ_T that had a slight increase.

- The death rates for the naïve d_N and stem memory cells d_{ST} have decreased by one order of magnitude. This means that their lifespans have increased considerably by 100 days.

- The death rates for the central memory d_C, transitional memory d_T and effector memory d_E cells have increased slightly.

As for the sensitivity analysis obtained over 200 days we observe the followings from Figures 4.1, 4.2 and 4.3:

Naïve Cell:

In Figures 4.1, we can see that although the thymic production rate λ has a significant effect soon after the treatment, but this effect started to fade 20 days post treatment. The proliferation rate of the naïve cell p_N, has a significant positive correlation with the naïve cell population. However, the transition rate ϕ_N has a negative correlation, where if the naïve cells transition to the memory stem cell in a higher rate the number of the naïve cell will decrease. These findings are expected as we had a decrease in the death rate d_N, thus the proliferation rate p_N is able to exhibit a positive correlation that results in an increase in the naïve cell count.

Memory Stem cell

The naïve cell proliferation rate p_N has a positive correlation with the memory stem cell count. This could be explained by the fact that as the number of the naïve cell increase, a higher number of naïve cells are going to transition to the memory stem cell state. The transition rate of the transitional memory cell ϕ_{ST} has a positive correlation. Moreover, for the first 20 days post treatment, the effector memory cell death rate d_E had a positive correlation with the memory stem cell count. This result could indicate a possible transition between the effector, transitional memory cell and the memory stem cells. Finally the transition rate of the memory stem cell population ϕ_{ST} has a negative correlation with the stem cell population count which is an expected correlation as the more the cells leave the memory stem state its population count will decrease.

Central Memory Cell

The central memory cells had similar results to the memory stem cell population where the naïve proliferation rate p_N had a positive correlation. This could indicate a possible backward transition between the memory stem and central memory cells. Similarly to the above cell populations, the transition rate of central memory cells phi_C has a negative correlation with the central memory population count.

Transitional Memory cell

Consistent with the results of memory stem and central memory cells, proliferation of naïve cells p_N is positively correlated with transitional cell population. Similarly the transition rate of the transitional memory cell population ϕ_T is negatively correlated with its cell population which indicates that the more the cells are leaving the transition state the lower the population count is becoming.

Effector Memory cell

Similarly, the effector cells were positively correlated with the naïve cell proliferation rate p_N. As for the first 50 days post injection of the down-modulated memory CD4 CCR5 cells, both the transition rates of the transitional memory ϕ_T and central memory cells ϕ_C have a positive correlation with the effector memory cell population. This indicates that the transitional and central memory cells play an important role in increasing the effector memory cells population count. Lastly, the death rate of the effector memory cell d_E has a significant negative correlation with the effector memory cell count. This result is consistent with the parameter estimation as the effector memory cell death rate increased after the initiation of the clinical trial in this patient.

Total naïve an memory CD4 T-cell

The total number of naïve and memory CD4 T-cells were negatively correlated with the naïve cell transition rate ϕ_N and positively correlated with the naïve cell

proliferation rate p_N and the effector memory cell death rate d_E. This indicates that for patient 103, the naïve cell population plays an important role in the reconstitution of the total CD4 T-cell. Furthermore, an increase in the effector memory cell death d_E can indicate a possible increase in their activity against the HIV virus.

Parameters	**Baseline**	**Non-modified T-cell**	**SATN fit**	**Min-Max**
λ	9.171294443	6.13933	2.7117	2.7117-9.1712
p_N	0.001154901	0.000201	0.0109	0.000201-0.0109
p_{ST}	0.011031316	0.00640	0.2590	0.0064-0.2590
p_C	0.010900323	0.08209	0	0-0.09
p_T	0.021644028	0.06390	0.0887	0.021644-0.0976
p_E	0.033551917	0.41340	0	0-0.483
ϕ_N	0.215314186	0.21732	0.0973	0.0973-.21732
ϕ_{ST}	0.463596696	0.46060	0.3590	0.3590-0.46359
ϕ_C	0.105280147	0.11722	0.0373	0.0373-0.117
ϕ_T	0.247796071	0.33989	0.1886	0.18866-0.33989
d_N	0.01040803	0.00608	0.0136	0.00565-0.0136
d_{ST}	0.01108572	0.00685	0	0-0.011085
d_C	0.003333333	0.00564	0.0627	0.00249-0.0627
d_T	0.023361168	0.03274	0	0-0.03274
d_E	0.375021515	0.49776	0.1	0.1-0.49776

Table 4.1: Illustrates the data fitting results and the range used in the LHS for pat 103. These values are the rates per day^-1.

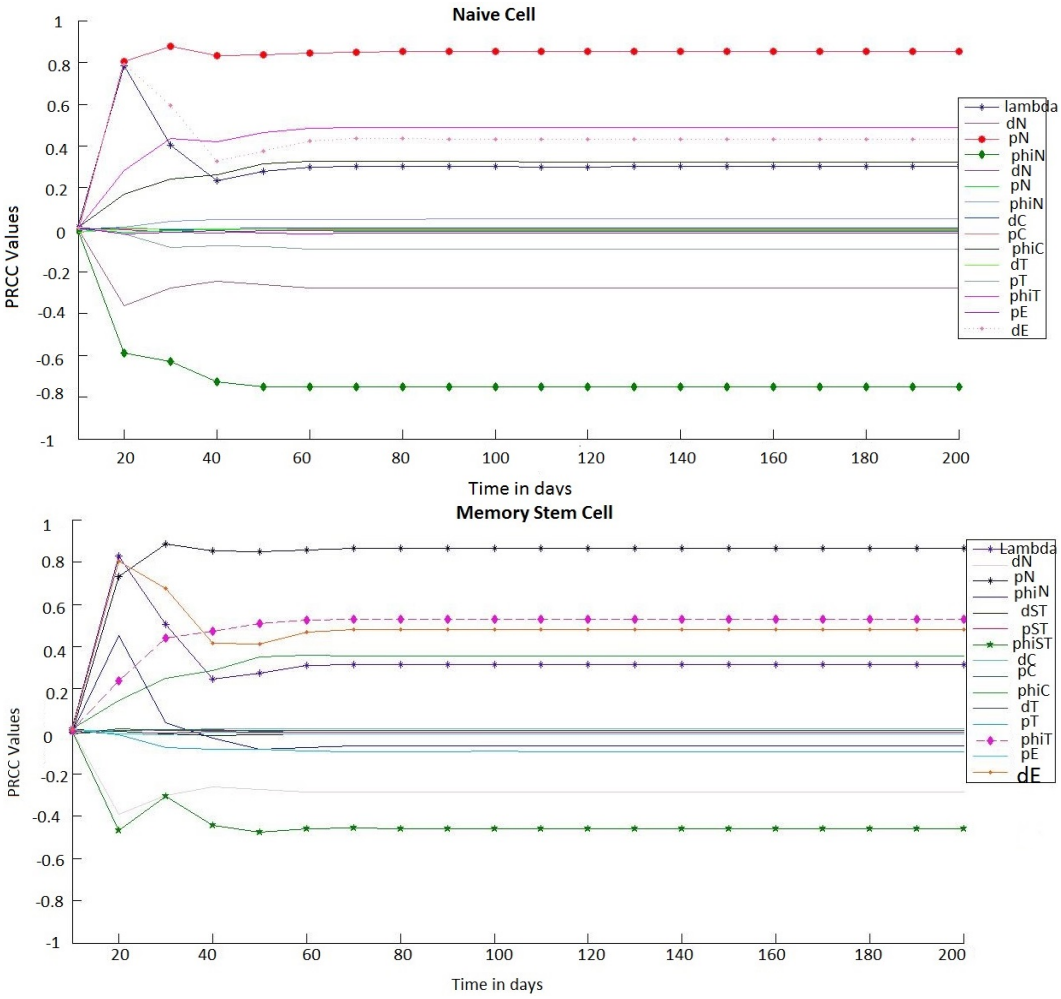

Figure 4.1: Correlation between the N, STM T-cell and the 15 parameters over 200 days for pat 103.

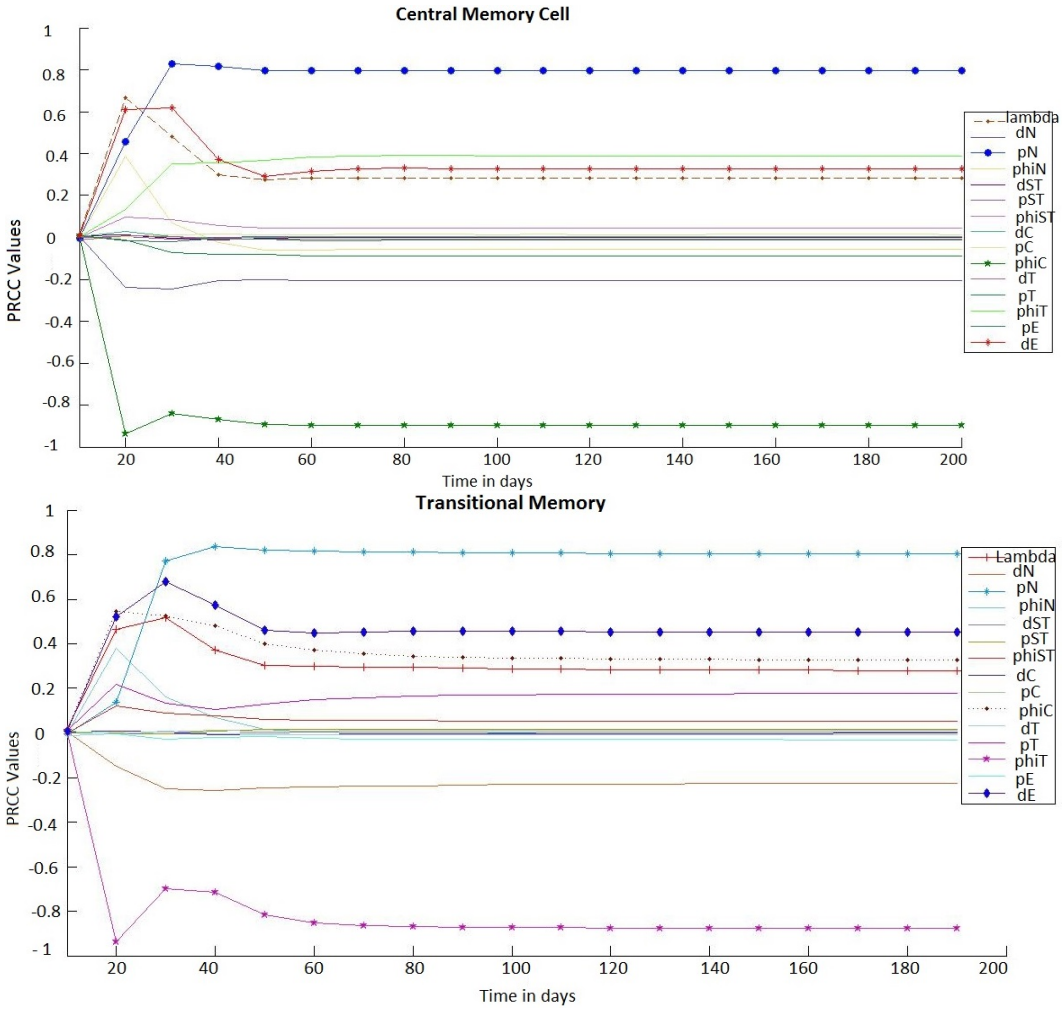

Figure 4.2: Correlation between the CM and TM T-cell and the 15 parameters over 200 days for pat 103.

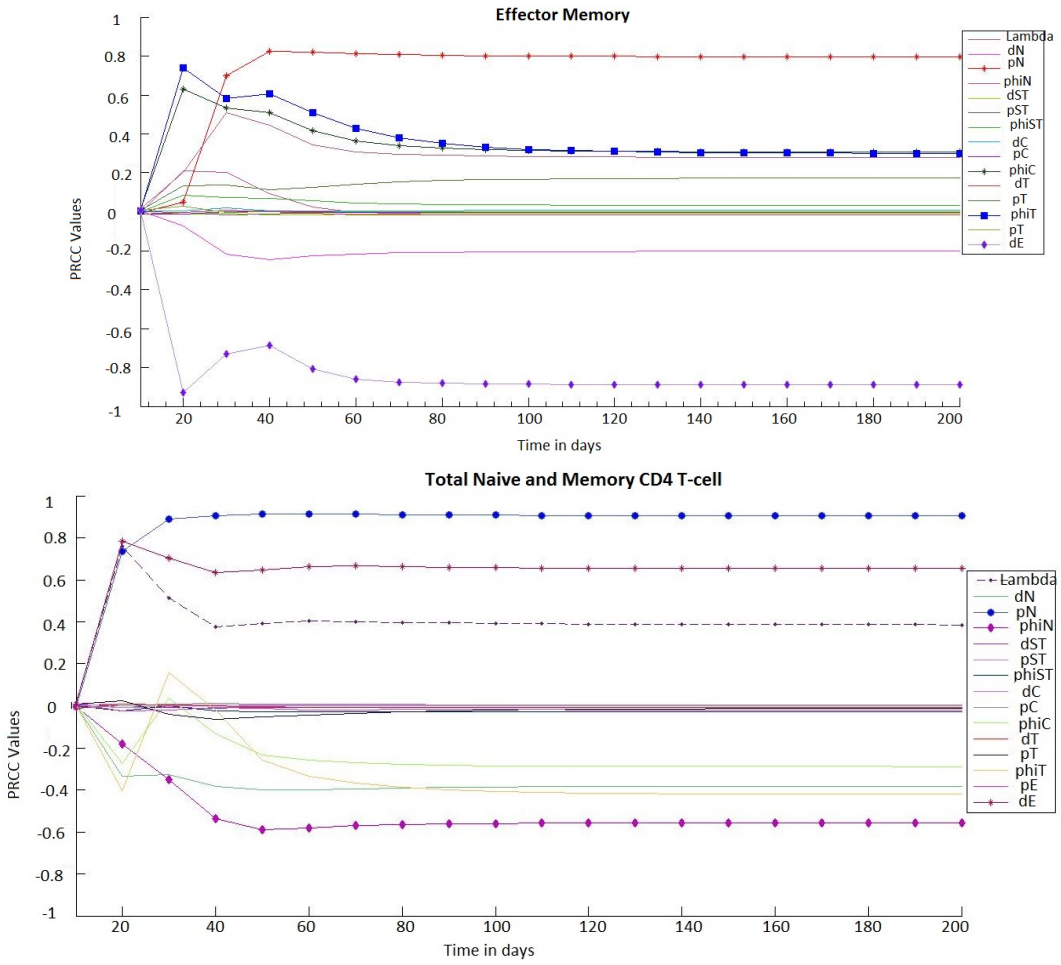

Figure 4.3: Correlation between the EM and total CD4 T-cell and the 15 parameters over 200 days for pat 103.

4.3.1.2 Patient 102

At baseline, this patient had a somewhat high count of 439 CD4 T-cell per μL. Looking at Table 4.2 we can note the following observations about each of the five sub-population behaviors:

- The number of naïve CD4 T-cells produced by the thymus λ increased.
- The proliferation rates for the naïve p_N, memory stem p_{ST} and effector memory cells p_E increased by one order of magnitude.
- The proliferation rates for the transitional memory cell p_T decreased by one order of magnitude.
- The transition rate of the central memory cells ϕ_C decreased by one order of magnitude.
- The transition rate of the memory stem cell ϕ_{ST} increased by half its value.
- The death rates for the naïve d_N and stem memory celld_{ST} decreased by one order of magnitude. This indicates that their lifespan increased by 100 days considerably.
- The death rate for the effector memory cell d_E increased by one order of magnitude.

As for the sensitivity analysis obtained over 200 days we observe the followings from Figure 4.4 and 4.5 and 4.6:

Naïve Cell:

The sensitivity analysis agrees well with the parameters estimations results presented in Table 4.2. The parameters λ and p_N have a positive correlation with the number of naïve cell population. The transition rate of the naïve cell ϕ_N has a negative correlation.

Memory Stem cell

While the naïve cell thymic production rate λ, and both the memory stem p_{ST} and naïve cell proliferation p_N rates are positively correlated with the memory stem cell population, the transition rate ϕ_{ST} has a negative correlation. This indicates that the naïve cells play an important role in the memory stem cell population.

Central Memory Cell

The importance of the naïve cell production rate from the thymus λ continues to exhibit a positive correlation with the central memory cell population in this patient. The proliferation rate for the naïve p_N and memory stem cell p_{ST} started to display a positive correlation with the central memory cell population about 30 days post injection of the clinical trial. This could be justified by the delayed effect

of the modified memory CD4 T-cell, as before the treatment patient 102 had a low proliferation rates for the naïve p_N and memory stem cell populations p_{ST}. Both the death and transition rates for the central memory cells d_C, ϕ_C have a negative effect on the central memory cell count.

Transitional Memory cell

Similar to the naïve cell production rate λ, along with the naïve and memory stem cell proliferation rates p_N and p_{ST} have a positive effect on the transitional cell population count. Moreover, the transition rate of the central memory cell ϕ_C has a positive correlation as well, indicating the importance of the central memory count in replenishing the transitional memory cell population, which explains the negative correlation between the central memory cell death rate d_C and the transitional memory cell population. Finally, as the proliferation rate p_T decreased after the infusion of the modified memory CD4 T-cells, the transition rate of the transitional cell ϕ_T displayed a negative correlation with the transitional memory cell population.

Effector Memory cell

The importance of the naïve, memory stem and central memory cells persist as well in the effector memory cell count. This is shown by the positive correlation of the naïve cell production rate λ, the central memory cell transition rate ϕ_C and prolif-

eration of both the naïve p_N and memory stem p_{ST} cell with the effector memory cell population. In addition to the above positive correlations, the importance of the central memory cells is evident in the negative correlation between the central memory death rate d_C and the effector memory cell population. These observations indicate that the central memory, naïve and memory stem cell play an important role in the effector memory cell population. As the death rate of the effector memory cell d_E has a significant increase, its negative correlation with the effector memory population is also observed.

Total naïve an memory CD4 T-cell

The important effect of the naïve and memory stem cells on the replenishment of the total number of CD4 T-cell is apparent by the positive correlation of the naïve cell production rate λ, along with the proliferation rates of both the naïve p_N and memory stem cell p_{ST}. This justifies the negative correlation between the transition rates of the memory stem cell ϕ_{ST}, naïve cell ϕ_N and the central memory cell ϕ_C with the total number of CD4 T-cells. Lastly, the death rate of the central memory cell d_C was negatively correlated with the total number of CD4 T-cell illustrating the importance of the central memory cell. These results indicates that the memory stem cell, central memory and naïve cell play an important role in increasing the total number of CD4 T-cell population.

Parameters	**Baseline**	**Non-modified T-cell**	**SATN fit**	**max-min**
λ	9.171294443	13	7.4924	7.4924-18.2
p_N	0.001154901	0.087105	0.0200	0.0011549-0.0871
p_{ST}	0.011004862	0.10033	0.1375	0.01100-0.159
p_C	0.010980541	0.03179	0	0-0.0511
p_T	0.021644028	0.0018056	0.0880	0.000167-0.0880
p_E	0.033551917	0.18993	0	0-0.268
ϕ_N	0.215314186	0.19645	0.1171	0.1171-0.227
ϕ_{ST}	0.463596696	0.23481	0.2375	0.235-0.463596696
ϕ_C	0.105280147	0.096345	0.0281	0.0281-0.10528
ϕ_T	0.247796071	0.36969	0.1880	0.1880-0.319
d_N	0.01040803	0.0071514	0.0028	0.0028-0.010408
d_{ST}	0.01108572	0.004742	0	0-0.011085
d_C	0.003333333	0.004742	0.0719	0.00333-0.0719
d_T	0.023195675	0.020894	0	0-0.0408
d_E	0.09096808	0.31554	0.1	0.0909-0.31554

Table 4.2: Illustrates the data fitting results for pat 102 along with the PRCC range.

These values are the rates per $day^{-}1$.

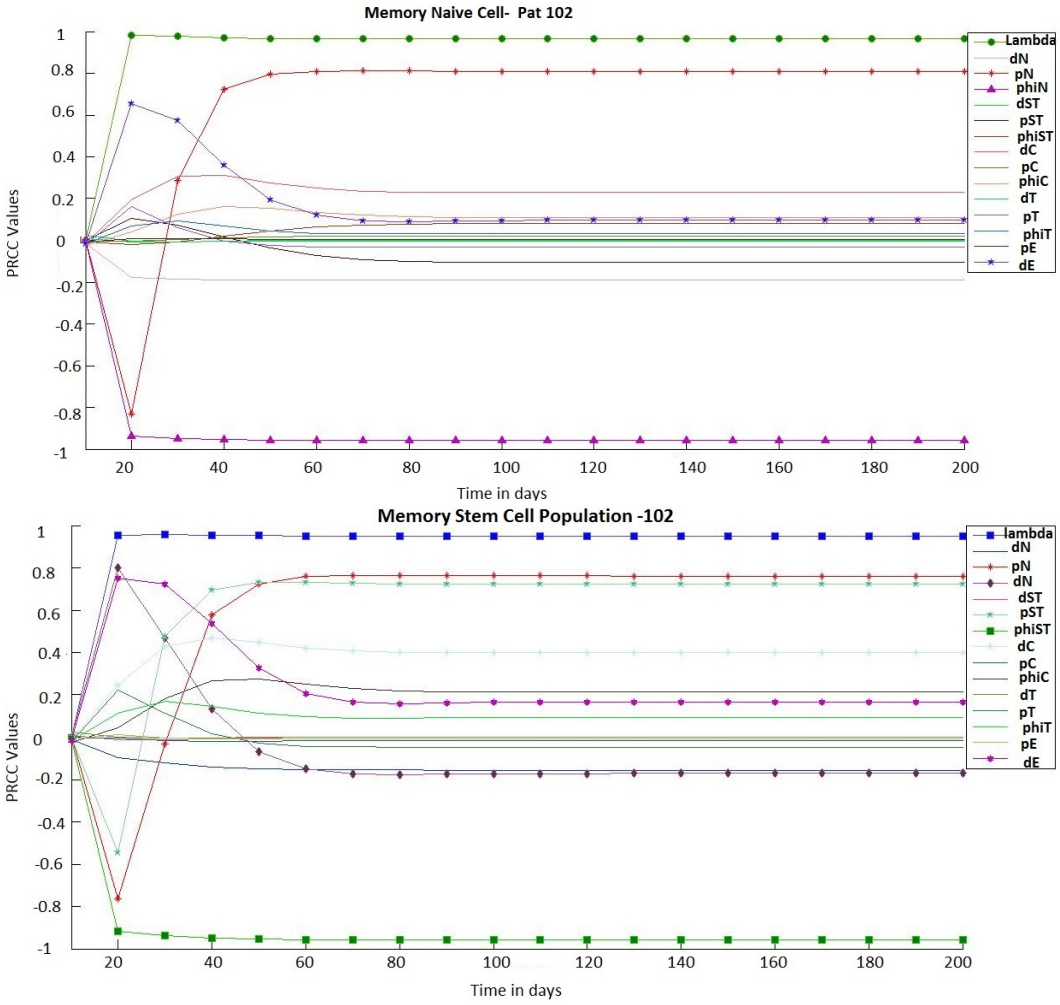

Figure 4.4: Correlation between the N and STM T-cell and the 15 parameters over 200 days for pat 102.

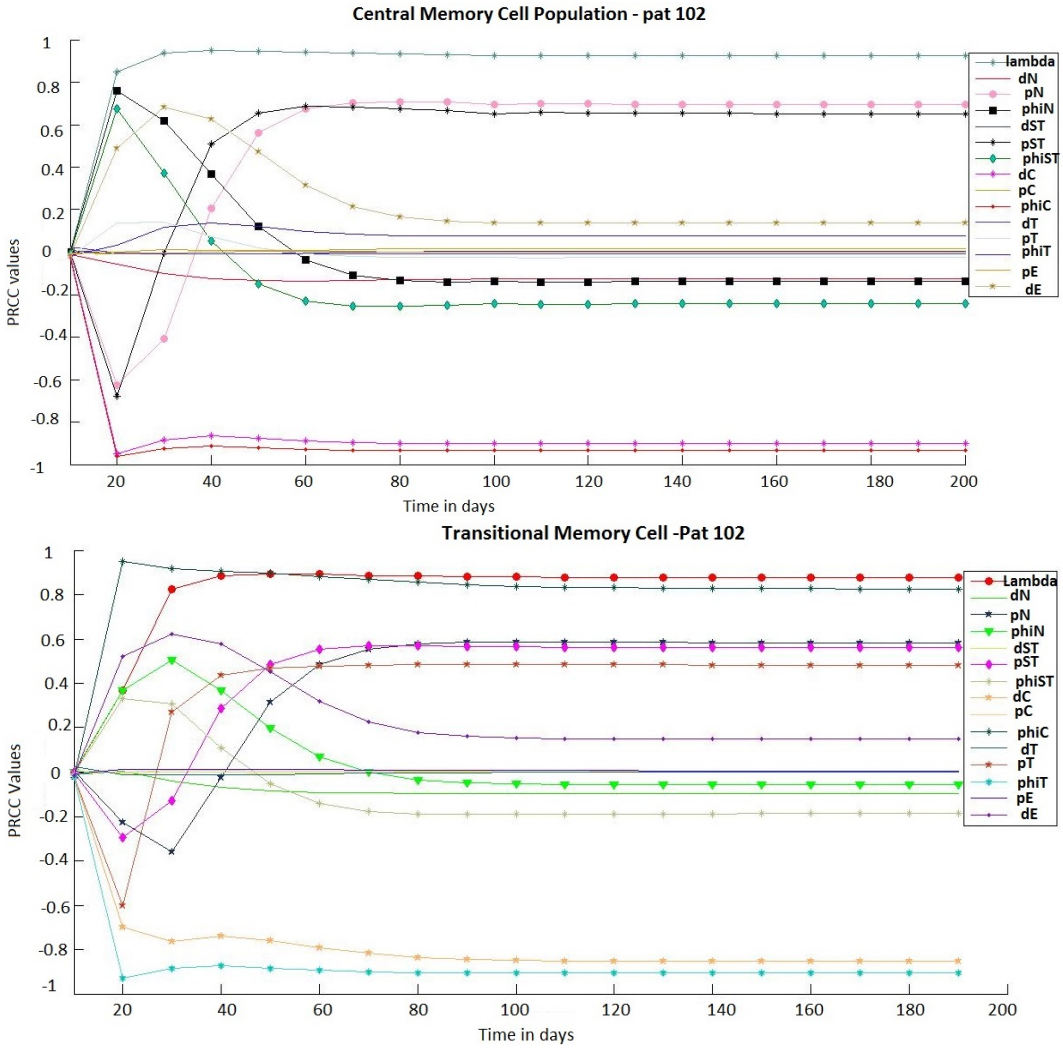

Figure 4.5: Correlation between the CM and TM T-cell and the 15 parameters over 200 days for pat 102.

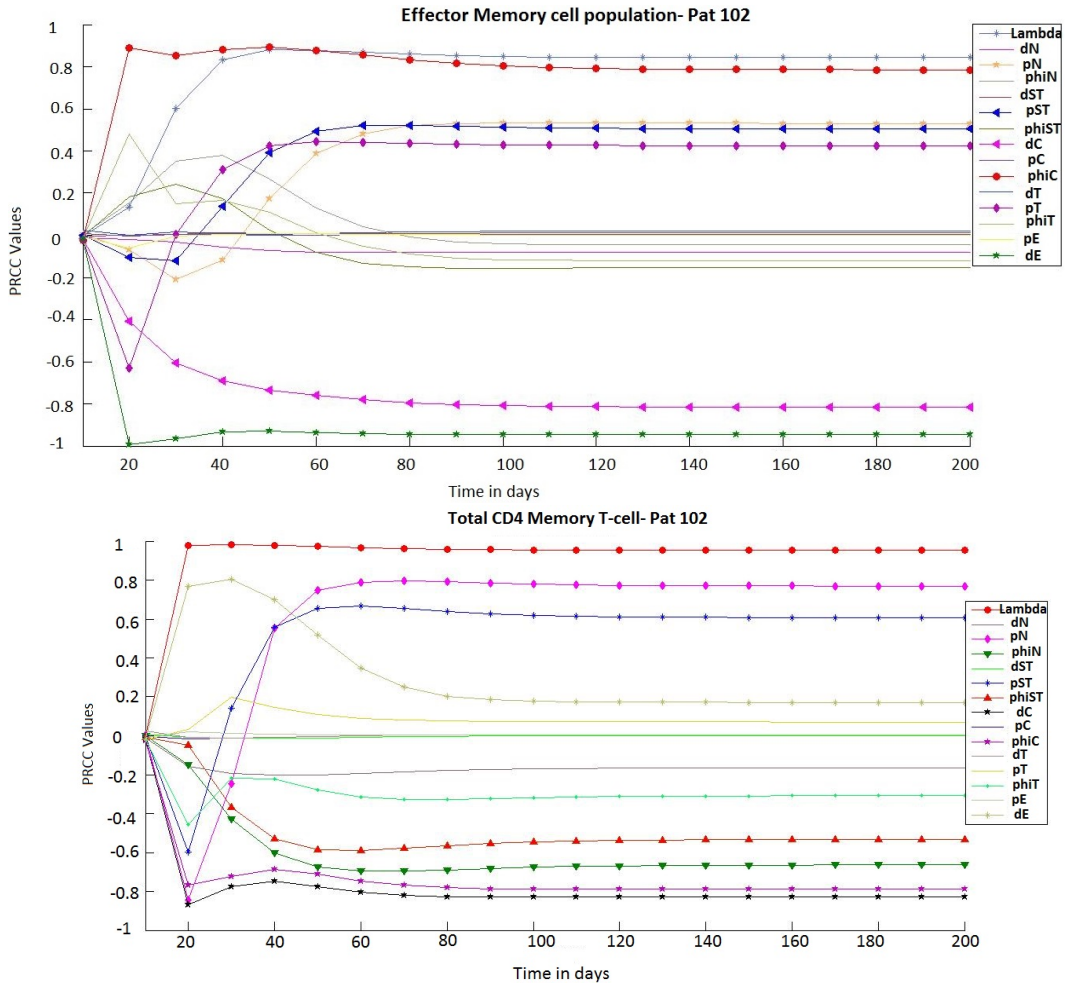

Figure 4.6: Correlation between each the EM and total CD4 T-cell and the 15 parameters over 200 days for pat 102.

4.3.1.3 Patient 104

At baseline, this patient had a low count of 261 CD4 T-cell per μL. Looking at Table 4.3 we can note the following observations about each of the five sub-population behaviors:

- The number of naïve CD4 T-cells produced by the thymus λ slightly increased.
- The proliferation rates for the naïve p_N, memory stem p_{ST} and central memory cells p_C have decreased.
- The transition rates for the naïve ϕ_N, memory stem $\phi_{ST}$and central memory cells ϕ_C decreased by one or two order of magnitude.
- The death rate for the memory stem cell d_{ST} decreased by one order of magnitude. In addition, the death rates for both the naïve p_N and transitional memory cells p_T slightly decreased.
- The central memory d_C and effector memory cells d_E have a doubled death rate from baseline.

As for the sensitivity analysis obtained over 200 days we observe the following from Figure 4.7, 4.8 and 4.9:

Naïve Cell:

While the thymic production rate λ has a positive correlation with the naïve cell population, the naïve cell transition rate ϕ_N has a negative correlation.

Memory Stem cell

Whereas the thymic production rate λ and the memory stem cell proliferation p_{ST} rate both display a positive correlation with the memory stem cell population, the transition rate of the stem cell ϕ_{ST} had a negative correlation. As the memory stem cell death rate d_{ST} decreased by one order magnitude after the initiation of the clinical trial, it was expected that a negative correlation with the memory stem cell population would be observed.

Central Memory Cell

The significant effect of the thymic production rate λ on the central memory cell population is not detected as the PRCC value is 0.4 indicating a trivial positive correlation. This is due to the low increase in λ. The proliferation rate of the central memory cell p_C and the transitional cell transition rate ϕ_T both display a positive correlation with the central cell population. The death rate of the central memory cell d_C have a negative correlation with the central memory cell as the death rate increased after the initiation of the treatment.

Transitional Memory cell

Proliferation rates of both the transitional memory p_T and central memory cell p_C were positively correlated with the the transitional cell population. The central memory death rate d_C is shown to have a negative correlation with the transitional cell population. However, as the death rate of the transitional cell d_T did not exhibit any change after the initiation of the treatment it did not have any significant negative correlation with the transitional memory cell population. In addition, the transition rate of the transitional cell ϕ_T displays a negative correlation with the transition cell population.

Effector Memory cell

Both the central d_C and effector memory cells d_E death rates have a significant negative correlation with the effector memory cell population as these rates increased after the introduction of the down modulated CCR5 CD4 T-cells. The proliferation rates of both the central memory p_C and transitional memory p_T cell display a positive correlation with the effector cell population. These results highlight the importance of both the transitional and central memory cells on the reconstitution of the effector memory cell population.

Total naïve an memory CD4 T-cell

The central memory cell population is shown to play an important role in most of the memory CD4 T-cell populations as illustrated above. Hence, it is expected that we continue to observe such importance in the total number of CD4 T-cell. The sensitivity analysis highlights the importance of the central memory cell population on the reconstitution of the CD4 T-cell where the proliferation rate of the central memory cell p_C is positively correlated with the total number of CD4 T-cell and the central memory cell death rate d_C has a negative correlation.

Parameters	**Baseline**	**Non-modified T-cell**	**SATN fit**	**max-min**
λ	7.696589511	8.73264	6.4731	6.4731-9.44
p_N	0.001155414	0	0.0413	0-0.0413
p_{ST}	0.010973075	0.00551	0.0213	0.004171-0.0213
p_C	0.051117	0.0036500	0	0-0.05117
p_T	0.021644028	0.00181	0.0590	0.00181-0.0590
p_E	0.033596587	0.01390	0	0-0.03359
ϕ_N	0.267280553	0.12537	0.1401	0.125-0.26728
ϕ_{ST}	0.467331743	0.00679	0.1158	0.00679-0.467331
ϕ_C	0.105280147	0.07149	0.0256	0.0256-0.10258
ϕ_T	0.249489254	0.32044	0.1584	
d_N	0.008914073	0.00261	0.0012	0.0012-0.008914
d_{ST}	0.01108572	0.00836	0.0054	0.0054-0.011085
d_C	0.003333333	0.0715	0.0744	0.003333-0.0744
d_T	0.023187029	0.01873	0.0007	0.0007-0.01873
d_E	0.288952838	0.47502	0.1	0.1-0.47502

Table 4.3: Illustrates the data fitting results for pat 104. These values are the rates per $day^{-}1$.

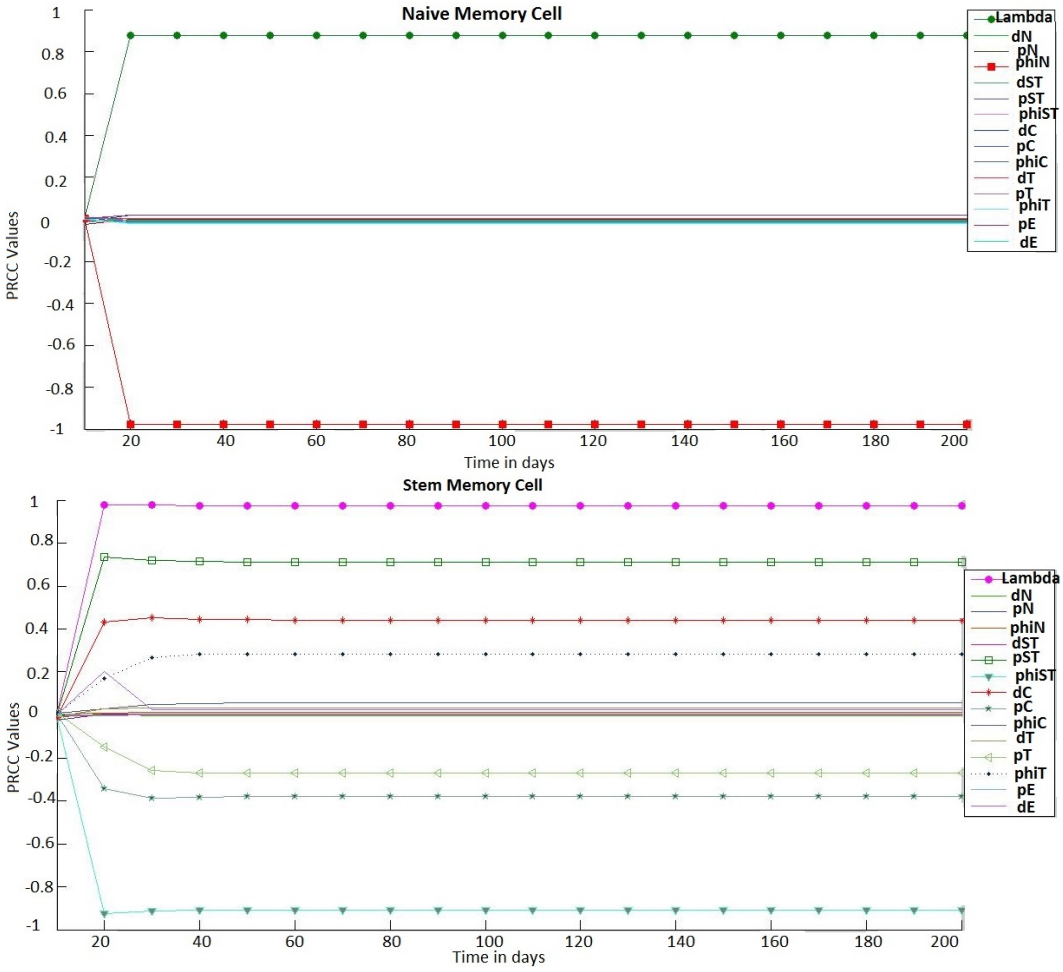

Figure 4.7: Correlation between the N and STM T-cell and the 15 parameters over 200 days for pat 104.

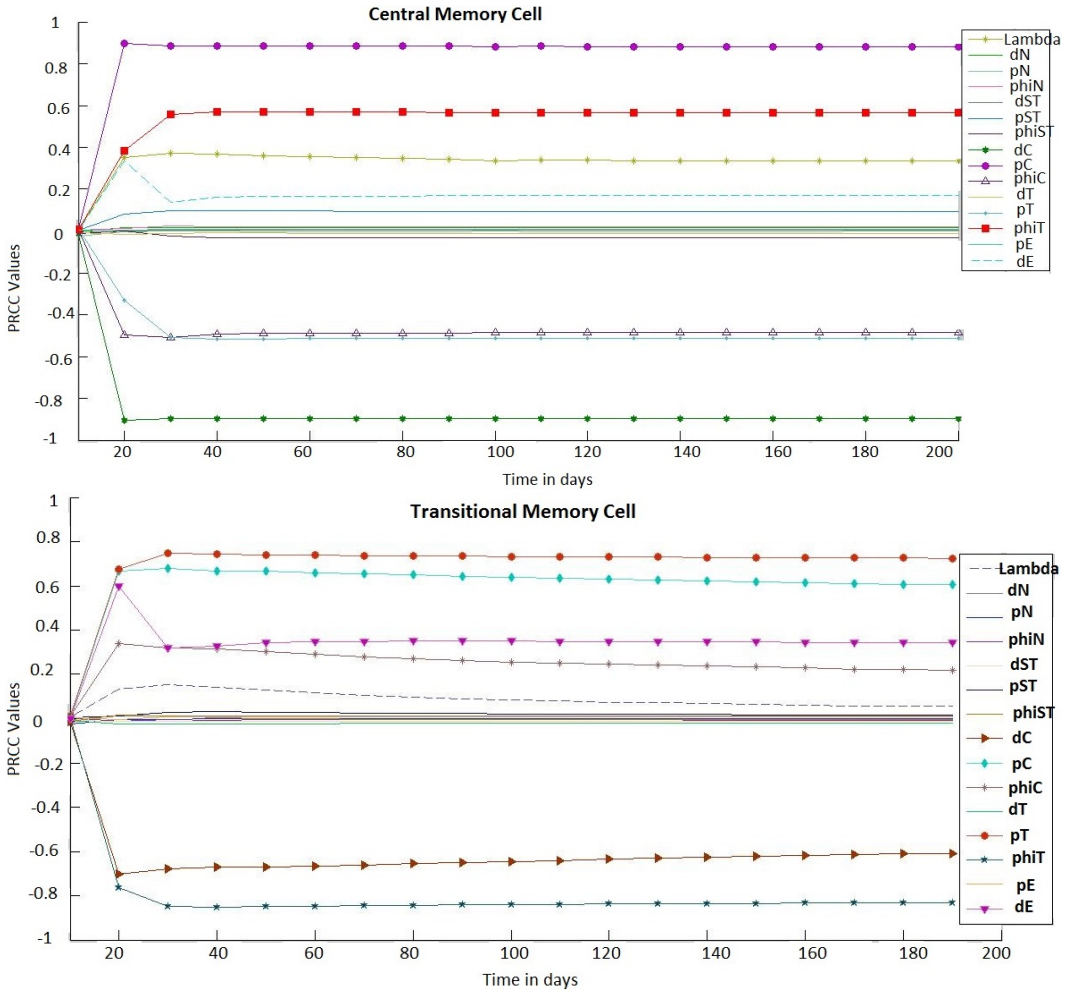

Figure 4.8: Correlation between the CM and TM T-cell and the 15 parameters over 200 days for pat 104.

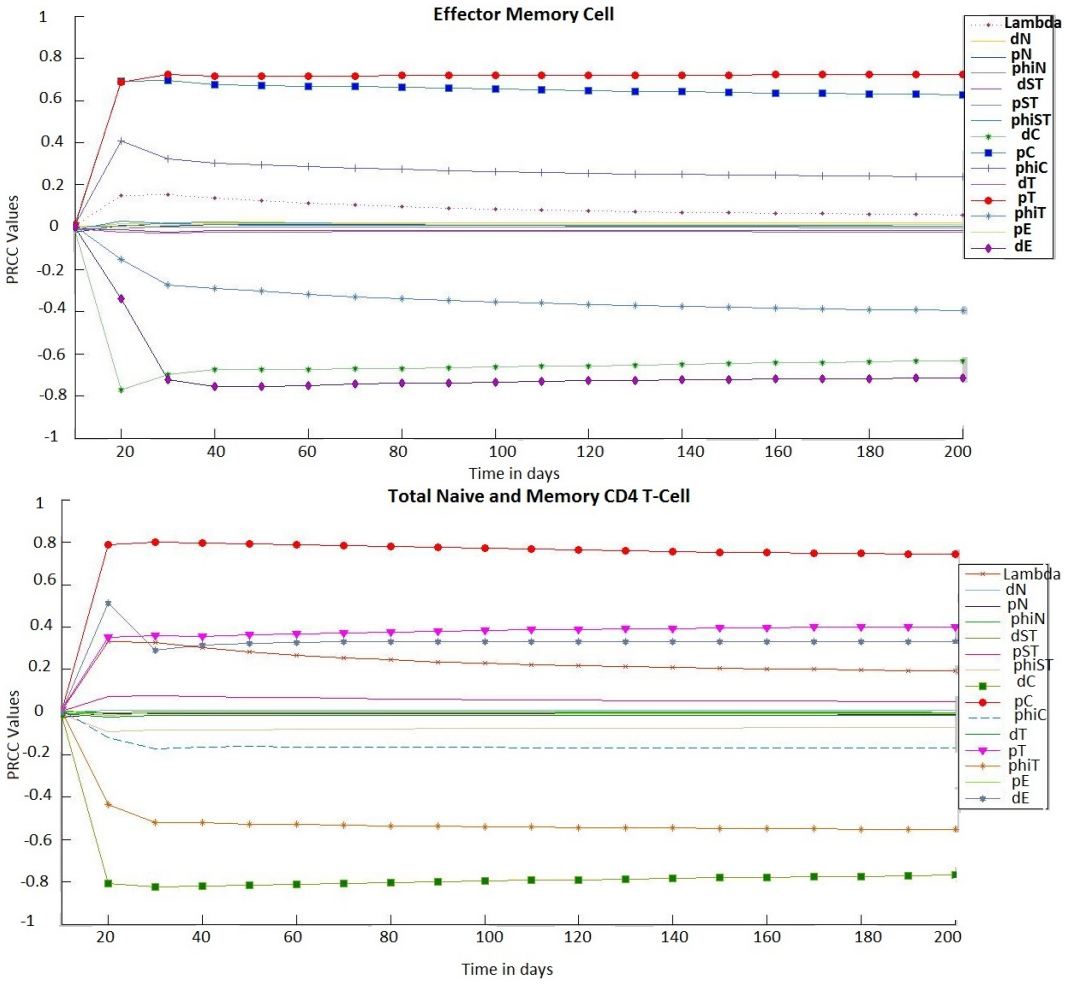

Figure 4.9: Correlation between the EM and total CD4 T-cell and the 15 parameters over 200 days for pat 104.

4.3.2 Medium Dose Cohort

The Medium dose cohort represents the three patients that received a single infusion of autologous CCR5-modified (SB-728-T) 2.0 x 10^{10} cells. All of the baseline measurement are measured seven days prior to the infusion of the CCR5 modified T-cells.

4.3.2.1 Patient 203

At baseline, this patient had a low count of 294 CD4 T-cell per μL. Looking at Table 4.4 we can note the following observations about each of the five sub-population behaviors:

- The number of naïve CD4 T-cells produced by the thymus λ did not have a large increase.
- While the proliferation rates for the naïve p_N and memory stem p_{ST} increased by at least one order of magnitude, the proliferation for the rest of the memory CD4 T-cell subsets decreased.
- The transition rate for the transitional memory cell ϕ_T has a significant increase.

- The death rate for the naïve d_N, memory stem cell $d_S T$, central d_C and transitional memory cells d_T have a very significant decrease.

- The effector memory cell death rate d_E doubled.

As for the sensitivity analysis obtained over 200 days we observe the followings from Figure 4.7 and 4.8:

Naïve Cell:

Similar to patient 104, the naïve cell population was positively correlated with the naïve cell production rate λ and negatively correlated with the naïve cell transition rate ϕ_N. It is important to mention that patients 104 and 203 have similar baseline count of the CD4 T-cell before they received the experimental treatment. As the death of the naïve cell d_N did not experience any significant change after the initiation of the clinical trial, in the sensitivity analysis we observed that it did not have any significant correlation rate with the naïve cell population.

Memory Stem cell

The naïve cell production rate λ display a positive correlation with the stem cell population. The transition rate of the naïve cell ϕ_N has a positive correlation only for the first 30 days post treatment. This observation could be justified by the fact

that the transition rate of the naïve cell population did not exhibit any significant increase post treatment. The memory stem cell transition rate ϕ_{ST} is negatively correlated with its cell count. The death rate of the memory stem cell d_{ST} did not show any negative correlation as this rate decreased notably from baseline.

Central Memory Cell

The naïve cell production rate λ maintains its positive correlation even with the central memory cell population. Similarly both the transition rates for the naïve ϕ_N and memory stem cell ϕ_{ST} were only positively correlated for 30 days post treatment. This is a valid observation as both of those transition rates did not have any significant increase from the baseline. The transition rate of the central memory cell ϕ_C had a negative correlation with its population. The death rate d_C did not express any negative effect as it decreased after the introduction of the down modulated CCR5 CD4 T-cells. The transitional and effector memory population have similar results to the central memory cells. However, the death rate of the effector memory cell d_E showed a strong negative correlation with the effector memory cell population as this rate increased notably after the infusion of the modified CD4 T-cell.

Total naïve an memory CD4 T-cell

As a result of the above observations, only the production rate of the naïve cell by the thymus λ has a positive effect on the reconstitution of the total number of CD4 T-cell. The transition rates for central memory ϕ_C and naïve cells ϕ_N had a strong negative correlation with the total CD4 T-cell count. This indicates that the central memory and naïve cell population are the main sub-populations that contribute to the total CD4 T-cells in this patient. Even though patients 203 and 104 had a similar CD4 T-cell count at baseline, both had different outcomes when it came to the importance of the central memory and naïve cell on the reconstitution of the total number of CD4 T-cell. While for patient 104, who had half of the modified CD4 T-cell dose compared to patient 203, the central memory cell played an important role in the total CD4 T-cell, as the proliferation rate for the central memory cellp_C was very small. However, this patient had a significantly higher increase compared to the one from the low dose cohort. This indicates the importance of having a higher dose of modified CD4 T-cell injected.

Parameters	**Baseline**	**Non-modified T-cell**	**SATN fit**	**max-min**
λ	9.108148985	10.6	6.94229	6.94229-10.28664
p_N	0.0011482771	0.159	0.0094	0-0.18386
p_{ST}	0.011020353	0.132	0.1243	0-0.1243
p_C	0.010895753	0.000514	0	0-0.01089
p_T	0.021563362	0.00117	0.0030	0-0.144
p_E	0.033205325	0.0221	0	0-0.033205
ϕ_N	0.164397745	0.251	0.1041	0.1041-0.251
ϕ_{ST}	0.464899662	0.337	0.2243	0.2243-0.751
ϕ_C	0.105562369	0.145	0.0225	0.0225-0.145
ϕ_T	0.249269451	0.638	0.0780	0.07801-0.638
d_N	0.010241378	0.000576	0.0080	0.000576-0.10241
d_{ST}	0.011141735	0	0	0-0.011141
d_C	0.003333333	0	0.0775	0-0.0775
d_T	0.023361168	0.00224	0.0249	0.00224-0.0667
d_E	0.375021515	0.0221	0.1	0.1-0.0221

Table 4.4: Illustrates the data fitting results for pat 203 and the PRCC ranges.

These values are the rates per $day^{-}1$.

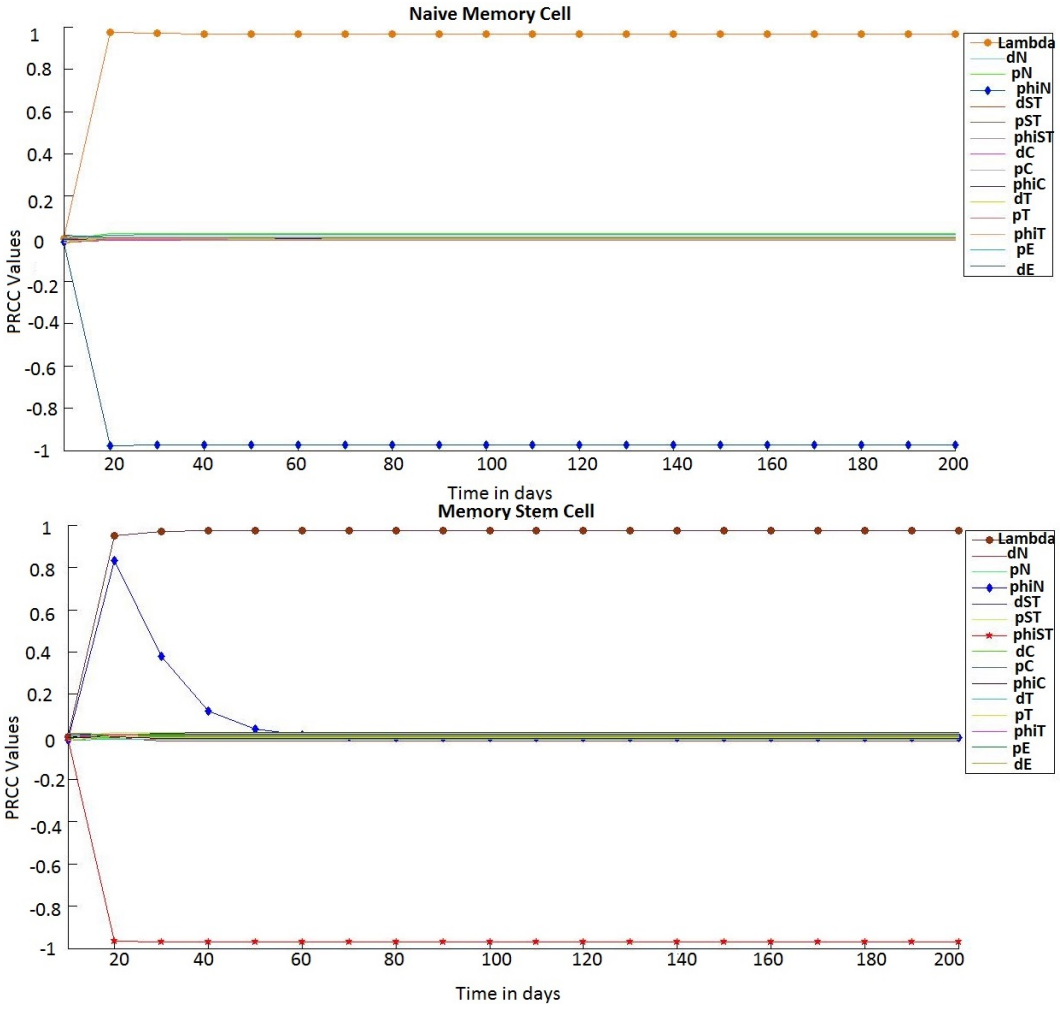

Figure 4.10: Correlation between the N and STM T-cell and the 15 parameters over 200 days for pat 203.

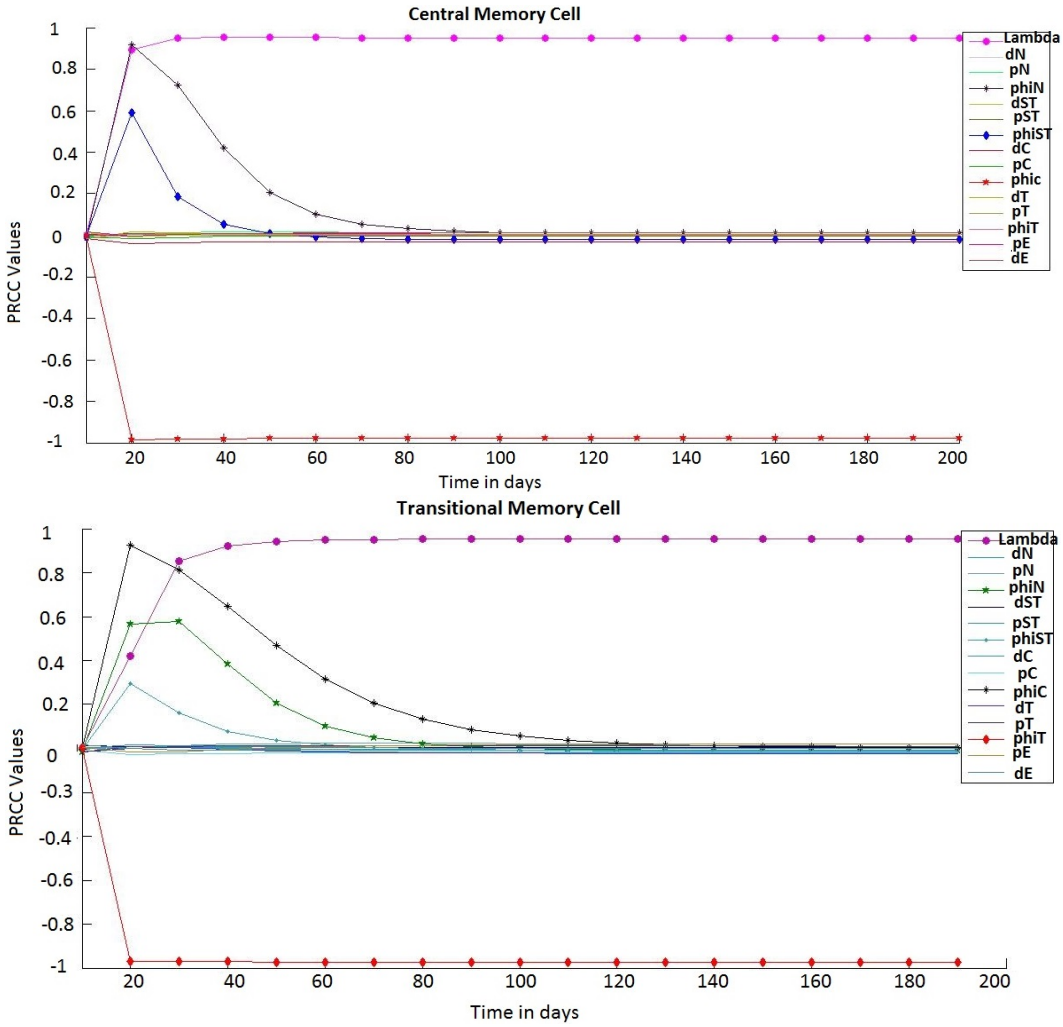

Figure 4.11: Correlation between the CM and TM T-cell and the 15 parameters over 200 days for pat 203.

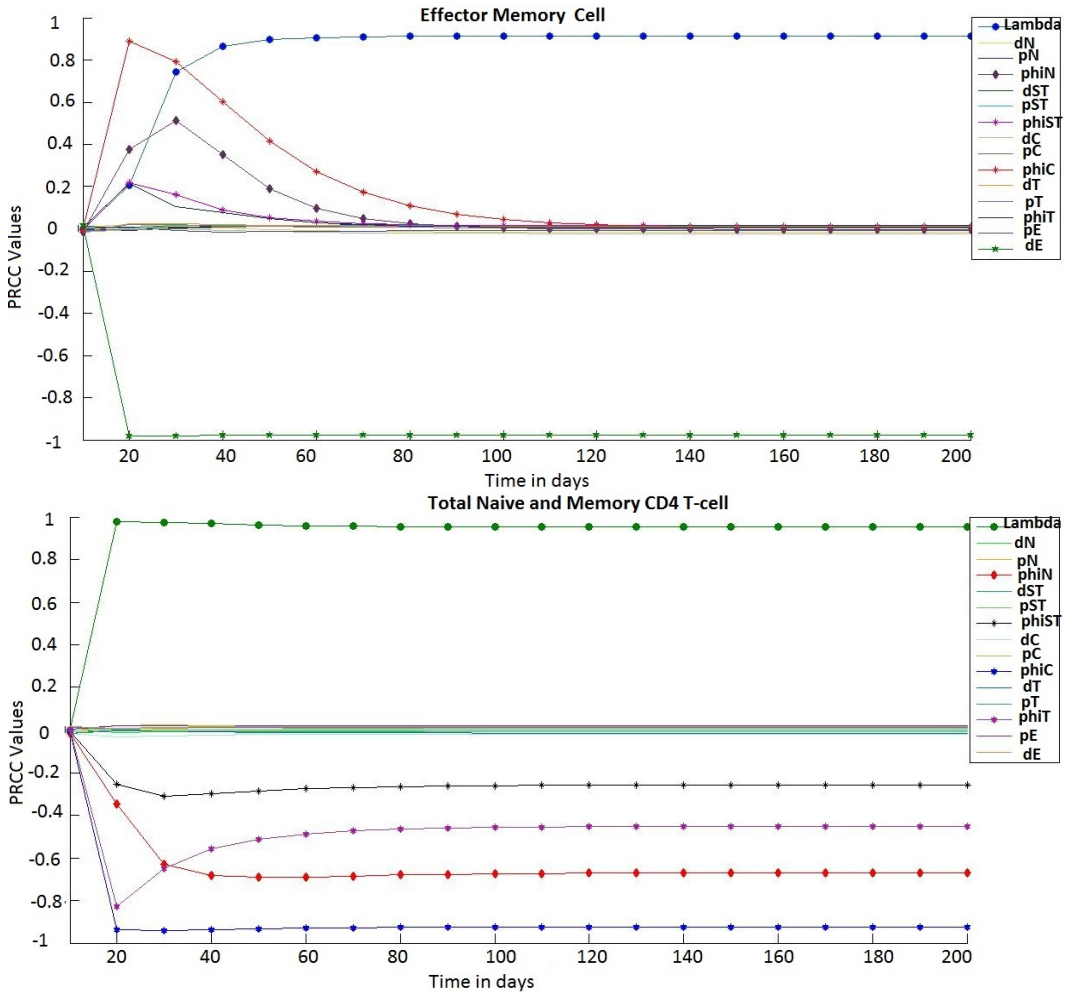

Figure 4.12: Correlation between the EM and total CD4 T-cell and the 15 parameters over 200 days for pat 203.

4.3.2.2 Patient 201

At baseline, this patient had the highest count of 525 CD4 T-cell per μL compared to the other eight patients. Looking at Table 4.5 we can note the following observations about each of the five sub-population behaviors:

- The number of naïve CD4 T-cells produced by the thymus,λ doubled in count.
- The proliferation rates for the naïve p_N, central memory p_C, transitional memory p_T and effector memory cells p_E decreased.
- The proliferation rates for the memory stem cell p_{ST} increased by one order of magnitude.
- The transition rates for the memory stem ϕ_{ST} and central memory cells ϕ_C decreased by one order of magnitude.
- The death rate of the naïve d_N, memory stem d_{ST} and central memory cells d_C decreased by one order of magnitude.
- The death rate of the effector memory d_E and transitional memory cells d_T increased.

As for the sensitivity analysis obtained over 200 days we observe the followings from

Figure 4.13, 4.14 and 4.15:

Naïve Cell:

Similar to patient 203, the production rate of the naïve cell from the thymus λ had a positive correlation with the naïve cell population. Only the naïve cell transition rate ϕ_N had negative correlation on the naïve cell population count. The death rate d_N did not have any effect as it decreased significantly after the treatment initiation.

Memory Stem cell

The naïve cell production rate λ started to show a significant effect on the memory stem cell count about 50 days post treatment. The naïve cell transition rate ϕ_N had a positive correlation with the memory stem cell population for only about 50 days post treatment. This is because the naïve cell transition rate did not increase by much after the initiation of the clinical study. The memory stem cell proliferation rate p_{ST} showed a negative correlation with the memory stem population. Death rate of the memory stem cell d_{ST} did not show any negative correlation as it decreased notably after the introduction of the down-modulated CCR5 CD4 T-cells.

Central Memory Cell

Likewise, the central memory cells show a similar result to patient 203 where the

transition rates of both the naïve ϕ_N and memory stem ϕ_{ST} cells had a positive correlation with the central memory cell count for about 20 and 50 days respectively. In addition, the central memory cell transition rate ϕ_C exhibits a negative correlation. Death rate of the central memory cell d_C did not show any negative correlation as it did have a significant decrease after the introduction of the down-modulated CCR5 CD4 T-cells as shown in the parameter estimation results.

Transitional Memory cell

The central memory cell death rate d_C and the transition rate of the transitional memory cell ϕ_T both had a negative correlation with the transitional memory cell population. In addition, the naïve cell production rate from the thymus λ along with the transition rate of the central memory cell ϕ_C were positively correlated with the transitional memory cell population. This indicates the importance of the central memory cell on the transitional memory cell population.

Effector Memory cell

While the naïve cell production rate λ, started to show a significant positive correlation (PRCC greater than 0.5) about 80 days post injection of the treatment, the significant positive correlation of the transitional memory cell started to decrease about 80 days post treatment. The transition rates for both the naïve ϕ_N and central

memory cell ϕ_C had a strong positive correlation with the effector memory cell count. The death rate of the naïve cell d_N started to exhibit a negative correlation with the effector cell population 30 days post treatment initiation. These results indicates that the naïve and central memory cells play a very important role in the reconstitution of the effector memory cell population.

Total naïve an memory CD4 T-cell

The sensitivity analysis results carried on the total number of CD4 T-cell for patient 201 emphasizes the importance of the naïve and central memory cell populations on the reconstitution of the total CD4 T-cell count. This was illustrated in the negative correlation between the central memory cell death rate d_C and the naïve cell transition rate ϕ_N on the total CD4 T-cell count. In addition the naïve cell production rate λ had a positive correlation on the total CD4 T-cell count. It is essential to mention that this patient had the highest count of the total CD4 T-cell after three years post treatment between all the nine patients.

Parameters	**Baseline**	**Non-modified T-cell**	**SATN fit**	**min-max**
λ	10.92205104	20.36874	19.3183	10.92205-24.9
p_N	0.001137926	0	0.0576	0-0.0576
p_{ST}	0.011052071	0.52791	0.0012	0-0.52791
p_C	0.011060716	0	0	0-0.044
p_T	0.021671103	0	0.0155	0-0.916
p_E	0.033205324	0	0	0-0.0332
ϕ_N	0.074641585	0.09417	0.1576	0.0746-0.1576
ϕ_{ST}	0.464899665	0.09583	0.0984	0.09583-0.567
ϕ_C	0.102277319	0.09125	0.0276	0.0271-0.102227
ϕ_T	0.244401048	0.26178	0.0429	0.0429-0.262
d_N	0.011824865	0.00297	0	0-0.01182
d_{ST}	0.011087322	0.00029	0.0058	0-0.01108
d_C	0.003333333	0.00004	0.0724	0.00004-0.0724
d_T	0.023222656	0.07113	0.0571	0-0.07113
d_E	0.44672007	0.61700	0.1	0-0.617

Table 4.5: Illustrates the data fitting results for pat 201 and the PRCC ranges. These values are the rates per $day^{-}1$.

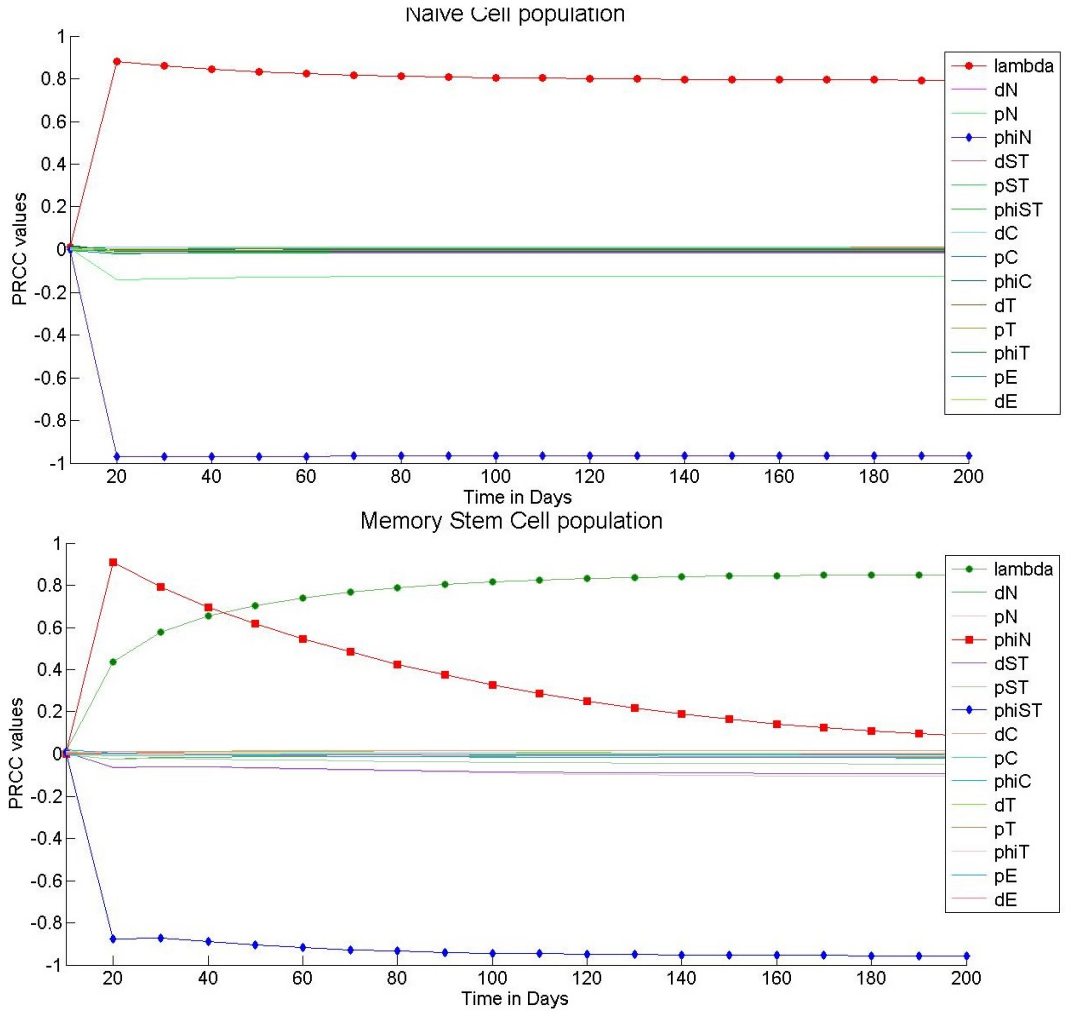

Figure 4.13: Correlation between the N and STM T-cell and the 15 parameters over 200 days for pat 201.

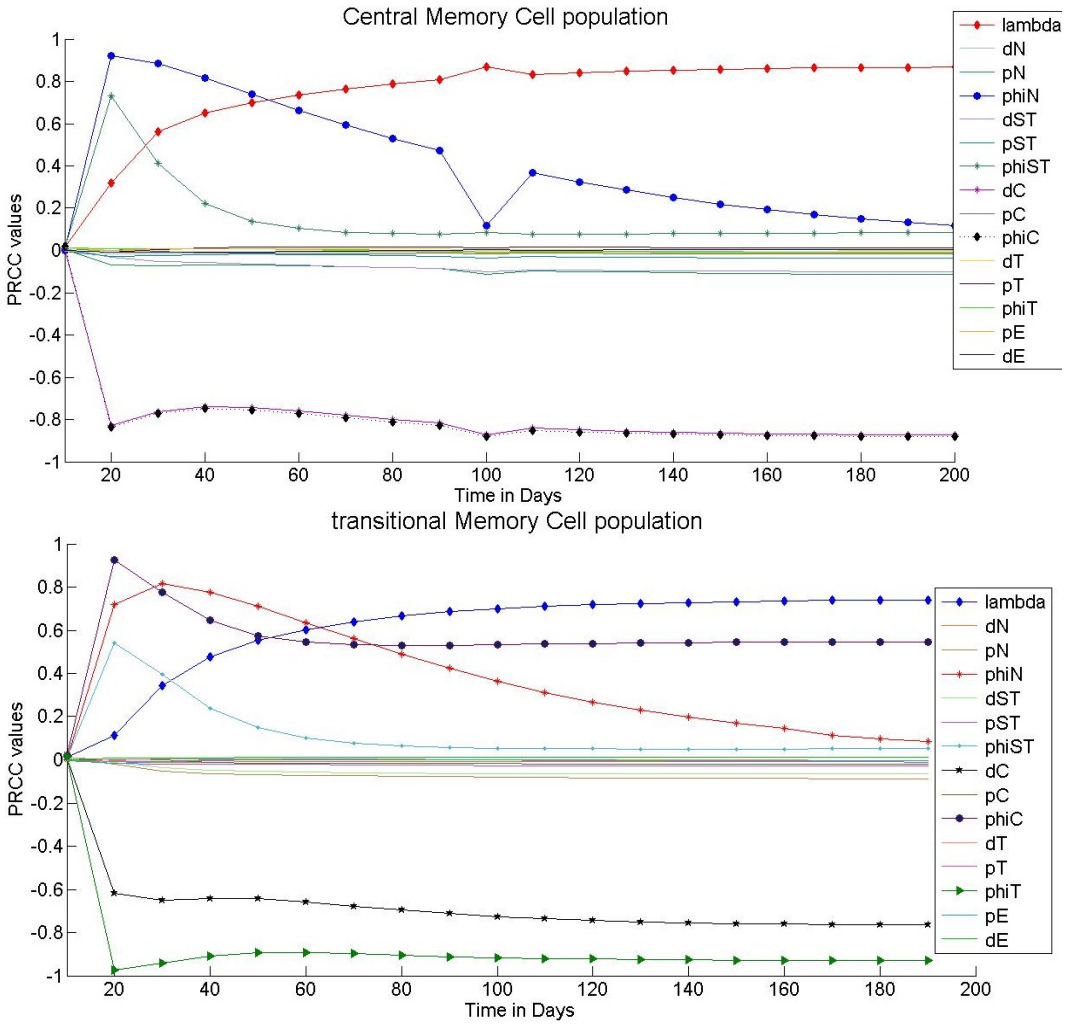

Figure 4.14: Correlation between the CM and TM T-cell and the 15 parameters over 200 days for pat 201.

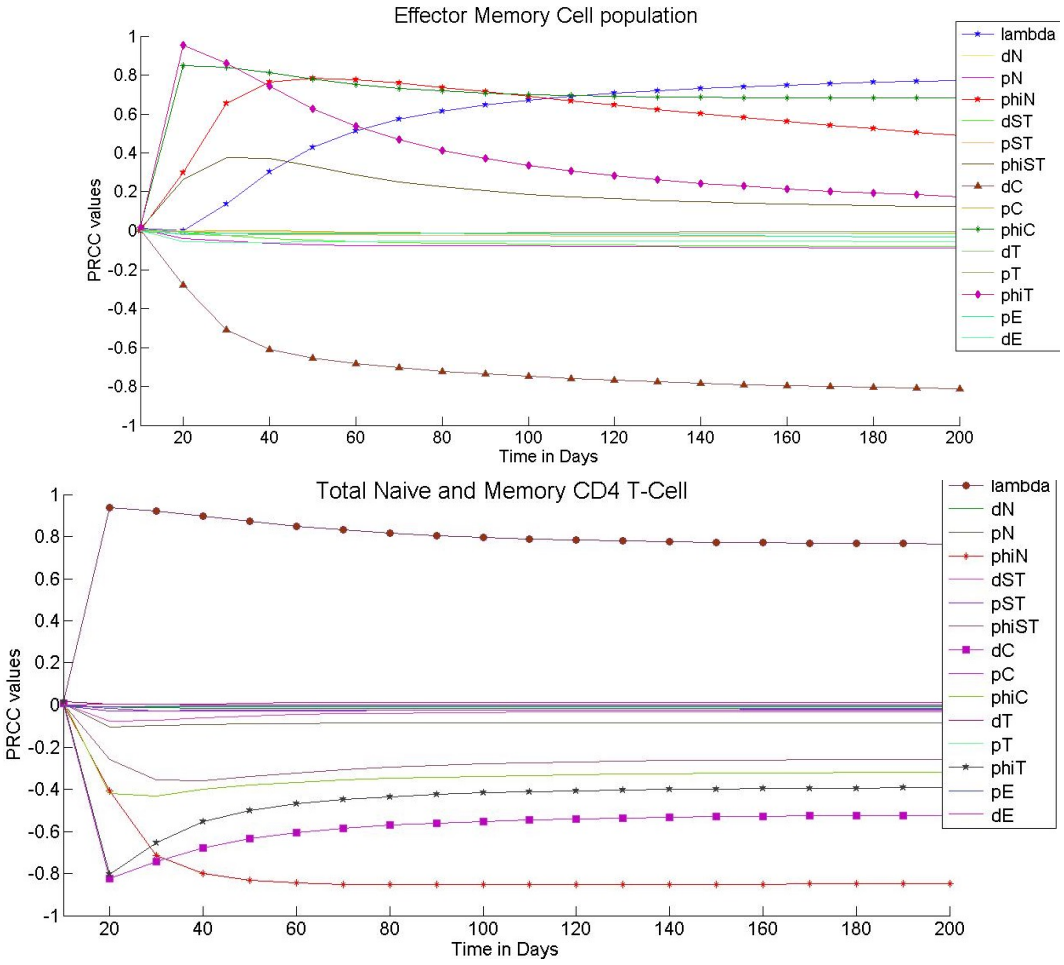

Figure 4.15: Correlation between the EM and total CD4 T-cell and the 15 parameters over 200 days for pat 201.

4.3.2.3 Patient 302

At baseline, this patient had a high count of 413 CD4 T-cell per μL. Looking at Table 4.6 we can note the following observations about each of the five sub-population behaviors:

- The number of naïve CD4 T-cells produced by the thymus λ had a slight increase.
- The proliferation rates for the naïve p_N and memory stem cell p_{ST} increased by one order of magnitude.
- The transition rate for the naïve ϕ_N and transitional memory cell ϕ_T increased.
- The death rate of the naïve d_N, stem memory d_{ST}, central memory d_C and transitional memory d_T cells decreased.
- The death rate of the effector memory cells d_E increased.

As for the sensitivity analysis obtained over 200 days we observe the followings from Figure 4.16, 4.17 and 4.18:

naïve Cell:

Similar to patient 201, while λ had a strong positive correlation with the naïve cell

population, the transition rate of the naïve cell ϕ_N has a negative correlation. The death rate d_N did not exhibit any negative correlation as this rate decreased after the injection of the down-modulated CCR5 CD4 T-cell.

Memory Stem cell

Similar to patient 201, the memory stem population is positively correlated with λ, and for the first 30 days post treatment the naïve cell transition rate ϕ_N shows a significant positive effect on the memory stem cell population. The stem cell transition rates ϕ_{ST} and death rate d_{ST} both have a negative correlation on the stem memory cell count.

Central Memory Cell

The naïve cell production rate λ started to have a positive correlation on the central memory cell population at about twenty days post treatment initiation. While the memory stem cell transition rate ϕ_C has a positive correlation with the central memory cell, the death rate of the memory stem cell d_{ST} has a negative correlation. This indicates the importance of the memory stem cell on the central cell population. In addition, the transition rate for the central memory cell ϕ_C is negatively correlated. This is an expected result as the higher the transition rate the more the cells are leaving the central state. Lastly, the transition rate of the naïve ϕ_N cell has

a positive correlation with the central memory cell for only 30 days post treatment initiation.

Transitional Memory cell

The transitional memory cell population has an equivalent result to the central cell, where the naïve and memory stem cell have an important effect on the reconstitution of the transitional memory cell. This finding is evident as the sensitivity analysis revealed that the rate λ and the memory stem cell proliferation rate p_{ST} both have a positive correlation with the transitional memory cell population and the memory stem cell death rate d_{ST} is negatively correlated. Moreover, as the previous memory sub-populations, the transition rate of the transitional memory cell ϕ_T has a negative correlation with the transitional memory cell population.

Effector Memory cell

The naïve and transitional cells have a positive correlation with the effector memory cell population where the naïve cell production rate λ and stem cell transition rate ϕ_{ST} are strongly positively correlated with the effector memory population. In addition, the effector memory proliferation p_E rate has a very significant effect on increasing the effector population count. The naïve and central memory transition rates, ϕ_N and ϕ_C, have positive correlations with the effector memory cell popula-

tions for about 30 days post treatment. This is a likely result to be observed as both types of cells do not have a long lifespan compared to the memory stem cell. Lastly, the death rate of the effector d_E and memory stem cell d_{ST} have a negative correlation with the effector memory population.

Total naïve an memory CD4 T-cell

From the observations drawn above, it is evident that the naïve cell production rate by the thymus λ plays an important role in increasing the total number of CD4 T-cell. Moreover, the transition rates for the naïve ϕ_N and transitional memory cells ϕ_T had a negative effect on the total CD4 T-cell count. The importance of the stem cell population continues to be evident where its death rate d_{ST} has a negative effect on the increase of he CD4 T-cell population. This concludes that both the naïve and memory stem cells both are important in the reconstitution of the total CD4 T-cell count for patient 302.

Parameters	**Baseline**	**Non-modified T-cell**	**SATN fit**	**min-max**
λ	8.764999372	13.6	16.1964	8.76499-16.1964
p_N	0.001154028	0.189	0	0-0.189
p_{ST}	0.01097524	0.0914	0.0107	0.0107- 0.0914
p_C	0.010877232	0.106	0	0-0.106
p_T	0.021672918	0.00575	0.0356	0-0.0356
p_E	0.033205325	0.274	0	0-0.0.274
ϕ_N	0.049752123	0.191	0.0517	0.0497-0.191
ϕ_{ST}	0.465153919	0.459	0.1024	0.102-0.459
ϕ_C	0.110786451	0.325	0.0314	0.0314-0.325
ϕ_T	0.247755424	0.832	0.1341	0.1341-0.0.832
d_N	0.010007563	0.000541	0.0483	0-0.0483
d_{ST}	0.011147981	0.00057	0.0083	0-0.01147
d_C	0.003333333	0.106	0.686	0-0.686
d_T	0.023187029	0.19	0.0015	0.00052-0.19
d_E	0.288952838	0.796	0.1	0.1-0.796

Table 4.6: Illustrates the data fitting results for pat 302 and the LHs ranges. These values are the rates per $day^{-}1$.

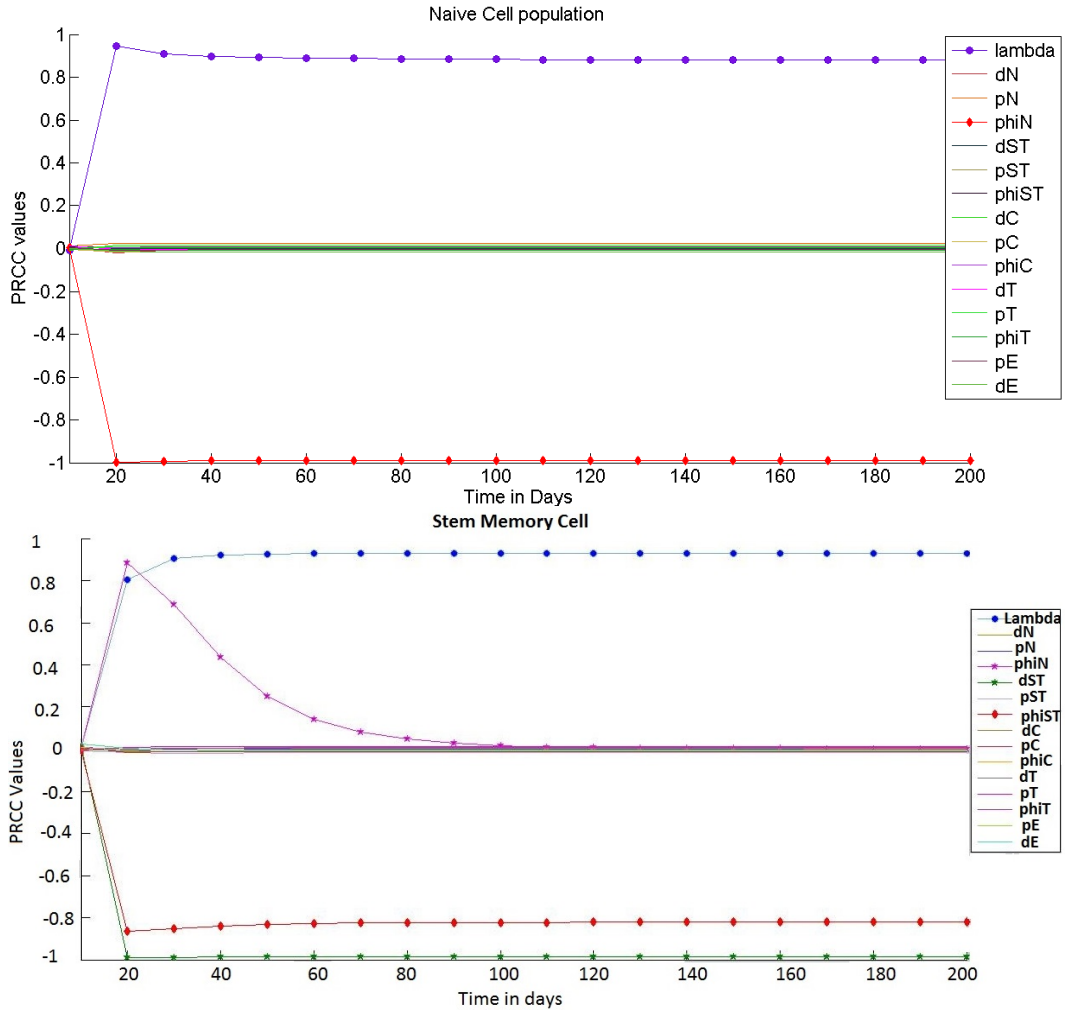

Figure 4.16: Correlation between the N and STM T-cell and the 15 parameters over 200 days for pat 302.

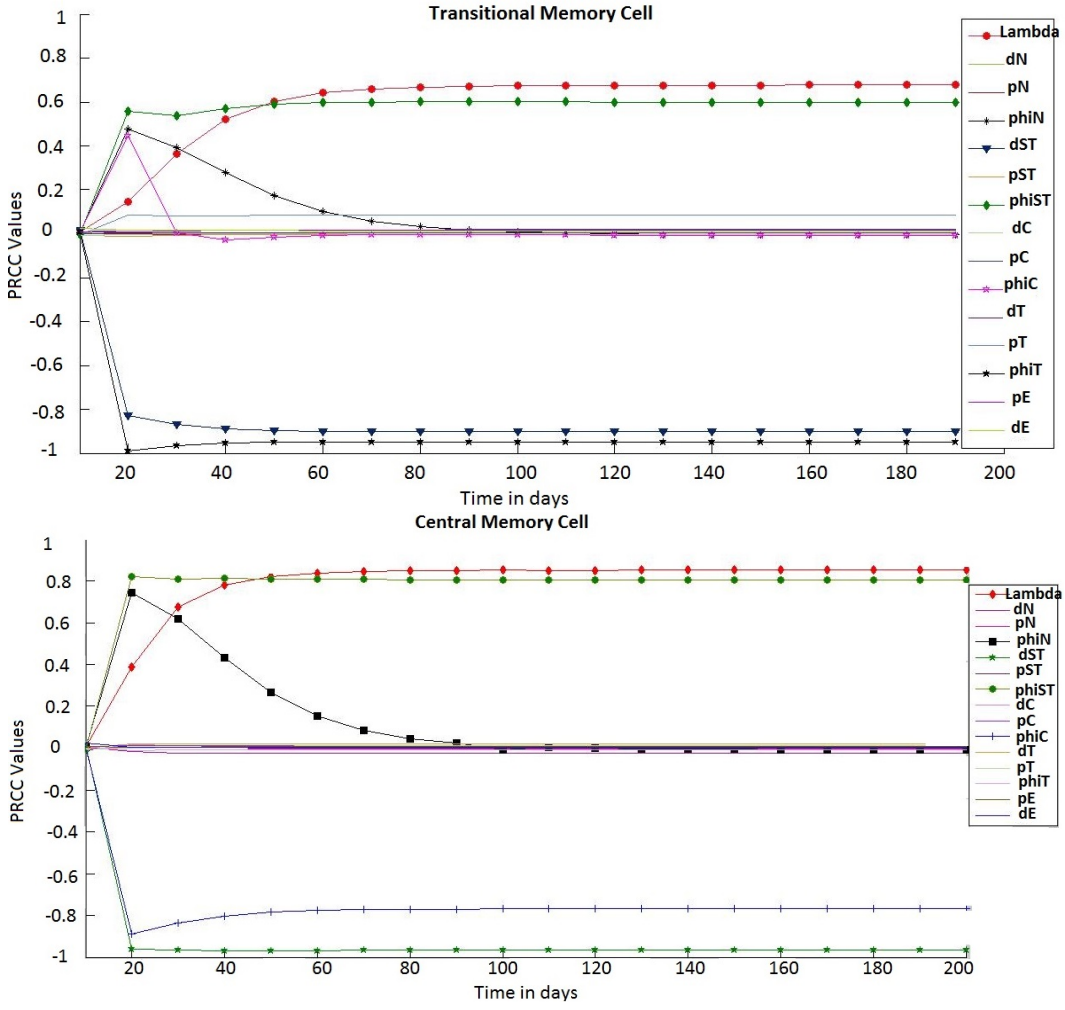

Figure 4.17: Correlation between the CM and TM T-cell and the 15 parameters over 200 days for pat 302.

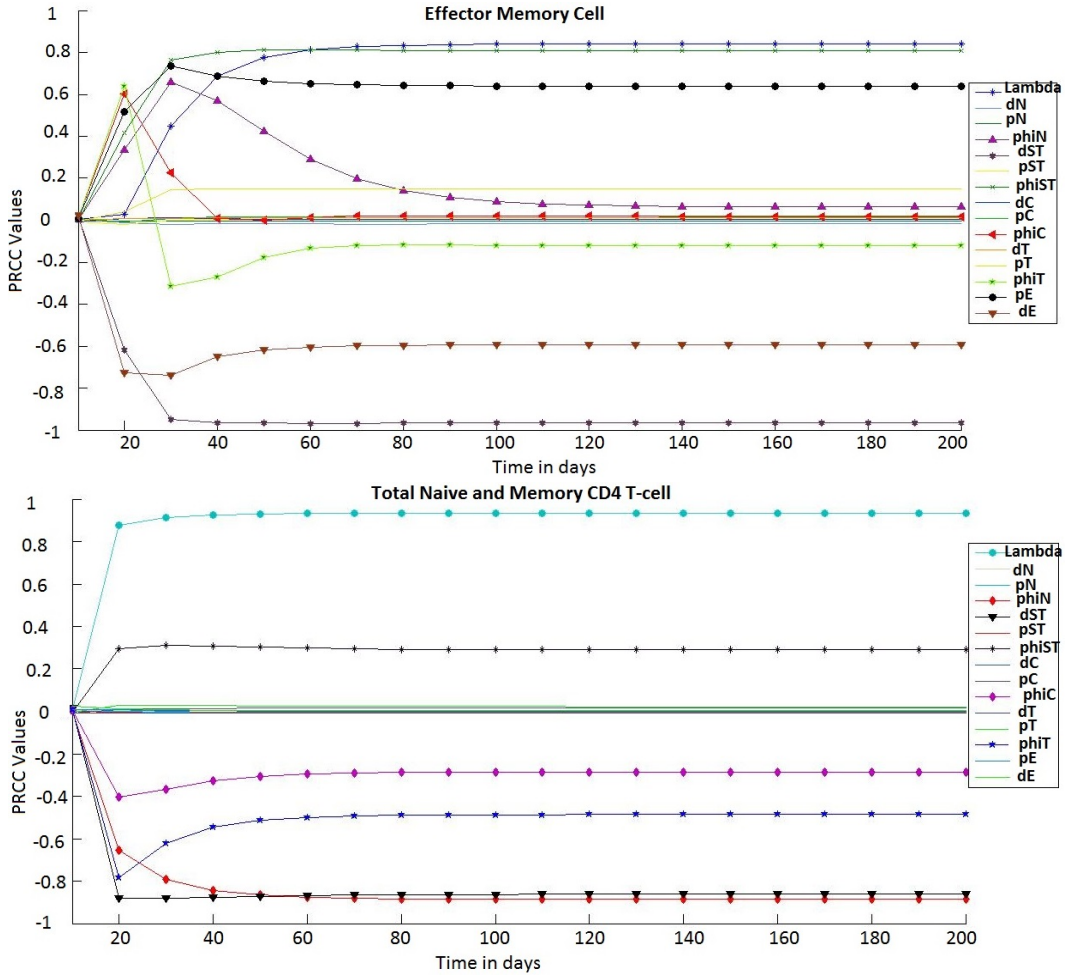

Figure 4.18: Correlation between the EM and total CD4 T-cell and the 15 parameters over 200 days for pat 302.

4.3.3 High Dose Cohort

The high dose cohort represents the three patients that received a single infusion of autologous CCR5-modified (SB-728-T) 3.0 x 10^{10} cells. All of the baseline measurements are measured seven days prior to the infusion of the CCR5 modified T-cells. The three patients had a range of high to medium CD4 T-cell count at baseline.

4.3.3.1 Patient 305

At baseline, this patient has a somewhat high count of 480 CD4 T-cell per μL. Looking at table 4.7 we can note the following observations about each of the five sub-population behaviors:

- The number of naïve CD4 T-cells produced by the thymus λ slightly decreased.
- The proliferation rate for the naïve cell p_N has increased by one order of magnitude.
- The proliferation rates for the memory stem p_{ST}, central memory p_C, transitional memory p_T and effector memory cells p_E decreased.
- The transition rates for all of the five subsets remained somewhat constant.

- The death rate of the naïve d_N and memory stem cells p_{ST} decreased. However, the death rate for the transitional memory p_T and effector memory cells p_E remained constant.

As for the sensitivity analysis obtained over 200 days we observe the followings from Figure 4.19, 4.20 and 4.21:

Naïve Cell:

Similar to most of the patients, the naïve cell production rate λ has a positive correlation with the the naïve cell population. Only the naïve cell transition rate ϕ_N has a negative correlation on the naïve cell population count. The death rate d_N did not have any effect as it decreased significantly after the treatment initiation.

Memory Stem cell

The naïve cell production rate λ continues to have a positive effect on the memory stem cell population count. The transition rate of the naïve cell ϕ_N displays a positive correlation with the memory stem count only for about 30 days post treatment initiation. This is a result of the decrease in the transition rate for the naïve cell after the treatment as illustrated in the Table 4.7. The transition rate of the memory stem cell ϕ_{ST} is negatively correlated with its population.

Central Memory Cell

Similarly to the memory stem cell, the naïve cell production rate by the thymus λ continues to exhibit a positive correlation with the central memory cell count. Moreover, the transition rates of the naïve ϕ_N and memory stem cells ϕ_{ST} both have a positive correlation with the central memory population for about 30 and 20 days post treatment initiation respectively. The death and transition rates of the central memory cell, d_C and ϕ_C have a negative effect on the central memory cell population count as these rates did not exhibit a decrease after the start of the clinical trial.

Transitional Memory cell

The central memory cell death rate d_C and the transition rate of the transitional memory cell ϕ_T both have a negative effect on the transitional memory cell population count. In addition, the naïve cell production rate from the thymus λ along with the transition rate of the central memory cell ϕ_C are positively correlated with the transitional memory cell population. This highlights the importance of the central memory cell on the transitional cell population. The naïve cell transition rate ϕ_N has a positive correlation with the transitional memory cell population for about 50 days post treatment.

Effector Memory cell

The importance of the central memory cell continues to be observed in the effector memory population, where the central memory death rate d_C has a negative correlation with the effector memory population and the central memory transition rate ϕ_C has a positive correlation. In addition, the naïve cell production rate λ has a positive correlation. Similarly to the transitional cell population, the transition rate for the naïve ϕ_N, and transitional memory ϕ_T cells have a positive correlation for about 50 days post treatment. The death rate of the effector memory cell d_E has a negative correlation as this value increased significantly after the initiation of the experimental treatment.

Total naïve an memory CD4 T-cell

The sensitivity analysis performed on the total CD4 T-cell count highlights the importance of the central cell on the reconstitution of the CD4 T-cell, where the death and transition rates of the central memory cells, d_C and ϕ_C, both have a negative correlation with the total CD4 T-cell count. The naïve cell population as well has some effect on the total CD4 T-cell count where the production rate λ has a positive effect and the transition rate ϕ_N has a negative correlation with the CD4 T-cell count. These results illustrates that for patient 305, the central memory and naïve cells played an important role in the reconstitution of the total CD4 T-cell count population.

Parameters	**Baseline**	**Non-modified T-cell**	**SATN fit**	**min-max**
λ	10.51654578	8.96240	11.2318	8.96240-12.56407
p_N	0.001136919	0.07369	0.0003	0-0.07369
p_{ST}	0.011035824	0.00396	0.1360	0.00396-0.1360
p_C	0.010987609	0	0	0-0.010987609
p_T	0.021796329	0.00141	0.0155	0-0.021796
p_E	0.033211285	0.02939	0	0-0.033211285
ϕ_N	0.215314186	0.13602	0.0492	0.0492-0.215314
ϕ_{ST}	0.464899665	0.48367	0.2358	0.2358-0.51269
ϕ_C	0.105961431	0.12000	0.0342	0.0342-0.1200
ϕ_T	0.233399786	0.32529	0.1082	0.182-0.3259
d_N	0.011498048	0.00918	0.0511	0-0.0511
d_{ST}	0.011057469	0	0.0002	0-0.011057
d_C	0.003333333	0.00148	0.0658	0.00082-0.0658
d_T	0.023025425	0.01267	0.0073	0.0073-0.023025
d_E	0.321528667	0.31500	0.1	0.1-0.32152

Table 4.7: Illustrates the data fitting results for pat 305 and the LHS ranges. These values are the rates per $day^{-}1$.

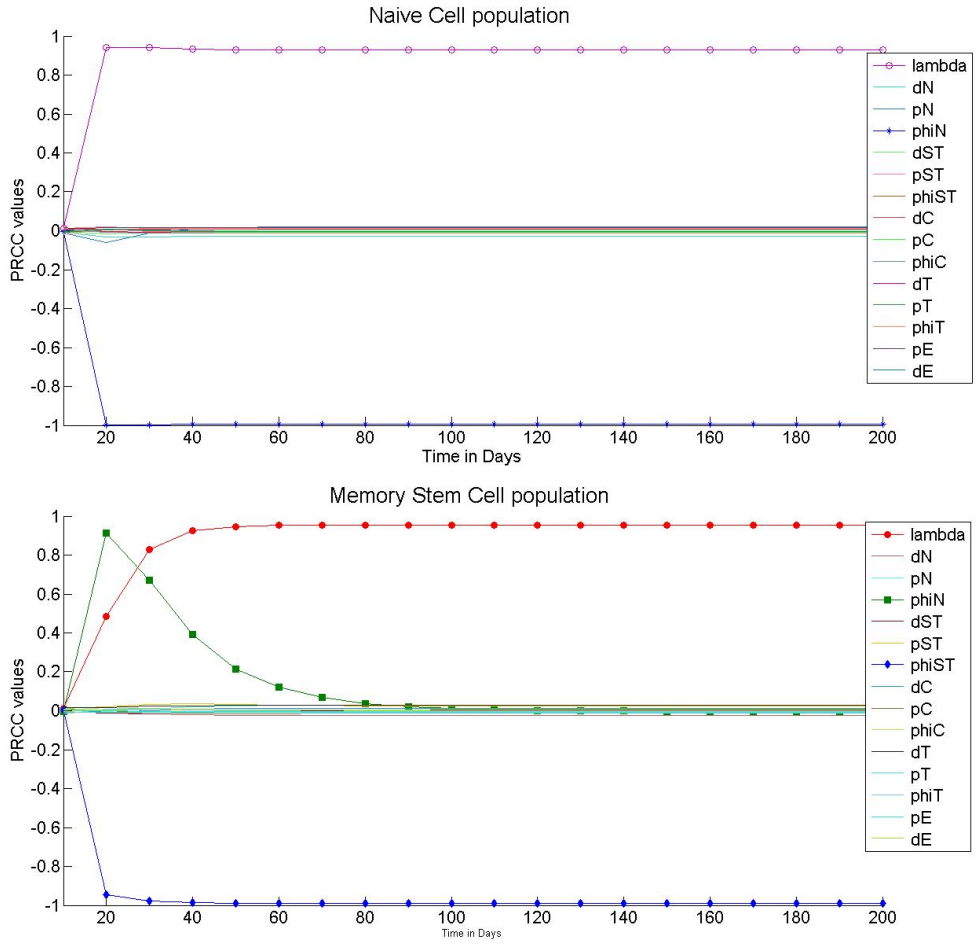

Figure 4.19: Correlation between the N and STM T-cell and the 15 parameters over 200 days for pat 305.

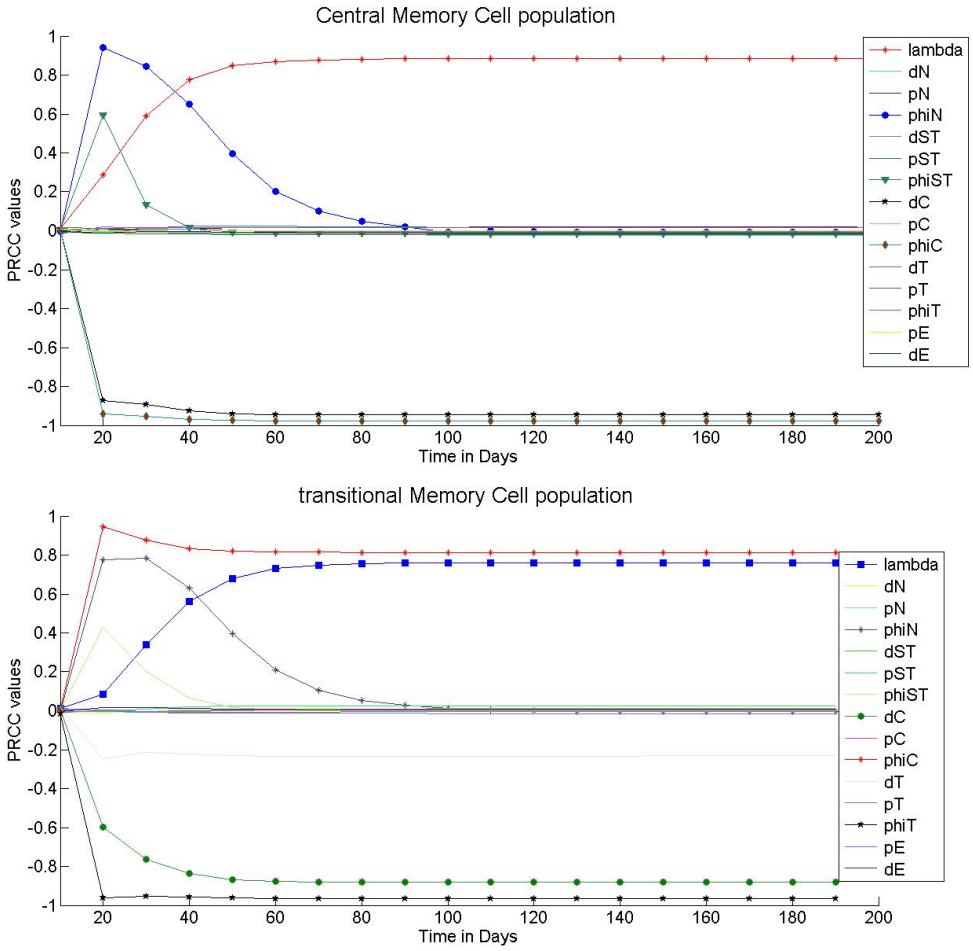

Figure 4.20: Correlation between the CM and TM T-cell and the 15 parameters over 200 days for pat 305.

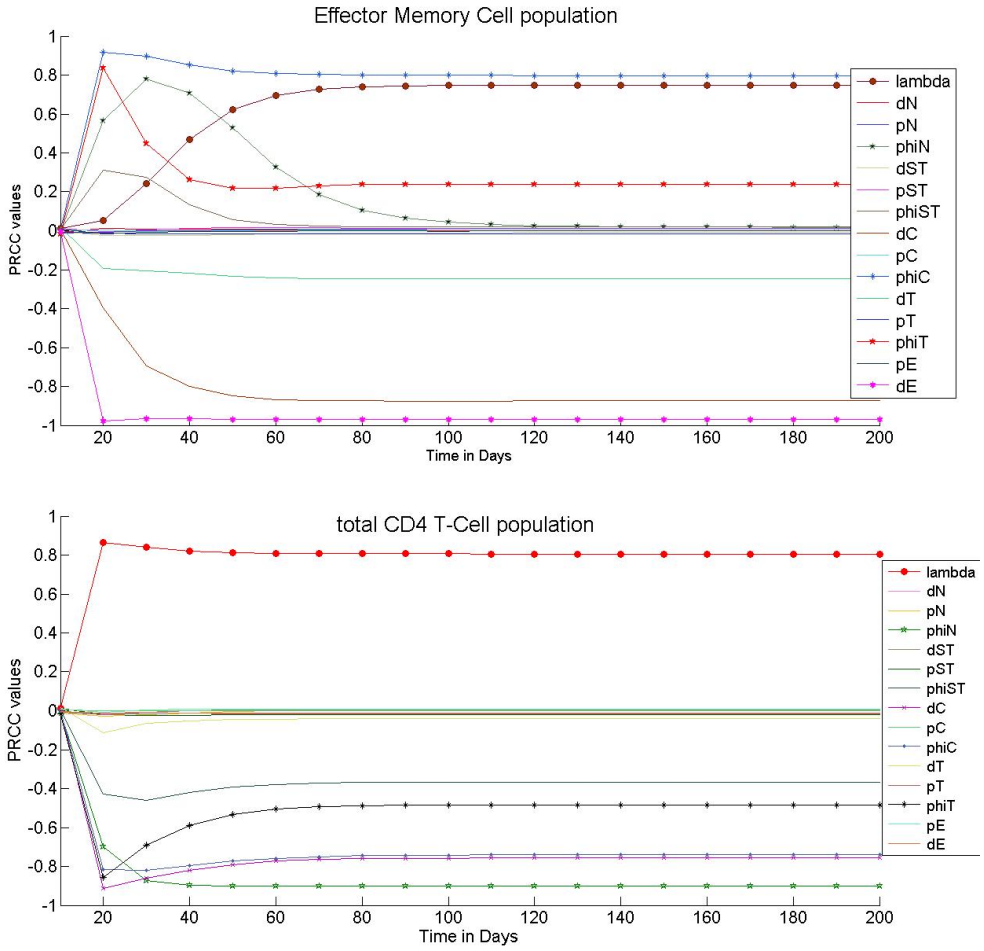

Figure 4.21: Correlation between the EM and total CD4 T-cell and the 15 parameters over 200 days for pat 305.

4.3.3.2 Patient 303

At baseline, this patient had a somewhat medium level of CD4 T-cell count of 330 cells per μL. Looking at Table 4.8 we can note the following observations about each of the five sub-population behaviors:

- The number of naïve CD4 T-cells produced by the thymus λ slightly increased.
- The proliferation rates for the naïve p_N and transitional memory cell p_T increased by one order of magnitude.
- The proliferation rate for the central memory cell p_C decreased by one order of magnitude.
- The transition rate for the naïve cell ϕ_N increased by one order of magnitude.
- The transition rate for both the memory stem ϕ_{ST} and central memory cells phi_C decreased.
- The death rates of the naïve d_N, transitional memory d_T and memory stem cell d_{ST} decreased.
- The death rate of the central memory cell d_C increased by one order of magnitude.

As for the sensitivity analysis obtained over 200 days we observe the followings from Figures 4.22, 4.23 and 4.24:

Naïve Cell:

Similar to patient 303, the naïve cell production rate λ has a positive correlation with the naïve dell population and the naïve transition rate ϕ_N has a negative correlation. The death rate d_N did not have any effect as it decreased significantly after the treatment initiation.

Memory Stem cell

The memory stem cell proliferation rate P_{ST} has a strong positive correlation with the memory stem cell population. However, unlike the previous patients, the thymic production rate of naïve cell λ did not have a strong positive correlation on the memory stem cell population. This could be explained by the fact that the naïve cell death rate d_N and production rate λ did not have a significant change from baseline. The central memory death rate d_c has a weak positive effect on the memory stem cell population and the transition rate of the memory stem cell ϕ_{ST} has a negative correlation.

Central Memory Cell

The importance of the memory stem is still noticeable in the central memory popu-

lation where its proliferation rate p_{ST} has a very strong positive correlation with the central memory cell count. The naïve cell birth rate λ continues to exhibit a weak positive correlation. the central memory cell death and transition rates, d_C and ϕ_C both are negatively correlated with the central cell population. This could be a result of the increase in the central memory death rate d_C after the initiation of the treatment as shown in Table 4.8.

Transitional Memory cell

The memory stem cell continue to maintain its importance in the transitional memory population, where its proliferation rate p_{ST} is strongly positively correlated with the transitional memory cell population. In addition, while the central memory cell death rate d_C has a negative correlation with the transitional memory cell populations, transition rate ϕ_C is positively correlated. This highlights the importance of both the memory stem and central CD4 T-cells on the reconstitution of the effector memory cell. The naïve cell production rate by the thymus λ does not exhibit a very strong positive effect. Lastly, the transition rate of the transitional cell ϕ_T has a negative correlation with the transitional memory cell count.

Effector Memory cell

While the effector memory cells have similar results to the transitional cells when it

comes to the importance of the central and memory stem cell, the transition rate of the transitional memory cell ϕ_T is positively correlated with the effector memory cell population. This is an expected result as the more cells are transitioning from the transitional memory to effector memory state, the effector memory population is exhibiting an increase in its count.

Total naïve an memory CD4 T-cell

It is expected to observe that the memory stem and central cell will play an important role in the reconstitution of the total CD4 T-cell count. The memory stem cell proliferation rate p_{ST} and the naïve cell production λ both have a positive correlation on the total memory CD4 T-cell count. This indicates as we increase these model parameters, the total CD4 T-cell population will increase. In addition, the death rate of the central memory cell d_C and the transition rate of the memory stem cell ϕ_{ST} are both negatively correlated with the total CD4 T-cell count. These results indicates that the memory stem cell population have a very important role in increasing or maintaining a high level of CD4 T-cell count in patient 303.

Parameters	**Baseline**	**Non-modified T-cell**	**SATN fit**	**min-max**
λ	8.445704953	10.75440	10.2921	8.44570-32.5
p_N	0.001153966	0.01191	0.0078	0-0.01191
p_{ST}	0.011042777	0.01122	0.0425	0.01104-0.991
p_C	0.010995681	0.00737	0	0-0.010995
p_T	0.021578612	0.27041	0.0007	0.00
p_E	0.033205325	0.02316	0	0-0.03320
ϕ_N	0.082314245	0.10208	0.0986	0.0823-0.304
ϕ_{ST}	0.464899661	0.11758	0.1419	0.11758-0.628
ϕ_C	0.107979023	0.09603	0.0196	0.0196-0.096
ϕ_T	0.247968152	0.24667	0.0575	0.0575-0.397
d_N	0.009633584	0.00559	0.0093	0-0.0093
d_{ST}	0.011193147	0.00538	0.0006	0-0.01119314
d_C	0.003333333	0.44393	0.0804	0.00333-0.44393
d_T	0.023148944	0.01795	0.0431	0-0.0431
d_E	0.515148304	0.49886	0.1	0.1-0.705

Table 4.8: Illustrates the data fitting results for pat 303 and the LHS ranges. These values are the rates per $day^{-}1$.

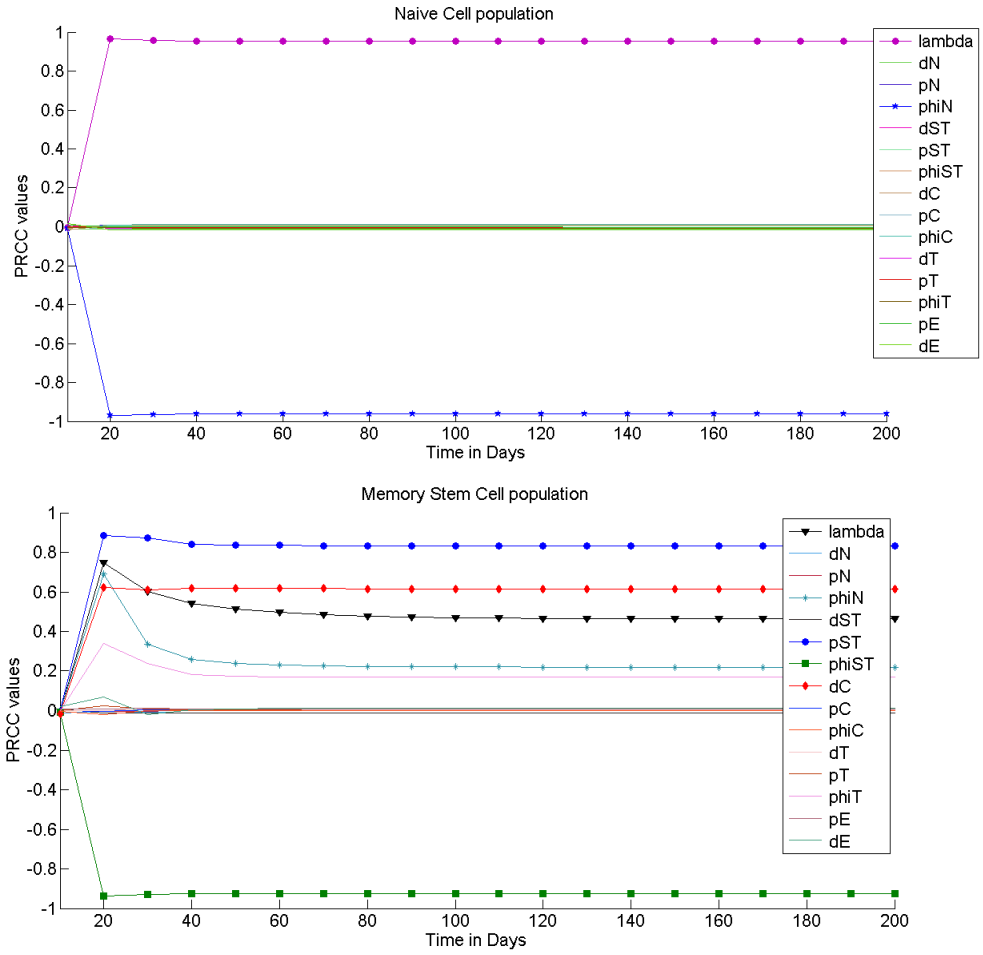

Figure 4.22: Correlation between the N and STM T-cell subsets and the 15 parameters over 200 days for pat 303.

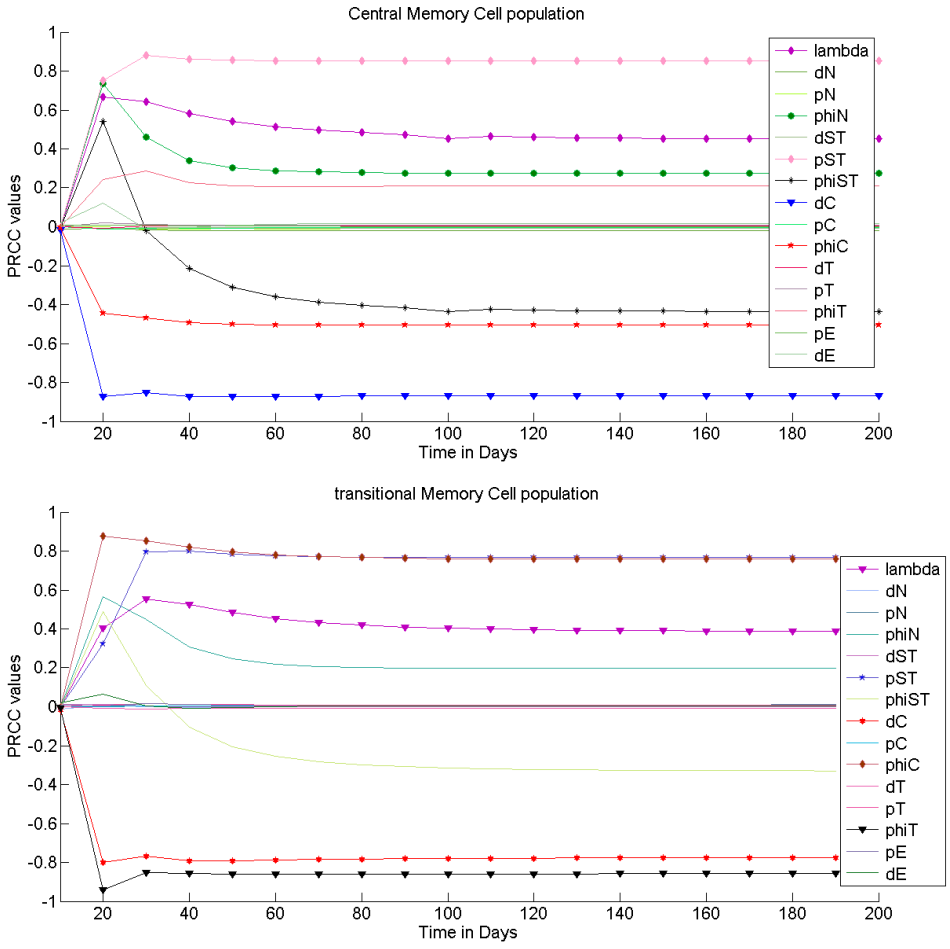

Figure 4.23: Correlation between the CM and TM T-cell subsets and the 15 parameters over 200 days for pat 303.

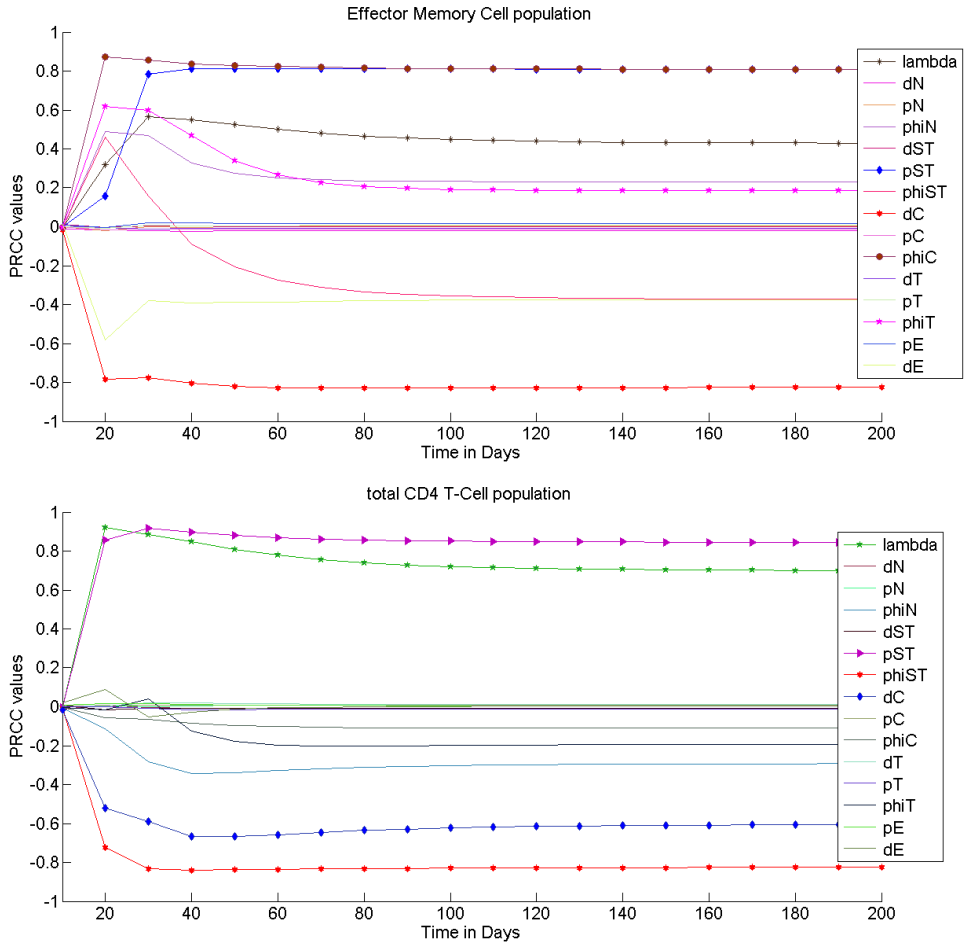

Figure 4.24: Correlation between the EM and total CD4 T-cell and the 15 parameters over 200 days for pat 303.

4.3.3.3 Patient 304

At baseline, this patient had a somewhat medium level of CD4 T-cell count of 306 cells per μL. Looking at Table 4.9 we can note the following observations about each of the five sub-population behaviors:

- The number of naïve CD4 T-cells produced by the thymus λ increased.
- The proliferation rates for the naïve p_N and memory stem cell p_{ST} increased by at least one order of magnitude.
- The proliferation rates for the central memory p_C, transitional memory p_T and effector memory cells p_E decreased.
- The transition rate for all the four subsets of cell increased.
- The death rate of the effector memory d_E and central memory cells d_C increased.
- The death rate for the naïve d_N, transitional memory d_T and memory stem cell d_{ST} decreased.

As for the sensitivity analysis obtained over 200 days we observe the following from Figures 4.25, 4.26 and 4.27:

naïve Cell:

While the naïve cell production rate λ is positively correlated with the naïve cell count, the transition rate of the naïve cell ϕ_N has a negative correlation. The same results were observed in both patients 305 and 303 from the same cohort. The death rate d_N did not have any effect as it decreased significantly after the treatment initiation.

Memory Stem cell

The naïve cell production rate λ has a strong positive correlation with memory stem cell population. As the transition rate of the memory stem ϕ_{ST} cell had an increase after the introduction of the down modulated CCR5 CD4 T-cells, it showed to have a negative correlation with the memory stem population. The transition rates of the naïve cell ϕ_N has a positive correlation with the memory stem cell population for about 30 days post treatment initiation.

Central Memory Cell

Similar to the naïve and memory stem cell, the naïve cell production rate λ has a positive correlation with the central memory cell population. While the transition rate of the central memory cell ϕ_C has a negative correlation with the central memory cell, the transition rate of the memory stem ϕ_{ST} is positively correlated for

about 40 days post treatment.

Transitional Memory cell

As The transitional memory cell has a positive correlation with the naïve cell production rate λ, the transition rates for both the central memory and naïve cells, ϕ_C and ϕ_N, both have a significant positive correlation for the first 40 days. The death rate of the transitional cell d_T along with the transition rate ϕ_T are negatively correlated with the transitional memory cell count.

Effector Memory cell

The importance of the naïve and transitional cell are evident in the effector cell population, as the naïve production rate by the thymus λ and the transitional cell transition rate ϕ_T both have a strong positive correlation with the effector memory population. Both the death rate for the transitional and effector memory cells, d_E and d_T, are negatively correlated with the effector memory cell population.

Total naïve an memory CD4 T-cell

In the sensitivity analysis results obtained for the total number of CD4 T-cell, the importance of the naïve cell population is evident where the naïve production rate λ is the only parameter that has a positive effect on the total T-cell count. In addition,

the transition rates for both the naïve and central memory cells, ϕ_N and ϕ_C, both have a negative effect on the total CD4 T-cell count. The sensitivity analysis results indicates that for patient 304 the naïve and central memory cells play an important role in the reconstituting the total CD4 T-cell count as seen in both patients 303 and 305.

Parameters	**Baseline**	**Non-modified T-cell**	**SATN fit**	**min-max**
λ	9.086416508	14.34899	8.3187	8.3187-48.2
p_N	0.001154028	0.22369	0.0227	0-0.22369
p_{ST}	0.01097524	0.67151	0.1237	0-0.67151
p_C	0.010877232	0	0	0-0.010877
p_T	0.021672918	0	0	0-0.02167
p_E	0.033205325	0.0232	0	0-0.0333205
ϕ_N	0.049752123	0.26594	0.1135	0.04975-0.304
ϕ_{ST}	0.465153919	0.61831	0.2234	0.2234-0.61831
ϕ_C	0.110786451	0.26555	0.0533	0.0533-0.2655
ϕ_T	0.247755424	0.39893	0.0466	0.0466-0.491
d_N	0.010007563	0	0.0092	0-0.010007563
d_{ST}	0.011147981	0	0.0004	0-0.0111479
d_C	0.003333333	0.44	0.0467	0-0.44
d_T	0.023187029	0	0.0534	0-0.0534
d_E	0.288952838	0.99672	0.1	0.1-0.99672

Table 4.9: Illustrates the data fitting results for pat 304 and the LHs ranges. These values are the rates per $day^{-}1$.

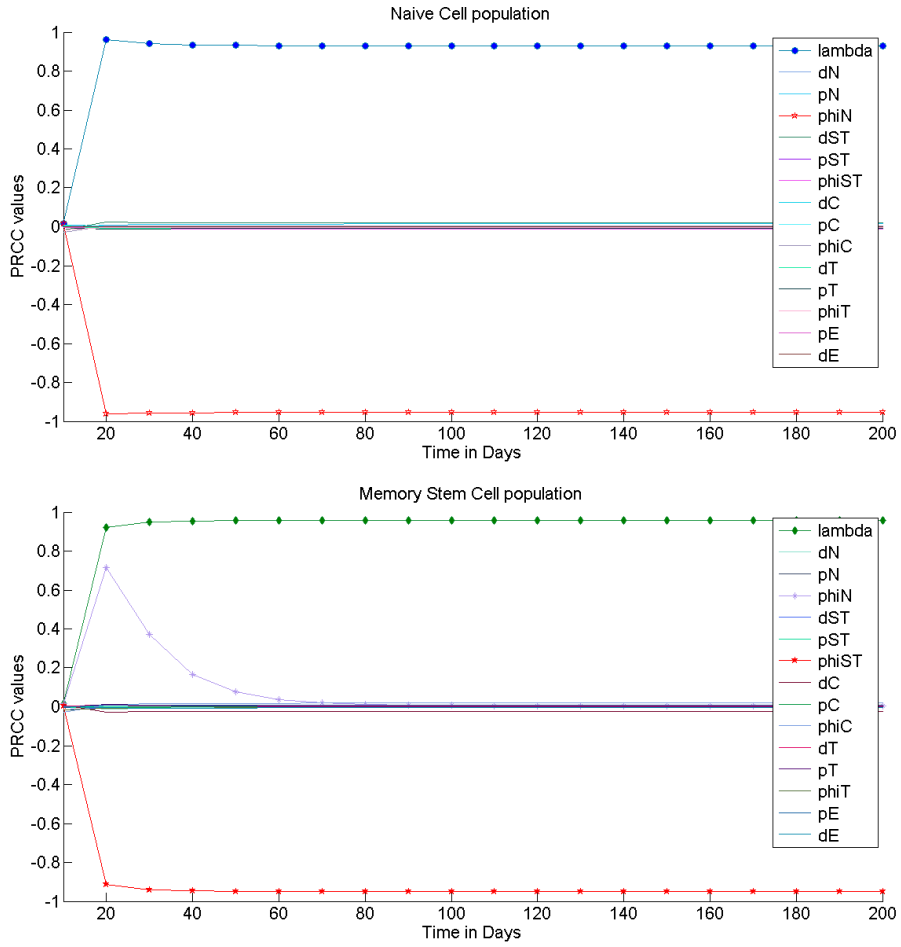

Figure 4.25: Correlation between the N and STM T-cell and the 15 parameters over 200 days for pat 304.

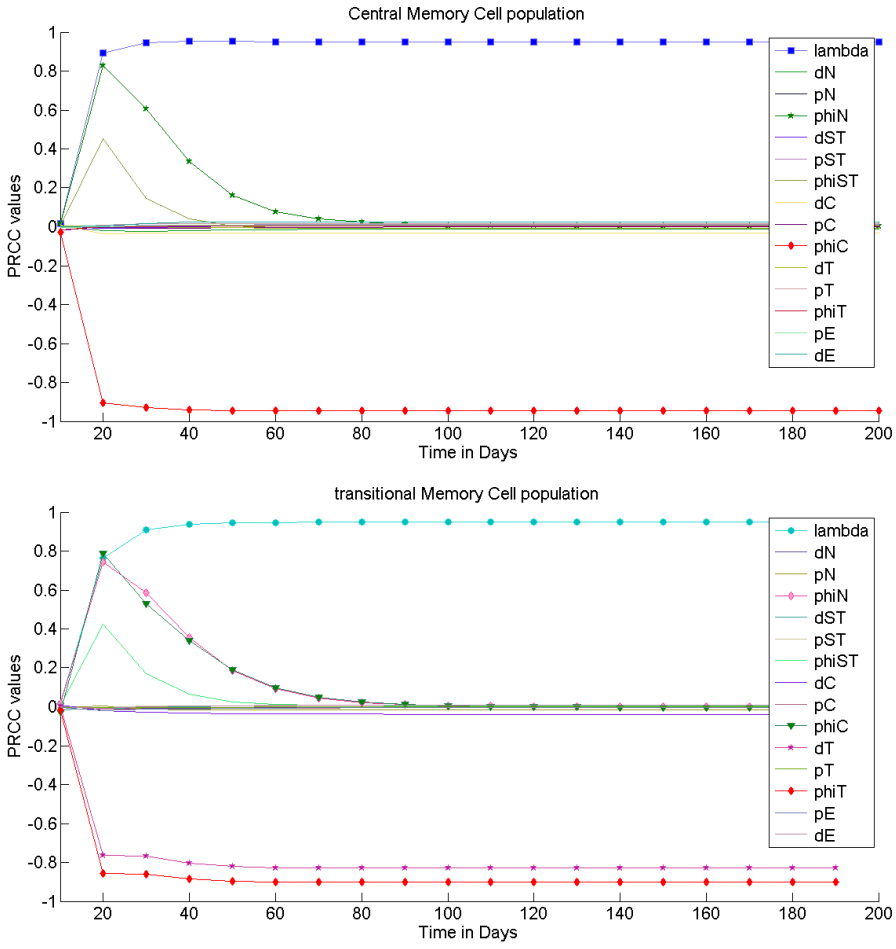

Figure 4.26: Correlation between the CM and TM T-cell and the 15 parameters over 200 days for pat 304.

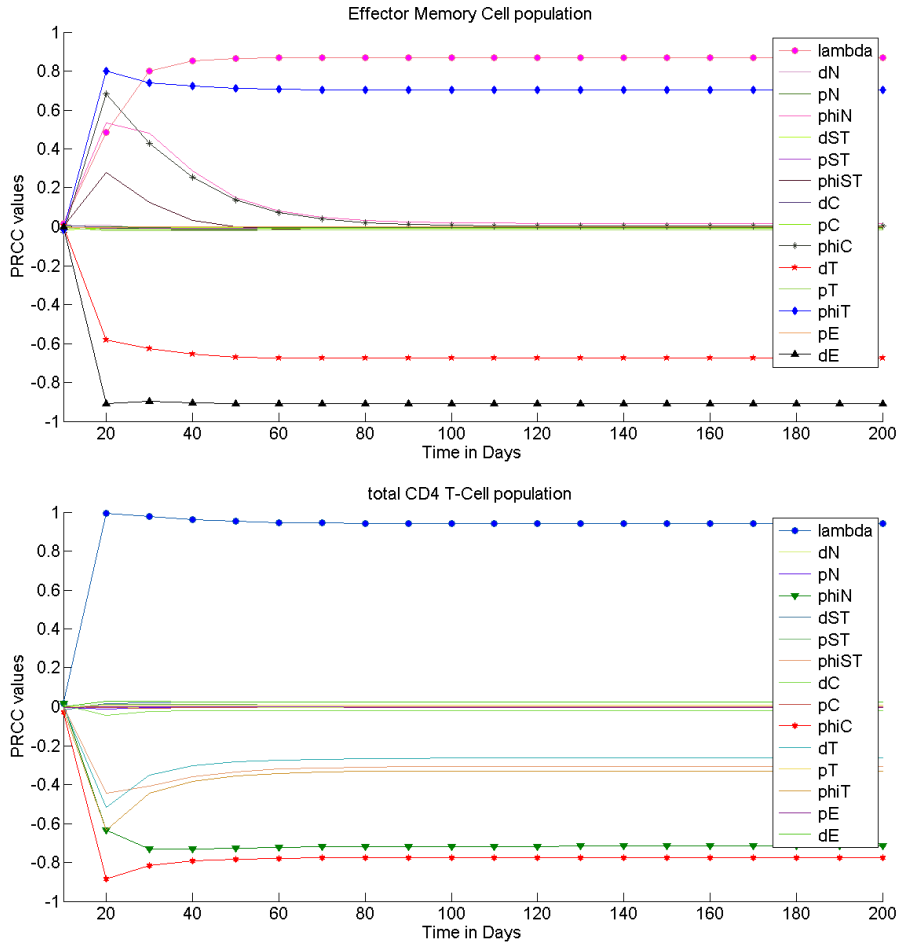

Figure 4.27: Correlation between the EM and total memory CD4 T-cell and the 15 parameters over 200 days for pat 304.

4.3.4 Brief Summary

In this section we will summarize the results that we have presented in this chapter. First we will summarize the interesting changes in some model parameters by comparing the fits obtained at baseline and after the injection of the CCR5-down-modulated memory CD4 T-cells, we have seen that the death rate of the memory stem cells d_{ST} and the naïve cell d_N have always decreased in magnitude and for some patients it was 0. This finding further indicates that the injection of the CCR5-down-modulated memory CC4 T-cells has succeeded in increasing the lifespan of the naïve and memory stem cell subset population in all the three HIV positive cohorts. In addition, patients who are categorized under the immunodiscordant group where the CD4 T-cell count is < 350 [31] and received a low dose of the CCR5-down-modulated CD4 T-cells injection had a decrease or no change in the naive cell production rate λ which was seen in patients 103 and 104. However, when immunodiscordant patients(203,304,303) received a higher dose of the injection double or triple of the low dose cohorts, they experienced an increase in the naive cell production λ as seen in patients 203, 304 and 303. As for the immunoconcordant individuals, patients with CD4 T-cell count > 400 [31], no matter what the injection dose was they experience an increase in the naive cell production λ. The central memory cell death rate d_C have increased in all

immunodiscordant patients(103,104,203,303,304) along with two of the low and medium dose cohorts that are immunoconcordant which had < 500 cells/μL. When an immunoconcordant medium dose cohort patient had > 500 cell/μL and a high dose immunoconcordant patients(305), the death rate of the central memory cells d_C decreased. The death rate of the transitional memory cell d_T have increased in low dose immunodiscordant individuals along with the immunoconcordant medium dose cohort(201,302). However, it decreased in immunodiscordant medium and high dose cohort(203,303,304) and high dose immunoconcordant individual (305). The effector memory cell death rate d_E, decreased in immunoconcordant high dose cohort(305) and in immunodiscordant high and medium dose cohort(303,304,203). d_E increased in all low dose cohort(102,103,104) and immunoconcordant medium dose cohort(201,302). As for the memoy stem cell proliferation rates p_{ST}, immunodiscordant low dose cohort(103,104) and immunoconcordant high dose cohort(305) experienced a decrease. It is important to note that patient 305 is the only patient who experienced a decrease in the total CD4 T-cell count 3 years after the treatment. The immunodiscordant medium and high dose cohort (203,303,304) along with the immunoconcordant low and medium dose cohort(102,201,302) have expereinced an increase in the memory stem cell proliferation 3 years after the treatment initiation.

As for the uncertainty and sensitivity analysis results we obtained the following result summary which can be summarized in the Tables 4.10, 4.11, 4.12,4.13 and

4.14 below:

- The naïve cell production rate λ had a positive correlation with all the naïve and memory CD4 T-cells in all immunoconcordant individuals no matter what the injected dose of the CCR5-down-modulated memory CD4 T-cells was along with the high and medium dose immunodiscordant individuals(203,303,304). However, the immunodiscordant low dose individuals did not experience that significant importance of the naïve cell production rate λ on the naïve and memory CD4 T-cell populations (patients 103 and 104).

- The proliferation rate of the central memory cell p_C had a positive correlation on the CD4 T-cell sub-populations when immunodiscordant low dose patients did not have an increase in their naïve production rate after the treatment as for pat. 104. In addition, immunoconcordant low dose pat 102 transitional memory CD4 T-cell population had a positive correlation with the central memory cell population. This further indicates that the central memory cell proliferation rate was more important in low dose immunodiscordant individuals that did not experience an increase in the naïve cell production rate λ.

- The proliferation rates for memory stem cell p_{ST} and naïve cells p_N had a positive correlation withe the CD4 T-cell subsets for low dose immunodsicordant

and immunoconcordant individuals(102,103). Moreover, memory stem cell proliferation rate p_{ST} had a positive correlation with the CD4 T-cell subsets of high dose immunodiscordant individual 303. As for the immunodiscordant low dose patient 104 where the naïve and memory stem cell proliferation rates along with the naïve cell production rate λ did not play an important role, the proliferation of the central memory cell p_C had a positive correlation with the CD4 T-cell sub-populations.

- The central memory death rate d_C was negatively correlated with the CD4 T-cell population of the low dose immudiscordant and immunoconcordant individuals(102,103,104). In addition we observed the same effect in all immunoconcordant individuals regardless of the injected dose of CCR5-down-modulated memory CD4 T-cell except for medium dose patient 302.

- The death rate of the memory stem cell population had a negative correlation on the CD4 T-cell subsets only in the medium dose immunoconcordant individual 302.

- The transition rate of the central memory cell ϕ_C had a negative correlation on the total CD4 T-cell count and a positive correlation with the effector and transitional memory CD4 T-cells in all high dose cohorts despite their CD4 T-cell levels. In addition the same was observed in immunodisocrdant

medium dose cohort patient 203. The naïve cell transition rate ϕ_N had a negative correlation with the total CD4 T-cell count in medium and low dose immunodiscordant individuals and one immunoconcordant high dose individual that experienced a decrease in the total CD4 T-cell count after three years as illustrated in table 1.1 in Chapter 1.

Now comparing the patients according to the injected CCR5-down-modulated memory CD4 T-cells we can draw the following conclusions:

- For the low dose cohort, when the patient had a fairly high CD4 T-cell count at baseline, at the end of the treatment the patient experienced a small amount of increase in the Total CD4 T-cell count. The naïve and memory stem cells played an important role in this increase. However, when the patient had a fairly low amount of CD4 T-cell at baseline, after the reinstatement the count of CD4 T-cell almost doubled. In these patients we observed two important points. If the patients had an increase in the number of naïve cell after the treatment, the naïve cells were the only population that played an important role in this increase. But when the number of naïve did not have any significant change, the central memory cell is the population that played an important role in the increase of the T-cell count.

- For the medium dose cohort, when the patient had a low CD4 T-cell count

at baseline, the naïve and central memory cell played an important role in increasing the total CD4 T-cell count by more than double the count. When having a somewhat high baseline count(400-500) at baseline, the patients experience an increase by a 1.5 factor where the memory stem cell and naïve cells played an important role in the CD4 T-cell count increase. However, it is important to note that when the number of naïve cell produced by the thymus did not have a significant increase from baseline value, the memory stem played an important role. But when the number of naïve cell produced by the thymus had a significant increase the naïve and central memory cells played an important role in the reconstitution of the CD4 T-cell count as the low dose cohort. These patients experienced a higher increase in their CD4 T-cell count compared to the other two cohorts.

- For the high dose cohort, the patient who experienced the most significant increase in the CD4 T-cell was due to an increase in the naïve cell production rate λ and a decrease in the memory central cell population. When the naïve cell production did not have any increase, the memory stem cell was the population that contributed to the increase in the total CD4 T-cell. However, when one of the patients had a decrease in the production rate of the naïve cell, the total number of CD4 T-cell at the end the three years decreased significantly. This points out the importance of the naïve cell which was

evident in all the three cohorts. This last result supports previous finding where researchers found that the CD4 T-cells are highly resistant to the HIV infection [53].

Parameters	N_{103}	N_{104}	S_{103}	S_{104}	C_{103}	C_{104}	T_{103}	T_{104}	E_{103}	E_{104}	TOT_{103}	TOT_{104}
λ		+		+							+	
p_N	+		+		+		+		+		+	
p_{ST}	+		+	+	+		+		+		+	
p_C		+				+	+	+		+		+
p_T		+								+		
p_E		+										
ϕ_N	-	-									-	
ϕ_{ST}			-	-								
ϕ_C			-		-				+		-	
ϕ_T							-	-	+			
d_C					-	-	-	-	-	-	-	-
d_E							+		-	-	+	

Table 4.10: Correlation between the model parameters and the CD4 T-cell population. Where + = positive correlation and - = negative correlation for low dose immuodiscordant patients 103 and 104.

Parameters	N_{304}	N_{303}	S_{304}	S_{303}	C_{304}	C_{303}	T_{304}	T_{303}	E_{304}	E_{303}	TOT_{304}	TOT_{303}
λ	+	+	+	+	+	+	+	+	+	+	+	+
p_{ST}				+		+		+		+		+
ϕ_N	-	-									-	
ϕ_{ST}			-	-								-
ϕ_C					-			+		+	-	
ϕ_T							-	-	+			
d_C				+		-		-		-		-
d_T							-		-			
d_E									-			

Table 4.11: Correlation between the model parameters and the CD4 T-cell population. Where + = positive correlation and - = negative correlation for immunodiscordant high dose patients(303,304).

Parameters	N_{102}	S_{102}	C_{102}	T_{102}	E_{102}	TOT_{102}	N_{305}	S_{305}	C_{305}	T_{305}	E_{305}	TOT_{305}
λ	+	+	+	+	+	+	+	+	+	+	+	+
p_N	+	+	+	+	+	+						
p_{ST}	+	+	+	+	+	+						
p_C				+								
ϕ_N	-					-	-					-
ϕ_{ST}		-						-				
ϕ_C			-		-	-			-	+	+	-
ϕ_T				-						-		-
d_C			-	-	-	-			-	-	-	-
d_E					-						-	

Table 4.12: Correlation between the model parameters and the CD4 T-cell population. Where + = positive correlation and - = negative correlation for immunoconcordant Low dose(102) and high dose patient (305).

Parameters	N_{201}	N_{302}	S_{201}	S_{302}	C_{201}	C_{302}	T_{201}	T_{302}	E_{201}	E_{302}	TOT_{201}	TOT_{302}
λ	+	+	+	+	+	+	+	+	+	+	+	+
p_E										+		
ϕ_N	-	-							+		-	-
ϕ_{ST}			-	-		+		+				
ϕ_C					-	-	+		+			
ϕ_T							-	-		+		-
d_{ST}				-		-		-		-		-
d_C					-		-		-		-	
d_E										-		

Table 4.13: Correlation between the 15 model parameters and the 5 T-cell subsets along with the total number of T-cell for immunoconcordant medium dose patients(201,302).

Parameters	N_{203}	S_{203}	C_{203}	T_{203}	E_{203}	TOT_{203}
λ	+	+	+	+	+	+
ϕ_N	-					-
ϕ_{ST}		-				
ϕ_C			-			-
ϕ_T				-		-

Table 4.14: correlation between the 15 model parameters and the 5 T-cell subsets along with the total number of T-cell for immunoconcordant medium dose patient 203

5 Variability in CD4 T-cells population

5.1 Introduction

In addition to parameter uncertainty, model variability is an important matter to address when modeling in host biological phenomenon. As our main focus in this work is the dynamics of the memory and naïve CD4 T-cells, it is important to address the natural variability that arises from the stochastic nature of the T-cells in the body in each of the nine patients.

5.2 Stochastic Model

In the previous chapter, we presented our stochastic model that we derived using the Gillespie's algorithm. In this section we will compare the results of the steady state obtained by both the deterministic model Eq's 2.1 and the stochastic model Eq's 2.4. Using a stochastic model will allow us to estimate the natural variability in the naïve and memory CD4 T-cell model that could not be captured in

a deterministic setting. To compare the ODE and stochastic model, we will simulate the five T-cell subsets along with the total number of CD4 T-cell in a single μL of plasma. We will be using the parameter estimates obtained from the baseline fit in Matlab (section 3.3.1), Stan fit (3.3.2) and the best fit from Monolix. This methodology will allow us to further investigate to what extent each of the fit can capture the data measurements by the fluctuations that arise from the stochastic model. Later, we will calculate the steady state value of the five cell subsets using both the ODE(Eq's 2,1) and MCMC(Eq's 2.4), along with the standard error of the mean. The standard error of the mean, SEM, will be obtained by dividing the standard error by the square root of the number of runs performed in the stochastic model. The results for each of the three cohorts is represented separately.

In Tables 5.1, 5.2 and 5.3 it was shown that for each of the parameter estimates, the steady state values obtained from both the ODE and MCMC were similar with a small SEM.

Low Dose Cohort

In Table 5.1, it was shown that for the three patients, the naïve cell steady state after the initiation of the treatment was the same as the steady state achieved at baseline. This indicates that the naïve cell did not increase in count in the low dose

cohort. As for the memory stem cell, patients 103 and 102 did not have an increase in the stem cell count. However, patient 104 showed a significant increase where the data value was in 1 standard deviation with the steady state obtained from the parameter estimation in Monolix. The central memory cell population estimated count by the ODE and stochastic model using the Monolix parameter fit was in close proximity to the real data value at 36 months post treatment for all the three patients which indicated an increase from the baseline value. Patient 104 and 103 steady state value was in close proximity to the steady state given by the baseline fit. This indicates that this population of cell did not exhibit any significance increase. However, patient 102 steady state value was closer to the parameter fit obtained from Monolix. For this patient the transitional memory cell steady state increased from the baseline value. In patient 104, the effector memory data measurement at 36 months, was closer in value to the steady state obtained by the MCMC model using Stan fitting results which resulted in a decrease in the population from the baseline value. Patient 103 and 102 data measurements at 1080 days was closer to the steady state calculated using the MCMC from the fitting obtained in Monolix. As for the total number of CD4 T-cell, all of the patients had an increase from the baseline value where the Monolix estimation using the MCMC model was closer to the data measurement at 36 months.

From Table 5.1, we can conclude that the patient that had the most significant

increase in the total number of CD4 T-cell from baseline is the one that had an increase in the memory stem cell population as a result of the experimental treatment. This was evident in patient 104 where the baseline for the total number of T-cell was 261 cell per μL that increased to 455 cell per μL three years post-treatment. This again highlights the importance of the memory stem cell population.

Figure 5.1 shows that most of the CD4 T-cell population for the three patients were in at least 1 standard deviation from the steady state calculated from the MCMC model using the Monolix data fitting values. This indicates that our stochastic model was able to capture the dynamics of the CD4 T-cell. In addition, a realization of the stochastic model (Figure 5.2) using the Monolix fit (LN) shows how the fluctuations of our model was able to capture the data measurements for all the three patients. However, the Total CD4 T-cell data measurement was not captured, this is because we defined the total CD4 T-cell by simply adding the five T-cell subsets differently measured unlike how our collaborator measured those cells.

		Naïve	Memory stem	Central memory	Transitional memory	Effector memory	Total CD4 T-cell
102 Baseline	ODE	40.77	18.8	86.23	34.45	129.7	310.8
	MCMC	40.6417	18.63455	86.1430	35.7022	130.5261	311.6530
	±SEM	0.3686	0.2396	0.5207	0.3406	0.6820	0.9460
102-MON	ODE	79.98	80.39	219.8	54.34	89.3	523.6
	MCMC	80.7210	80.1730	216.8873	53.5757	88.9752	520.3506
	±SEM	0.5734	0.60255	1.1515	0.4781	0.6274	1.7880
102 STAN	ODE	70.25	56.44	133	28.68	53.42	341.4
	MCMC	69.8927	55.9081	134.6426	29.4694	54.7132	344.6428
	±SEM	0.5158	0.5657	0.9534	0.3652	0.6331	1.5329
Val. at 36 months		51.6	28.4	197.6	48.2	102.3	518
104 Baseline	ODE	27.98	15.89	71.89	30.53	73.68	220.2
	MCMC	27.4404	15.5202	68.8210	30.3845	73.022	215.1937
	±SEM	0.30686	0.2308	0.5779	0.3419	0.5881	1.1155
104 MON	ODE	67.8813	119.9	104.7	22.79	15.38	331.2
	MCMC	67.8813	116.8126	103.2355	22.8445	15.3936	326.2180
	±SEM	0.4862	0.8149	0.7338	0.2989	0.2561	1.51126
104 STAN	ODE	58.11	77.01	89.24	18.54	29.32	270.3
	MCMC	57.8486	77.4645	88.9932	17.9478	28.9492	271.2814
	±SEM	0.4869	0.6107	06717	0.2717	0.3962	1.2750
Val. at 36 months		76.7	119.3	124.6	30.6	27.5	455
103 Baseline	ODE	37.8	12.64	58	24.86	14.35	147.6
	MCMC	37.3896	12.7074	56.3290	24.1237	14.1110	144.6629
	±SEM	0.3434	0.1981	0.4422	0.2848	02194	0.7199
103-MON	ODE	27.51	12.92	98.36	35.58	65.8	240.10
	MCMC	28.4891	13.2551	101.6508	36.9183	68.2782	248.5790
	±SEM	0.3206	0.2152	0.7425	0.4233	0.7867	1.2938
103-STAN	ODE	26.63	17.98	64.52	20.96	39.61	169.7
	MCMC	26.3436	18.1230	63.8209	19.7328	38.0755	166.0909
	±SEM	0.3111	0.3786	0.7337	0.3553	0.4916	1.2115
Val. at 36 months		38	9.6	81.9	28.6	104	315

Table 5.1: Low dose cohort's steady state for each of the CD4 T-cell subsets using the ODE and stochastic models. 300 runs are used in the MCMC model.

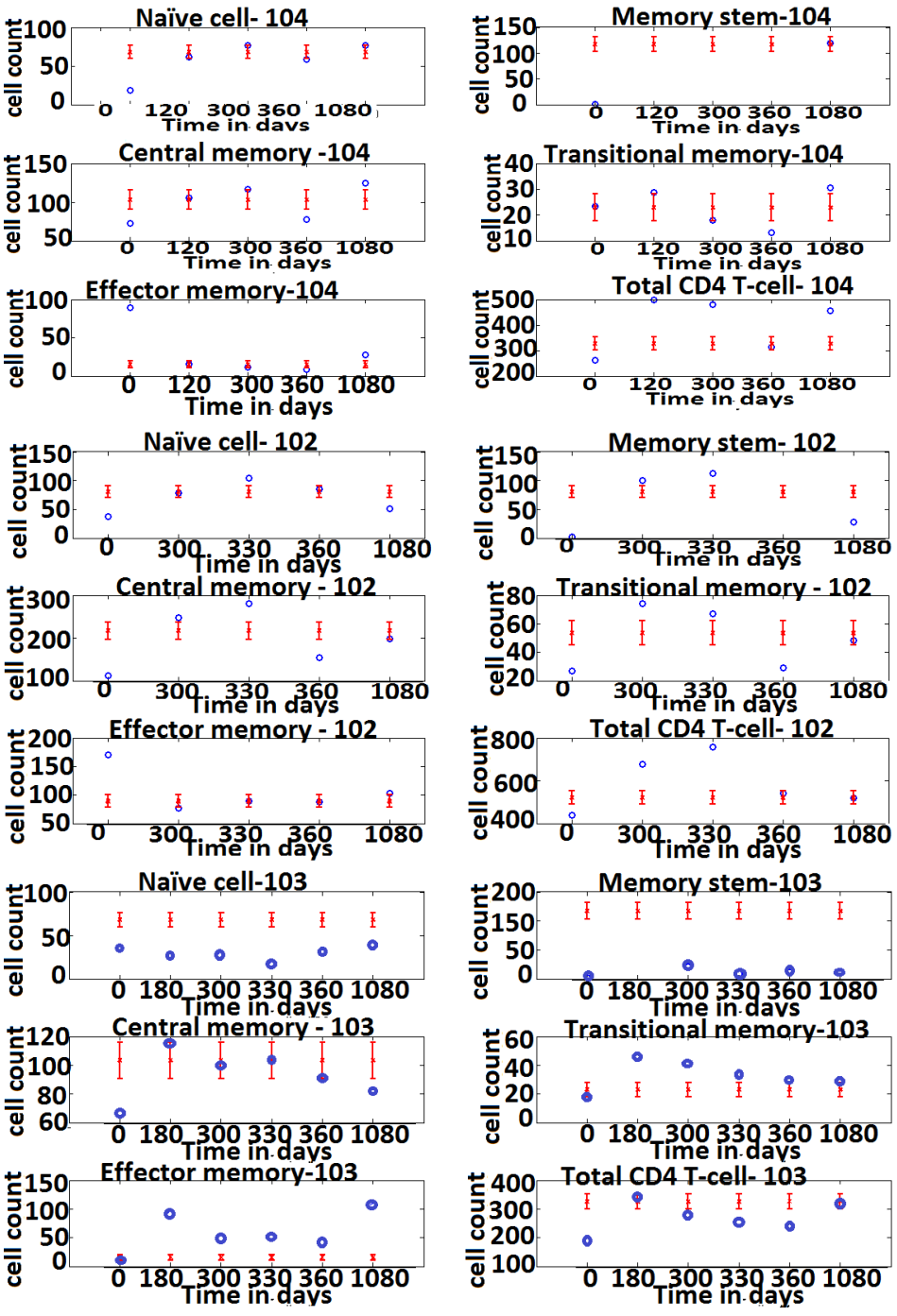

Figure 5.1: Error bar low dose cohort of the steady with experimental data points.

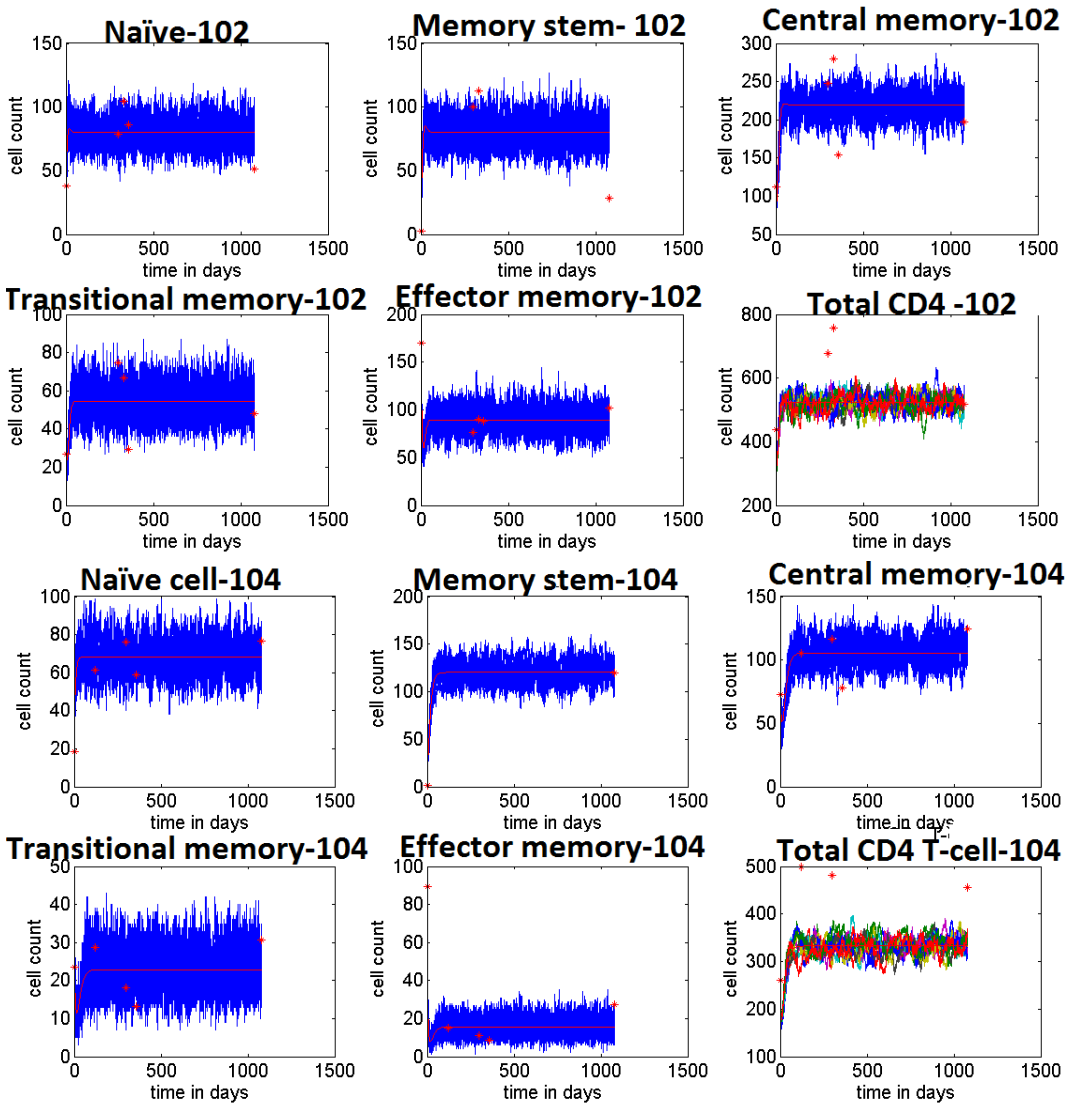

Figure 5.2: 10 stochastic realizations for low dose cohort.

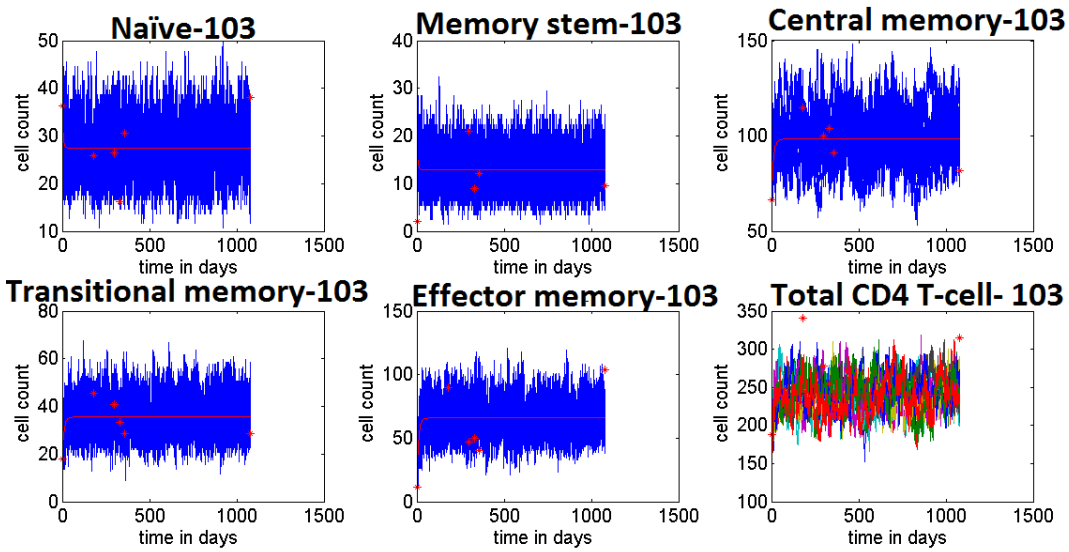

Figure 5.3: 10 stochastic realizations for low dose cohort.

Medium Dose Cohort

Looking at Table 5.2 and Figures 5.3 and 5.4 we observe the observation below :

In patient 203, the data measurements for the six population of cells at 36 months post-treatment was in close proximity to the steady state obtained in the MCMC by using data estimates from Monolix. The memory stem cell along with the naïve and central memory cells showed a significant increase from baseline. The total number of CD4 T-cell have increased form 294 to 617 cells per μL. The importance of the memory stem, central memory and naïve cells play an essential role in the reconstitution of the CD4 cell count in this patient. It is important to note that the increase in the total CD4 T-cell count was one time an a half greater than the patients from low dose with similar baseline count for the total CD4 T-cell.

As for patient 302, the data measurements for the memory stem, transitional memory and effector memory were in at least one to two standard deviation from the steady state calculated by the MCMC model as shown in Figure 3.3. However, the MCMC model for patient 201, showed a small variability between the estimated steady state and the data measurements after three years, indicating that the MCMC model does a fairly good job in describing the dynamics of the CD4 T-cell for patient 201. All of the three patients had an increase in their total CD4 T-cell populations. Figure 5.4 shows how the stochastic model suing the Monolix parameter estimates was able to capture all the data measurements. This further indicates the validity of

our model and the importance of using stochastic models to capture variably within and between individuals.

		Naïve	Memory stem	Central memory	Transitional memory	Effector memory	Total CD4 T-cell
203 Baseline	ODE	52.42	18.44	85.43	35.26	25.18	216.7
	MCMC	52.9382	18.3199	85.2483	35.0252	25.4318	216.9625
	±SEM	5.4730	0.3026	0.6717	0.4101	0.36628	1.0479
203-MON	ODE	70.66	69.55	162.4	36.95	26.68	366.2
	MCMC	71.6534	70.8923	165.0647	37.2156	26.9858	371.8109
	±SEM	0.7155	0.7410	1.0988	0.4539	0.37901	1.8823
203 STAN	ODE	68.14	34.74	124.7	36.01	67.91	329.5
	MCMC	68.7948	36.2339	128.4525	36.2193	68.2193	337.7407
	±SEM	0.61518	0.5812	1.13985	0.50911	0.6505	1.75928
Val. at 36 months		108.7	70.7	202.5	46.7	34.2	617
302 Baseline	ODE	148.7	15.79	69.03	29.92	27.94	291.4
	MCMC	148.2785	16.0114	68.4546	30.0212	28.16658	290.9211
	±SEM	0.8061	0.29998	0.608	0.4186	0.4044	1.15965
302 MON	ODE	70.55	69.47	162.3	36.93	26.65	365.9
	MCMC	70.535	69.4977	158.6232	36.6385	26.1904	361.4967
	±SEM	0.7976	0.7311	1.11157	0.45113	0.36911	1.8087
302 STAN	ODE	162	81.35	11.62	3.353	4.506	262.9
	MCMC	161.317	81.1099	11.4666	3.0770	4.6135	262.4085
	±SEM	0.4869	0.6107	0.6717	0.2717	0.3962	1.2750
Val. at 36 months		236	42	98.8	39.9	51.7	606
201 Baseline	ODE	127.5	20.32	96.46	39.06	22.53	305.8
	MCMC	127.5483	20.4968	97.3861	39.3882	23.032	307.8527
	±SEM	0.81034	0.3337	0.7566	0.4511	0.3394	1.3010
201-MON	ODE	210	297.9	313.11	85.87	36.4	943.3
	MCMC	209.5010	299.0.79	311.4175	85.0449	36.2554	941.2943
	±SEM	0.9630	1.429	1.5924	0.6943	0.451135	2.630
201-STAN	ODE	140.5	213.4	210.1	61.24	26.25	651.5
	MCMC	139.5236	211.5900	207.9940	61.9867	26.2680	647.2886
	±SEM	0.7608	1.1228	1.2247	0.5133	0.4030	2.3122
Val. at 36 months		176.7	143.8	248.9	54.8	27.2	848

Table 5.2: Medium dose cohort's steady state for each of the CD4 T-cell subsets using the ODE and stochastic models. 300 runs are used in the MCMC model.

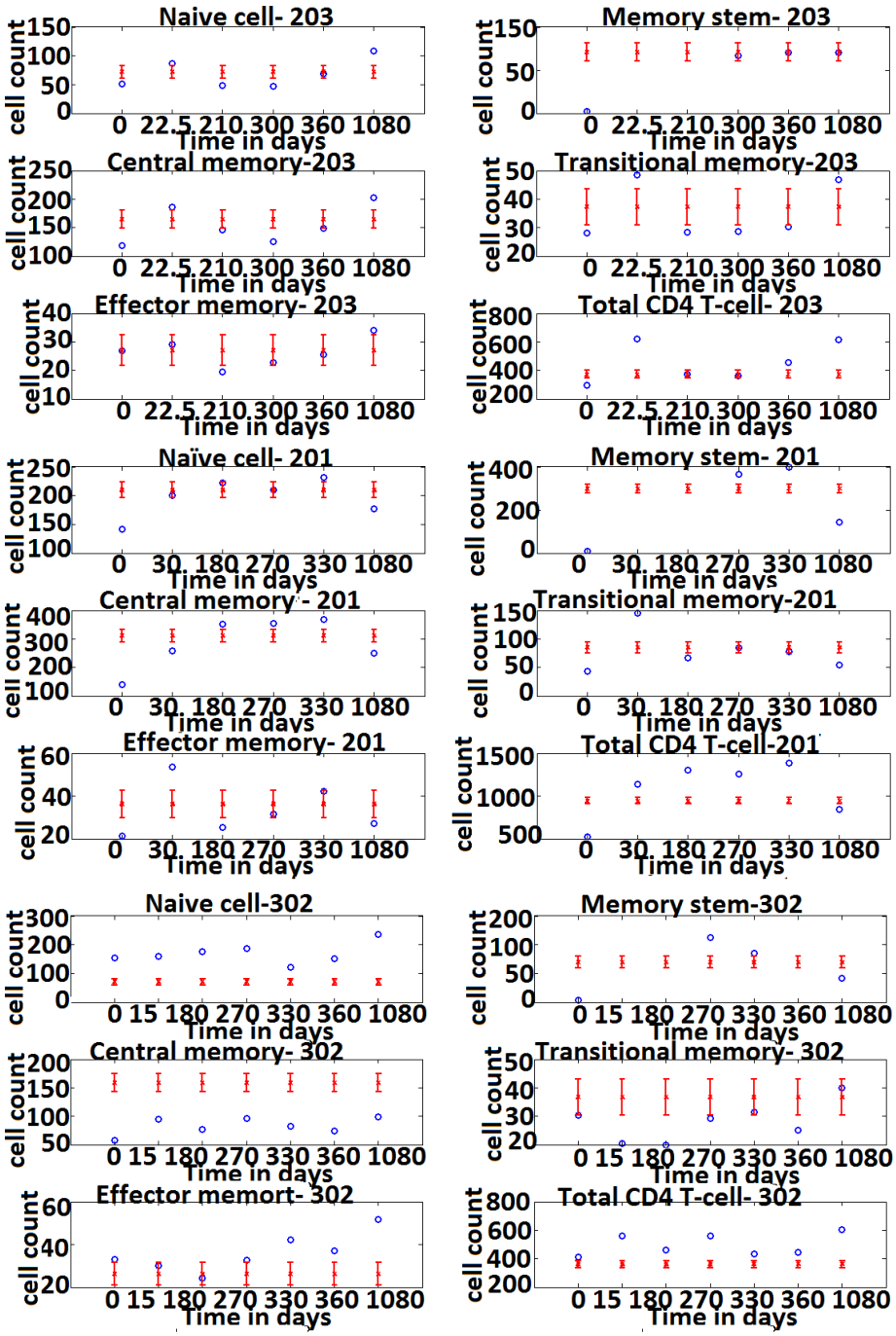

Figure 5.4: Error bar for medium dose cohort of the steady state of the MCMC model

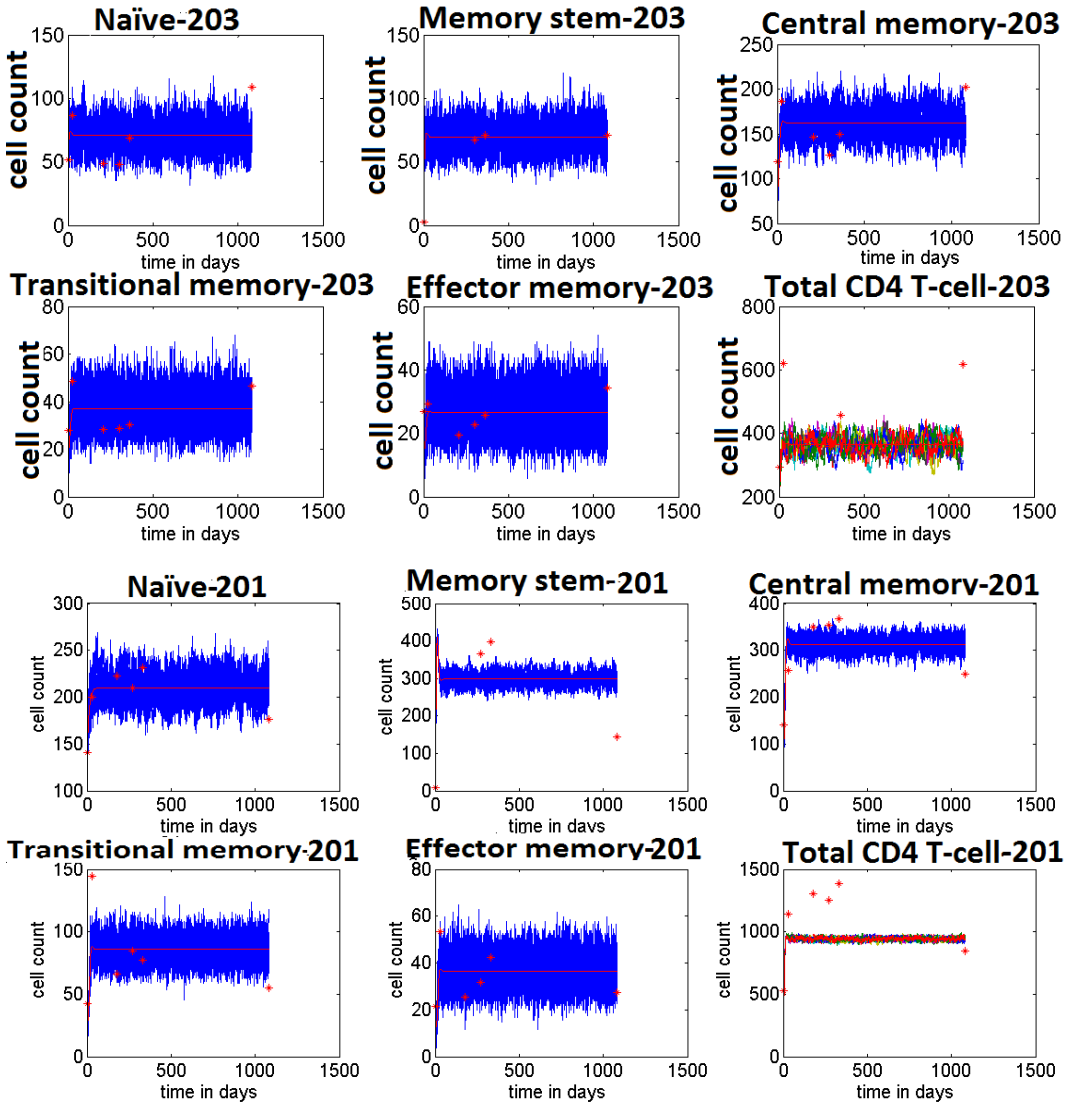

Figure 5.5: 10 stochastic realization for the medium dose cohort.

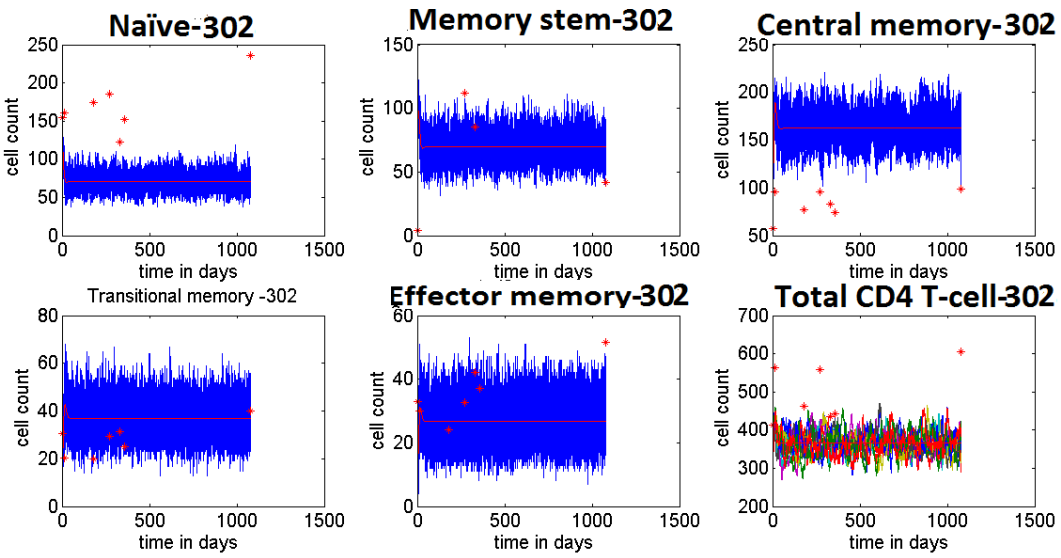

Figure 5.6: 10 stochastic realization for the medium dose cohort.

High Dose Cohort

Looking at Table 5.3 and Figures 5.5 and 5.6 we observe the observations below. In patient 305, the data measurements for the six population of cells at 36 months post-treatment was in close proximity to the steady state obtained in the MCMC by using data estimates from Monolix. The number of central memory, naïve and transitional memory cells decreased as shown in Figure 5.5 which led to a decrease in the total number of CD4 T-cell. The MCMC model was able to capture the CD4 T-cell dynamic for this patient. This further validates the goodness of our model. Patient 303 CD4 T-cell population confirmed the validity of our MCMC model where most of the data measurement where within the standard deviation bar at the steady state as shown in Figure 5.5. This patient did not show a significant increase in all of the memory cell count which led to having a very small increase in the total number of CD4 T-cell.

As for patient 304, the data measurements for all of the CD4 T-cell subsets were in at least one to two standard deviation from the steady state calculated by the MCMC model as shown in Figure 5.5. Similarly, as this patient did not experience an increase in their memory stem and transitional memory, it did not result with an increase in the total CD4 T-cell count. Nevertheless, this patient was not followed for a period of three years hence we are not able to firmly understand the effect of the clinical trial on their CD4 T-cell count. Figure 5.6 further shows that fluctuations

from our stochastic model are able to capture the data measurements. However, as the previous patients, the CD4 T-cell memory is not well captured as the definition between the model and the experimental data of the total CD4 T-cell varies. this indicates that our model underestimates the experimental data measurements for the the total number of CD4 T-cell.

		Naïve	**Memory stem**	**Central memory**	**Transitional memory**	**Effector memory**	**Total CD4 T-cell**
304 Baseline	ODE	50.41	18.44	83.86	36.23	24.25	213.2
	MCMC	50.8774	18.7154	83.9669	36.3757	23.8453	213.7767
	±SEM	0.03258	0.0217	0.0511	0.0315	0.258	0.0804
304-MON	ODE	104.3	119.7	104.5	69.63	28.23	425.8
	MCMC	106.5673	119.3142	104.5175	70.1099	27.9563	428.4660
	±SEM	0.92936	1.083	0.8993	0.67025	0.4052	2.1189
304 STAN	ODE	97.43	97.51	179.7	122.1	43.98	540.4
	MCMC	96.2795	99.3641	179.0027	122.8336	44.1077	541.5196
	±SEM	0.6986	1.190	1.13987	0.98446	0.5126	2.69107
Val. at 36 months		114.5	NaN	203.8	84.4	41.8	757
303 Baseline	ODE	92.75	16.33	73.81	31.32	15.86	230.1
	MCMC	92.3040	1.7585	75.2461	31.2035	16.0882	231.6068
	±SEM	0.0452	0.0213	0.0477	0.0261	0.0187	0.0806
303 MON	ODE	108.6	96.56	21.14	30.15	14.91	271.4
	MCMC	109.2415	94.0503	21.7402	31.2544	15.3950	271.6823
	±SEM	0.05176	0.06091	0.0239	0.0513	0.0274	0.1034
303 STAN	ODE	100.1	86.06	122.1	23.88	13.74	345.5
	MCMC	99.7113	84.3362	121.2785	22.3015	13.3420	341.0045
	±SEM	0.0547	0.0548	0.0628	0.0251	0.0176	0.1091
Val. at 36 months		129.2	64.4	147.7	30.4	19.8	525
305 Baseline	ODE	131.4	19.45	88.91	39.03	30.51	309.3
	MCMC	132.0991	18.9786	88.2353	38.5513	30.2379	308.0919
	±SEM	0.0509	0.0221	0.0471	0.0301	0.0289	0.0871
305-MON	ODE	95.02	26.85	107.21	38.14	42.07	309.2
	MCMC	95.9471	27.1896	107.2144	38.1235	42.2247	310.6977
	±SEM	0.0559	0.0278	0.0528	0.0322	0.0354	0.1074
305-STAN	ODE	112.2	39.03	91.95	30.04	32.49	305.7
	MCMC	113.1999	39.5414	91.0155	29.7937	32.0592	305.5785
	±SEM	0.0481	0.0434	0.0592	0.0271	0.0269	0.1035
Val. at 36 months		65.5	27.9	88.4	35.9	54.1	340

Table 5.3: High dose cohort's steady state for each of the CD4 T-cell subsets using the ODE and stochastic models. 300 runs are used in the MCMC model.

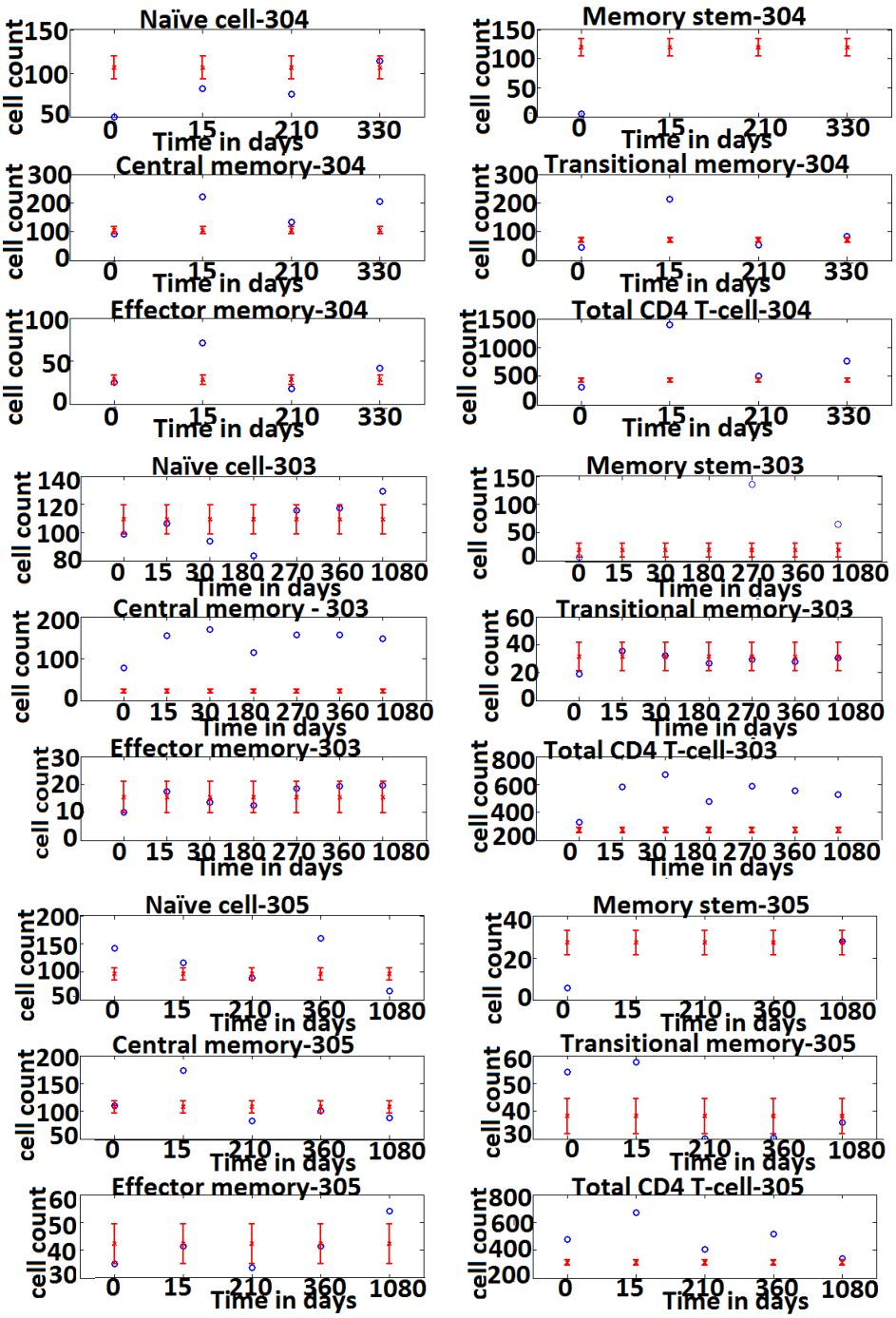

Figure 5.7: Error bar graph for high dose cohort of the steady state of the MCMC model.

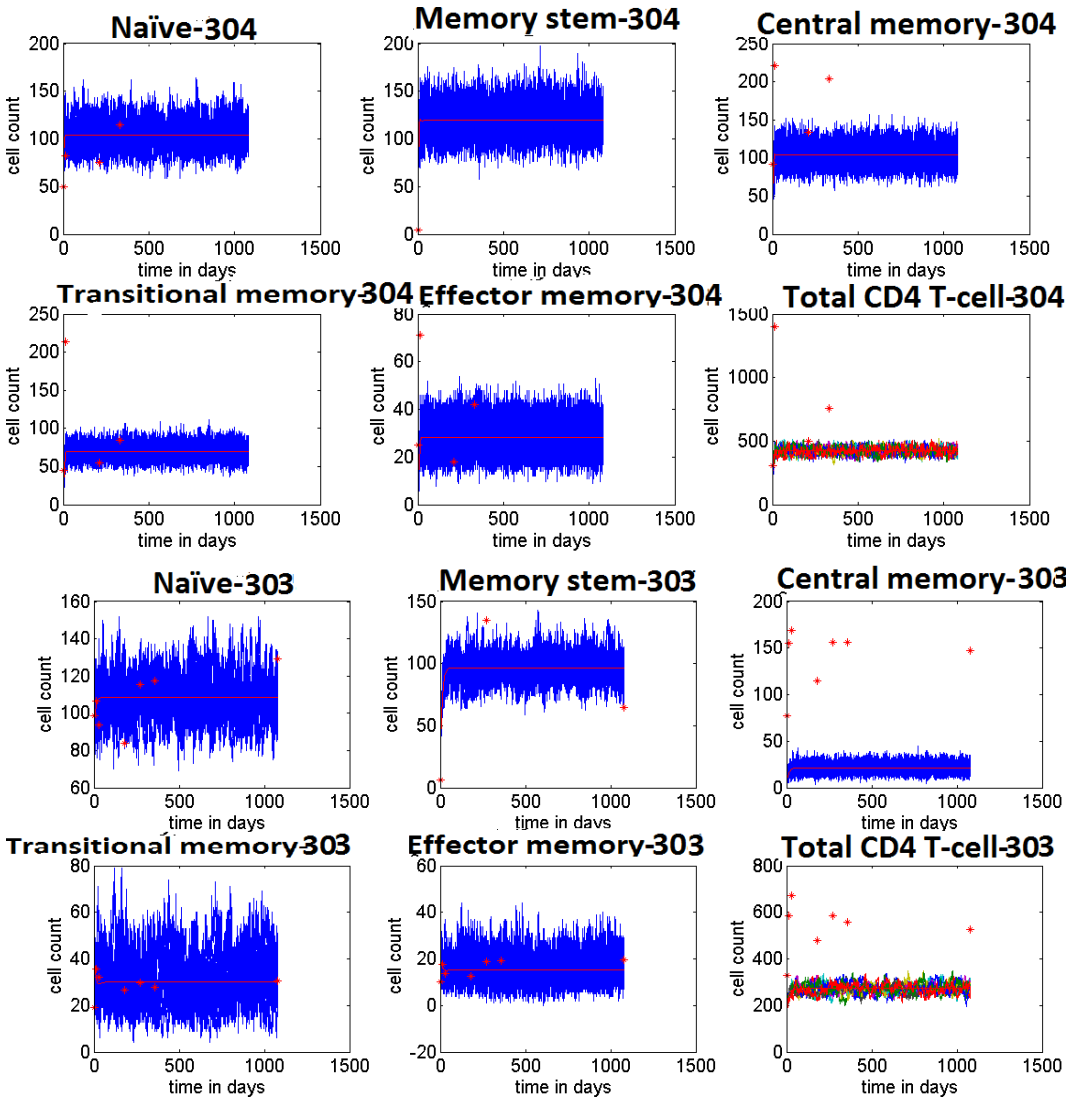

Figure 5.8: 10 stochastic realization for the High dose cohort.

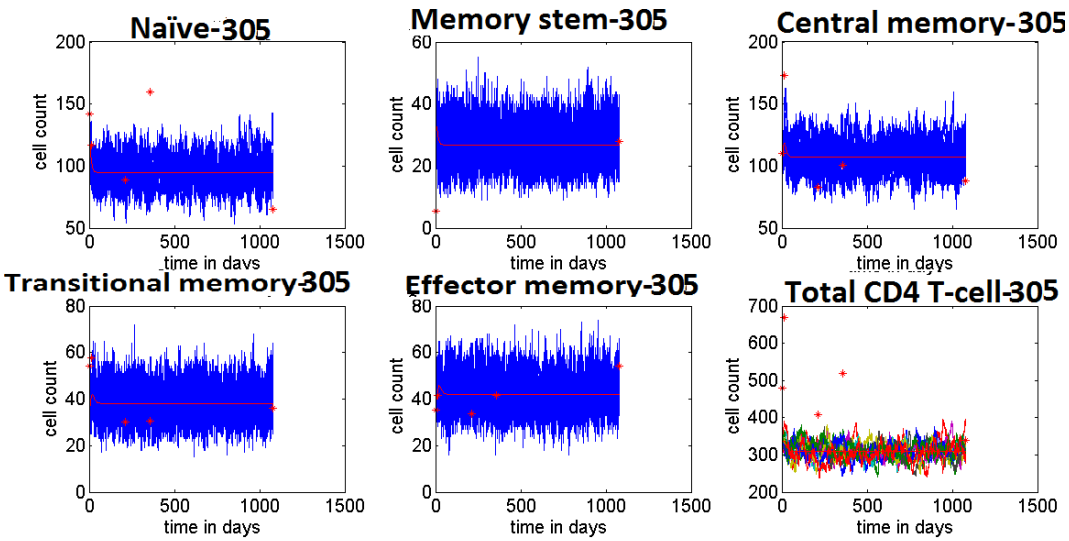

Figure 5.9: 10 stochastic realization for the High dose cohort.

6 Conclusion and Future Directions

In mathematical biology, the main purpose of mathematical models is to gather the proper information and analysis with the aim of informing medical researchers and health care professionals about the efficacy of a drug.

Highly active antiretroviral therapy (HAART) have been used to suppress and prevent the HIV replication with the aim of delaying and slowing down the disease progression. However one of the main drawbacks and challenges of the HAART are the toxicity and the emergence of the drug-resistant strains of the virus as HAART is a life long therapy. Due to these facts and the nature of this virus, researchers have switched their focus from making a vaccine to using immunotherapy. HIV immunotherapy is used to boost and utilize the body's own immune system in order to suppress the HIV progression. The success of this approach remains unclear to the scientific community.

This thesis focused on investigating the effect of the insertion of the CCR5-down-modulated memory CD4 T-cell on reconstituting the immune CD4 T-cell population

in nine different chronically infected HIV patients. The key challenge in analyzing the effect of such study, is to build a sufficiently exact model that describes the immune system dynamics and estimating important biological parameters from the experimental data.

In Chapter 2, we have introduced a deterministic model using an ODE set of equations to describe the inward and outward flow within and between the 5 CD4 T-cell sub-populations. Later, a stochastic model was developed using the Gillespie algorithm.

In chapter 3, we have showed several fitting routines used to estimate the model parameters. The best fit model was chosen based on some statistical criterion such as the AIC and the standard error. Plots for the individual fits and the observed vs prediction graphs were used to asses the goodness of the fits. Parameter identifiablity and over-fitting issues were addressed. It is important to mention that some of the proliferation rates were 0, indicating that the number of CD4T-cell have increased to a maximum capacity that proliferation cannot take place. This further indicates the positive effect of the treatment. In addition, we observed that the death rates for the stem cell was 0 for a lot of patients, indicating that the stem memory are having an increase in their lifespan. This further points out the importance of such cell populations.

In Chapter 4, in order to highlight the change in the fifteen biological parameters

(birth, death, proliferation and transition), after the initiation of the treatment we compared the estimation of the best fitted model to the baseline. We realized that the naïve cell and memory stem lifespan have increased in all the nine patients. In addition, the naïve cell production rate λ increased in all immunoconcordant patients despite the infused dose but increase only in immundiscordant patients that received a medium and high dose of CCR5-down-modulated memory CD4 T-cells. However, λ decreased for all low dose immunodicordant individuals. Further, using the uncertainty sensitivity analysis techniques, we studied the importance of these model parameters to each of the 5 sub-population and the total CD4 T-cell populations. From these results we are able to conclude that the most important cell populations were the naïve, memory stem and central memory cells. This was evident by observing the positive effect of the thymic production rate λ, memory stem cell and central memory cell proliferation rates p_{ST} and p_C along with the negative correlation with the naïve, memory stem and central memory death rates d_N, d_{ST} and d_C. In addition, the medium dose cohort experienced on average a higher significant increase in the total number of CD4 T-cell. In a nutshell, from this experiment we were able to show that not just the naïve cell are important in reconstitution the CD4 T-cell memory as proven in previous studies, but when infusing the CCR5-down-modulated memory CD4 T-cells, the central memory and memory stem cells had an essential role in this reconstitution.

Moreover, in Chapter 5, our MCMC simulations have shown a great agreement with the empirical data measurements which indicate the model and the parameter estimation obtained using Monolix describes to a good extent the dynamics of the CD4 T-cell in chronically infected HIV patients. This validates the conclusions we obtained in chapter 4 about the importance of some of the naïve and memory CD4 T-cell populations. Further more, the stochastic fluctuations were shown to successfully capture all the experimental data points.

6.0.1 Future Work

As this work showed that the naïve, memory stem cell and central memory cell played an important role in the CD4 T-cell reconstitution, future work could focus on further investigating these types of CD4 T-cell. Moreover, as an extension for this work, we can include the integrated viral DNA in our CD4 T-cell dynamical model. This will allow us to further understand the HIV reservoir in the nine patient so we are able to better understand the change in behavior for these sub-populations after the initiation of the treatment. In this work we considered a linear CD4 T-cell model. However, it was shown that backward transitions between the memory stem, central and transitional memory cells is possible. This might be able to help us get a more realistic insight about the real dynamics of the CD4 T-cell and the virus in chronically infected HIV patients. This step is our future direction to

further elaborate on in collaboration with Dr. Sekaly et al. Laboratory at Case Western University. Moreover, when fitting the data, we have faced several issues that affected the goodness of our results. We are planning to elaborate on this area considering the following issues, with possible collaboration with Dr. Giles Hooker at Cornell University.

- Allow some backward transition between the memory stem , central, transitional and effector memory cells. This will allow us to capture hidden dynamics that the current model was not able to do so. This issue was evident when trying to estimate some of the proliferation and transition rates for these cells in most of the patients.

- Perform a multiple imputation to better deal with the missing points.

- Have more observations per cell population for all of the three cohorts.

- Consider a hierarchical mixed approach allowing us to better understand variability within and between individuals to better understand the effect of injecting CCR5-down-modulated memory CD4 T-cell on the three different cohorts.

7 Appendix

Appendix A

Monolix: SEAM Algorithm- A Simulated Annealing version

This section is a summary of the algorithm presented in (Lavielle, 2015) [50].

The Stochastic approximation expectation-maximization also known as the SEAM which computes an estimate to maximum likelihood (ML).It is divided into two simulation stages:

1. The first stage is to get a neighboring solution through only few iteration from the initial values that was given by the user in the software. In this stage a simulated annealing is preferred to be used in to reduce the simulation time if the given initial values are far from the real solution.

2. The second stage is when the occurrence of the convergence to the located maximum in a deterministic behavior similar to the gradient algorithm.

Before presenting the SEAM algorithm it is important to describe the EM algorithm

that is used in the SEAM algorithm.

First we will consider a general model where we have a vector of parameters θ, set of observations y= y_i, unobserved parameters sets z = z_i, where $1 \leqslant$ i $\leqslant$ N and z is normally distributed. The maximum likelihood estimator of θ is given by the following formula:

$$L_y(\theta) = p(y;\theta) = \int p(y,z;\theta)dz \tag{7.1}$$

with p(y,z;θ) being the joint distribution.

As z_i is unobserved, the EM algorithm replaces it by its conditional expectation, where θ_{k-1}^{EM} gets updates at each iteration step k to θ_{k-1}^{EM} given an initial value θ_0 by the following steps:

- **E-step** Where the expectation is evaluated by :

$$Q_k^{EM}(\theta) = E(logp(y,z:\theta)|y;\theta_{k-1}^{EM}) \tag{7.2}$$

- **M-step** In this step the expectation calculated in the E-step is maximized:

$$\theta_k^{EM} = arg_\theta max Q_k^{EM}(\theta). \tag{7.3}$$

As we are dealing with a non linear mixed effect model, the expectation cannot be quantified explicitly using the EM algorithm as the observations y and the parameters z are nonlinear. However, this problem can be solved by using a Monte Carlo approximation using a large number of independent simulations of z. In the

SEAM algorithm a single simulation of z is used.

The iterative steps of the SEAM algorithm are:

- **The simulation step** For i= 1....N, draw z_i^k from the conditional distribution given by p(z_i| y_i;θ_{k-1}), where z_i are the individual parameters, y_i are the observations given by a data set. Monolix uses a Hasting Metropolis-Hastings(MH) algorithm to simulate the individual parameters with a limiting distribution p($z_i|;\mu_{k-1}, \omega_{k-1}$).

- **Stochastic Approximation** In this step will updateQ_k in the following way:

$$Q_k = Q_{k-1} + \gamma_k(logp(y, z^k; \theta) \tag{7.4}$$

  such that γ_k being a decreasing sequence of positive numbers with γ_1=1.This indicates that the stochastic approximation has no memory of previous approximations.

- **Maximization step** In this step we will update $\theta k - 1$ using the M-step in the EM algorithm where:

$$\theta_k = arg_\theta max Q_k(\theta) \tag{7.5}$$

Calculation of the Fisher Information Matrix

The observed Fisher Information Matrix also known as FIM, is defined as the

measurement of the amount of information that an observable random variable posses about an unknown parameter which depends on a specific probability distribution. it is expressed as a function of θ where :

$$I_y(\theta) = -\partial^2_\theta LL_y(\theta) \tag{7.6}$$

Where $\mathrm{L}L_y$ is the log-likelihood defined in equation 2.1 as :

$$LL_y = l(y;\theta) = log(p(y;\theta)). \tag{7.7}$$

As it was mentioned above, due to the complexity of some models, the likelihood does not have a closed from solution which often suggests that the observed FIM cannot have a closed form solution either. Hence it will be approximated using the following methods:

- **Stochastic approximation** Using a Monte Carlo procedure based on Louis' formula(louis 1982), Kuhn and Lavielle estimated the FIM using the following formula:

  $$\partial^2_\theta l(y;\theta) = E(\partial^2 l(y,z,\theta)|y;\theta) + cov(\partial_\theta l(y,z,\theta)|y;\theta)) \tag{7.8}$$

  Where:

  $\partial^2_\theta \mathrm{l}(\mathrm{y};\theta)$ is the conditional expectation that is calculated using an MH Monte Carlo method.

cov(∂_θ l (y,z,θ)|$y;\theta$))= E((∂_θ l(y,z;θ))(∂_θl(y,z;θ))^t|$y;\theta$) -E((∂_θ l(y,z;θ))|$y;\theta$) E ((∂_θ l(y,z;θ))|$y;\theta$)^t

In here it is important to mention that the MH algorithm used for the SEAM, the sequence of estimated parameters θ_k is remained fix at the estimated maximum likelihood $\hat{\theta}$.

- **Model Linearization** Consider the following model:

$$y_{ij} = f(t_{ij}, z_i) + g(t_{ij}, z_i)\epsilon_{ij} \tag{7.9}$$

Where f is the structural model, g is the residual error model and ϵ_i is the standardized residual error.

In this fit, we have used a constant error model (in both the stochastic and linearization methods) g= a so

$$y_i = f_j + a\epsilon_j \tag{7.10}$$

The error model selection was based on the BIC and AIC information. The model is linearized around the predicted vector of parameters $\tilde{z}_i$ in the following way:

$$y_{ij} = f(t_{ij}, \tilde{z}_i) + \partial_z f(t_{ij}, \tilde{z}_i)(z_i - \tilde{z}_i) + g(t_{ij}, \tilde{z}_i)\epsilon_{ij} \tag{7.11}$$

$$= \mathrm{f}(t_{ij}, \tilde{z}_i) + \partial_z f(t_{ij}, \tilde{z}_i)(z_{pop} - \tilde{z}_i) + g(t_{ij}, \tilde{z}_i)\epsilon_{ij} + \partial_z f(t_{ij}, \tilde{z}_i)\eta_i.$$

Where the marginal distribution of the observation vector defined as y_{ij} is approximated using a Normal Distribution with the following mean and standard deviation:

$$\mu_i = f(t_i, \tilde{z}_i) + \partial_z f(t_i, \tilde{z}_i)(z_{pop} - \tilde{z}_i) \tag{7.12}$$

$$\Gamma = \partial_z f(t_i, \tilde{z}_i)\omega\partial_z f(t_i, \tilde{z}_i)^t + g(t_i, \tilde{z}_i)\Sigma_i g(t_i, \tilde{z}_i)^t \tag{7.13}$$

The model selection is based on the Bayesian Information Criterion also known as the BIC where a minimum BIC indicates a better model if the BIC is defined as:

$$BIC = -2LL_y(\hat{z}) + log(n)d. \tag{7.14}$$

Where n is the number of observation and d is the dimension.

Appendix B

Stan Fit

This is based on a communication from Georges Monette (2015), the author of this section.

Six fitted models

Working out a model using linearization as suggested by Giles Hooker(2015) has been very helpful in clearly revealing the large differences in information available for different portions of the parameter space. We will refer to the five types of cells by number: 1, 2,..,5. The model has 15 independent parameters.

- λ_i the rate of autonomous production (per microlitre) of type 1, (Naive cells).
- τ_i i=1..4, the probability per day that a cell of type transforms into a cell of type I+1.
- δ_i i=1..5, the probability per day of death for a cell of type i.
- β_i i=1..5, the probability per day that a cell of type replicates.

let θ_i denote the net depletion probability per day for a cell of type i :

$\theta_i = \delta_i + \tau_i - \beta_i$, i= 1..4

$\theta_5 = \delta_5 - \beta_5$

The stationary expectation for the density per microlitre of each type of cell is:

$\mu_1 = \frac{\lambda}{\theta_1}$

$\mu_2 = \frac{\lambda\tau_1}{\theta_1\theta_2} = \mu_1\frac{\tau_1}{\theta_2}$

$\mu_3 = \frac{\lambda\tau_1\tau_2}{\theta_1\theta_2\theta_3} = \mu_2\frac{\tau_2}{\theta_3}$

$\mu_4 = \frac{\lambda\tau_1\tau_2\tau_3}{\theta_1\theta_2\theta_3\theta_4} = \mu_3\frac{\tau_3}{\theta_4}$

$\mu_5 = \frac{\lambda\tau_1\tau_2\tau_3\tau_4}{\theta_1\theta_2\theta_3\theta_4\theta_5} = \mu_4\frac{\tau_4}{\theta_5}$

Let the sampling volume be n microlitres and the sampling fraction $\phi = \frac{n}{V}$, where V is the total body blood volume in microlitres (approximately 5 litres). Different transformations of the 15 dimensional parameter space are estimable with very different orders of precision. The transformations for which there is the most information are those that are functions of μ_i i=1..5. For example, the $\mu's$ themselves or the four ratios: $\rho_i = \frac{\mu_{i+1}}{\mu_i} = \frac{\tau_i}{\theta_{i+1}}$, i=1..4.
The variance of estimates of μ_i is approximately $\frac{\mu_i}{n}$ for a stable process.

The next set of parameter functions is that involving θ_i, or, combining θ_i with μ_i, e.g :λ and τ_i. The relative variances of estimators of these parameters compared with μ is ϕ^{-1}, considerably larger than those for μ. Thus the standard errors would be in the order of 2000 times greater. There is an intuitive explanation for this. Let Y_{t1} and Y_{t2} represent vectors in R^5 consisting of the total body cell counts at two times t_1 and t_2. Let X_{t_1} and X_{t_2} be the vectors of densities per microliter obtained

by counting cells in a sample of n microlitres at each time if we could observe Y_{t1} and Y_{t_1}. The stationary expectation of Y_{t_1} and Y_{t_2} and would allow us to estimate μ with considerable precision. The exact precision depends on the parameters and would be smaller if the process is close to the stationary boundary. For a relatively stable process the precision is relatively large with standard error of the order $\sqrt{\frac{u}{u}}$. The distribution of Y_{t_1} given Y_{t_2} would provide information with a similar order of precision on θ and observing additional values of provides information on the conditional variance of the process which allows the separate identification of δ and β with a lower precision than that available for μ and θ. However, with our data we don't have the full body cell count available but only samples assumed to be well mixed Poisson samples.

Assuming a stable process, the variance of a relatively small sample will be primarily due to sampling and the variance of estimates of μ_i will be approximately $\frac{\mu_i}{n}$. To estimate the θ_i parameters we need to use the conditional distribution from one time to the next. We would have a good estimate of θ_i from Y_{t1} and Y_{t2} but, using X_{t1} and X_{t2} instead introduces variance in both the response and the predictor of the conditional distribution. The result is to increase variance by a factor of ϕ^{-1} relative to the variance for the estimation of μ. Finally there is less information (with relative variance of order greater than ϕ^{-2}) for β_i versus δ_i, i.e. estimating the

separate contributions to depletion contritubed by the partly cancelling components of birth versus death. Altogether, this means that, assuming the model is valid, different values in the full 15 dimensional parameter space will produce dynamics for the full body cell counts whose sample would be substantially equivalent to those observed, as long as the μ parameters of the 15 dimensional parameters are similar. One can generate sets of nearly equivalent parameters by inverting the transformation from the 15 dimensional parameter space to μ . Taking estimated values of μ , one can first choose a range of reasonable values for θ which leads to values for λ and τ.

$\lambda = \theta_i \ \mu_i$.

$\tau_1 = \frac{\phi_{i+1}\mu_{i+1}}{\mu_i} = \rho_i$ i =1..4.

From:

$\theta_i = \delta_i + \tau_i$ - β_i , i=1..4

$\theta_5 = \delta_5$ - β_5

We see that the θ and τ determine the net death rate:

δ_i- $\beta_i = v_i = \theta_i$-τ_i

δ_5- $\beta_5 = v_5 = \theta_5$

Since $\delta_i \geq 0$, and $\beta_i \geq 0$ and since the variance of the population process is monotonic in $\delta_i + \beta_i$, the choice of δ_i and β_i that satisfies non-negativity and

minimizes variance of the population process is :

$\delta_i = v_i$ and $\beta_i = 0$, if $v_1 \leq 0$,

$\delta_i = 0$ and $\beta_i =$ - v_i, if $v_1 \geq 0$,

The remaining 5 degrees of freedom allow adding the same quantity to both δ and β. This does not affect marginal expectations but increases the marginal variance. Letting $\alpha_i \geq 0$, i=1..5 denote the additional quantity, we have:

$\delta_i = \alpha_i + v_i$, $\beta_i = \alpha_i$, if $v_i > 0$

$\delta_i = \alpha_i$, $\beta_i = \alpha_i$- v_i, if $v_i \leq 0$

The parameters μ_1 and ρ_i are estimated using MCMC and using a selection of values for θ_i (0.01, 0.05 and 0.10) and α_1 (0.0 and 0.1). The tables show the distribution between individuals with means $mu1 - mean$, $ratio - mean[i]$ and standard deviations $mu1 - sd$ and $ratio - sd[i]$. Also shown are the individual estimates for λ, $birth[i]$, $death[i]$ and $tran[i]$, as well as an estimate pop-var which is an estimate of the population variation contribution to variance. In every case pop-var is verified to be very small in comparison with sampling variance. The model also uses a parameter for overdispersion: to what extent is the observed variance smaller or larger than expected if the true sampling volume is 1 microlitre. The estimate of the parameter is close to 0.10 which has a number of possible

interpretations that are not mutually exclusive:

- the effective sample count might be based on less than a microlitre, e.g. 0.1 microlitres
- the model is inadequate on the variation from time to time does not reflect a relatively stable state of the model

8 Bibliography

[1] Alberts B, Johnson A, Lewis J, et al. (2002). Molecular Biology of the Cell. 4th edition. New York: Garland Science; Chapter 24, The Adaptive Immune System.

[2] Services., U. D. (2015). AIDS Gov. Retrieved 2016, from www.aids.gov: $https://www.aids.gov/hiv-aids-basics/hiv-aids-101/global-statistics/$

[3] Mahnke, Y.D, Brodie, T.M., Sallusto,F., Roederer,M. and Lugli, E.(2013). The Who's who of T-cell differentiation: human memory T-cell subsets. Eur J immunology.

[4] Baveja, U.K.,Rewari, B.(2004). Diagnosis and Management of HIV/AIDS A Clinician's Perspective. New Delhi: B.I publication.

[5] U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES NATIONAL INSTITUTES OF HEALTH. (2003). Understanding The Immune System, How It Works. NIH Publication.

[6] Pressemolo, T., Guggino, G., La Manna, M.P., Di Liberto, D., Dieli, F., and Caccamo,N. (2014). Functional Signatures of human CD4 and CD8 T-cell responses to Mycobacterium tuberculosis. Frontiers in Immunology.

[7] Flynn, J.K., Gorry, P.R. (2014). Stem memory T cells (TSCM)- their role in cancer and HIV immunotherapies. Clinical and Transitional Immunology.

[8] Macallan, D.C., Wallace, D., Zhang, Y., De lara, C., Worth., A.T., Ghattas, H., ..., Tough, D.F.(2004). Rapid Turnover of Effector-Memory CD4 T cells in Healthy Humans. The Journal of Experimental Medicine, 255-260.

[9] Heffernan, J.M. (2011). Mathematical Immunology of infectious Diseases.

Mathematical Population Studies,47-54.

[10] Roberts, M.G., Heesterbeek, J.A.P., (2003). Mathematical Models In Epidemiology. EOLSS.

[11] Bernoulli,D. and Blower, S. (2004). An attempt at a new analysis of the mortality caused by smallpox and of the advantages of inoculation to prevent it. Rev. Med. Virol., 275-288.

[12] Kermack,W.O and McKendrick, A.G. (1927). A Contribution to the Mathematical Theory of epidemics. The Royal Society of London,700-721.

[14] Louzoum,Y. (2007). The Evolution of Mathematical Immunology. Immunological Reviews, 9-20.

[15] Perelson, A.S. and Nelson, P.W (2011). Modeling Viral Infections. An Introduction to Mathematical in Physiology, Cell Biology and immunology. New Orleans, Louisiana : American Mathematical Society.

[16] Nowak, M.A. and May, R.M (2000). Virus Dynamics, Oxford University Press, Oxford.

[17] Wei,X., Gosh, S.K, Taylor, M.E, Hohnson, V.A., Emmi, E.A., Deutsch, P., ..., Hahn, B.H. (1995). Viral Dynamics in HIV-1 infection. Nature, 117-122.

[18]Kirschner, D., and Webb, G.F. (1998). A mathematical model of combined drug therapy of HIV infection. J Theor Med., 25-34.

[19] Heffernan,J.M., Keeling, M.J. (2008). An in-host model of acute infection: measles as a case study. Theoretical population biology, 134-47.

[20] Smith, R.J., Li, J., Mao, J. and Sahai, B. (2015). Using within-host Mathematical modeling to predict the long-term outcome of Human Papillomaravirus vaccine. Canadian Applied Mathematics Quarterly, 281-299.

[21] Qesmi, R., ElSaadany,S., Heffernan, J.M., Wu, J. (2011). Model of Hepatitis B Virus with Age since Infection that Exhibits Backward Bifurcation. SIAM J. Appl. Math.,1509-1530.

[22] Laurie, K.L., Guarnaccia T.A., Carolan, L.A., Yan., A.W.C., Aban, M., Petrie, S., ..., Barr, I.G. (2015). Interval Between Infectious and Viral Hierarchy Are Determinants of Viral Interference Following Influenza Virus Infection in a Ferret Model. The Journal of Infectious Diseases, 1701-1710.

[23] Asquith, B., Bangham, C.R.M (2003). The dynamics of T-cell fratricide: application of a robust approach to mathematical modelling in immunology. Journal of Theoretical Biology, 53-69.

[24] Karlsson, I., Malleret, B., Brochard, P., Delache, B., Calvo, J., Le Grand, R., and Vaslin, B. (2007). Dynamics of T-Cell Responses and Memory T-cell during Primary Simian Immunodeficiency Virus in Cynomolgus Macaques. American Society for Microbiology, 13456-13468.

[25]Perelson, A.S., Essunger, P., Cao, Y., Vesanen, M., Hurley, A., Saksela, K., ..., Ho, D.D. (1997). Decay characteristics of HIV-1-infected compartments during combination therapy. Nature, 188-191.

[26] Douek, D.C., Picker, L.J., Koup, R.A. (2003). T-cell dynamics in HIV-1 infection. Annu. Rev. Immunology,265-304.

[27] Conway, J., and Coombs, D. (2011). A Stochastic model of latently infected cell reactivation and viral blip generation in treated HIV patient. PLOS.

[28]Ostrowski, M., Chun, T.W., Justement, S., Motola, I., Spinelli, M., Adelsberger, ..., Fauci, A. (1999). Both Memory and CD45RA1/CD62L1 naïve CD41 T Cells Are Infected in Human Immunodeficiency Virus Type 1-Infected Individuals. Journal of Virology, 6430-6435.

[29]Johnson, L.F, White, P. (2011). A review of mathematical models of HIV/AIDS interventions and their implications for policy. NCBI, 629-634.

[30] Ho, D.D., Neumann, A.U., Perelson, A.S., Chen, W., Leonard, J.M, Markowitz, M. (1995). Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. Nature,123-126.

[31] Massanella, M., Gomez-Mora,E., Carillo, J., Curriu, M., Ouchi, D., Puig, J., ..., Blanco, J. (2015). Increased ex vivo cell death of central memory CD4 T-cell in treatedHIV infected individuals with unsatisfactory immune recovery. Journal of

Translational Medicine.

[32] Global HIV and AIDS Statistics.(2014). Retrieved from AVERT: http://www.avert.org/professionals/hiv-around-world/global-statistics.

[33]Giles Hooker, (2015). Personal communication.

[34] Wang, K. (2014, December 30). Yale Scientific . Retrieved from Yale Scientific : http://www.yalescientific.org/2014/12/dragging-hiv-into-the-limelight-exposing-the-anatomy-of-the-worlds-leading-infectious-killer/

[35] National Institute on Drug Abuse. HIV/AIDS and Drug Abuse: Intertwined Epidemics Retrieved from https://www.drugabuse.gov/publications/drugfacts/hivaids-drug-abuse-intertwined-epidemics on April 28,2016.

[36] Hazenberg, M.D., Otto, S.A., Benthem, B.V., Roos, M.T., Coutinho, R.A., Lange, J.M.,..., Miedema, F. (2003). Persistent Immune Activation in HIV-1 infection is associated with progression to AIDS. AIDS, 1881-1888.

[37] Georges Monette, (2015), MCMC fit using Stan. Personal communication.

[38] Perelson, A.S., Neumann, A.U., Markwoitz, M., Leanord, J., Ho, D. (1996). HIV-1 dynamics in vivo:vrion clearance rate, infected cell lifespan, and viral generation time. Science, 1582-1585.

[39] Hefferan, J.M. Wahl, L. (2005). Monte Carlo estimates of natural variation in HIV infection. Journal of Theoretical Biology, 137-153.

[40] Merrill, S.(1989). Modeling the interaction of HIV with the cells of the immune system. Mathematical and Statistical approaches to AIDS Epidemiology, Lecture Notes in Biomath. Vol 83, Springer-verlag, New York.

[41] Marino, S., Hogue, I.B., Ray, C.J., Kirschner, D.E.(2008). A Methodology For Performing Global Uncertainty ad Sensitivity analysis In System Biology. Journal of Theoretical Biology, 254:178-196.

[42] Wang, L., Cao, J., Ramsay, J.O., Burger, D.M., Laporte, C.J.L., Rockstroth,

J.K. (2012). Estimating Mixed Effects Differential Equation Models. Stat Comput.

[43] Pinheiro, J. and Bates, D. (2000). Mixed Effects Models in S and Splus. New York: Springer.

[44] Gelman, A., Lee, D., Guo, J. (2015). Stan: A probabilistic programming language for Bayesian inference and optimization. Journal of Educational and Behavioral Statistics.

[45] Sheiner, L.B. (1994). A new approach to the analysis of analgesic drug trials, illustrated with bromfenac data. Clinical Pharmacology and Therapeutics, 309-22.

[46] Fisher, D., Shafer, S. (2007). Pharmacokinetic and Pharmacodynamic Analysis with NONMEM.

[47] Mathworks. (n.d.). Mixed Effect Models.Retrieved July 2016, from Mathworks: $http://www.mathworks.com/help/stats/mixed-effects-models.html$

[48] Savic, R.M., Mentre, F., Lavielle M. (2011). Implementation and Evaluation of the SAEM Algorithm for Longitudinal Ordered Categorical Data with an Illustration in Pharmacokinetics Pharmacodynamics. NCBI, 44-53.

[49] Allen, L.(2010). An Introduction to Stochastic Process with Application to Biology. Chapman and Hall New Jersey.

[50] Lavielle, M. (2015). Mixed Effect Models for the Ppopulation Approach . Orsay: Chapman and Hall,CRC Biostatistics Series.

[51] (n.d.). Retrieved 2016, from Lixoft:$http://lixoft.com/$

[52] Sekaly, P., Zeidan, J. et al. (2013). Infusion of a small number of CCR5-edited CD4+ T-cells reshapes the immune system, restores T cell homeostasis and reduces the HIV reservoir. Case Western University.

[53] Bajar, S., Webb, G., Cloyd, M., Kirschner, D. (2002). Dynamics of Naive and Memory CD4 T-cell Lymphocyte in HIV-1 Disease Progression. JAIDS, 41-58.

[54] Perelson, A.S., Kirschner, D.E., De Boer, R. (1993). Dynamics of HIV Infection of CD4 T cells. Math Biosci., 81-125.

[56] Nabatov, A.A., Pollakis, G., Linnemann, T., Paxton, W.A., De Baar, M.P. (2007). Statins Disrupt CCR5 and RANTES Expression Levels in CD4+ T Lymphocytes In Vitro and Preferentially Decrease Infection of R5 Versus X4 HIV-1. PLos.

[57] Downing, R. (n.d.). Studio Macbeth. Retrieved 2016, from $http://www.studiomacbeth.com/images/hiv-life-cycle-9.html$