Universiteit Utrecht

University College Utrecht

UCSCIRES32

Science Research Thesis

Mathematically Capturing HIV

Quantifying CTL killing of HIV-infected cells using mathematical models

A thesis submitted in fulfilment of the requirements
for the degree of Bachelor of Science
in the
Theoretical Biology and Bioinformatics group
Utrecht University

Author:
Tim Coorens

Supervisor:
Prof. dr. Rob de Boer

September 4, 2015

Contents

List of Tables

List of Figures

Abstract

The dynamics of HIV infected cells are of the utmost important in understanding the disease, and mathematical modeling can help us quantify and understand the processes at work. One parameter that is heavily debated is the contribution of CD8+ T cells in the killing of HIV infected cells; some attribute a negligible role to CD8+ T cells in this respects, others an indispensable role. Moreover, CD8+ T cell mediated killing can take place at different stages in the cycle of HIV infected cells. Here, we used a two-stage mathematical model, distinguishing between non-productively and productively infected cells. The killing rates were estimated in two ways. First, the viral replication rate was calculated based on six different studies using CD8+ depletion studies. From this replication rate, we then calculated the killing rates for the early, late, and equal killing scenarios. An interesting, novel finding on the side is that the viral replication rate of elite controller rhesus monkeys is higher than that of progressor monkeys, corresponding to a higher killing rate in the former group. Secondly, we fitted our mathematical model to a dataset of an experiment using adoptive transfer of CD8+ T cells using a script written in R. These fits were done with fully free, partially fixed, and mostly fixed parameters. Both estimation methods suggest that the role of CD8+ T cells in the killing of HIV infected cells is crucial and not negligible.

Foreword and Acknowledgements

This thesis is written to fulfill the requirements for Bachelor of Science at University College Utrecht, where I majored in biology, chemistry, physics, mathematics and medical science. I wrote the thesis in the field of theoretical biology, mainly because of my dual interest in the both the medical/biological sciences and the physical/mathematical sciences. This thesis also serves as a preparation for my master's degree: computational biology at Cambridge University.

Every student of biology, no, simply everyone comes across HIV in one way or another. But the use of mathematical models to analyse HIV was for me a very novel way of trying to solve this biological problem and I thoroughly enjoyed performing the estimates and fits, and writing this thesis.

But of course, some gratitude is more than due. First of all, I would like to thank my supervisor, Rob de Boer, for guiding me through the process of writing this thesis and giving me advice whenever needed. I have learnt a great deal while doing this and am enormously grateful for that. I would also like to thank Guido Silvestri, Nichole Klatt, and Ankita Chowdhury for sharing their data with me. Without those data, this thesis would have been a shadow of what it now is.

Thank you for reading this thesis. I hope you find the read interesting and enjoyable.

- Tim

1 Introduction

The Human Immunodeficiency Virus (HIV) is one of the most studied viruses in the world, partly due to its recalcitrance to vaccination. Since HIV is a retrovirus, its viral RNA is reverse-transcribed into viral DNA, which is incorporated into the host genome. HIV infects CD4+ T helper cells and enters these cells via their CD4 receptor. Progressive HIV infection causes the immune system to fail and is known as acquired immunodeficiency syndrome (AIDS). Hence, the cause of death in HIV infected individuals often consists of opportunistic infections.

Cytotoxic T lymphocytes (CTLs) or CD8+ T cells are an essential part of the adaptive immune system by specifically targeting and killing cells infected by viruses or bacteria and even tumor cells. The role of CTLs in the killing of HIV-infected cells is a topic of debate in the current literature. Some researchers conclude that the lytic killing of HIV-infected cells by these CTLs is negligibly small and that CTLs influences these HIV-infected populations non-lytically [1–5]. Other researchers do argue that cytotoxic killing is indispensable in the destruction of HIV-infected cells [6, 7].

Due to the complex mechanisms by which the virus and the immune system influence each other, it is hard to measure the effects of CTLs directly on the killing of infected cells. A much performed experiment is the depletion of CD8+ cells in SIV infected rhesus macaques and the subsequent observation of the behaviour of the number of viral particles in the blood [4, 5, 8–10]. The increase in the viral load can be parametrised by its slope, here called the upslope. Since there is a steep increase in viral load after CD8+ depletion, this argues for an important role of CD8+ T cells. Alternatively, one can also observe the behaviour of the viral load after assigned therapy, such as antiretroviral therapy (ART), in both the presence and absence of CD8+ cells [4, 5]. This slope is here called the downslope, and is remarkably equal in the case with and without CTLs. This suggests an almost negligible rate of killing attributable to CTLs. How can we account for this discrepancy?

Next to experimental techniques, mathematical modelling provides us with tools to make sense of the data observed, to predict the behaviour of infected cells, and to estimate parameters that cannot be directly estimated via experiments. Sometimes conclusions make hidden assumptions that are not automatically valid, and mathematical models can pinpoint these assumptions. One of these is that the difference in downslopes would automatically reflect the CTL-mediated killing. From the mathematical model, it will be shown that this is not the only possible scenario.

First, different mathematical models of HIV dynamics will be considered, and one of them will serve as the basis for the rest of this thesis. Previous estimates of the upslopes, downslopes, and other parameters will be discussed, after which possible estimations of the CTL-mediated killing rate will be reviewed. The experimental part of the thesis consists of estimating the upslope from a variety of different datasets. A novel finding in this thesis is the different value of the upslope in elite controller monkeys and progressors monkeys. Next to that, a mathematical model of HIV dynamics will be fitted to data of an experiment with injected HIV-specific CTLs [11].

Part I
Literature Review

2 The Mathematical Models

There are various different models for the description of the dynamics of HIV infected cells. Here, a general, one-stage model of HIV dynamics will be discussed, as well as a two-stage model [7,12]. A crucial difference between the general model and the two-stage model is the incorporation of the eclipse phase, the delay between the HIV virus infecting the helper T cell and the cell producing new viral particles. The general model does not distinguish between the infected stage and the producing stage and is therefore a one-stage model. In this section, each of these models will be discussed and their differences will be highlighted. In addition, a possible three-stage model will be briefly considered and reflected upon.

2.1 The One-stage Model

The general model contains a population of target cells (T) and a population of infected cells (I), who readily produce new viral particles. In addition, equations are also given for the number of viral particles (V) and the number of effector T cells against the virus in question (E_i). The following system of differential equations represents the one-stage model.

$$\begin{aligned} \frac{\mathrm{d}T}{\mathrm{d}t} &= F(T) - bTV \qquad \frac{\mathrm{d}I}{\mathrm{d}t} = bTV - (d_I + \sum_i^n k_i E_i)I \\ \frac{\mathrm{d}V}{\mathrm{d}t} &= pI - cV, \qquad \frac{\mathrm{d}E_i}{\mathrm{d}t} = G(E_i, V) - d_E E_i, \text{ for } i = 1, 2, \ldots, n, \end{aligned} \tag{1}$$

Here $F(T)$ is an (unknown) function for the production of target cells, b the infection rate, d_I the death rate of the infected cells and $\sum_i^n k_i E_i$ the killing rate by CTLs. p is the coefficient of the production of viral particles by I and c is the fraction of viral particles perishing (not infecting). $G(E_i, V)$ is a function describing the production of HIV-specific CTLs and d_E is the death rate of these immune cells.

Assuming a quasi steady state of the viral load ($dV/dt = 0$) and replacing the target cell population and killing by the constants $\bar{T}$ and K, respectively, allows us to inspect the behaviour of the model in chronic infection. The latter assumption can be made since the process of generation of target cells and immune cells is much slower than the dynamics of the infected cells [6]. This reduces equation (1) to:

$$\frac{\mathrm{d}I}{\mathrm{d}t} = (\beta\bar{T} - d_I - K)I \tag{2}$$

With $\beta = bc/p$. The solution to this ordinary differential equation is given by

$$I(t) = I_0 e^{(\beta\bar{T} - d_I - K)t} = I_0 e^{\lambda t} \tag{3}$$

The population of infected cells stays constant if the exponent of equation (3), λ equals zero. We can define the term $f\beta\bar{T} - d_I$ as the effective replication rate of the virus, ρ. The exponent does indeed equal zero if replication equals killing, i.e. $\rho = K$.

2.2 The Two-stage Model

The previously described general model is a crude approximation to the actual viral dynamics. Various important aspects are left out of the model, such as the notion of the eclipse phase. This phase denotes the time between the infection of a CD4+ T lymphocyte by HIV and the production of new viral particles by the infected cell. In a model incorporating this eclipse phase, we need to distinguish between the population of infected cells which carry the viral DNA but are not yet actively translating this, I, and productively infected cells, P. Again, T denotes the population of target cells, V the number of viral particles, and E the number of (cytotoxic) immune effector cells, for which the index i denotes the specific clone, up until n clones. Such a model of HIV viral dynamics has already been explored by Klenerman *et al.* approximately two decades ago [12]. Mathematically, the model looks as follows and is equal to the model presented in [6].

$$\begin{aligned}
\frac{\mathrm{d}T}{\mathrm{d}t} &= F(T) - bTV, \qquad \frac{\mathrm{d}I}{\mathrm{d}t} = fbTV - (d_I + \gamma + \sum_i^n k_{I_i} E_i)I \\
\frac{\mathrm{d}P}{\mathrm{d}t} &= \gamma I - d_P P - P\sum_i^n k_{P_i} E_i, \qquad \frac{\mathrm{d}V}{\mathrm{d}t} = pP - cV \\
\frac{\mathrm{d}E_i}{\mathrm{d}t} &= G(E_i, V) - d_E E_i, \text{ for } i = 1, 2, \ldots, n
\end{aligned} \tag{4}$$

In this model, γ is the eclipse phase parameter for the transition from I to P. $1/\gamma$ yields the time period of the eclipse phase. Note that immune effector cells can both kill cells in I and P.

It is conventionally assumed that $\mathrm{d}V/\mathrm{d}t = 0$, which further simplifies the model [6]. This quasi steady state assumption allows us to isolate V as $V = (c/p)P$ and eliminate it from (4). This yields the following system of differential equations.

$$\begin{aligned}
\frac{\mathrm{d}T}{\mathrm{d}t} &= F(T) - \beta TP, \qquad \frac{\mathrm{d}E_i}{\mathrm{d}t} = G(E_i, V) - d_E E_i \\
\frac{\mathrm{d}I}{\mathrm{d}t} &= f\beta TP - (d_I + \gamma + \sum_i^n k_{I_i} E_i)I, \qquad \frac{\mathrm{d}P}{\mathrm{d}t} = \gamma I - d_P P - P\sum_i^n k_{P_i} E_i
\end{aligned} \tag{5}$$

Where again, $\beta = bp/c$. In addition, it is convenient to assume that the target cell population only varies relatively slowly compared to the entire model, so that it can be assumed to be constant, i.e. T becomes the constant $\bar{T}$. Similarly, the terms $\sum_i^n k_{I_i} E_i$ and $\sum_i^n k_{P_i} E_i$ can be replaced by the constants K_I and K_P, respectively, as we also did in the case of the general model. In this way, we arrive at a very neat version of the mathematical model, which can be denoted as a compact linear system.

$$\begin{pmatrix} \mathrm{d}I/\mathrm{d}t \\ \mathrm{d}P/\mathrm{d}t \end{pmatrix} = \begin{pmatrix} -d_I - \gamma - K_I & f\beta\bar{T} \\ \gamma & -d_p - K_P \end{pmatrix} \begin{pmatrix} I \\ P \end{pmatrix} \tag{6}$$

If the entire system is in a steady state, i.e. the values for the entire system do not change (which would be the case in a chronic infection), the determinant of the above matrix is required to be equal to zero. This is because the eigenvalue of the matrix needs to be zero in order to ensure that $\mathrm{d}I/\mathrm{d}t = 0$ and $\mathrm{d}P/\mathrm{d}t = 0$.

$$(d_I + \gamma + K_I)(d_P + K_P) - f\beta\bar{T}\gamma = 0 \tag{7}$$

Alternatively, this can be written as

$$d_p + K_P = \frac{f\beta\bar{T}\gamma}{\gamma + d_I + K_I} = \frac{bp}{c}\bar{T} \cdot f \cdot \frac{\gamma}{\gamma + d_I + K_I} \tag{8}$$

From above equation, it becomes apparent that there is indeed a steady state, in the sense that all the newly formed productively infected cells are balanced out by being killed 'naturally' (d_p) or by CTLs (K_P). The newly formed cells of P are given by the right-hand side, since it is the fraction of cells progressing from I to P ($\gamma/(\gamma + d_I + K_I)$) of the fraction of cells that progressed from T to I (f) of the total newly infected cells ($bp\bar{T}/c$).

In general, we can find a solution to equation 6 by finding its eigenvalues. These are as follows:[1]

$$\lambda_{1,2} = \frac{1}{2}\left(-(d_I + \gamma + d_p + K_P + K_I) \pm \sqrt{\Delta^2 + 4f\beta\bar{T}\gamma}\right) \tag{9}$$

Where $\Delta = d_I + \gamma + K_I - d_P - K_P$. Both λ_1 and λ_2 are fully real eigenvalues, where λ_1 is the overall positive eigenvalue and λ_2 the negative one. The positive eigenvalue dominates over the negative one and the eigenvector associated with λ_1 is

$$\mathbf{v}_1 = \left(\frac{-\Delta + \sqrt{\Delta^2 + 4f\beta\bar{T}\gamma}}{2\gamma}, 1\right) \tag{10}$$

The above eigenvalues and eigenvector will be utilised further when looking into experimentally verified parameters with respect to these models of HIV dynamics.

2.3 Models incorporating more stages

In reality, there are most likely more cellular stages involved in the dynamics of HIV than discussed so far. For example, the mathematical model used by Wick *et al.* featured a population Q of quiescent target cells [13]. Recently, Althaus *et al.* identified and described eleven distinct sub-populations of CD4+ cells, which all have a role in the dynamics of HIV [14]. An eleven-stage model would most likely be too cumbersome to adequately work with and would probably produce unclear results. However, as the introduction of the eclipse phase forms a major break between the one- and two-stage model, it might be worthwhile to consider the implications of adding additional stages to our model of HIV dynamics.

[1] Eigenvalues and eigenvectors were obtained using Wolfram Mathematica.

The next step from a two-stage model is of course a three-stage model. We propose to also consider the very early phase after infection. In this phase, many of the newly infected cells die because of pyroptosis or another mechanism [15]. Viral molecules can rapidly induce CD8+ T cell-mediated killing even before the viral DNA is being translated [16]. Hence, for the three-stage model, we introduce a population I_0, consisting of cells infected with HIV but not yet transcribing the viral genes. Cells in I_0 either transition to I at a rate γ_0, die at a rate d_0 or are killed by the immune system at a rate of K_0.[2] This changes the model from equation (5) into:

$$\begin{aligned}
\frac{\mathrm{d}T}{\mathrm{d}t} &= F(T) - \beta TP, \qquad \frac{\mathrm{d}E_i}{\mathrm{d}t} = G(E_i, V) - d_E E_i \\
\frac{\mathrm{d}I_0}{\mathrm{d}t} &= \beta TP - (d_0 + \gamma_0 + \sum_i^n k_{0_i} E_i) I_0 \\
\frac{\mathrm{d}I}{\mathrm{d}t} &= \gamma_0 I_0 - (d_I + \gamma + \sum_i^n k_{I_i} E_i) I, \qquad \frac{\mathrm{d}P}{\mathrm{d}t} = \gamma I - d_P P - P \sum_i^n k_{P_i} E_i
\end{aligned} \tag{11}$$

A first assumption about this model can be inferred from the observation that a large fraction of the newly infected cells are killed. Because of this, we may assume that $d_0 + K_0 >> \gamma_0$. From this we can define the fraction of cells progressing to I as $f = \gamma_0/(\gamma_0 + d_0 + K_0)$. Furthermore, this early stage is known to be relatively short compared to I and P, which makes it justifiable to assume a steady state of this phase, $\mathrm{d}I_0/\mathrm{d}t = 0$.

These assumptions effectively reduce the three-stage model into the two-stage model of equation (5), where f is now clearly defined in terms of the parameters of the I_0 population. If there were no CTL-mediated killing of I_0 at all, the two-stage model would continue to be equal to this three-stage model.

[2] K_0 is the constant version of $\sum_i^n k_{0_i} E_i$.

3 Previous parameter estimations

3.1 The Downslope, δ

In chronic infection with HIV or SIV, the number of viral particles in blood plasma will decay when antiretroviral therapy (ART) is administered. The rate at which the number of viral particles decreases, is referred to as the downslope, generally denoted as δ. ART experiments were first done in 1995, by Ho *et al.* [17]; from these data it was found that $\delta \approx 0.45$/day [18]. Subsequent research by the same group has shown that a better estimate is $\delta \approx 1$/day [19], which is now the consensus among researchers in this field.

3.2 The Upslope, ρ

As beautifully invariant as δ is, estimates of the upslope parameter, ρ, also known as the effective viral replication rate, are much more varied. In general, one can distinguish a viral replication rate in the acute phase of HIV infection or the chronic phase. In this thesis, we will only consider chronic infections. Various estimates of the upslope in chronic infections after CD8+ T cell depletion include $\rho = 0.25$ d^{-1} [20], $\rho = 0.4$ d^{-1} [5], and $\rho = 0.8$ d^{-1} [8]. The variety in these estimates could be due to biological differences in the used monkeys, noise in measurement, or even to the use of different approximation methods. Therefore, to increase the consistency of estimation, ρ will be estimated from different datasets in section 5 of this thesis.

3.3 Miscellaneous parameters

As is evident from equation (4), the two-stage mathematical model has many different parameters. The fitting in the second part of this thesis would be improved if some of these parameters would be locked to a known value. Furthermore, since the fitting will be done on data from three different patients, locking parameter values grants a way of making the claims more biologically sound; we can expect certain parameters to be independent of the patient, and merely dependent on the (strain of) virus itself.

Table 1 provides an overview of values of parameters either estimated or used by other studies. These are merely a guideline to see whether the parameter values found in the fittings are realistic and consistent with previous findings.

Parameter	Value	Reference
γ	1.0	[21, 22]
d_T	0.33	[23]
d_P	1 - 2	[6]
d_I	0.1 - 1	[6]
β	6.3 - 9.1	[6, 24]

Table 1: Previous estimates of parameter values relevant for the two-stage model.

4 Estimating the CTL Mediated Killing Rate

4.1 Killing rates from ART data: Downslopes

In the previous section, we found that the conventional estimate of the downslope is $\delta \approx 1$. But what does δ actually reflect in light of the previously discussed mathematical models of HIV dynamics? ART prevents HIV from infecting new cells, which turns the differential equation modelling the virus count and the productively infected cells from (1) into:

$$\frac{\mathrm{d}V}{\mathrm{d}t} = pP - cV \qquad \frac{\mathrm{d}I}{\mathrm{d}t} = -(d_I + K)I \tag{12}$$

Assuming that $\mathrm{d}V/\mathrm{d}t = 0$, the load of viral particles becomes proportional to the amount of virus producing cells. Hence, in the general model, this δ reflects any process that can lead to the death of the virus producing cells. This can be non-specific lysis, clearance, but also death due to HIV-specific CTLs. In short, $\delta = d_I + K$.

In the case of the two-stage models, it becomes less evident what δ exactly reflects. However, it has been shown that in the case of infected populations with multiple stages, the slowest transition rate within the model defines the δ parameter [12]. This is quite clear when looking at the eigenvalues from (9) under ART conditions. If we assume perfect ART and set $\beta = 0$, we arrive at the following eigenvalues.

$$\lambda_1 = -(d_P + K_p), \qquad \lambda_2 = -(d_I + \gamma + K_I) \tag{13}$$

In the case that $|\lambda_1| >> |\lambda_2|$, [3] δ will reflect the slowest rate, i.e. $d_I + \gamma + K_I$. In the case that $|\lambda_1| << |\lambda_2|$, δ will take the value of $d_P + K_P$.

4.1.1 ART after CD8+ cell depletion

Recently, the same ART experiments were done after CD8+ cell depletion in model organisms by Klatt *et al.* [4] and Wong *et al.* [5]. In terms of our mathematical models, this CD8+ cell depletion sets the value of K (or K_I and K_P) equal to zero. The viral decay rate was studied in both control rhesus macaques and CD8+ cell depleted rhesus macaques, both groups chronically infected with SIV. Interestingly, both studies found that δ was not significantly different between the control and the CD8+ cell depleted macaques. If we attempt to come to terms with this finding through the general model of HIV dynamics, we can only arrive at the conclusions that $\delta \approx d_I$ and $K \approx 0$, since removing the CTL-mediated killing does not influence the value of the downslope.

However, when the eclipse phase is introduced in the mathematics and the model consists of two stages, there is more to δ than meets the eye. What is more, there exists the option that δ does not reveal much at all about the CTL-mediated killing rate. What δ exactly signifies in the different models can be found in Table 2.

What can we infer about the killing rates with the observation that δ is invariant upon depletion of CD8+ cells? As mentioned before, δ reflects the minimum value of the transition rates, which is - in case of the two-stage model - either $d_I + \gamma + K_I$ and $d_P + K_P$. Because of the invariance of δ, it has to be true that either $d_I + \gamma + K_I = d_I + \gamma$ (so $K_I = 0$) or $d_P + K_P = d_P$ (so $K_P = 0$). This is also known as a late-killing scenario (when $K_I = 0$) or an early-killing scenario (when $K_P = 0$).

[3] Since we defined δ as a positive quantity, the absolute values of the eigenvalues are used.

	Predicted viral decay rate, δ	
Model	Control	After CD8 depletion
General	$\delta = d_I + K$	$\delta = d_I$
Two Stage	$\delta = \min(\gamma + d_I + K_I, d_P + K_P)$	$\delta = \min(\gamma + d_I, d_P)$

Table 2: Summary of the meaning of δ, the viral decay rate, in the general one-stage model of HIV dynamics, and the two-stage model with early or late killing. Note that min(x,y) denotes the minimum value of x and y.

4.2 Killing rates from CTL depletion studies: Upslopes

Another way to probe into the effects of CTLs on viral load count is to simply fully or partially deplete the CTLs in SIV infected model organisms, such as rhesus macaques. This depletion is normally done by administering anti-CD8 antibodies. Because of this depletion, the viral count increases. The rate at which the viral count grows is referred to as the upslope, denoted as ρ. There have been a variety of studies researching the effect of CD8+ depletion directly on the viral load, in acute [25,26] and chronic SIV/HIV infection [4, 5, 8–10]. We can define this ρ as the eigenvalue of the system (see equation (9)) when the killing rates equal zero.

$$\rho \equiv \lambda_1\big|_{K=0} = \frac{1}{2}\left(-(d_I + \gamma + d_p) + \sqrt{(d_I + \gamma - d_P)^2 - 4f\beta\bar{T}\gamma}\right) \tag{14}$$

We can further simplify this equation using elementary algebra (along the lines of [6]).

$$2\rho + d_I + \gamma + d_p = \sqrt{(d_I + \gamma - d_P)^2 + 4f\beta\bar{T}\gamma} \tag{15}$$

$$(2\rho + d_I + \gamma + d_p)^2 = (d_I + \gamma - d_P)^2 + 4f\beta\bar{T}\gamma \tag{16}$$

$$4\rho^2 + 4\rho(d_I + \gamma + d_p) + 2\gamma d_P + 2d_I d_P = -2\gamma d_P - 2d_I d_P + 4f\beta\bar{T}\gamma \tag{17}$$

$$\rho^2 + \rho(d_I + \gamma + d_p) + \gamma d_P + d_I d_P = f\beta\bar{T}\gamma \tag{18}$$

$$f\beta\bar{T}\gamma = (\rho + d_P)(\rho + d_I + \gamma) \tag{19}$$

Why bother rewriting the expression for ρ into this compact equation above? We can combine equation (19) with equation (7) to yield an expression for the killing rates in terms of the viral upslope, namely:

$$K_I = \frac{\rho(d_I + \gamma + d_P + \rho) - (d_I + \gamma)K_P}{d_P + K_P} \text{ and } K_P = \frac{\rho(d_I + \gamma + d_P + \rho) - d_P K_I}{d_I + \gamma + K_I} \tag{20}$$

Based on the estimates of ρ from the previous section, this would put the killing rates at considerably higher than the estimates based on the general model and the downslope δ. In the early killing scenario, K_I would be in between 0.8 d^{-1} and 4 d^{-1}, and in the case of late killing, K_P would be between 0.75 d^{-1} and 3.7 d^{-1} [6]. Both estimates are for $0.25 < \rho \leq 1$ d^{-1}.

Part II
Data Analysis and Fitting

5 Estimating the Upslope

In this section, the viral growth rate will be estimated from various datasets of measurements of the plasma viral load in chronically SIV infected rhesus macaques. The value of ρ is estimated by calculating the slope between two points of measurement after CD8+ T cell depletion, as well as calculate the slope of linear regressions when more than two data points are involved. Before such calculations are allowed, it must be ensured the values of the viral load used are the natural logarithms of the actual values. In this way, the exponential growth of the viral load in the linear scale is converted into a linear growth in a logarithmic scale. This directly allows us to estimate ρ as the slope of the linear growth and use linear regressions as an approximation method.

A possible complication in estimating ρ is that in all these experiments depletion treatments were given over multiple days. However, as there is no perfect binary switch between turning the CTL-mediated killing on or off, these multiple depletion treatments are the best possible approximation. The first day of depletion is taken as the most significant depletion event. For the sake of clarity, the day corresponding to the first CD8+ cell depletion treatment is consistently referred to as day 0.

Nichole Klatt, Ankita Chowdhury and Guido Silvestri kindly shared the data from their papers with us [4, 27]. The other data points have been incorporated from the figures in the respective papers using the computer program PlotDigitizer.

5.1 From Klatt *et al.* Data

The first dataset that will be studied in this thesis comes from a research conducted by Klatt and colleagues [4]. In this study, there were two groups of SIVmac239 infected rhesus macaques, denoted as group A (RRf6, RAj7, RLi6, RPp6, RZl5) and group B (RSq8, RUe7, RWf7, XHB) for convenience. Group A underwent a first CD8+ cell depletion on day 0. Note that there were no measurements of the viral load available on day 0, hence we shifted the value of the viral load on day -2 artificially to day 0. Subsequent data was available on day 1 and day 4. Group B was subject to similar conditions, receiving the first depletion treatment on day 0 as well. The data of day -6 were shifted to day 0, and data of day 1 and day 5 were available.

The estimates for the upslopes for the individual macaques and their mean can be found in Table 3. The mean of the slopes obtained via linear regressions seems the most reliable estimate for ρ in this study, which amounts to $\rho = 0.45/\text{day}$.

5.2 From Wong *et al.* Data

Wong *et al.* performed similar experiments on a total of eight rhesus macaques infected with SIVmac251 [5]. Three of these (MMU32906, MMU27562, and MMU33580) received a full depletion of their CD8+ cells prior to ART, while two others received only a partially depleting treatment (data not shown nor used). It was observed that depletion caused the transition from one steady state of viral load levels to another within 5 days. The natural

Upslopes of group A	Day 0-1	Day 1-4	Linear regression
RRf6	1.59	0.73	0.85
RAj7	0.63	0.108	0.18
RLi6	2.05	0.31	0.56
RPp6	-0.14	0.16	0.12
RZl5	2.27	-0.13	0.21
Mean group A	1.28	0.23	0.39
Upslopes of group B	Day 0-1	Day 1-5	Linear regression
RSq8	2.55	0.11	0.46
RUe7	1.58	0.57	0.72
RWf7	1.97	-0.27	0.047
XHB	2.03	0.76	0.95
Mean group B	2.04	0.29	0.54
Mean both groups	1.62	0.26	**0.45**

Table 3: Estimates of the viral growth rates, ρ for the Klatt data [4]. Note that for group A, the datapoints were artificially shifted from day -2 to day 0, and for group B from day -6 to day 0.

logarithms of these values are shown in Table 4. The steady state at day 0 and at day 5 represent the viral count before and after CD8+ depletion, respectively. I then proceeded to calculate the slopes from these data, simply by dividing the difference in value over the time step. Calculations of the slopes between the two steady states for each monkey can be found in Table 4. The mean value of these slopes gives a very similar estimate of the viral replication rate, $\rho = 0.54$/day.

	V(0)	V(5)	Slope
MMU32906	15.20	17.04	0.37
MMU27562	13.01	14.92	0.38
MMU33580	8.80	13.13	0.87
Mean			**0.54**

Table 4: Values of the viral load count in three rhesus macaques infected with SIVmac251. The steady states at day 0 reflect the viral load before CD8+ cell depletion, the one at day 5 after the depletion. The slope between these two points was calculated, as well as the average of these slopes.

5.3 From Schmitz *et al.* Data

Already in 1999, Schmitz *et al.* performed CD8+ depletion experiments on rhesus macaques chronically infected with SIVmac [9]. There were three rhesus macaques (named A, B, and C). See Table 5 for the estimates of the slope and the linear regression. The estimated value of the viral replication rate was in accordance with previous estimates, $\rho = 0.57$/day.

	Day 0-3	Day 3-6	Linear Regression
A	0.31	1.49	0.90
B	0.74	0.18	0.46
C	0.67	0.02	0.35
Mean	0.57	0.56	**0.57**

Table 5: The calculated slopes of the viral load after CD8+ cell depletion in three SIV-infected rhesus macaques. Data from Schmitz *et al.* [9].

5.4 From Jin *et al.* Data

In 1999 as well, Jin *et al.* [8] conducted experiments on rhesus macaques with CD8+ depletion via antibody treatment on three consecutive days (days 0, 1, and 2). These experiments were performed on five rhesus macaques infected with SIVmac251 (AT-02, AR-68, AR-71, AR-93) or SIVmac239 (AT-22). Data of the viral load in these monkeys were collected on day 0, 1, and 2. The calculated slopes are presented in Table 6. As can be read from the table, the mean value of the slopes of the linear regressions amounted to $\rho = 0.58$/day.

	Day 0-1	Day 1-2	Linear Regression
AT-22	0.29	0.51	0.40
AR-68	0.48	1.12	0.80
AR-71	0.010	0.57	0.33
AR-93	0.037	0.40	0.39
AT-02	1.51	0.41	0.96
Mean	0.55	0.60	**0.58**

Table 6: The calculated slopes between values of viral RNA levels in five rhesus macaques, from the Jin data [8]. The slope of a linear regression of these three datapoints was also calculated.

5.5 From Fukuzawa *et al.* Data

As part of a larger study on B cell follicle sanctuaries, Fukuzawa *et al.* performed measurements of the SIV RNA count in the blood plasma of seven elite controller rhesus macaques after treatment with anti-CD8 antibodies [10]. These rhesus monkeys had been chronically infected with SIVmac239 or SIVmac251. Remarkably, the depletion treatment was not followed by ART, but the monkeys were able to regenerate their SIV-specific CTL population and keep the virus in check. This is likely due to the fact these monkeys are elite controllers of SIV. Note that this is a vital difference with the four previous datasets, where the rhesus macaques under investigation were no elite controllers.

The dataset obtained from the study and used in the estimate of ρ consists of data taken on days 0, 3, 7, and 10 after depletion treatments. The viral loads went down after day 10. The obtained datapoints are average values of all participating monkeys. See Table 7 for the estimates. As the last interval (day 7 - 10) showed less increase than the previous two intervals, linear regressions were performed both with and without this last

interval. The linear regression slope between day 0 and 10 yielded $\rho = 0.85$/day, and between day 0 and 7 $\rho = 0.96$/day.

	Day 0-3	Day 3-7	Day 7-10	Lin. reg. 0 -10	Lin. reg. 0-7
Estimated upslope	0.99	0.94	0.56	**0.85**	**0.96**

Table 7: Estimate of ρ based on the Fukuzawa data [10]. The data points used were at day 0, 3, 7, and 10. Only aggregate data was available, not data per rhesus macaque. Linear regressions were performed both from day 0 to day 10 and from day 0 to day 7.

5.6 Upslopes in elite controllers versus progressors

Very recently, Chowdhury *et al.* published their research on the differential behaviour of progressors and controllers with respect to the dynamics of HIV [27]. They observed a significantly higher increase in plasma viral load after CD8+ depletion in controllers than in progressors. This makes intuitive sense, but can also be inferred from the mathematical definition of ρ. As the term indicates, elite controllers are far better at controlling the virus than progressors are. In other words, the killing rate of CTLs is higher in controllers than in progressors. If this control is then taken away by depletion of these CTLs, the viral load will increase more in controllers than in progressors, simply because these CTLs did not play a role as significant as in controllers.

On the other hand, mathematically we have found that in order for the system to reach a steady state, ρ must balance out the killing rates. If the killing is higher, then the ρ must have been higher as well, which results in a steeper slope for controllers than for progressors.

This intuition is consistent with our estimations of the upslope parameter from the Chowdhury *et al.* data, which we very kindly received from the authors (see Table 8 and 9). The means of the slopes for all intervals and linear regression are higher in controllers than in progressor rhesus macaques. Note that both the slopes of the linear regression for day 0-2 and day 0-3 are shown. This is because the mean value of the slope between day 2 and day 3 is visibly much less than for the other two intervals. Since it is possible this lower value corrupts the estimate of ρ, both scenarios are shown.

Controllers	Day 0 - 1	Day 1 - 2	Day 2 -3	Lin. Regr. (0-3)	Lin. Regr. (0-2)
RAk9	2.60	1.85	-1.11	1.18	2.22
RJm7	0.78	-0.80	1.16	0.26	-0.08
RLf8	4.39	0.97	-0.31	1.61	2.68
RSk8	2.84	2.62	1.50	2.35	2.73
RZc9	1.43	1.47	1.02	1.32	1.45
Mean	2.41	1.22	0.45	**1.35**	**1.82**

Table 8: Calculated values of ρ for the controller monkeys of the Chowdhury *et al.* study [27]. The slopes of the linear regressions have been calculated for both the day 0 -2 interval and the day 0-3 interval.

To investigate whether this difference in upslopes between progressor and controller

Progressors	Day 0 - 1	Day 1 - 2	Day 2 -3	Lin. Regr. (0-3)	Lin. Regr. (0-2)
RWm8	-0.14	N/A	0.01	-0.07	N/A
RKv7	1.34	1.52	-0.35	0.91	1.43
RNw10	0.92	0.46	-0.07	0.44	0.69
RAy8	1.80	0.29	1.46	1.09	1.04
RLw10	1.32	0.37	-3.47	-0.49	0.85
RFn7	0.67	0.72	-0.39	0.37	0.70
RHd10	-0.07	0.33	0.11	0.14	0.13
RIc9	1.76	0.91	0.61	1.08	1.34
RDa10	1.92	1.43	0.19	1.21	1.68
RBc9	1.83	0.57	0.24	0.85	1.20
Mean	1.13	0.73	-0.18	**0.55**	**1.00**

Table 9: Calculated values of ρ for the progressor monkeys of the Chowdhury *et al.* study [27]. The slopes of the linear regressions have been calculated for both the day 0 -2 interval and the day 0-3 interval.

monkeys is significant, the non-parametric Mann-Whitney U test was used.[4] On the linear regressions on interval day 0-2, the test proved to be non-significant ($p = 0.10$). However, monkey RJm7 seemed to corrupt the data, since his viral load hardly increased at all upon CD8+ depletion. If taken out, the results of the test indicated a very significant difference ($p = 0.005$). The same test was done for the slopes of the linear regression on interval day 0-3 and both with and without monkey RJm7, the test indicated a significant difference ($p = 0.044$ with RJm7 and $p = 0.004$ without RJm7).

5.7 Discussion of Estimates: Consistency and Implications

Now that the upslope has been estimated from data of six different studies, we can look back at the values and determine whether they are consistent and what these values imply for the CTL-mediated killing rates. See Table 10 for an overview of the estimated values of ρ.

Out of the first five values, four seem to be very constistent, with values between 0.45/day and 0.6/day. The ρ from the Fukuzawa data lies outside of this interval. A noticeable difference in the setup of the experiments is that Fukuzawa *et al.* exclusively used elite controller rhesus monkeys, while the other studies did not. However, since the Fukuzawa data only contain aggregate data and not on individual monkeys, it is hard to make such claims valid. [5]

Fortunately, the Chowdhury data makes a distinction between progressor and elite controller monkeys. Indeed, when reviewing the results from the Chowdhury data, the elite controller monkeys do generate a higher estimate of ρ than their progressor counterparts. The value of 0.55/day for the progressor monkeys and 1.00/day for the controllers are very consistent with the previous five estimates. Taking all these estimates together, this would mean that in controllers $\rho \approx 1$/day and in progressors $\rho \approx 0.55$/day.

From these upslopes, we can estimate the killing rates K_I and K_P using equation

[4] A non-parametric test was used, since we could not assume the slopes of the linear regressions were normally distributed.

[5] Via e-mail correspondence, we tried to obtain the full dataset, but unfortunately failed.

Study	Estimate of ρ
Klatt *et al.* [4]	0.45
Wong *et al.* [5]	0.54
Jin *et al.* [8]	0.58
Schmitz *et al.* [9]	0.57
Fukuzawa *et al.* [10]	0.85 or 0.96
Chowdhury *et al.* [27]	
Progressors	0.55 or 1.00
Elite Controllers	1.35 or 1.82

Table 10: Summary of the estimates of ρ from the previously described studies. The two values per category of the Chowdhury data reflect the choice to either perform a linear regression on four or three datapoints, respectively.

(20). Using $\gamma = 1.0$/day, $d_I = 0.05$/day, and $d_P = 1.5$/day (see Table 1). For the progressor monkeys, we then find $K_I \approx 1.14$/day (when $K_P = 0$/day, so for the early killing scenario), $K_P \approx 1.62$/day for the early killing scenario, and $K_I = K_P = \rho = 0.55$/day for the equal killing scenario. The killing rates in controllers are higher, as expected, with $K_I \approx 2.32$/day for early killing, $K_P \approx 3.32$/day for late killing, and $K_I = K_P = \rho = 1$/day for equal killing. This is consistent with the estimates mentioned after (20).

As a side note, more rigorous statistical analysis might be appropriate in the future to arrive at even more precise estimate for ρ. Some later intervals in the datasets might already contain a viral slope that is leveling off and corrupting the total picture. However, a practical limitation to these estimates is of course the difference in experimental setups and parameters between all these studies.

Overall, these estimates of the killing rates derived from the aforementioned datasets suggest an important role for killing by CD8+ T cells in the dynamics of HIV infected cells. In the next section, these killing rates will be estimated in a different way, by fitting our previously discussed mathematical to experimental data.

6 Model fitting

To estimate the killing rate in another way, we fitted the model of equation (5) to a dataset provided by Brodie *et al.* [11]. This concerns measurements of the population of productively infected cells, as well as injected CTLs specific for HIV from patients with chronic HIV infection. These CTLs expressed neomycin phosphotransferase (*neo*+) and were hence distinguishable from endogenous CTLs. These patients were injected with two doses of these *neo*+ CTLs.

A fitting to this data has already been performed by Wick *et al.* [13]. However, this study used a different mathematical model for the dynamics of HIV, incorporating a population of quiescent target cells. This seems to complicate the model unnecessarily and hence here we try to fit the data to our model without this quiescent population. The mathematical model used for the fitting here is as follows

$$\begin{aligned} \frac{\mathrm{d}T}{\mathrm{d}t} &= s - d_T T - \beta T P, & \frac{\mathrm{d}I}{\mathrm{d}t} &= f\beta T P - (\delta_I + \gamma + k_1 C) I \\ \frac{\mathrm{d}P}{\mathrm{d}t} &= \gamma I - (\delta_P + k_2 C) P, & \frac{\mathrm{d}C}{\mathrm{d}t} &= -d_C C \end{aligned} \tag{21}$$

Where C is the population of injected *neo*+ T cells. Furthermore, there are two injection events at time $t = 0.5$ and $t = 7.5$ with dose amount C_0.[6] For the generation of target cells, the function $F(T) = s - d_T T$, was used, where s is the constant generation of target cells (thymopoiesis rate) and d_T the death rate of these cells. Since the model with only free parameters would not be able to distinguish d_I and K_I mathematically, we use $\delta_I = d_I + K_I$ here. The same holds for d_P, K_P, and δ_P. Injected, *neo*+ CTLs will decay with a rate of d_C, and kill cells of the I and P population at rate k_1 and k_2, respectively. All other parameters are as described in the standard two stage model provided by equation (5).

The y-axis of the graphs represent the amount of P or C. The exact units are copied from Brodie *et al.*'s original data [11]. The numbers on the y-axis for P are the percentage of productively infected cells in the total CD4+ cell population. The total CD4+ T cell count was stable for all three patients, at around 259, 224, and 261 cells per μl for patient 1, 2, and 3, respectively [11, 13]. This makes it straightforward to convert the percentage of P to an actual number of productively infected cells per μl: simply multiply by a factor of 2.59 for patient 1, 2.24 for patient 2, and 2.62 for patient 3.

The y-axis for the graph of the *neo*+ CD8+ T cell are also a percentage of the total CD8+ cell count, but unfortunately neither Brodie *et al.* nor Wick *et al.* report what was used as a total CD8+ count. However, when looking at a typical ratio of CD4+/CD8+ cells in HIV infected individuals, we can approximate the CD8+ T cell counts. If we use a CD4+/CD8+ ratio of 0.75 [28], the CD8+ T cell counts of our patients becomes 345, 299, and 348, so we can convert the percentage of C to a cell count by multiplying it with a factor of 3.45 for patient 1, 2.99 for patient 2, and 3.48 for patient 3. These numbers become important when we want to convert the units of our estimated killing rates to conventional units so that we can compare it with previous findings.

Unlike Wick *et al.* [13], here, the sum of the different CD4+ T cell populations (T, I, and P) is not forced to add up to the total population of CD4+ T cells. This is because

[6] Note that C_0 is called '*inj*' in the R script.

there may compartments of these cells that are not taken into account by the mathematical model (such as quiescent target cells, which Wick *et al.* do take into account).

6.1 The Fitting Script

The fitting procedure was performed using a script in the programming language R (see appendix A). This script uses the deSolve package (for numerically solving the system of ordinary differential equations) [29] and the FME package (for the fitting itself) [30] as key components. The script is based on an example of parameter fitting provided by Soetaert and Petzoldt [31]. See the script itself for smaller comments regarding specific settings and commands. The primary data were imported as an excel-file, after being obtained via the computer program PlotDigitizer.

The first rough fit was produced by invoking the PORT method (analogous to the NLMINB command in R), after which the MARQ method was used for a more precise localization of the minimum of the cost function. This MARQ method utilizes the Levenberg-Marquardt algorithm, equivalent to the NLS.LM command from the Minpack package [32]. To ensure the positivity of the parameters, and to make the solver and fitter able to deal with the sizes of parameters, the natural logarithm of the parameters was used for the fitting, after which they were transformed back by means of the exponential function.

6.2 Free parameters

The first fit was done without any constraints on the parameters. See Figure 2 for the result of the fitting procedure, and Table 11 for the parameter estimates per patient. Most parameters are in fair accordance with the previous estimates from Table 1.

Parameter	Patient 1	Patient 2	Patient 3
d_T	0.31	0.12	0.32
δ_P	1.38	1.18	1.29
β	5.36	2.62	8.82
k_1	0.51	1.03	0.47
k_2	0.66	0.27	0.80
s	3.08	1.36	2.89
C_0	3.30	1.90	4.68
δ_I	0.01	0.98	0.02
d_C	0.57	0.54	0.39
γ	1.00	0.62	1.03
f	0.96	0.97	0.96

Table 11: Parameters from the fitting of our two-stage model to the Brodie *et al.* data

However, to ensure that the fit is reliable and the parameter values are trustworthy, we must look at the collinearity between the various parameters. The collinearity is a measure for the linear dependence between parameters. When too large, the collinearity tells us that the parameters cannot be estimated uniquely based on the data [31]. In other words, (many) different combinations of parameter values can reproduce the data. In R, this value can be calculated by the COLLIN command. A value of 10 to 15 for the collinearity is the borderline [31].

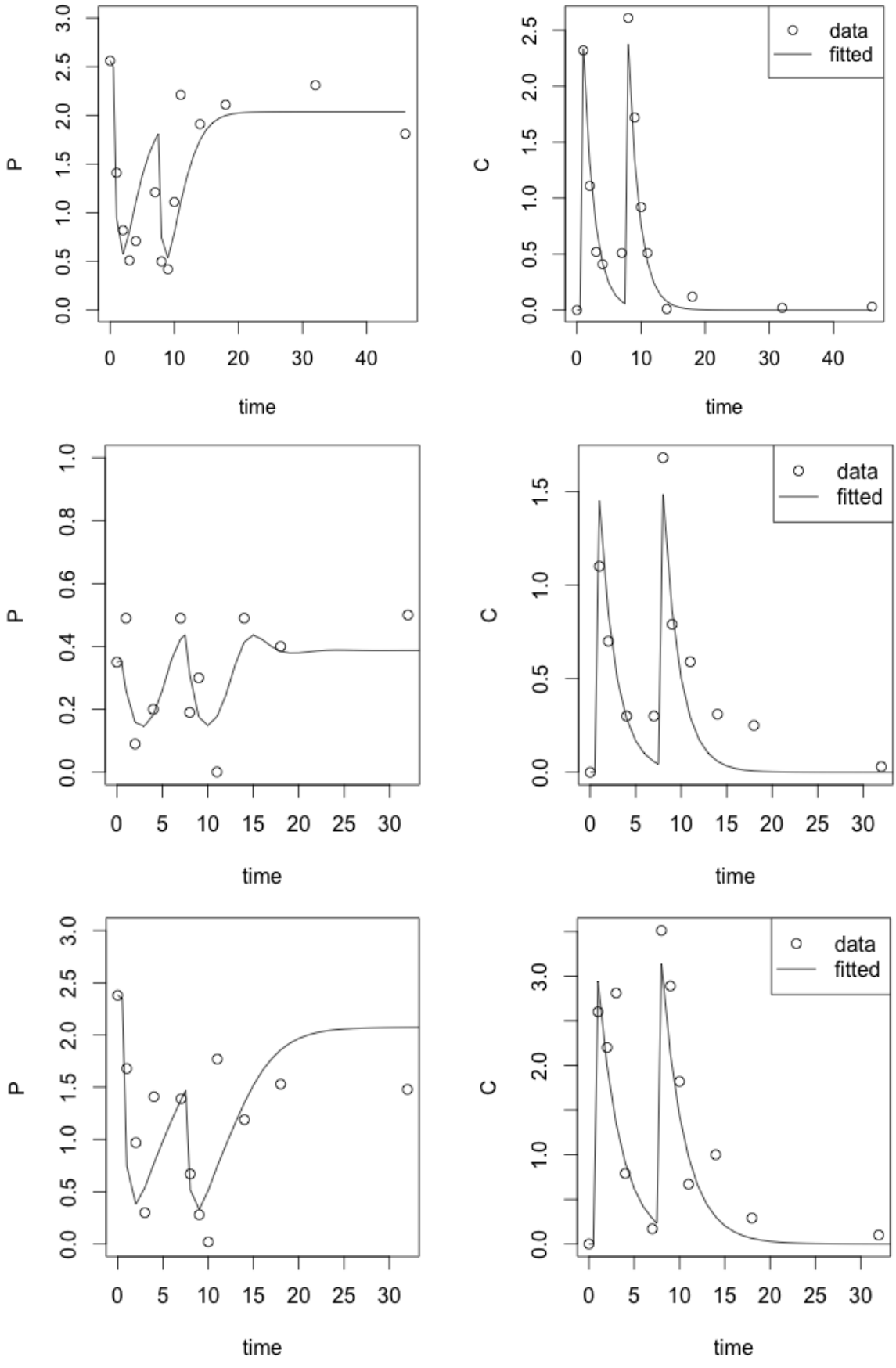

Figure 2: Fitting of model to data by Brodie *et al.* with no constraints on the parameters for patient 1, 2, and 3, respectively.

For example, the collinearity of all 11 parameters was well beyond the limit, approximately 300. This makes intuitive sense as well, of course; there are many different sets of parameters that could account for the observed behaviour. And a slight increase in target cell death could be overcome by a small increase in target cell generation, to a certain extent. This is the main reason to fix some parameters and perform the fit over others, which will be the subject of the next subsection.

```
  k1 k2 d_P d_I Beta d_T s d_C inj gamma f  N collinearity
1  1  1   1   1    1   1 1   1   1     1 1 11          303
```

6.3 Constrained parameters

In reality, some parameters will be characteristic of the virus and independent of the specific patients. Hence, it would be realistic to lock the values of most of the parameters, also because otherwise the collinearity of the fit would be too large. This in turn would render the fitting unreliable. Therefore, some parameters will be fixed in subsequent fits. In one fit, five of the eleven parameters will be made constant. We can expect the constants that parametrize the progression of the virus through its stages to be invariant among the patients (infection β, eclipse phase progression f and γ). Furthermore, to focus on the fitting of the model to the productively infected cells, d_C and C_0 are now fixed. C_0 is patient-specific and its value is based on earlier fits. For the exact values of these parameters, see the left-hand column of Table 12.

Next, we additionally set the death rates for each of the cell populations constant for the different patients (d_T, δ_I, and δ_P). Also, the generation of CD8+ T cells will be fixed, s, but will also be patient-specific. Please see the right-hand column Table 12 for an overview of these parameters. The values were based on Table 1 and Table 11.

Parameter	Value	Parameter	Value
γ	1.0	d_T	0.30
f	0.96	δ_I	0.10
β	8.0	δ_P	1.20
C_0, pat. 1	3.30	s, pat. 1	3.08
C_0, pat. 2	1.90	s, pat. 2	1.36
C_0, pat. 3	4.68	s, pat. 3	2.89
d_C	0.54		

Table 12: Fixed values for the parameters used in subsequent fitting. C_0 and s are patient-specific. Only the left column of fixed parameters was used for the partially fixed fitting. The right column of fixed parameters was introduced when only the killing rates were being fitted.

6.3.1 Fit with five constrained parameters

With the parameters f, γ, β, s, and C_0 fixed, we have improved the reliability of the fit. The collinearity of the remaining parameters is logically less than the collinearity of all parameters, but its value is 93 and hence not yet below the limit of 15. Nevertheless, it is fruitful to provide an intermediate step between fully free and almost fully fixed parameters to see whether the quality of the fit changes visibly.

```
> collin(Sfun, parset=c("k1","k2","d_T","d_I","d_P","s"))
   k2 d_T s k1 d_P d_I Beta d_C inj gamma f N collinearity
1  1    1 1  1    1   1    0   0   0     0 0 6          93
```

The fits can be found in Figure 3 and the found values for the remaining parameters can be found in Table 13.

Parameter	Patient 1	Patient 2	Patient 3
k_1	0.72	0.88	0.37
k_2	0.52	1.23	0.32
d_T	0.32	0.90	0.05
δ_I	0.15	0.10	0.12
δ_P	1.29	1.59	1.85
s	3.38	1.17	3.88

Table 13: Values of fitted parameters in the case of five fixed parameters (f, γ, β, s, and C_0) for all three patients.

6.3.2 Fit with nine or ten constrained parameters

The goal of this fitting procedure is to estimate the values of the CD8+ T cell mediated killing rates. In this last fitting, all parameters are fixed with the exception of the killing rates k_1 and k_2. The collinearity of this parameter combination is 1.7, well below the limit.

```
> collin(Sfun, parset=c("k1","k2"))
  k1 k2 d_P d_I Beta d_T s d_C inj gamma f N collinearity
1  1  1    0   0    0   0 0   0   0     0 0 2          1.7
```

Subsequently, the model fitting was performed with three different killing scenarios: early killing (setting $k_2 = 0$), late killing (setting $k_1 = 0$), or combined killing. The fits of all three scenarios were combined and are displayed in Figure 4 and the fitted values of the killing rates can be found in Table 14.

		k_1	k_2
Patient 1	Early	2.08	0
	Late	0	1.29
	Comb.	0.66	0.79
Patient 2	Early	4.32	0
	Late	0	3.83
	Comb.	1.49	1.71
Patient 3	Early	1.34	0
	Late	0	1.08
	Comb.	0.63	0.87

Table 14: Outcome of the fit values for the killing rates in the early, late, and combined killing scenarios.

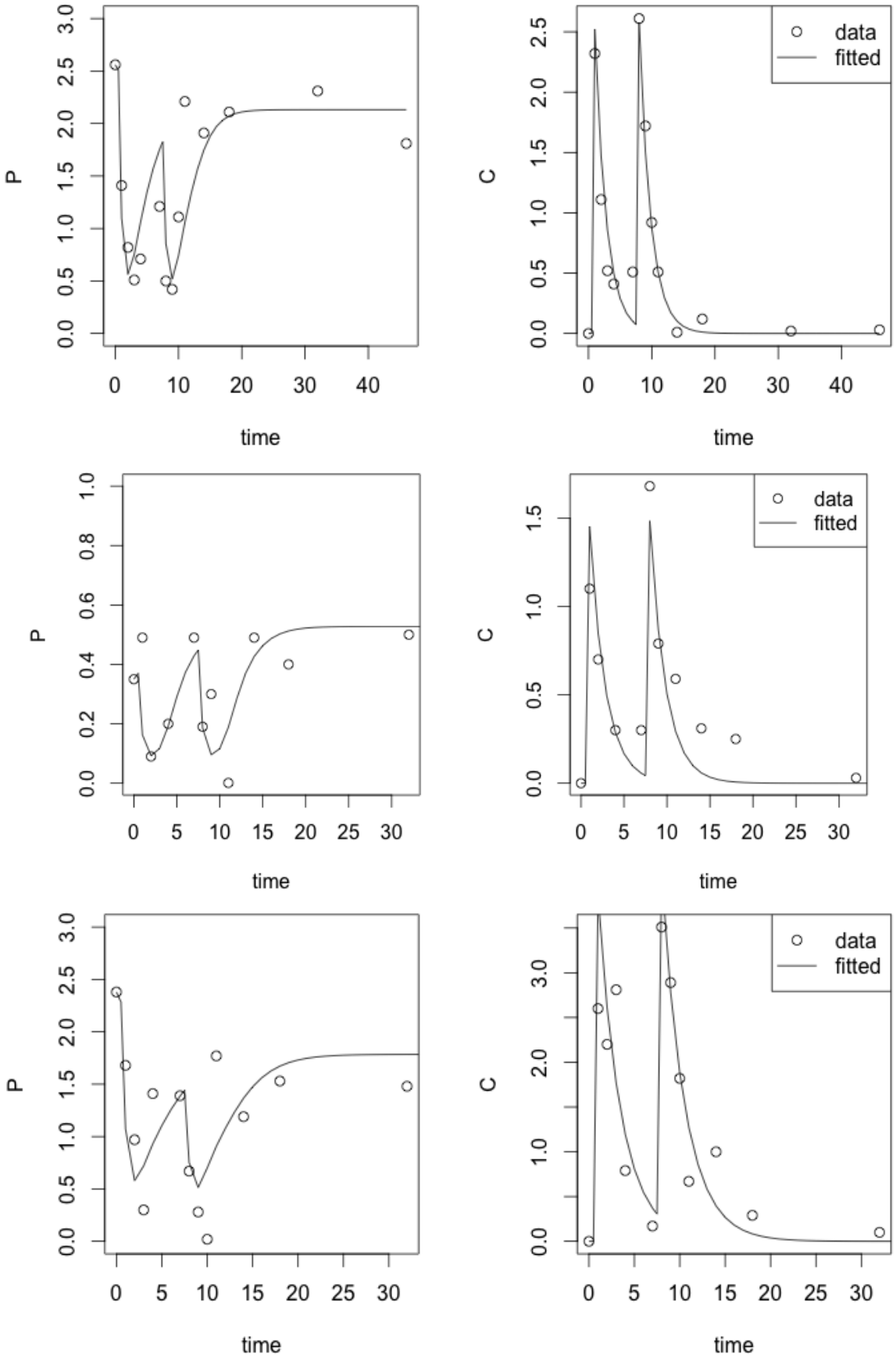

Figure 3: Fitting of model to data by Brodie *et al.* with five fixed parameters (β, γ, f, C_0, and d_C) for patient 1, 2, and 3, respectively.

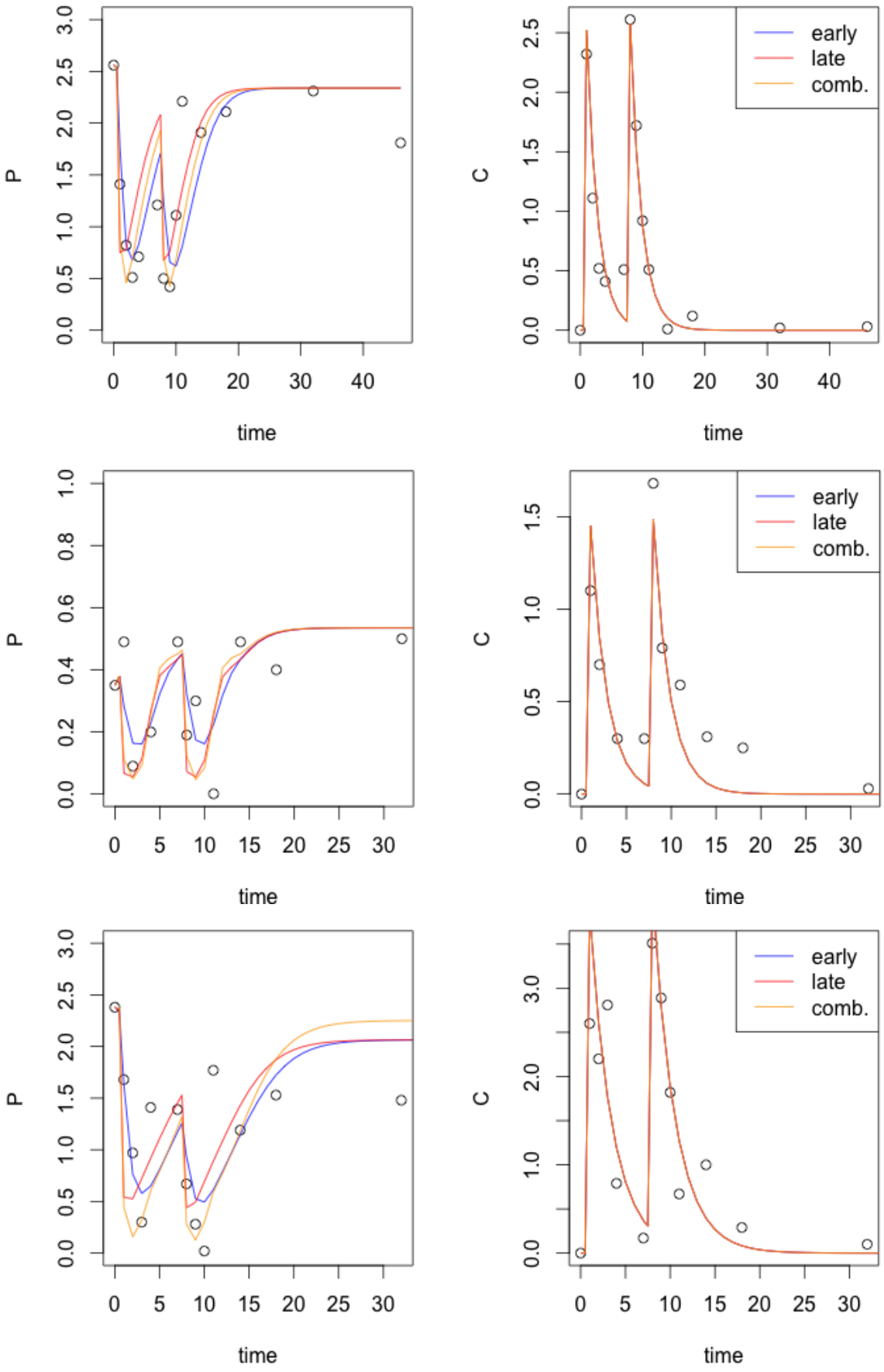

Figure 4: Fitting of model to data by Brodie *et al.* with constrained parameters and for different killing scenarios for patient 1, 2, and 3, respectively.

6.4 Discussion of Fits

The mathematical model has now been fitted against the data with three parameter situations: fully free, partially fixed, and almost fully fixed. Up to some small differences, the fits for every patient under the different parameter conditions are very similar. Freezing the values of some parameters will influence the value of the individual fitted parameters, but the overall shape of the curve stays more or less the same.

A peculiarity of these fits that was also present in the fits of Wick *et al.* is the relatively high value of the death rate of the target cells, d_T, both when the parameter is fitted and fixed. Remember that the reciprocal of this value yields the life time of the target cells. In the case of $d_T = 0.30$/day, this would mean CD4+ T cells would only live for approximately three days. Wick *et al.* justify this by introducing their population of quiescent target cells, Q, which has a much lower death rate in their fitting ($d_Q = 0.00014$ [33], yielding a life time of approximately 20 years). It is a possibility that HIV only infects the CD4+ T cells with a rapid turnover and a high death rate, which backs up the notion of a quiescent target cell population.[7]

Before discussing the estimated killing rates, it is important to convert their now slightly strange units to more conventional units. As of now, the killing rate represents the death of I or P due to C when 1% of C 'meets' 1% of P or I. To convert this to more logical units, i.e. the killing rate when one C cell meets one P or I cell, we simply need to multiply by the total CD4+ T cell count and divide by the total CD8+ T cell count. In other words, we need to multiply the killing rates with the used CD4+/CD8+ ratio of 0.75 [28]. In Table 15, the converted killing rates are shown.

An important difference with the mathematical model and fits of Wick *et al.* are the killng rates. First of all, Wick *et al.* only use a late killing model and assume the CD8+ T cells do not target population I [13]. Moreover, these fits produce values of k_2 that are equal to 0.14 for patient 1, 0.17 for patient 2, and 0.076 for patient 3. These estimates are consistently, though not much lower than all reported estimates here, for any parameter scenario. As a comparison, the lowest values of k_2 reported here are 0.38 for patient 1 (partially fixed fitting), 0.20 for patient 2 (fully free fitting), and 0.24 for patient 3 (partially fixed fitting). However, all of these k_2 values are estimated in presence of a k_1 that is of the same order of magnitude. This means that the killing of HIV infected cells is higher in these fits than in Wick *et al.*'s.

		Patient 1	Patient 2	Patient 3
Free Fitting	k_1	0.38	0.77	0.35
	k_2	0.50	0.20	0.60
Partially fixed	k_1	0.54	0.66	0.28
	k_2	0.38	0.92	0.24
Almost Fully fixed	k_1	0.50	1.12	0.47
	k_2	0.59	1.28	0.65

Table 15: Overview of the found killing rates in the fits with fully free parameters, partially fixed, and almost fully fixed parameters. Only the killing rates in case of combined killing are shown here. Units were converted from percentages to single cells (see text).

[7] However, notice that the introduction of a differential equation for quiescent target cells was not of vital importance for this model, as the produced fits are similar in shape to the fits of Wick *et al.* [13].

In Table 15, an overview of the estimated killing rate from the various fits can be found. While there is some heterogeneity in the reported rates, the overall values seems consistent enough to suggest that CD8+ mediated killing plays an important role in the death of both infected and productively infected cells. Since there is no steady state population of C, these rate constants cannot simply be converted into the killing constants estimated from the upslope data. However, both methods of estimation suggest an important role for CD8+ T cells in the battle against HIV infected cells.

To test the robustness and the consistency of outcomes of these fittings, a future effort could be directed towards fitting this model with associated parameters to other datasets. Especially data derived from ART with CD8+ depletion studies could provide an interesting case, since the invariance of the downslope suggests either an early or late killing scenario. A fitting of the mathematical model to these data could potentially identify the likelihood of either early or late killing. These datasets could be taken from the Klatt *et al.* [4] and Wong *et al.* [5] studies.

7 Conclusion

Estimates of the CTL-mediated killing rate of HIV infected cells are quite heterogeneous, with some suggesting a negligible role for CTL-mediated killing in HIV dynamics and others claiming this killing is of the utmost importance. As heterogeneous as these estimates are, experiments on the down- or upslope of the viral load after ART and CD8+ depletion are strikingly consistent. Here, we tried to estimate the killing of HIV infected cells by CD8+ T cells in two distinct ways: estimating it from the upslope from various datasets and estimating it from model fits to data from adoptive transfer studies of trained CD8+ T cells.

The estimates of the upslope from the six different datasets were rather consistent with one another, especially when the difference between progressor and elite controller monkeys was incorporated. This difference in upslopes between controller and non-controller monkeys is quite novel and a nice additional finding of this thesis. Furthermore, these estimates of ρ suggest an important role for CD8+ cells in the killing of HIV infected cells, since estimates of K_I and K_P are roughly between 0.5 and 3.

Afterwards, our mathematical model of HIV dynamics was fitted to data from an adoptive transfer study of injected *neo*+ CD8+ T cells trained to kill HIV. The fitting was performed under various parameter conditions and -when only the killing rates were allowed to vary -for the early, late and combined killing scenario. Again, the estimates of the killing rates from this fitting suggest an indispensable role for CD8+ cells in HIV dynamics. In addition, the fit was equally well for early, late and combined killing, and its quality did not change much for the different parameter fixations. The killing rates k_1 and k_2 vary from approximately 0.20 to 1.30 in the different patients. However, these can not directly be linked to K_I and K_P as no steady state concentration of *neo*+ CD8+ T cells was reached (other than zero).

Taken together, all this information emphasises the importance of killing of HIV infected cells by CD8+ T cells during chronic HIV infection. Though the results from these estimates is striking, further modelling and fitting is needed to verify the conclusions from this thesis. Especially of interest would be the fitting of our mathematical model to ART with and without CD8+ depletion studies. This could in the end lead to a consensus on the role of CD8+ T cells in the dynamics of HIV infected cells. Whatever the outcome, mathematical modelling and theoretical analysis of biological experiments and results remains a vital field, and shows much promise in understanding and grasping the full picture of HIV.

References

[1] Marjet Elemans, N-K Seich Al Basatena, Nichole R Klatt, Christos Gkekas, Guido Silvestri, and Becca Asquith. Why don't CD8+ T cells reduce the lifespan of siv-infected cells in vivo? *PLoS Comput Biol*, 7(9):e1002200–e1002200, 2011.

[2] Nafisa-Katrin Seich al Basatena, Konstantinos Chatzimichalis, Frederik Graw, Simon DW Frost, Roland R Regoes, and Becca Asquith. Can non-lytic CD8+ T cells drive HIV-1 escape? *PLoS pathogens*, 9:e1003656, 2013.

[3] Marjet Elemans, Arnaud Florins, Luc Willems, and Becca Asquith. Rates of CTL killing in persistent viral infection in vivo. *PLoS computational biology*, 10(4):e1003534, 2014.

[4] Nichole R Klatt, Emi Shudo, Alex M Ortiz, Jessica C Engram, Mirko Paiardini, Benton Lawson, Michael D Miller, James Else, Ivona Pandrea, Jacob D Estes, et al. CD8+ lymphocytes control viral replication in SIVmac239-infected rhesus macaques without decreasing the lifespan of productively infected cells. *PLoS Pathog*, 6(1):e1000747–e1000747, 2010.

[5] Joseph K Wong, Matthew C Strain, Rodin Porrata, Elizabeth Reay, Sumathi Sankaran-Walters, Caroline C Ignacio, Theresa Russell, Satish K Pillai, David J Looney, and Satya Dandekar. In vivo CD8+ T-cell suppression of siv viremia is not mediated by CTL clearance of productively infected cells. *PLoS Pathog*, 6(1):e1000748–e1000748, 2010.

[6] Saikrishna Gadhamsetty and Rob de Boer. We can estimate the rate at which HIV-1 infected cells are killed from the upslope of the viral load following CD8+ T cell depletion. Unpublished.

[7] Christian L Althaus and Rob J De Boer. Implications of CTL-mediated killing of HIV-infected cells during the non-productive stage of infection. *PLoS One*, 6(2):e16468, 2011.

[8] Xia Jin, Daniel E Bauer, Sarah E Tuttleton, Sharon Lewin, Agegnehu Gettie, James Blanchard, Craig E Irwin, Jeffrey T Safrit, John Mittler, Leor Weinberger, et al. Dramatic rise in plasma viremia after CD8+ T cell depletion in simian immunodeficiency virus-infected macaques. *The Journal of experimental medicine*, 189(6):991–998, 1999.

[9] Jörn E Schmitz, Marcelo J Kuroda, Sampa Santra, Vito G Sasseville, Meredith A Simon, Michelle A Lifton, Paul Racz, Klara Tenner-Racz, Margaret Dalesandro, Bernhard J Scallon, et al. Control of viremia in simian immunodeficiency virus infection by CD8+ lymphocytes. *Science*, 283(5403):857–860, 1999.

[10] Yoshinori Fukazawa, Richard Lum, Afam A Okoye, Haesun Park, Kenta Matsuda, Jin Young Bae, Shoko I Hagen, Rebecca Shoemaker, Claire Deleage, Carissa Lucero, et al. B cell follicle sanctuary permits persistent productive simian immunodeficiency virus infection in elite controllers. *Nature medicine*, 21(2):132–139, 2015.

[11] Scott J Brodie, Deborah A Lewinsohn, Bruce K Patterson, Daniel Jiyamapa, John Krieger, Lawrence Corey, Philip D Greenberg, and Stanley R Riddell. In vivo migration and function of transferred HIV-1-specific cytotoxic t cells. *Nature medicine*, 5(1):34–41, 1999.

[12] Paul Klenerman, Rodney E Phillips, Charles R Rinaldo, Linda M Wahl, Graham Ogg, Robert M May, Andrew J McMichael, and Martin A Nowak. Cytotoxic T lymphocytes and viral turnover in HIV type 1 infection. *Proceedings of the National Academy of Sciences*, 93(26):15323–15328, 1996.

[13] W David Wick, Otto O Yang, Lawrence Corey, and Steven G Self. How many human immunodeficiency virus type 1-infected target cells can a cytotoxic T-lymphocyte kill? *Journal of virology*, 79(21):13579–13586, 2005.

[14] Christian L Althaus, Beda Joos, Alan S Perelson, and Huldrych F Günthard. Quantifying the turnover of transcriptional subclasses of HIV-1-infected cells. *PLoS Comput. Biol.*

[15] Gilad Doitsh, Nicole LK Galloway, Xin Geng, Zhiyuan Yang, Kathryn M Monroe, Orlando Zepeda, Peter W Hunt, Hiroyu Hatano, Stefanie Sowinski, Isa Muñoz-Arias, et al. Cell death by pyroptosis drives CD4 T-cell depletion in HIV-1 infection. *Nature*, 505(7484):509–514, 2014.

[16] Jonah B Sacha, Chungwon Chung, Eva G Rakasz, Sean P Spencer, Anna K Jonas, Alexander T Bean, Wonhee Lee, Benjamin J Burwitz, Jason J Stephany, John T Loffredo, et al. Gag-specific CD8+ T lymphocytes recognize infected cells before AIDS-virus integration and viral protein expression. *The Journal of Immunology*, 178(5):2746–2754, 2007.

[17] David D Ho, Avidan U Neumann, Alan S Perelson, Wen Chen, John M Leonard, Martin Markowitz, et al. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature*, 373(6510):123–126, 1995.

[18] Alan S Perelson, Avidan U Neumann, Martin Markowitz, John M Leonard, and David D Ho. HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. *Science*, 271(5255):1582–1586, 1996.

[19] Martin Markowitz, Michael Louie, Arlene Hurley, Eugene Sun, Michele Di Mascio, Alan S Perelson, and David D Ho. A novel antiviral intervention results in more accurate assessment of human immunodeficiency virus type 1 replication dynamics and t-cell decay in vivo. *Journal of virology*, 77(8):5037–5038, 2003.

[20] Marjet Elemans, Rodolphe Thiébaut, Amitinder Kaur, and Becca Asquith. Quantification of the relative importance of CTL, B cell, NK cell, and target cell limitation in the control of primary siv-infection. *PLoS Comput Biol*, 7(3):e1001103, 2011.

[21] Narendra M Dixit, Martin Markowitz, David D Ho, and Alan S Perelson. Estimates of intracellular delay and average drug efficacy from viral load data of HIV-infected individuals under antiretroviral therapy. *Antivir. Ther*, 9:237–246, 2004.

[22] DS Dimitrov, RL Willey, H Sato, L-Ji Chang, R Blumenthal, and MA Martin. Quantitation of human immunodeficiency virus type 1 infection kinetics. *Journal of virology*, 67(4):2182–2190, 1993.

[23] David F Tough and Jonathan Sprent. Lifespan of lymphocytes. *Immunologic research*, 14(1):1–12, 1995.

[24] Rob J De Boer. Understanding the failure of CD8+ T-cell vaccination against simian/human immunodeficiency virus. *Journal of virology*, 81(6):2838–2848, 2007.

[25] Afam Okoye, Haesun Park, Mukta Rohankhedkar, Lia Coyne-Johnson, Richard Lum, Joshua M Walker, Shannon L Planer, Alfred W Legasse, Andrew W Sylwester, Michael Piatak, et al. Profound CD4+/CCR5+ T cell expansion is induced by CD8+ lymphocyte depletion but does not account for accelerated SIV pathogenesis. *The Journal of experimental medicine*, 206(7):1575–1588, 2009.

[26] Tetsuro Matano, Riri Shibata, Christine Siemon, Mark Connors, H Clifford Lane, and Malcolm A Martin. Administration of an anti-CD8 monoclonal antibody interferes with the clearance of chimeric simian/human immunodeficiency virus during primary infections of rhesus macaques. *Journal of Virology*, 72(1):164–169, 1998.

[27] Ankita Chowdhury, Timothy Lee Hayes, Steven E Bosinger, Benton O Lawson, Thomas Vanderford, Joern E Schmitz, Mirko Paiardini, Michael Betts, Ann Chahroudi, Jacob D Estes, et al. Differential impact of in vivo CD8+ T-lymphocyte depletion in controller versus progressor SIV-infected macaques. *Journal of virology*, pages JVI–00869, 2015.

[28] Jeremy MG Taylor, John L Fahey, Roger Detels, and Janis V Giorgi. CD4 percentage, CD4 number, and CD4: CD8 ratio in HIV infection: which to choose and how to use. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, 2(2):114–124, 1989.

[29] Karline Soetaert, Thomas Petzoldt, and R Woodrow Setzer. Solving differential equations in R: package deSolve. *Journal of Statistical Software*, 33, 2010.

[30] Karline Soetaert and Thomas Petzoldt. FME: A flexible modelling environment for inverse modelling, sensitivity, identifiability, monte carlo analysis. *R package version*, 1, 2009.

[31] Karline Soetaert and Thomas Petzoldt. Inverse modelling, sensitivity and monte carlo analysis in R using package FME. *Journal of Statistical Software*, 33, 2010.

[32] TV Elzhov, KM Mullen, and B Bolker. minpack. lm: R interface to the Levenberg-Marquardt nonlinear least-squares algorithm found in MINPACK. *R package version*, pages 1–1, 2009.

[33] Angela R Mclean and Colin A Michie. In vivo estimates of division and death rates of human T lymphocytes. *Proceedings of the National Academy of Sciences*, 92(9):3707–3711, 1995.

Appendices

A Code in R

```
#-----------------------------------------------------
# Fitting a Mathematical Model of HIV to Data
# Tim Coorens
#-----------------------------------------------------

#Set working directory
setwd("/Users/timcoorens/Documents/R Projects/Thesis")

#-----------------------------------------------------
# Load Libraries
#-----------------------------------------------------
library(gdata) #Import data from Excel-files
library(deSolve) # Solving differential equation
library(minpack.lm) # Least squares fit using Levenberg-Marquart
   algorithm, used in modFit
library(FME) # Fitting model to data using modFit-command

#-----------------------------------------------------
# Load Data and Plot
#-----------------------------------------------------

wick=read.xls("wick2.xlsx")
names(wick)=c("time","C", "sd")

wickPIT=read.xls("wickPIT2.xlsx")
names(wickPIT)=c("time", "P", "sd")
wickPIT=wickPIT[wickPIT$time %in% wick$time,]

#-----------------------------------------------------
# The Model
#-----------------------------------------------------
HIV <- function (parameters) {
  derivs <- function (time, state, parameters) {
    with(as.list(c(parameters, state)), {
      dQ <- s - d_T * Q - Beta * Q * P
      dI <- -(d_I + gamma + k1 * C) * I + f * Beta * Q * P
      dP <- gamma * I - (d_P + k2 * C ) * P
      dC <- -d_C * C

      return(list(c(dQ, dI, dP, dC)))
    })
  } #Note: Q is used instead of T to avoid errors

  # injection events of CD8+ T cells
  injectevents <-data.frame(var = "C",
```

```
                              time = c(0.5 , 7.5),
                              value = with(as.list(parameters), inj
                                 ),
                              method = "add")

  # Initial conditions, measured for P and C, dependent on
     parameters for I and T
  P_0 <- 0.35
  C_0 <- 0
  I_0 <- with(as.list(parameters), (d_P/gamma)*P_0)
  Q_0 <- with(as.list(parameters), s/(d_T + Beta * P_0))
  state <- c(Q=Q_0, I=I_0, P=P_0, C=C_0)

  times <- c(seq(0, 46, 1), wick$time)

  out <-ode(func = derivs, times = times,
            y = state, parms = parameters, method = "impAdams",
            events=list(data =injectevents))

  as.data.frame(out)
}

#-----------------------------------------------------
# Model parameters
#-----------------------------------------------------
parameters <- c(k1 = 0.76, #rate of killing of I by C
                k2 = 0.9, #rate of killing of P by C
                d_P = 1.30, #death rate of P
                d_I = 0.13, #death rate of I
                d_T = 0.21, # death rate of T
                s = 1.38, # generation of T
                Beta = 8.0, #beta=b*p/c, infection rate
                d_C = 0.54, #death rate of C
                inj = 1.90, #injection of neo CD8+ cells
                gamma = 1, #eclipse phase parameter
                f = 0.96) #fraction of infected T to I = gamma_0/
                   (d_I0+gamma_0)

#-----------------------------------------------------
# Model Cost Function
#-----------------------------------------------------

HIVcost <- function (parameters) {
  out <- HIV(parameters)
  cost <- modCost(model=out, obs = wickPIT, err="sd")
  return(modCost(model=out, obs=wick, err="sd", cost=cost))
}

HIVcost(parameters)$model

```

```
#-----------------------------------------------------
# Sensitivity analysis
#-----------------------------------------------------

Sfun <- sensFun(HIVcost, parameters, varscale=1)
summary(Sfun)

#-----------------------------------------------------
# Fitting the Model to the Data
#-----------------------------------------------------

HIVcost2 <- function(lpars) HIVcost(c(exp(lpars),
                                      Beta = 8.0,
                                      d_C = 0.54,
                                      inj = 1.90,
                                      gamma = 1,
                                      f = 0.96)) #parameters
                                         included here will not be
                                          fitted over

Pars <- parameters[1:6] #select parameters to be fitted over
Fit <- modFit(f = HIVcost2, p = log(Pars), method = "Marq")
exp(coef(Fit))

# method = "Marq" used Levenberg-Marquardt algorithm, "Port" uses
    nlminb command

#-----------------------------------------------------
# Plot Initial Guess vs. Fitted Result
#-----------------------------------------------------

ini <- HIV(parameters=c(Pars,
                        Beta = 8.0,
                        d_C = 0.54,
                        inj = 1.90,
                        gamma = 1,
                        f = 0.96))
final <- HIV(parameters=c(exp(coef(Fit)),
                          Beta = 8.0,
                          d_C = 0.54,
                          inj = 1.90,
                          gamma = 1,
                          f = 0.96))

par(mfrow =c(1,2))
plot(wickPIT$time, wickPIT$P, xlab = "time", ylab="P", ylim = c
   (0,1))
lines(ini$time, ini$P, lty=2)
lines(final$time, final$P, lty=1)
plot(wick$time, wick$C, xlab="time", ylab ="C")
```

```
138 lines(ini$time, ini$C, lty=2)
139 lines(final$time, final$C)
140 legend("topright", c("data", "fitted"),
141        lty = c(NA, 1, 1), pch = c(1, NA, NA))
142 par(mfrow = c(1,1))
```

model.R