Mathematical and Statistical Modelling of factors affecting CD4 T-cell Reconstitution in HIV-Infected Children starting Anti-retroviral Therapy

Adedeji Oluwaseun Majekodunmi

A thesis submitted to University College London for the degree of
Doctor of Philosophy

UCL Centre for Mathematics, Physics and Engineering in the Life Sciences and Experimental Biology
and
Infection, Inflammation and Rheumatology Unit,
UCL Great Ormond Street Institute of Child Health

University College London

March 8, 2018

I, Adedeji Oluwaseun Majekodunmi, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Thymic output in children increases to its peak during the first 1 year of life followed by a gradual decline to a steady state by 20 years of age. Immune reconstitution in HIV-infected children is a highly dynamic process driven by factors such as thymic output, age, proliferation and loss of T cells, viral load, anti-retroviral therapy (ART) and co-infection. Understanding the mechanisms driving immune reconstitution is critical to ensuring the best clinical outcomes in HIV-infected children and in shaping the future of medical intervention. This thesis focuses on the factors affecting the degree and extent of $CD4^+$ T cell reconstitution in HIV-infected children receiving ART.

Hepatitis C virus (HCV) co-infection in HIV-infected patients has been linked to an increased incidence of hospitalisation, rapid progression to end stage liver disease and hepatocellular carcinoma. Despite effective antiretroviral drugs, liver-related complications are now a leading cause of mortality and morbidity in HIV-infected patients. The first section of this work uses a monophasic asymptotic recovery model to investigate the impact of hepatitis C co-infection on $CD4^+$ T cell recovery in HIV-infected children. The model was fitted to age-adjusted $CD4^+$ T cell counts using the framework of non-linear mixed-effects regression and the effects of important covariates such as age, viral load and pre-ART hepatitis C status were investigated. The results indicate that HIV/hepatitis C co-infected children recover their $CD4^+$ T cells much more slowly compared to HIV-monoinfected individuals.

The next chapter uses the same model to identify predictors of poor immune reconstitution in 2204 HIV-infected children on ART and higher age at start of therapy was highlighted as a main factor associated with poor immune recovery.

Finally, a mechanistic model of naive $CD4^+$ T cell homeostasis was developed aiming to understand the recovery of $CD4^+$ lymphocyte subsets in HIV-infected children receiving ART. As $CD31^+$ T cells are an important component of naive T cell pool in children, the model was further adapted to understand dynamics of $CD31^+$ T cell homeostasis in HIV-infected children on ART. Reduced theoretical thymic output was associated with poor $CD31^+$ T cell recovery which is strongly related to poor overall $CD4^+$ T cell recovery.

Publications

- Majekodunmi AO, Thorne C, Malyuta R, Volokha A, Callard RE, Klein NJ, Lewis J; European Paediatric HIV/HCV Co-infection Study group in the European Pregnancy and Paediatric HIV Cohort Collaboration and the Ukraine Paediatric HIV Cohort Study in EuroCoord. **Modelling CD4 T Cell Recovery in Hepatitis C and HIV Co-infected Children Receiving Antiretroviral Therapy**. *Pediatr Infect Dis J.* 2017 May; 36(5):e123-e129. doi: 10.1097/INF.0000000000001478.

- Majekodunmi AO, Callard R, Klein N, Lewis J. **Re: Recovery of CD4 T Cells in HIV/HCV Coinfected Children: Is it Really Impaired?** *Pediatr Infect Dis J.* Mar; 37(3):278-279. doi: 10.1097/INF.0000000000001779.

- Adedeji Majekodunmi, EPPICC Cohort; **Predictors of poor immune recovery in HIV-infected children on anti-retroviral therapy** (First draft-August 2017).

Acknowledgements

Undertaking a research project leading to a doctoral thesis represents a unique opportunity to study in depth a specialist subject of interest. Over the last four years, I have been very fortunate to have worked with numerous talented individuals who have contributed to the success of this endeavour and to whom I am immensely grateful. I would like to express my appreciation as follows:

To the children and adolescents who participated in the ARROW and EPPICC studies and their families.

I am very grateful for the opportunity to work with my primary supervisor Nigel Klein during the course of my PhD at UCL Great Ormond Street Institute of Child Health. Thank you for providing direction, encouragement, support and valuable collaborations.

To my secondary supervisor Robin Callard for his expertise in immunology and modelling. Thank you for providing timely feedback and allowing me to grow in confidence in research methodology. Thanks for the inspiration I received from your research which led me to apply to UCL CoMPLEX in the first instance.

To my secondary supervisor Joanna Lewis for her patience and immense dedication especially with the mathematical and computational aspects of this work. Thank you for your availability and attentiveness in answering a broad range of questions over the past four years. Thanks for gently setting me off into this field of immune reconstitution and non-linear mixed-effects modelling. It has been a pleasure working with you and thanks for sacrificing your time and efforts in supervising my PhD.

To Joseph Standing for his expertise in non-linear mixed-effects modelling, mechanistic modelling and use of NONMEM software.

To the UCL Pharmacometrics group for the numerous constructive feedback that I have received over the years and for the opportunities to develop my presentation skills. In particular, I would like to thank Charlotte Barker, Eva Germovsek, Rollo Hoare, Dagan Lonsdale, Frank Kloprogge, Julie Bertrand and Joseph Standing for their feedback and for making the PAGE (Population Approach Group in Europe) conference a highlight of the year.

To Claire Thorne with whom I collaborated on the EPPICC hepatitis C/HIV project which materialised into the 3rd chapter of this work.

To Rollo Hoare whose preliminary work made it possible to develop the mechanistic model presented in chapter 5.

To fellow CoMPLEX students for their moral support, encouragement and good sense of humour.

To the EPPICC team for the opportunity to collaborate on interesting projects.

I would like to appreciate fellow PhD students, Post-Docs and staff at UCL Great Ormond Street Institute of Child Health who have made my PhD experience all the more enjoyable and fun: Ronan Doyle, Lisa Carter, Liam Shaw, Hannah Jones, Vania de Toledo, Ben Margetts, Teresa Attenborough, Athina Ghazi, Ebun Omoyinmi, Felicity Fitzgerald, Julia Kenny, Sira Nanthapisal and Keylock Annette.

To my parents Mr and Mrs Majekodunmi and siblings Opeyemi, Itunu and Bukky for their moral support.

To my wife Dolapo and daughter Isabelle for their perseverance, dedication and unwavering love and care.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 An overview of the vertebrate immune system

The human immune system is composed of the *innate* and *adaptive* arms which coordinate together in fighting pathogenic organisms.

The innate immune system which serves mainly as the first line of defence against invading pathogens is composed of physical barriers such as the skin, complement proteins, leucocytes and pattern recognition molecules which act swiftly in eliminating infectious organisms.

The *adaptive* immune system (also *acquired* or *cell-mediated* immune system) is primarily made up of B and T lymphocytes which provide a much more sophisticated and effective defence against pathogens. Both B and T lymphocytes possess cell surface receptors through which they are able to recognize molecular structures that are unique to invading pathogens (*antigens*). Through somatic recombination, they are able to generate an enormous repertoire which ensures that unencountered antigens are recognised and eliminated accordingly. B cells secrete antibodies which bind directly to antigens whilst T lymphocytes recognise antigens presented in association with a *major histocompatibility complex* (MHC) molecule. This MHC-antigen combination is usually processed by *antigen presenting cells* (APCs) of the innate immune system. One unique feature of the adaptive immune system is its ability to remember previously-encountered pathogens enabling it to mount a much more rapid immune response in cases of re-infection by the same organism. This property termed "immunological memory" is responsible for the global success of vaccination in controlling infectious diseases. Following the union of an antigen to a B cell receptor (BCR), the B cell divides and differentiates into plasma cells which secrete antibodies, the soluble form of the BCR. A naive T cell can differentiate into one of three effector T lymphocytes: the first two are differentiated by the *cluster of differentiation* (CD) molecules which they express on their cell surface whilst the third type are regulatory T cells which are able to inhibit the activity of other lymphocytes and hence play an important role in autoimmunity and immune surveillance. $CD4^+$ T cells are helper cells

which recognise antigens presented with MHC II molecules whilst $CD8^+$ T cells are cytotoxic cells which recognise antigens expressed with MHC I molecules. MHC I molecules are expressed ubiquitously on all cells whilst MHC II molecules are only expressed on B cells, dendritic cells and monocytes. Helper T cells are very important in the adaptive immune system because they enable the effector function of other cells including B cells and macrophages. As will be discussed later in this chapter, helper T cells are the main targets in HIV infection. Cytotoxic T cells directly kill cells that are infected with intracellular pathogen such as viruses.

The two arms of the immune system are able to communicate via cell-cell contact as well as through soluble cell signalling molecules known as *cytokines*. Following exposure to an infectious agent, the innate immune system is almost immediately activated to fight the infectious agent followed by activation of the acquired immune system a few days later. This collaboration ensures that pathogens that evade the innate immune system are mostly eliminated by the acquired immune system.

All the cellular components of the blood including erythrocytes (red blood cells), thrombocytes (platelets) and leucocytes (white blood cells) are derived from haematopoietic stem cells of the bone marrow (Figure 1.1 shows a full diagram of the lineages). These stem cells are said to be pluripotent due to their ability to differentiate into other cell types in the blood. They lose their pluripotency by differentiating into the myeloid and lymphoid lineages. The myeloid lineage gives rise to cells of the innate immune system including macrophages, granulocytes, mast cells and dendritic cells. They also give rise to other components of the blood such as platelets and erythrocytes. Due to their ability to directly engulf infectious agents, macrophages, granulocytes and dendritic cells are collectively known as "phagocytes". The lymphoid lineage gives rise to the B and T lymphocytes (adaptive immune system) and natural killer cells (innate immune system).

Organs of the immune system can be broadly divided into central (primary) and peripheral (secondary) lymphoid organs. The bone marrow and the thymus are the central lymphoid organs which produce early precursors and coordinate the maturation of immune cells respectively. The secondary lymphoid organs including the spleen, lymph nodes, specialized gut and mucosal tissues are sites where mature naive lymphocytes continually circulate until they encounter a foreign antigen. The lymph nodes are connected together within a network and are drained by lymphatics which eventually empty into the large veins of the circulatory system. Both B and T cells are produced in the bone marrow and whilst T cells migrate to the thymus for maturation, B cells mature in the bone marrow.

Lymphocytes yet to encounter foreign antigens are termed *naive* lymphocytes. These continue to

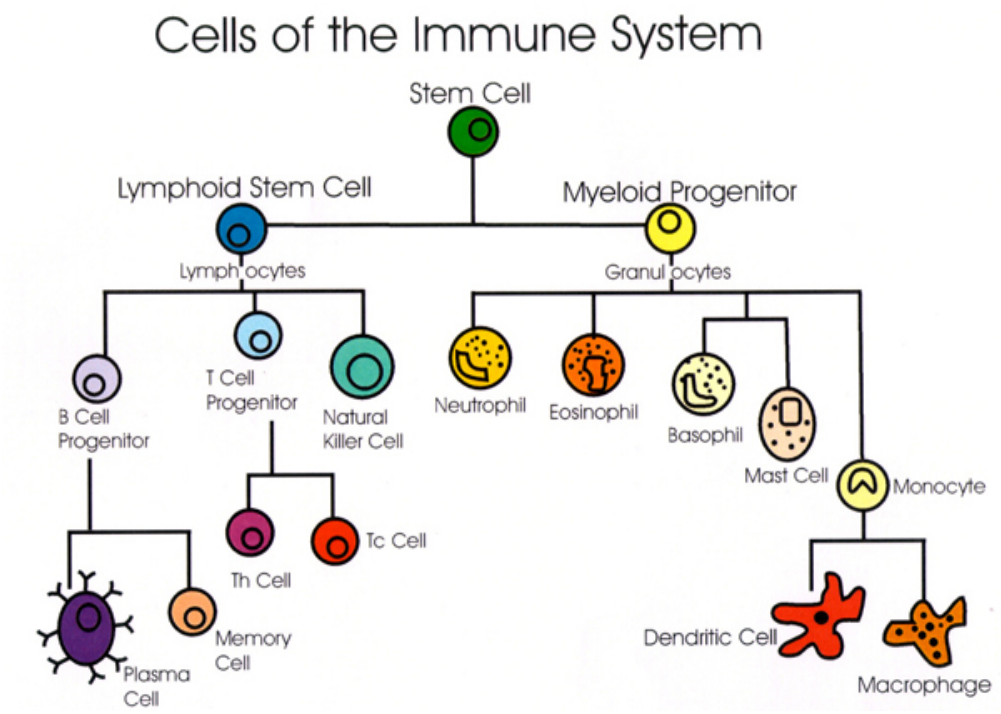


Figure 1.1: The two main lineages of the haematopoietic stem cells-lymphoid and myeloid. Lymphocytes (acquired immune system) emerge from the lymphoid progenitors whilst most cells of the innate immune system are from the myeloid progenitors [Kenneth Todar, 2006]

divide and circulate between the blood and the lymphatics until they encounter an antigen following which they differentiate into effector lymphocytes which eventually become *memory* cells. From the site of an infection, free antigens and antigen-laden dendritic cells migrate into the peripheral lymphoid organs via the draining lymph nodes. On arriving at the peripheral lymphoid organs, dendritic cells activate T cells whose receptors are specific to these antigens. These activated lymphocytes become effector cells and undergo rapid proliferation. Activated T cells are pushed along the lymphatics under the pressure of extracellular fluid drained directly from the blood vessels. After about 4-6 days, these activated antigen-specific T cells return to the site of infection via the blood stream.

The main hallmark of HIV infection is the depletion of $CD4^+$ T cells. Therefore, in the sections that follow, we shall examine the development of T cells in more detail.

1.2 Development and survival of T cells

The T cell receptor (TCR) is a membrane-bound molecule located on the T cell membrane which is capable of recognizing specific antigens presented with the histocompatibility complex molecule. The T cell receptor is made up of two distinct polypeptide chains (heterodimer) linked by a disulphide bond. Two

TCR heterodimers have been identified: the $\alpha\beta$ receptor and the less common $\gamma\delta$ receptor. In humans and mice, the $\alpha\beta$ lineage of T cells constitute about 80% of the total T lymphocyte pool. They are therefore a major part of the adaptive immune system which also includes B lymphocytes. The $\gamma\delta$ cells on the other hand are present in large numbers in the epithelium and mucosal-associated lymphoid tissue [Ciofani and Zúñiga-Pflücker, 2010, Pollock and Welsh, 2002]. This is necessary as the genetic locus encoding the TCR α-chain has variable (V), joining (J) and constant (C) regions whilst the β and δ chains both have an additional diversity (D) region. These gene segments undergo recombination during T cell development in the thymus to produce complete V-domain exons. The genes encoding for the various parts of each α and β polypeptide chains of the TCR do not occur together on the chromosome and as a result, unwanted pieces of DNA need to be excised in order to bring these genes closer together. T cell receptor excision circles (TRECs) are circles of DNA produced from this TCR *genetic re-arrangement* which are neither replicated nor degraded during mitosis and are abundant in recent thymic emigrants [Bains et al., 2009b, Bains et al., 2009a]. The role of TRECs in quantifying naive $CD4^+$ T cell homeostasis will be discussed further in chapter 5. The process of genetic re-arrangement generates an enormous diversity in the possible combinations of α and β chains leading to as many as 10^{10} unique TCR molecules in one individual. This diversity ensures that most antigens can be recognised and hence eliminated by the adaptive immune system.

1.2.1 Positive and negative selection of early T cell precursors

Selection of mature T lymphocytes in the thymus is a rigorous and costly process that results in the death of up to 98% of immature thymocyte emigrants from the bone marrow. In the thymus, the T cell receptor loci undergo re-arrangement into the $\alpha\beta$ and $\gamma\delta$ lineages. The immature thymocytes progress through stages of differentiation from double negative ($CD4^-CD8^-$ cells) and eventually into single positive T cells ($CD4^+$ or $CD8^+$). Immature thymocytes with weak affinity for self-peptide-MHC complex (p-MHC, about 1% of T cells) are selected to undergo maturation whilst thymocytes with a strong affinity for self-peptide-MHC complex (about 1%) are signalled by antigen-presenting cells (APCs) to undergo apoptosis [Boyman et al., 2009, Starr et al., 2003]. The remaining 98% of thymocytes die through "neglect" as a result of inadequate TCR signals. The thymocytes which emerge from the positive selection process have a very diverse repertoire of TCR which enables them to recognise foreign antigens whilst tolerating self antigens. Hence, the thymus is the main source of TCR diversity which ensures effective immune response to a broad range of pathogenic organisms.

1.3 $CD4^+$ and $CD8^+$ T cell subsets

1.3.1 $CD4^+$ T cell subsets

$CD4^+$ T cells potentiate the effector functions of other immune cells in eliminating infectious agents and are able to differentiate into five main subsets namely T helper cells 1 (Th1), Th2, Th17, T regulatory T cells (Treg) and Tfh (follicular helper T cells). This is made possible through cytokine production. For example, the differentiation of Th1 T cells from naive $CD4^+$ T cells is favoured by high concentrations of IL-12 and IFN-γ cytokines whilst Th2 development is via IL-4. Furthermore, individual Th subsets secrete cytokines which may have survival, protective and pro/anti inflammatory roles and are responsible for eradicating different categories of pathogens. Th1 cells secrete IFN-γ which activate macrophages in destroying intracellular pathogens and also favour production of more Th1 cells via a positive feedback loop. Th2 cells produce IL-4, IL-5 and Il-13 which are involved in eradicating extracellular parasites such as helminths. Th17 cells are involved in eradicating extracellular bacteria and fungi and also potentiate the activity of neutrophils against these organisms. T regulatory T cells are involved in immune regulation and suppression of T cells which is important in preventing autoimmunity. Two forms of Tregs have been recognised: natural Tregs develop within the thymus whilst adaptive Tregs are induced from existing naive $CD4^+$ T cells. The natural regulatory T cells are said to make up about 10-15% of the total circulating $CD4^+$ T cell population in humans . Other $CD4^+$ T cell subsets include Th9 and Th22 cells [Golubovskaya and Wu, 2016, Murphy and Weaver, 2017].

1.3.2 $CD8^+$ T cell subsets

Unlike $CD4^+$ T cells, $CD8^+$ T cells all differentiate into cytotoxic T lymphocytes. These cells play an important role in defence against intracellular pathogens which include viruses. They are able to recognise peptide:MHC complex displayed on the surfaces of virus-infected cells. $CD8^+$ T cells can be activated into cytotoxic T cells via two ways; through activated dendritic cells and via $CD4^+$ effector T cells. Cytotoxic T cells are able to control intracellular pathogens by releasing cytotoxic granules which perforate cell membranes (perforin) and activate apoptosis (granzymes). $CD8^+$ Cytotoxic T cells also secrete cytokines that directly inhibit viral replication and increase expression of MHC class I molecules (IFN-γ) and others that strengthen the host immune response (TNF-α) [Murphy and Weaver, 2017].

1.3.3 Naive T cell homeostasis

Maintenance of positively-selected T cells in the periphery is facilitated through continuous contact of the TCRs with self p-MHC complexes under low levels of IL-7 [Sprent and Surh, 2002]. This mechanism ensures that viable T cells capable of repopulating the T cell pool are retained in the periphery [Boyman et al., 2009, Surh and Sprent, 2008]. These signals enable naive T cells to remain in interphase for much longer via the upregulation of anti-apoptotic proteins (Bcl-2, Mcl-1) by prolonging their life spans. This also means that the naive T cell pool is maintained at a constant level with minimal proliferation especially after the onset of thymic involution at around 1 year of age. This is certainly the case when the naive T cell pool is normal in size [Boyman et al., 2009].

However, when the naive T cell pool is reduced as is the case in lymphopenic conditions, there is an amplification of the weak TCR signal delivered from the self p-MHC ligands. Since T cells competitively bind to IL-7 receptors (IL-7R), the reduction in the lymphocyte pool means that there are more IL-7 molecules available to bind to IL-7R [Jameson, 2005]. These events lead to the delivery of mitogenic signals via the self p-MHC ligand which forces the naive T cells out of interphase into cell division and hence restoring the naive T cell pool to its equilibrium level. The process is termed lymphopenia-induced proliferation (LIP) or homeostatic expansion [Almeida et al., 2005, Boyman et al., 2007, Freitas and Rocha, 2000, Surh and Sprent, 2000]. The combined signalling strength of the IL-7R and TCR determines the rate at which lymphopenia-induced proliferation takes place. As such, lymphocytes with strong affinity TCR for self p-MHC complex are seen to exhibit a much more rapid rate of proliferation compared to low affinity TCRs [Kieper et al., 2004]. Studies have suggested that competition-driven affinity for self p-MHC complexes by T lymphocytes is responsible for the generation of a diverse repertoire of TCR during lymphopenia-induced proliferation [Jameson, 2005].

Following export from the thymus, naive T cells continue to circulate between the secondary lymphoid organs (spleen and lymph nodes) via the blood and the lymphatics until they encounter a foreign antigen coupled to an MHC molecule. Upon recognition of a foreign antigen, they undergo activation and proliferation to become effector T cells. Once the pathogen is cleared, majority of these effector cells die due to the absence of sustaining TCR signals. However, about 5% of these effector T cells survive and are transformed into memory T cells which are identified by high expression of CD45RO [Sprent and Surh, 2002, Surh and Sprent, 2008, Boyman et al., 2007]. Naive T cells that do not encounter a foreign antigen will continue to bind to self p-MHC ligands and remain in a quiescent state. One difference between the

effector response and the homeostatic expansion is that homeostatic proliferation takes place more slowly with 1-5 cycles of cellular division over a period of 1-2 weeks. $CD4^+$ T cells divide more slowly when compared to $CD8^+$ T cells [Sprent et al., 2008]. Hence under normal physiological conditions, IL-7 is important for survival of naive T cells and at elevated levels, IL-7 augments homeostatic proliferation of naive T lymphocytes. Under lymphopenic conditions, most naive CD4 T cells are able to acquire phenotypic characteristics and physiological properties of memory cells without prior activation into effector cells [Surh and Sprent, 2008]. The implication of these observations is that memory-like phenotype or function cannot be reliably used as an indicator of antigen-experienced cells [Murali-Krishna and Ahmed, 2000, Oehen and Brduscha-Riem, 1999].

As the primary source of naive $CD4^+$ T cells in humans, the thymus plays an important role in T cell homeostasis. Steinmann et al. showed that thymic volume is at its maximum around age 1 year and then remains relatively unchanged throughout life [Steinmann et al., 1985]. At birth, the thymus is fully developed whilst the volume of the thymic epithelial space continues to increase to a peak at age 1 year and then decreases over an individual's life span. Notable morphological changes take place within the intra-thymic compartments in the first 20 years of life where the thymic epithelial space (TES) undergoes 70% involution with a concomitant expansion of the perivascular space, connective and adipose tissues. Since the TES is the main site of T cell maturation (thymopoiesis), it has been proposed that the age-related involution taking place in this region is likely to affect overall thymic functionality [Bains et al., 2009b, Rezzani et al., 2014]. This was later corroborated by evidence from studies involving TRECs suggesting that production of naive T lymphocytes declined with age [Starr et al., 2003, Sprent et al., 2008]. In spite of these changes, the population of peripheral naive T cells remains fairly constant throughout adulthood [Berzins et al., 2002, Kimmig et al., 2002, Livak and Schatz, 1996], suggesting that naive T cells are either continuing to divide after leaving the thymus without losing their naive phenotype or that they are very long-lived. Hence, two mechanisms have been proposed: continuous production of mature T cells by the thymus and peripheral expansion or death of recent thymic emigrants (RTEs). This indicates there must be at least two distinct subsets of naive $CD4^+$ T cells which can be distinguished by their proliferative history: the first subset are the recent thymic emigrants (RTEs) whilst the other are RTEs that have undergone further division in the periphery, so called central naive $CD4^+$ T cells [Kohler and Thiel, 2009]. These two subsets are now distinguishable in peripheral blood samples by the expression of the cell surface molecule CD31 (PECAM-1: platelet endothelial cell adhesion molecule or glycoprotein IIa) and concentration of T receptor excision circles. A number of studies have observed an increase in the

TCR excision circles content in the peripheral $CD4^+$ T lymphocyte in patients receiving antiretroviral therapy and in adult bone marrow transplantation which both support continuous production of mature T cells from the thymus even after thymic involution [Haynes et al., 2000, Kilpatrick et al., 2008, Ye et al., 2003].

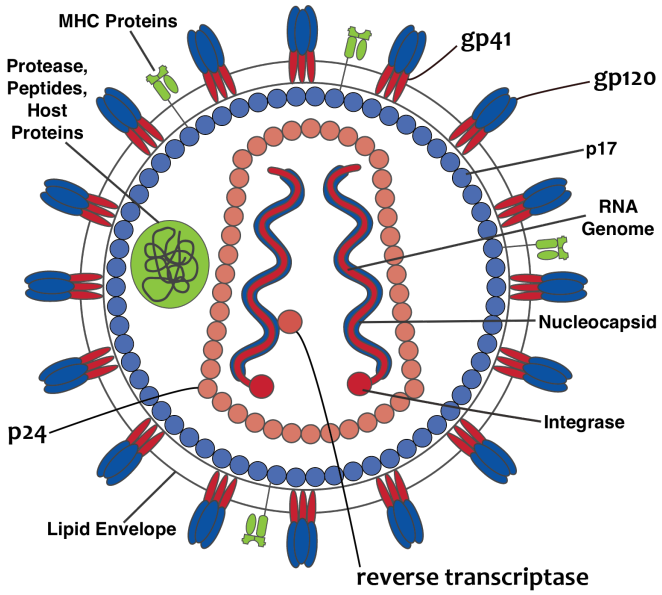


Figure 1.2: The above diagram illustrates the structural components of the Human immunodeficiency virus (HIV). p24: capsid protein, p17: matrix protein, gp41: transmembrane glycoprotein, gp120: docking glycoprotein. Original image taken from [International Genetically Modified Machine (iGEM) Foundation, 2016] and further annotated here for clarity

1.4 Pathogenesis and clinical course of HIV infection

The human immunodeficiency virus (HIV) is an enveloped RNA lentivirus which primarily infects activated $CD4^+$ T cells. Each virion has two copies of RNA and several copies of reverse transcriptase, integrase and proteases which are essential viral enzymes needed to produce new virions (Figure 1.2). HIV is able to gain entry into macrophages, dendritic cells and $CD4^+$ T cells due to expression of CD4 molecules on their surfaces. Once in an infected host, HIV invades uninfected cells by linking its gp120 glycoprotein with both the CD4 receptor and either of CCR5 or CXCR4 chemokine co-receptors [Berger et al., 1999]. The binding of the CD4 receptor to the chemokine co-receptor allows the gp41 glycoprotein to fuse the viral envelope to the cell's plasma membrane. On entering a host cell, HIV makes a complemetary DNA (cDNA) from its own RNA through its reverse-transcriptase enzyme. Using its integrase enzyme, this transcribed cDNA (*provirus*) is then integrated into the chromosome of the infected cell. New mRNA produced from transcription of the provirus are used in assembling new virions which bud from the cell's plasma

membrane to infect another cell. Due to lack of proof-reading mechanisms, the HIV reverse transcriptase enzyme introduces errors in the process of making copies of cDNA molecules from RNA. This produces a mutation rate of approximately 3×10^{-5} per nucleotide base pair per replication cycle [Abram et al., 2010, Roberts et al., 1988] leading to the production of several mutant viruses (quasi species) within the same individual which inevitably evade cytotoxic T cell killing of infected cells. This is due to the inability of TCRs to recognise antigens from the mutant species.

In untreated patients, the acute phase of HIV infection is characterised by a rapid decline in $CD4^+$ T cells and a corresponding increase in HIV viral load (Figure 1.3). During the acute phase, non specific symptoms such as coryzal illness are fairly typical in adult patients making it very difficult to diagnose early onset HIV infection. Cytotoxic T cells and antibodies are initially directed against infected cells and HIV respectively leading to a reduction in the viral load to a steady state level. The acute phase is followed by an asymptomatic period (*clinical latency*-Figure 1.3) during which the viral load remains low and $CD4^+$ T cells fall slowly and continuously. This is due to the fact that the rate of $CD4^+$ T cell production is much slower compared to the rate of $CD4^+$ T cell loss [Moir et al., 2011]. The latency period may last for a few years during which HIV continues to replicate within the T cells in the secondary lymphoid organs and this is accompanied by a high turnover of $CD4^+$ T cells, partly due to the ability of macrophages and dendritic cells to accommodate proliferating HIV virions without being destroyed by them. Therefore large numbers of macrophages and dendritic cells within the lymphoid tissues are an important reservoir through which HIV virions are able to re-infect $CD4^+$ T cells [Moir et al., 2011, Maartens et al., 2014]. Infected $CD4^+$ memory cells with long half lives (44 months) can also be re-activated by antigens causing them to produce new viral particles that can infect other activated $CD4^+$ T cells [Murphy and Weaver, 2017]. In the absence of antiretroviral therapy (ART) the CD4 count continues to decline until the development of an acquired immune deficiency syndrome (AIDS) which is clinically defined as a CD4 count < 200 cells/μL or the manifestation of certain AIDS-defining illnesses [Haase, 1999]. At this stage, the control of HIV viral load is lost which increases vulnerability to opportunistic infections such as *Pneumocystis jiroveci* pneumonia (previously *Pneumocystis carinii*), development of malignancies and eventually death. AIDS is marked by a failure of all immune defences normally facilitated by $CD4^+$ T cells.

1.5 Paediatric HIV Infection

Most infected children acquire HIV vertically from HIV-infected mothers especially during birth but also through breast feeding or in-utero [Prendergast et al., 2007]. In the absence of any intervention,

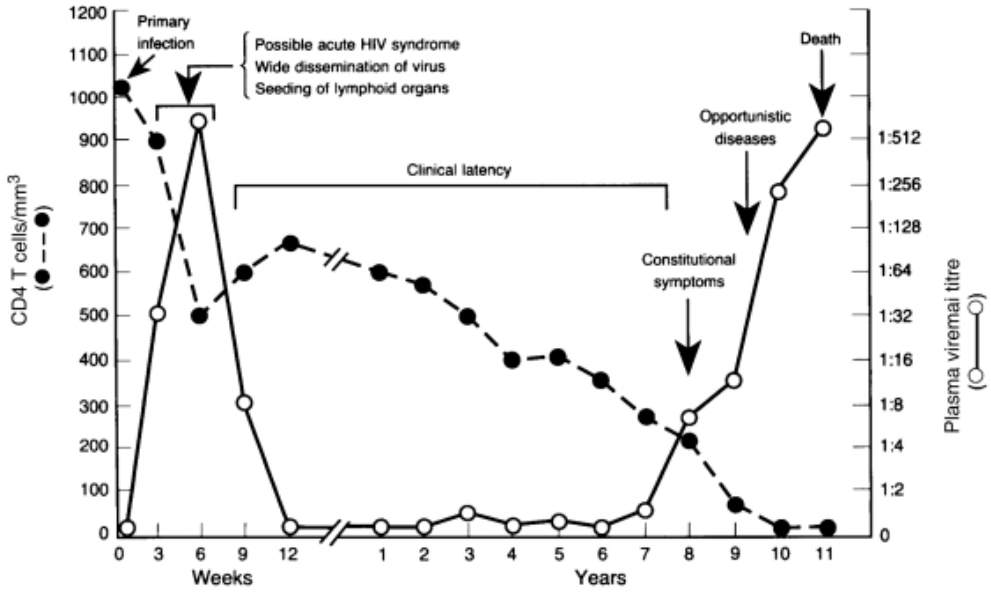


Figure 1.3: Clinical course of HIV infection illustrating the initial drop in $CD4^+$ T cell count associated with a sharp increase in HIV viral load. Diagram taken from [Pantaleo et al., 1993]

mother-to-child transmission rate of HIV infection is 15-45% [World Health Organisation, 2017b, Maartens et al., 2014]. However, when interventions such as treating the mother and baby with ART, avoiding breastfeeding and delivering baby via caesarian section are employed, the transmission rate can be reduced to under 5% [World Health Organisation, 2017b]. A maternal/paediatric study conducted in the UK and Ireland on children born between 2000 and 2011 found that mother-to-child transmission rate reduced significantly from 2.1% in children born in 2000-2001 to 0.46% in 2010-2011 due to a number of factors which included earlier administration of antenatal ART [Townsend et al., 2008]. It was noted that mothers who had viral load of <50 copies/ml had a much more reduced mother-to-child transmission rate (MTCT) of 0.09% (six of 6347) compared to a transmission rate of 1% (14/1349) in mothers who had viral load of 50-399 copies/ml. Risk of transmission was 0.26% (two of 777) amongst mothers with viral load of 50-399 copies/ml who had elective caesarian compared to 1.1% (two of 188) in women who had planned vaginal delivery [Townsend et al., 2014].

HIV infection tends to present with non-specific illnesses in children including pneumonia, hepatosplenomegaly, oral thrush, failure to thrive, etc. Unlike in adults, cytotoxic T cell response is much weaker in the first year of life which allows HIV viral load to remain very high in the acute phase of the infection (Figure 1.4). Hence, the natural history of HIV infection is much more severe in children as exemplified by higher mortality rates and more rapid progression to AIDS. By 2 years of age, mortality rate is around 45-59% in untreated children living in resource-limited settings [Dabis et al., 2001]. This is due to

the relatively underdeveloped immune system of younger children compared to adults [Prendergast et al., 2007]. As a result, younger children are unable to mount appropriate cytotoxic T lymphocyte-mediated response needed to control viral replication (Figure 1.4). However, with effective ART, HIV can be managed as a chronic condition with most children surviving into adulthood. Administration of ART leads to a reduction in HIV viral load and a gradual increase in $CD4^+$ T cell numbers to a steady state level.

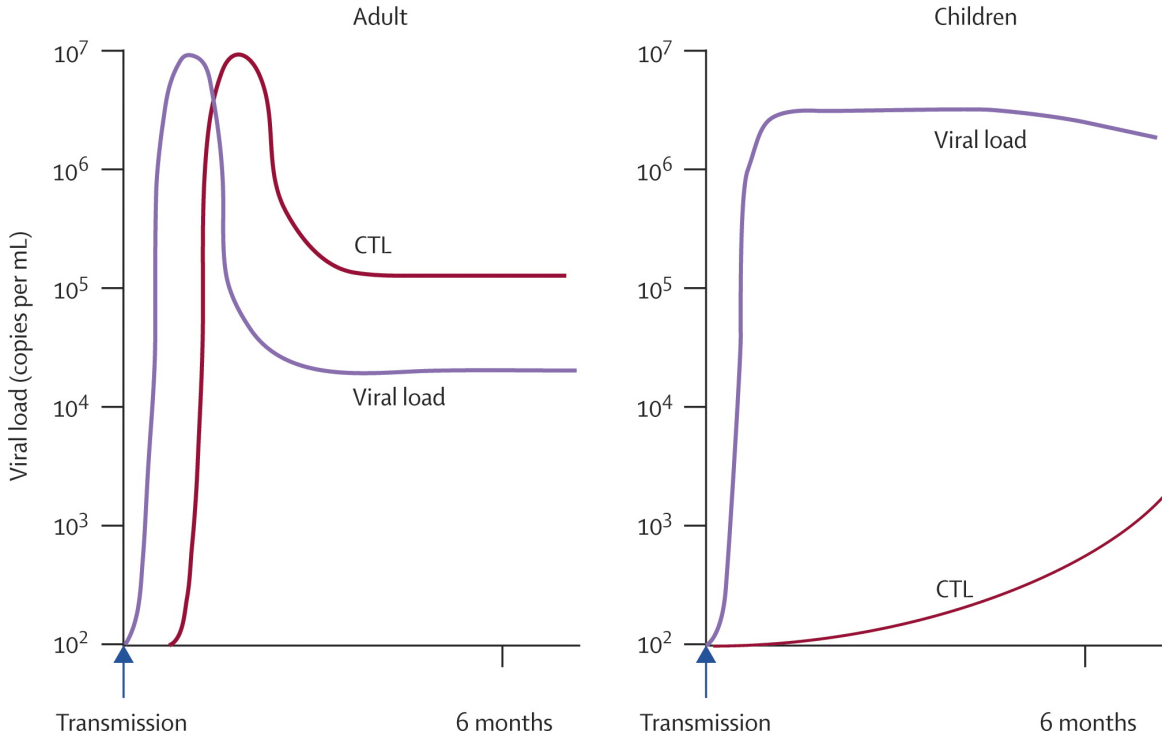


Figure 1.4: HIV-specific cytotoxic T cells (CTL) response and viral load changes during the acute phase of HIV infection in untreated adults (left) and children (right). Image taken from [Prendergast et al., 2007]

Improvements in CD4 count in response to ART are the result of viral suppression, which allows the body's homeostatic mechanisms to repopulate the $CD4^+$ T cell pool. These mechanisms vary with age, and in young children are thought to rely largely on a highly active thymus which produces large numbers of naive $CD4^+$ T cells before undergoing involution in early adulthood [Steinmann et al., 1985, Hazenberg et al., 2004]. Thymic activity in HIV-infected children has also been observed to improve on ART [Douek et al., 2000, Ometto et al., 2002, Sandgaard et al., 2014]. The process of immune recovery in children differs from adults. Perhaps as a result of different mechanisms of reconstitution, children seem to experience a monophasic CD4 recovery originating mainly from the naive T cell pool, whereas in adults there is an initial recovery of the memory cell count, followed by a slower growth of the naive pool [De Rossi et al., 2002]. This CD4 reconstitution in children differs both empirically and mechanistically from recovery in adults and requires separate investigation.

1.6 How are CD4$^+$ T cells lost in HIV infection?

One of the mechanisms responsible for CD4$^+$ T cell decline in untreated HIV-infected individuals is apoptosis of infected T cells. However, it is well known that only a small percentage of CD4$^+$ T cells are infected at any one time whilst a much larger proportion of T cells are proliferating and dying in HIV infection [Anderson et al., 1998]. Hence, death of HIV-infected cells is not enough to explain the decline in CD4 T cells and another mechanism proposed for T cell loss is death of uninfected T cells mostly due to chronic immune activation and is driven by HIV and cytokines. Increased levels of activated CD4$^+$ T cells have been seen in the peripheral blood of HIV-infected individuals [Hazenberg et al., 2000b, Ribeiro et al., 2002, Mohri et al., 2001] and the degree of T cell activation has been shown to be a strong predictor of the rate of clinical deterioration to acquired immune deficiency syndrome [Giorgi et al., 1999].

Other proposed mechanisms for CD4$^+$ T cell depletion include dysfunctional T cell production in the bone marrow and thymus [McCune, 2001, De Rossi et al., 2002], damage to secondary lymphoid organs and change in distribution of CD4$^+$ lymphocytes between the blood and the lymphatics [Bujdoso et al., 1989, Zeng et al., 2012, Schacker et al., 2002].

1.7 Antiretroviral therapy and immune reconstitution

The introduction of the first antiretroviral agent zidovudine (AZT) in 1987 was a remarkable breakthrough in the management of AIDS. This was followed by introduction of didanosine (ddl) and zalcitabine (ddC) in 1991 and 1992 respectively [Cihlar and Fordyce, 2016]. However, the ability of HIV to develop resistance to single antiretroviral agents led to the recommendation of combined ART (also HAART) in the late 1990s. The use of a combination of drugs with different mechanisms of action increased the overall efficacy of treatment and helped with combating drug resistance. This resulted in an immediate decrease in morbidity and mortality from AIDS. The different classes of ARTs reduce HIV viral load by targeting essential HIV enzymes needed to make new virions at various points in its replication cycle. The main classes of anti-retroviral drugs include the nucleoside reverse transcriptase inhibitors (NRTIs), non-nucleoside reverse transcriptase inhibitors (NNRTIs) and protease inhibitors (PIs). Others include DNA polymerase inhibitors, integrase inhibitors, neuraminidase inhibitors and inhibitors of viral coat disassembly, viral entry inhibitors, immunomodulators and immunoglobulins.

There are two classes of drugs that inhibit the HIV reverse transcriptase. The nucleoside reverse transcriptase inhibitors (NRTIs) inhibit the activity of the reverse transcriptase enzyme by competing with

host substrates needed for pro-viral DNA synthesis. There are several drugs in this group and examples include zidovudine, abacavir, adefovir, lamivudine, didanosine. The non-nucleoside analogues (NNRTIs) are drugs which inactivate the reverse transcriptase enzyme by binding directly to it. Drugs in this group include nevirapine and efavirenz.

Protease inhibitors (PIs) bind to the site of cleavage of the HIV protease enzyme thereby leading to the production of defective or non-infectious virus particles. An example is ritonavir. DNA polymerase inhibitors terminate the nucleotide chain by inhibiting HIV's DNA polymerase. An example of a drug in this class is aciclovir. Likewise, integrase inhibitors target the activity of the HIV enzyme integrase.

The treatment of HIV infection involves a combination of three antiretroviral drugs ideally from at least two different classes. The aim of treatment is threefold: to reduce morbidity, improve survival and reduce transmission of HIV from infected individuals. As effective as ART may be, they are not able to eradicate the virus completely from the host and consequently, most infected individuals will need to remain on life-long therapy. Patients who respond to ART usually experience a rapid reduction in their plasma viral load often to undetectable levels by current assays. This is matched by a concomitant increase in $CD4^+$ T cell population indicating some restoration of immune function and a reduction in opportunistic infections and AIDS [Haase, 1999]. However, despite anti-retrovial therapy, not all HIV-infected individuals see the expected recovery in their $CD4^+$ T cell pool. This is the subject of investigation in chapter 5 where I explored predictors of poor immune reconstitution in children receiving ART.

Antiretroviral drugs do have many unwanted side effects which may impact on compliance especially in children. According to the 2016 WHO guidelines, ART is indicated in all HIV-infected individuals irrespective of CD4 count or age. Priority should be given to children $\leq$ 2 years or children <5 years with severe or advanced HIV clinical disease (WHO stage 3 or 4) or CD4 count $\leq$750cells/mm^3 or CD4 percentage $<$ 25% and children above 5 years with stage 3 or 4 disease or CD4 count <350cells/mm^3 [World Health Organisation, 2016a]. According to the 2016 WHO guidelines, children in the 3-10 years category are recommended to be started on a first line ART combination which includes two NRTIs (abacavir or zidovudine + lamivudine) and one NNRTI (efavirenz or nevirapine). In addition to the backbone of two NRTIs, children below the age of 3 years are recommended lopinavir/ritonavir or a nevirapine [World Health Organisation, 2016a]. According to the 2015 WHO guidelines, all HIV-infected pregnant and breastfeeding women are now advised to receive ART irrespective of disease stage or CD4 count [World Health Organisation, 2017b].

1.8 CD4$^+$ counts and z-scores

Z-scores provide a way to compare outputs from a test to a "normal" population. A z-score represents the number of standard deviations a data point is from the mean. In fitting models to CD4 counts and due to rapid fluctuations in CD4 count data, z-scores provide a useful way to normalise these measurements in children. As such a CD4 count of 2000cells/mm^3 in a 2 year old cannot be directly compared to a count of 2000cells/mm^3 in a 6 year old without some normalization. However, CD4 z-scores can be compared irrespective of age and physiological development: a z-score of 2 indicates that a child has a CD4 count that is 2 standard deviations above the population mean whilst a negative value indicates a CD4 count below the population mean. Hence z-scores tell us where a child's CD4 count sits in comparison to the population average. The formula for calculating z-scores is given below:

$$z_{ij} = \frac{x_{ij} - \bar{x}}{s} \tag{1.1}$$

where z_{ij} represents the z-score for an individual CD4 count, $\bar{x}$ represents the sample mean and s represents the sample standard deviation. In chapter 3, z-scores are used in fitting CD4 T cell data from HIV/Hepatitis C co-infected children.

1.9 Mathematical modelling in paediatric immunology

Mathematical modelling helps to simplify complex immune interactions into a set of equations containing parameters with which predictions and inference can be made about the behaviour of CD4$^+$ T cells. For example, it will not be possible to model the behaviour of each individual T cell trafficking between the systemic circulation and secondary lymphoid organs. However, we can study the overall population dynamics of immune cells such as loss, division and production. Given a number of variables, modelling can be used to augment clinical decision making in predicting long-term CD4$^+$ T cell counts in HIV-infected children. Modelling has also been used in generating hypotheses, quantifying immunological processes such as thymic output and identifying variables to measure given an experimental goal [Castro et al., 2016].

The work undertaken in this thesis relies largely on routine longitudinal laboratory measurements taken from HIV-infected children receiving ART. This type of data can be highly variable and sparse between children. Population modelling allows us to draw as much information in determining average rates of recovery as well as initial and long-term trends in CD4$^+$ T cell reconstitution. Interactions with other variables of interest such as viral load and age at start of therapy can also be investigated.

Immune recovery in children is much more complex with $CD4^+$ T cell concentrations changing rapidly and non-linearly with age on a background of a developing immune system. For this reason, there often isn't enough data to compare children like for like.

Mathematical models have been extensively used in investigating the impact of various co-infections between HIV and other pathogens including malaria, hepatitis C and *mycobacterium tuberculosis*. In a recent article by Birger *et al*, mechanistic models were developed to investigate the impact of HIV co-infection on clearance and sustained viral response in patients receiving treatment for hepatitis C infection [Birger et al., 2015].

One key advantage of modelling is the fact that it enables us to made predictions regarding future outcomes in HIV-infected children much early on. This can have a greater impact in the long-term management of these children who are likely to live with HIV throughout their adult lives. A combination of empirical and mechanistic models have been explored in this work. Mechanistic models have the added advantage of enabling us to capture relevant underlying biology whilst making predictions on long-term trends. They can also be used to obtain estimates of parameters of interest which cannot be obtained through laboratory work (e.g. thymic output).

1.10 Aims of the thesis

The aim of this thesis is to use mathematical models to investigate factors affecting $CD4^+$ T cell recovery in HIV-infected children receiving therapy. In chapter 2, the statistical framework of non-linear mixed-effects regression used in fitting the models is described more formally. In the subsequent chapters, mathematical modelling is used to investigate the impact of hepatitis C co-infection on overall $CD4^+$ T cell recovery (chapter 3), to determine other factors responsible for poor immune reconstitution in HIV-infected children (chapter 4) and also to further understand homeostasis of naive $CD4^+$ T cell subpopulations in HIV-infected children receiving therapy (chapter 5).

Chapter 2

Methods

2.1 Modelling CD4$^+$ T cell recovery using Non-linear mixed-effects regression

Non-linear mixed-effects models are a population-based statistical framework appropriate for analysing longitudinal data and are widely used in population pharmacology where the statistical methodology was originally developed [Sheiner and Beal, 1980, Marie Davidian, David M. Giltinan, 1995]. Under this framework, pharmacological or similar repeated datasets can be modelled as inter- (*between subject*) and intra-subject (*within subject*) variabilities where *between-subject* differences are handled with the fixed effects (and covariates) and *within-subject* variabilities are modelled using random effects. The structure of mixed models is hierarchical where observations between subjects are independent whilst observations from the same patient are dependent. Mixed-effects models can have multiple levels but for the purpose of this work, we shall be considering three levels. The first level (outermost) fits a parametric regression model to the entire population (structural model or *fixed effect*) with unknown subject-specific estimates. At the second level, patient-specific parameters are estimated and assumed to be randomly sampled from normal or log-normal distribution. As such, any variability that is not explained by the fixed effect is incorporated through *random effects*. At the third level of the hierarchy, the model incorporates *residual errors* which account for variability that cannot be explained by either the fixed effect or the random effect.

In mathematical notation, an individual CD4 count, y_{ij} may be described as follows:

$$y_{ij} = f(\mathbf{\Phi_i}, \mathbf{v_{ij}}) + \epsilon_{ij}, \quad i = 1, \ldots, M, \quad j = 1, \ldots, k_i, \tag{2.1}$$

where f is a non-linear function of real values which contains child-specific deterministic parameter vector $\mathbf{\Phi_i}$ and a covariate vector $\mathbf{v_{ij}}$. The variable M represents number of HIV-infected children whilst k_i denotes number of observations k in the ith child. If carefully chosen, the structural and random effect models should be able to describe the observations as much as possible. However, there is the need to

Table 2.1: A summary of parameters for the general mixed-effects model described in equations 2.1 and 2.2

Variables	Definition
Known variables	
f	Non-linear function
A_{ij}	Design matrix for fixed effects
B_{ij}	Design matrix for random effects
Unknown variables	
ϵ_{ij}	residual error
$\mathbf{\Phi_i}$	parameter vector
$\boldsymbol{\eta_{ij}}$	random effects vector
$\mathbf{\Omega}$	variance-covariance matrix
$\mathbf{v_{ij}}$	covariate vector

account for other sources of variability in the data which may not be related to the choice of fixed and random effects. Examples of such variabilities include model mis-specification, measurement errors and random variations and these can be modelled using residual errors denoted by ϵ_{ij} and are assumed to be normally distributed:

$$\epsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$$

As discussed, each child-specific parameter vector $\mathbf{\Phi_i}$ can be decomposed into fixed effects ($\boldsymbol{\theta}$) and random effects ($\boldsymbol{\eta_i}$) using a number of relationships including additive, proportional and exponential. An additive linear relationship between the fixed effects and random effects is described below:

$$\mathbf{\Phi_i} = \mathbf{A_{ij}}\boldsymbol{\theta} + \mathbf{B_{ij}}\boldsymbol{\eta_i} \tag{2.2}$$

where $E(y_{ij}) = \mathbf{A_{ij}}\boldsymbol{\theta}$, $\boldsymbol{\theta}$ is a fixed effects vector and $\boldsymbol{\eta_i}$ is a random effects vector. The random effects vector has a variance-covariance matrix $\mathbf{\Omega}$ and is assumed to be normally distributed and independent of the residual errors.

$$\eta_{\mathbf{i}} \sim \mathcal{N}(0, \mathbf{\Omega}), \quad Cov(\epsilon_{ij}, \boldsymbol{\eta_i}) = 0 \tag{2.3}$$

In the case of a non-linear mixed-effects models, the function f given in equation 2.1 is non-linear in at least one of the parameters contained in the parameter vector Φ given as follows (exponential case):

$$\mathbf{\Phi_i} = \mathbf{A_{ij}}\mathbf{e}^{\boldsymbol{\theta}} + \mathbf{B_{ij}}\mathbf{e}^{\boldsymbol{\eta_i}} \tag{2.4}$$

An example of a non-linear case is given in equation 3.3 where the function used is non-linear in parameter c_i. **A** and **B** are design matrices which associate the $CD4^+$ count to the fixed and random effects. In summary, three hierarchies of effect influence the $CD4^+$ T cell count, y_{ij}: the fixed effect, the random effect and the residual errors. This way, mixed-effects models are able to provide a picture of the $CD4^+$ T cell response to treatment both in the population as a whole, on average, and in each child individually and enables quantification of the differences between children attributable to differences in physiology or environment. Table 2.1 gives a summary of model parameters and their meanings.

2.2 Estimation methods for non-linear mixed-effects models

One of challenges of fitting non-linear mixed-effects models is that in most cases, there are no analytical solutions for the best parameter estimates. In particular, there is no exact solution for the integral that describes the marginal likelihood of $\mathbf{\Phi}$ and $\mathbf{\Omega}$ for the $CD4^+$ count data y_i, $L_i(\Phi, \mathbf{\Omega}|y_i)$ (also marginal probability density, $P(y_i|\Phi, \mathbf{\Omega})$) as given below [Wang, 2007]:

$$P(y_i|\Phi, \mathbf{\Omega}) = \int P(y_i, \eta_i|\Phi, \Omega) d\eta_i \tag{2.5}$$

The objective of most algorithms is to maximize either the likelihood function (or log-likelihood) or the empirical Bayesian posterior probability of the data given the parameters. In an attempt to optimize the likelihood function numerically, a number of approximations have been implemented in various softwares.

Most of the pioneering work in the field of non-linear mixed-effects (NLME) modelling was undertaken by Lewis Sheiner and Stuart Beal within the context of population-based pharmacokinetic and pharmacodynamic analysis [Beal and Sheiner, 1980, Sheiner et al., 1972]. In the late 1970s, they successfully implemented the first order approximation method ("FO") in the fortran-based software, NONMEM [Sheiner and Beal, 1980, Beal et al., 2014]. The FO method approximates the NLME model as a first-order Taylor series expansion of the model function, f about $\eta = 0$. Model parameters are then estimated using a linear approximation to the non-linear model. One limitation of the FO method is that it does not produce inter-individual estimates and hence can give very inaccurate assessments in the presence of large residual errors or inter-subject variability. To overcome this, a first-order conditional estimation method, (FOCE) was implemented in NONMEM. With FOCE, estimates for η_i are computed conditionally on given values for Φ and Ω that maximize the posterior density. The FO and FOCE methods are suitable for estimating parameters from very rich datasets with the FOCE method producing more accurate estimates.

However, FOCE requires more computational time and is not always suitable for very complex random effect models with complex likelihood surfaces: FOCE has the tendency to run into local minima.

Another category of estimation method implemented in NONMEM are the Monte-Carlo expectation maximization (EM) methods. The EM algorithms were originally developed for MLE in problems with missing data [Dempster et al., 1977]. As the name suggests, they estimate parameters in a two stage process namely the expectation step and the maximization steps. Two implementations of this method in NONMEM include the Markov-Chain Monte Carlo (MCMC) stochastic approximation expectation-maximization (SAEM) and the Monte Carlo (MC) importance sampling (IMP) expectation maximization. The SAEM algorithm was developed by Marc Lavielle who first implemented it in the MONOLIX software [Dalyon et al., 1999]. The Monte Carlo algorithms estimate the maximum likelihood in two steps namely the expectation step (Monte Carlo integration) and the maximization step (update of fixed effect parameters) and have been widely used in modelling T cell recovery in adults starting anti-retroviral therapy [Dalyon et al., 1999]. Importance sampling EM algorithm was first implemented in PDx-MCPEM and S-ADAPT by Serge Guzy and Robert Bauer. Although these methods are more computationally intensive, they have the advantage of a greater optimization success than FOCE especially with increasingly complex models which have numerous parameters with random effects and large variance-covariance matrices. They are also very successful at estimating parameters from very sparse datasets and are less likely to run into local minimum problems since they are based on sampling rather than gradient estimation.

FOCE and IMP have the added advantage in that they can be implemented without the maximization step (expectation only step). This is very helpful for calculating the objective function value after SAEM whilst retaining the population estimates.

The third category uses Bayesian estimation by MCMC simulation. Fitting non linear models using maximum likelihood is not without its own challenges. As explained in the preceding paragraphs, statistical inference for NLMEs is made to depend on linear approximation methods and assumptions of normality. Therefore, the uncertainty of parameter estimates and predictions is not always easy to assess under the frequentist approach especially when model assumptions are relaxed and the population is heterogeneous. Although not explored in this work, Bayesian methods allow the addition of priors such as variance covariance matrices which can be useful when applying a fitted model to new datasets. This could allow for incorporation of prior information regarding biological relationships between parameters in modelling $CD4^+$ T cell recovery following therapy. The Bayesian approach can generate full posterior distribution of estimated parameters which makes it easy to calculate various statistics from samples obtained during the

model fitting [Lu and Huang, 2014]. Furthermore, the Bayesian methodology does not impose additional constraints or assumptions of normality on posterior distributions as is the case in the maximum likelihood method. Hence, parameter estimates are allowed to deviate from the standard normal distribution and parameter estimates are more realistic for models fitted to skewed data [Marie Davidian, David M. Giltinan, 1995]. Bayesian approach is therefore less biased compared to NLME maximization methods such as FOCE which use linearization approximation methods. One potential disadvantage of the Bayesian method is in the introduction of errors in the event of wrong priors although it is likely that faster samplers will be implemented for non-linear hierarchical models in the near future. They can also take excessively long run times due to the incorporation of priors. Implementations are available in NONMEM and WinBUGS software [Spiegelhalter et al., 1999].

Model fitting in this work was done through maximum likelihood as opposed to the Bayesian method. Implementations of the maximum likelihood method have been thoroughly and extensively tested in the NONMEM software which is used for model fitting in this work. NONMEM remains a gold standard in the pharmaceutical industry which is also highly regulated by the food and drug administration (FDA). As there was no access to a dedicated computer cluster for this work, the longer run times often encountered in the Bayesian approach made maximum likelihood approach much more practical [Jonsson et al., 2007]. Although the Bayesian method has been shown to be more robust, for problems of sufficient data, the sample mean parameters are similar to maximum likelihood values [Robert J Bauer, Brian Sadler, 2015]. Discrimination between models in the Bayesian approach is not always straight forward and there are several limitations to standard diagnostic plots in the case of hierarchical non-linear models [Yuan and Johnson, 2012].

The work undertaken in this thesis does not involve the use of non-parametric models such as quantile regression methods and splines fitted in the context of linear mixed effect models. Non-parametric mixed effect models make use of non-parametric fixed-effects and random effect functions as well as a measurement error process in describing longitudinal data. The non-parametric components are usually fitted using various smoothing techniques including smoothing spline, regression spline, penalized spline and local polynomial methods [Ke and Wang, 2001, Wu and Zhang, 2002, Keonker, 2004, Wu and Zhang, 2006]. Some advantages of non-parametric methods include flexibility in fitting to repeated data and robustness against model misspecification. Non-parametric models may also be useful in cases where a structural model for a particular longitudinal dataset is unknown and where data does not fit the assumption of normality. Also, non-parametric methods such as the quantile regression are able to extend the regression

for the mean (typically used in mixed effects models) to a full range of conditional distribution of the dependent variable [Geraci, 2014]. However, these methods can be computationally intensive and can only handle very few covariates.

2.2.1 Objective Function Value (OFV)

In fitting non-linear mixed-effects models using NONMEM, the user is required to specify an appropriate mathematical model, random effects model and a residual error model. Suitable starting estimates are needed for all the above methods and can influence the time taken to maximize the likelihood. In NONMEM, maximum likelihood, L, is achieved by minimizing the objective function value (OFV) which is given as $-2\log(L)$. In this context, "likelihood" is defined as the probability of obtaining the observations given the parameter estimates. Given that Y is a measured observation, $\hat{Y}$ is the predicted observation from the model and σ^2 is the variance of the model, the likelihood of this observation given the model is defined by (assuming normality):

$$L = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(Y-\hat{Y})^2} \tag{2.6}$$

The joint probability of obtaining n observations given the model parameters is given as a product of individual probabilities:

$$L = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\frac{1}{2\sigma_i^2}(Y_i-\hat{Y}_i)^2} \tag{2.7}$$

To avoid rounding off errors, it is easier to sum the above product by taking the logarithms:

$$\log(L) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\sum_{i=1}^{n}\left(\log(\sigma_i^2) + \frac{(Y_i-\hat{Y}_i)^2}{\sigma_i^2}\right) \tag{2.8}$$

Multiplying both sides of 2.8 by -2 gives us $-2\log(\text{likelihood})$ which is the objective function value:

$$\text{OFV} = -2\log(L) = n\log(2\pi) + \sum_{i=1}^{n}\left(\log(\sigma_i^2) + \frac{(Y_i-\hat{Y}_i)^2}{\sigma_i^2}\right) \tag{2.9}$$

Hence, to maximize likelihood, we minimize the objective function value (OFV) which is given as $-2\log(\text{likelihood})$. Since $n\log(2\pi)$ is a constant term, most of this minimization task is focused on the sum in equation 2.9. This sum is sometimes referred to as the "extended least squares" objective function. Hence, model fits were compared using the OFVs, a goodness-of-fit measure related to log(likelihood).

If we have one model nested within another, the significance of the additional parameter(s) in the larger model can be tested by considering the ratio of their likelihoods, L. The difference between their OFVs (denoted ΔOFV later in this thesis) follows approximately a χ^2 distribution, where a change in OFV of 3.84 corresponds to a *p-value* < 0.05 for one degree of freedom. This is the likelihood ratio test (LRT) which forms the basis of the stepwise covariate selection in the PsN software ("Perl Speaks NONMEM") [Jonsson and Karlsson, 1998, Lindbom et al., 2005]. The NLME algorithm is also implemented in other softwares including R ("nlme" and "lme4" packages) [Pinheiro et al., 2013], SAS [SAS Institute, 2017] and Phoenix [Phoenix WinNonlin Certara, 2016].

2.3 Estimation of the variance-covariance matrix (Ω)

In NONMEM, the default method used in estimating variance-covariance matrix is the sandwich matrix computation method ($\mathbf{R}^{-1}\mathbf{S}\mathbf{R}^{-1}$). This is the method used in estimating all variance-covariance matrices in this work. The $\mathbf{R}$ matrix is the Hessian matrix (matrix of second derivatives) of objective function which is evaluated at the final estimate of the model parameters $(\theta, \Omega, \epsilon)$. The $\mathbf{S}$ matrix represents the sum of matrices, S_j, one matrix for each child. The individual matrices, S_j, are each equivalent to $\nabla_j \nabla_j'$ where ∇_j is the gradient vector of the contribution to the OFV from the jth child, computed at the final estimate of the model parameter. Assuming random effects are normally distributed,

$$\lim_{\mathrm{M}\to\infty} \frac{\mathbf{R}}{\mathrm{M}} = \lim_{\mathrm{M}\to\infty} \frac{\mathbf{S}}{\mathrm{M}} = \mathrm{X} \tag{2.10}$$

where M is the total number of children and X is a matrix. Under normality assumptions of random effects, either $\mathbf{R}^{-1}$ or $\mathbf{S}^{-1}$ estimates the true covariance matrix. However, when this assumption of normality is not imposed, the matrix $\mathbf{R}^{-1}\mathbf{S}\mathbf{R}^{-1}$ estimates the true covariance matrix (default estimate in NONMEM). The $\mathbf{R}$ matrix does not always have a positive eigenvalue. For Monte Carlo assessed information matrices, the NONMEM software subjects the matrix through a positive definiteness filter that makes minor adjustments to the eigenvalues of $\mathbf{R}$ if necessary. Formal mathematical derivations of the above methods and much more are available in the NONMEM 7 technical guide by Robert Bauer from the ICON plus development team (Senior Director, Pharmacometrics Research and Development, ICON Early Phase) [Beal and Sheiner, 1980, Beal and Sheiner, 1988, Bauer, 2013].

2.4 Covariate modelling

In non-linear mixed-effects modelling, a covariate is an individual-specific variable capable of explaining some of the variability in a parameter value which would otherwise be attributed to random effects or residual errors. Covariates which can either be continuous or categorical can be incorporated within the mixed effect framework in a nested fashion and their significance tested using the likelihood ratio test. To this end, the PsN software was designed to be used alongside NONMEM for automating covariate analysis. It starts with a base model which is pre-specified by the user and works by selecting one covariate at a time in a stepwise covariate model building process (SCM). Initially, all covariates are added one at a time for each parameter in a forward process and the covariate which causes the largest improvement in the overall model fit (lowest *p-value*) is selected using the likelihood ratio test (LRT) and provided they meet the *p-value* criterion of $p < p_f$ (where p_f is typically in the region of 0.05). This process is carried out for all the remaining covariates and for all parameters. The full covariate model from the forward process is then subjected to a backward process where each of the selected covariates is omitted one at a time and significance tested using the LRT. Of all the omitted covariate relations that satisfy a *p-value* criterion of $p > p_b$, the covariate associated with the largest *p-value* is excluded. The value of p_b (typically 0.01) is usually lower that p_f. This way, the backward process eliminates all covariate relationships that do not improve the model significantly using the specified *p-value* criterion. A number of relationships between parameters and covariates are tested (where appropriate) including linear, piece-wise linear, exponential and power. If a linear relationship is significant, a non-linear covariate relationship is also considered.

2.5 Diagnostic Plots

Diagnostic plots are used for assessing the overall model fit to data and the validity of the statistical and distributional assumptions upon which mixed effect models are based [Karlsson and Savic, 2007]. As such, diagnostic plots can give information about model bias, mis-specification and overall predictive capability.

Residual Plots

Residuals are the differences between the individual observed values and predicted values for each measurement. Normalisation of residuals can be obtained by dividing the residuals by the standard deviation obtained from a particular individual's data (weighted residuals). These weighted residuals are estimated from a linear Taylor series approximation that is conditioned around the average or post-hoc

patient-specific empirical Bayes estimates of the between-subject random effects estimates. Hence, the standard deviation used in the normalisation is dependent on the FOCE estimation method, generating conditional weighted residuals (CWRES) [Hooker et al., 2007] given by the following equation:

$$\text{CWRES} = \frac{y_i - E_{\text{FOCE}}[y_i]}{\sqrt{\mathbf{Cov}_{\text{FOCE}}(y_i)}} \in \mathcal{N}(0,1) \tag{2.11}$$

where y_i is the given CD4$^+$ count for individual i, $E_{\text{FOCE}}[y_i]$ represents individual model predictions for individual i and $\mathbf{Cov}_{\text{FOCE}}(y_i)$ represents the variance of the between-subject random effects. As given in equation 2.11, the CWRES should follow the standard normal distribution and 95% of the residuals should be within two standard deviations of mean, 0. Since CWRES are assumed to be independent, they can be plotted against time, population-level predictions or other factors to evaluate model biases with any of these variables.

Population-level and individual predictions

Plots of individual predictions (IPRED in NONMEM) against observed data (DV in NONMEM) can be useful in detecting any model mis-specifications especially at the extreme regions of the observed data. In good models, IPREDs usually have a good agreement with the observed values. They can also be used in checking for non-normality and systematic under or over-prediction of CD4 counts.

Smoothers

Throughout this work, smoothers fitted through data are loess (locally weighted polynomial regression) unless stated otherwise [Cleveland, 1979, Cleveland and Devlin, 1988].

Visual predictive checks

The visual predictive check (VPC) is a simulation based method which assumes that simulated datasets from the model should agree with the original data from which the model was developed [Holford, 2005, Karlsson and Holford, 2008]. To perform a VPC, a large number of datasets are simulated (typically 300-1000) based on the model parameters. Random variability is incorporated into the dataset using the variance-covariance matrix of the random effects from the final model. The result of VPCs are best displayed as a plot showing percentiles of the original dataset overlaid by prediction intervals from the simulated dataset.

Shrinkage

Shrinkage is a term used to describe deviation of estimated individual level parameters (random effects) from the population mean. A dataset that is uninformative about a parameter will produce high shrinkage on the variable indicating that the individual level random effects are shrinking towards the population average. Mathematically, shrinkage (sh) is defined by the following equation:

$$sh = 1 - \frac{\sigma(\hat{\eta_{EBE}})}{\hat{\omega}}, \tag{2.12}$$

where $\hat{\eta_{EBE}}$ are estimated random effects and $\hat{\omega}$ is the estimated standard deviation for the random effect [Xu et al., 2012]. NONMEM produces shrinkage values on all estimated parameters and shrinkage above 20% is considered high and could indicate that the model needs to be re-formulated, re-parameterised or a richer dataset needs to be sought after.

Chapter 3

Mathematical modelling of $CD4^+$ T cell recovery in HIV/Hepatitis C co-infected children receiving anti-retroviral therapy

3.1 Introduction

Co-infections with hepatitis B, C, cytomegalovirus and *Mycobacterium tuberculosis* play a major role in the clinical management of HIV-infected patients. In this chapter, a mathematical modelling approach is used to explore the impact of hepatitis C co-infection on immune recovery in HIV-infected children receiving anti-retroviral therapy.

3.1.1 Epidemiology of hepatitis C and HIV co-infection

The introduction of highly active antiretroviral drugs in the 1980s/90s resulted in a significant reduction in the overall burden of AIDS-related complications amongst HIV-infected individuals. However, within the HIV-infected population, it has also indirectly led to an increased prevalence and burden of secondary infections such as hepatitis C. An estimated 71 million individuals worldwide are thought to be chronically infected with hepatitis C virus (HCV) [World Health Organisation, 2017a]. Of these, 2.3 million are estimated to be co-infected with HIV [Platt et al., 2016] and 1 million individuals under 18 years are thought to be mono-infected with chronic HCV [Yeung et al., 2001].

In Western Europe, about 150-300 children are born annually to women known to be co-infected with hepatitis C virus (HCV) and HIV [England et al., 2006, Majekodunmi et al., 2017]. The global prevalence of paediatric HCV infection varies geographically from 0.05% in the western world to 5.8% in resource-limited settings including Egypt and sub-Saharan Africa [El-Shabrawi and Kamal, 2013]. Most children with HCV are vertically infected and an estimated 60,000 HCV-infected infants are born yearly to HCV-infected mothers [Bortolotti et al., 1998]. The risk of mother-to-child transmission for HCV seropositive, RNA-positive women is about 4-8%, increasing substantially to 10.8-25% in HIV-co-infected

mothers [Thomas et al., 1998, Benova et al., 2014, Mast et al., 2005, Selvapatt et al., 2015, Floreani, 2013]. Although highly debated and only demonstrated in a few studies, additional factors contributing to increased risk of transmission may include high hepatitis C viral load and the presence of HCV in maternal peripheral mononuclear cells [Indolfi et al., 2013, Durmaz, 2012]. In addition, HIV/HCV co-infected mothers receiving combination ART are less likely to transmit HCV to their unborn children [Conte et al., 2000, Snijdewind et al., 2015]. The incidence of spontaneous HCV clearance in children varies from 7.5% to 25% depending on virus genotype [Bortolotti et al., 2008, Yeung et al., 2007, European Paediatric Hepatitis C Virus Network, 2005]. However, in the presence of HIV/HCV co-infection the likelihood of spontaneous clearance is much reduced in children [Indolfi et al., 2015] as in adults [Villano et al., 1999]. Evidence from the literature and other cohort studies suggests that HIV/HCV co-infected patients have an increased risk of developing cirrhotic liver disease at a much early stage when compared to HCV mono-infected patients [Hernandez and Sherman, 2011]. A meta-analysis revealed that HIV/HCV co-infected patients were twice more likely to develop cirrhosis and 6 times more likely to develop decompensated liver disease when compared to patients only infected with HCV [Tuyama et al., 2010, Thein et al., 2008]. Whilst these complications are not common in childhood, infected children become exposed to similar risks as they become adults.

Although HCV-related complications can take up to 20 years to manifest clinically, such secondary infections are increasingly being diagnosed due to improved HIV care [England et al., 2006]. Alongside cardiovascular disease, hepatitis C is now a leading cause of morbidity and mortality in co-infected patients receiving antiretroviral therapy [Reiberger et al., 2011]. Over a period of 20-30 yrs, about 20% of adults with chronic HCV will go on to develop cirrhosis of the liver [Monga et al., 2001]. Co-infected individuals are reported to have an accelerated progression to liver fibrosis and higher HCV RNA levels, and to be at greater risk of developing hepatocellular carcinoma than those infected with HCV alone [Sulkowski et al., 2007]. A Swiss study reported hepatocellular carcinoma (HCC) as the most prevalent non-AIDS malignancy in the HIV-infected population [Hasse et al., 2011]. The underlying mechanisms driving these complications are poorly understood and may include impaired T lymphocyte response and HIV-mediated hepatotoxicity [Tuyama et al., 2010, Kim and Chung, 2009].

3.1.2 Pathogenesis and treatment of hepatitis C mono-infection in children

Hepatitis C is a single stranded RNA virus of the *Flaviviridae* family and has 6 identified genotypes and more than 100 subtypes [Durmaz, 2012, Hochman and Balistreri, 2007]. Genotype 1 has the highest

global prevalence whilst genotype 4 is exclusively found in Egypt, Middle East and sub-Saharan Africa. Genotypes 5 and 6 are prevalent in South Africa and Asia respectively whilst genotype 3a is common amongst intravenous drug users within Europe [Esteban et al., 2008]. Hepatitis C virus is capable of causing both acute and chronic hepatitis. Acute HCV infection is relatively uncommon in childhood whilst the chronic form is clinically asymptomatic and rarely causes severe liver disease. The vast majority of children are infected through the vertical route and remain completely asymptomatic in the first 4 years of life. Following exposure, spontaneous resolution is much more likely in comparison to adult patients. In a European study involving 266 vertically infected children, about 20% cleared the virus spontaneously whilst the rest became chronically infected. Despite a transmission rate of 14-17% in children whose mothers are chronically infected with HCV, spontaneous resolution is commonly seen at a median age of 15 months. This spontaneous clearance has been shown to be associated with genotype *IL28B* and all other genotypes excluding genotype 1 [Durmaz, 2012].

In the long-term, untreated chronic hepatitis C can eventually cause cirrhosis, liver failure and hepatocellular carcinoma. About 5% of untreated children with chronic HCV infection will go on to develop hepatic fibrosis and cirrhosis in adulthood [Abdel-Hady et al., 2014]. Hepatic cirrhosis is a premalignant condition with 1-5% of HCV-infected adults going on to develop hepatocellular carcinoma [Tosone et al., 2014, Perz et al., 2006].

Following an acute episode, antibodies against HCV are produced which continue to circulate throughout life. Diagnosis of HCV infection in children is through the detection of these circulating anti-HCV antibodies as well as HCV-RNA. The detection of anti-HCV antibodies in a child only indicates a prior exposure to HCV and is not a confirmation of active viral replication. In order to confirm an active infection, all children with positive anti-HCV antibody are tested for the presence of HCV-RNA. The HCV-RNA assay can also be used for monitoring treatment efficacy as well as the viral genotype. In vertically co-infected children, screening is delayed until after the first 12 months of life in order to allow for elimination of maternal antibodies [Durmaz, 2012]. Chronic HCV infection can be confirmed through detection of HCV-RNA in the serum (Figure 3.1) [Gerlach et al., 2003, Thomson et al., 2011].

In treating paediatric HCV infection, the objective of therapy is to achieve a sustained virological response (SVR) which is defined as a persistently negative HCV-RNA 6 months after completing the treatment regime [Karnsakul et al., 2009]. A sustained virological response is defined as complete viral eradication where the virus is not detectable even within hepatocytes. This can be confirmed by absence of HCV-RNA in hepatic or mononuclear blood cell samples [Marcellin et al., 1997, Marinho et al., 2014]. The

standard therapy in children aged 3-17 years consists of a combination of pegylated interferon (peg-IFNα-2a or b) with ribavirin for 6-12 months [Wirth, 2012]. However since most children are asymptomatic carriers, active treatment is often difficult to justify. Furthermore, there are no anti-HCV treatment recommendations for children under the age of three. In a UK study involving 75 children with chronic HCV infection, a treatment course of pegylated interferon-α 2a or b (peg-IFN) and ribavirin for 24-48 weeks was reported to have achieved a response rate of 65%, 89% , 89% and 100% in genotypes 1, 2, 3 and 4 respectively [Abdel-Hady et al., 2014]. Overall sustained viral response rate was 76% and children with *IL-28B* polymorphisms had a better response to therapy. The efficacy of ribavirin and peg-IFN tends to be genotype-dependent with genotypes 2 and 3 reported to have the highest treatment success [Wirth, 2012]. There is currently no vaccine against hepatitis C infection. This is partly due to the diversity of strains that emerge from viral mutation which poses a barrier to the development effective vaccines.

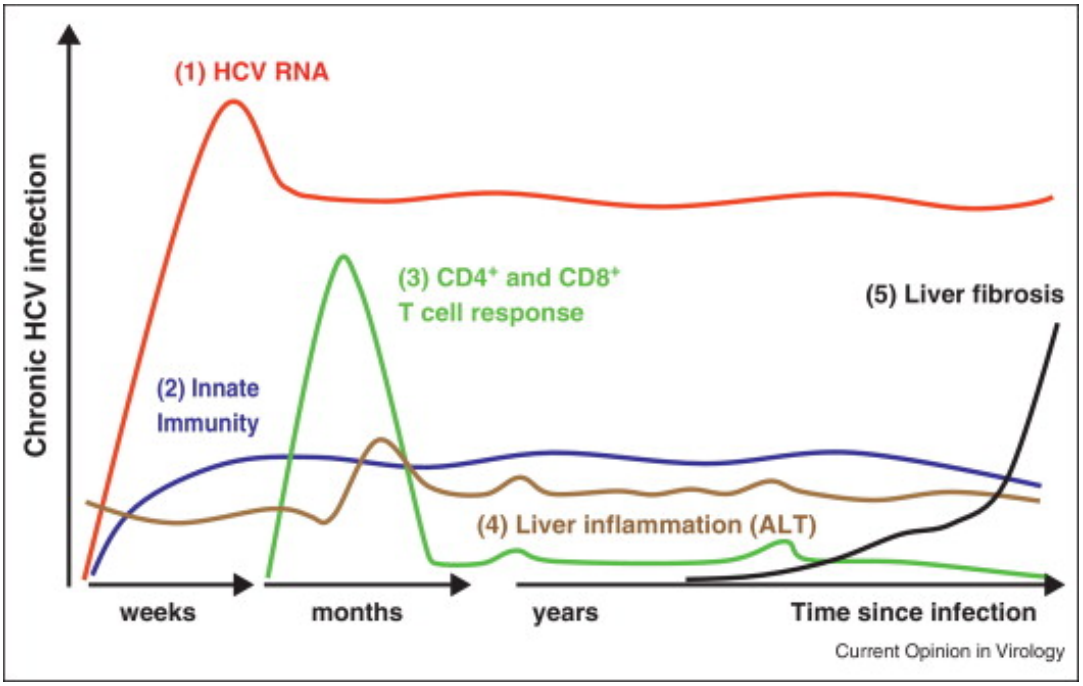


Figure 3.1: Immune response to chronic hepatitis C virus. Figure was taken from [Claassen et al., 2013]

3.1.3 Immune response to hepatitis C infection

Hepatitis C virus mainly infects hepatocytes and following initial exposure, the vast majority of infected individuals (80%) are unable to clear the virus. Following exposure to HCV, the innate immune system is promptly activated leading to the secretion of type 1 interferons and cytokines in the liver. Production of type 1 interferon is usually preceded by activation of toll-like receptor 3 (TLR-3) and retinoid acid-inducible

gene-1 (RIG-1). These two signals induce the release of IFN-β followed by IFN-α [Terilli and Cox, 2013]. Interferons are thought to create an antiviral environment within the liver which facilitates clearance of the virus and collectively, it is these properties that are exploited in the use of pegylated-interferon in treating HCV. In addition, natural killer cells (NK cells) and natural killer T cells (NKT cells) are also activated and recruited to infected hepatocytes where they produce IFN-γ and induce direct cytotoxic killing of infected cells [Jang et al., 2013, Kokordelis et al., 2014].

Due to the early development of virus escape mutants, the innate immune response directed against HCV becomes ineffective at suppressing viral replication [Dowd et al., 2009]. Consequently, there is a rapid rise in HCV viral load a few weeks following infection (Figure 3.1). Furthermore, HCV has developed mechanisms for inhibiting the activation of both TLR signalling and RIG-1 [Gale and Foy, 2005]. As a result, about 10% of hepatocytes are infected before T cell-mediated immune response can be detected at around week 8-12 following initial infection [Claassen et al., 2013]. This delayed cell-mediated immunity against HCV has been attributed to a deficiency in the priming of naive B and T cells [Rehermann, 2009]. The importance of T cell immune response is further corroborated by studies in chimpanzees demonstrating that *in-vivo* depletion of CD4$^+$ or CD8$^+$ T cells disrupts HCV clearance from infected hosts. In particular, it has been shown that robust CD8$^+$ T cell response with diverse epitope-specificity boosted by CD4$^+$ T cells is important for spontaneous resolution of acute hepatitis C infection whilst a weak immune response is linked to the acquisition of chronic HCV infection [Thimme et al., 2002, Lechner et al., 2000, Shin et al., 2013, Shin et al., 2016]. Thus, patients who manage to spontaneously clear the infection in the first 4 weeks of exposure usually have a much more sustained CD8$^+$ T cell response [Mehta et al., 2002, Rehermann and Nascimbeni, 2005]. On the other hand, incomplete suppression of HCV replication by cytotoxic lymphocytes leads to the emergence of viral escape mutants and hence latency of HCV [Terilli and Cox, 2013]. Furthermore, patients with viral persistence have been observed to have a decreased CD4$^+$ T cell division and reduced IL-2 production [Semmo et al., 2005, Schulze Zur Wiesch et al., 2012]. Hence, the lack of CD4$^+$ T cell immunity can result in a dampened HCV-specific cytotoxic lymphocyte response leading to a chronically infected state [Thimme et al., 2002, Thimme et al., 2001].

Dysfunctional HCV-specific T cell response enables the virus to persist within hepatocytes for several years without any symptoms [Schulze Zur Wiesch et al., 2012]. Unlike HIV, the dysfunctional T cell response seen in HCV infection is only restricted to HCV-specific T cells leaving immunity against other pathogens preserved [Claassen et al., 2012]. Two mechanisms responsible for the T cell dysfunction include T cell exhaustion from persistent antigenic stimulation and suppression by regulatory T cells [Rehermann, 2009,

Dustin and Rice, 2007, Klenerman and Semmo, 2006]. Evidence from chronic lymphocytic choriomeningitis virus (LCMV) suggests that exhausted T cells express inhibitory receptors such as PD-1 making them functionally impaired [Penna et al., 2007].

3.2 Objectives

In this study, I aimed to improve understanding of the impact of HCV co-infection on CD4$^+$ T cell recovery in children receiving ART. By using a larger study population than has been available, and analysis methods specifically suited to longitudinal CD4$^+$ T cell measurements in children [ARROW Trial team et al., 2013, Paediatric European Network for Treatment of AIDS (PENTA), 2002], I aim to present a fuller picture of CD4 recovery in HIV/HCV co-infected children.

3.3 Methods

3.3.1 Data sources and eligibility

The dataset used in this analysis came from the European Pregnancy and Paediatric HIV Cohort Collaboration (EPPICC). Data from mono-infected children came from the Ukraine Paediatric HIV Cohort Study. This study was established in January 2011 and enrolls HIV-infected infants, children, and adolescents (from birth up to 18 years) being cared for in six HIV/AIDS Centres in Ukraine. It collects anonymized demographic, clinical and laboratory data on children according to a standard protocol, with informed consent and is a member of EPPICC.

Data from co-infected children came from a study of HIV/HCV co-infection within EPPICC, in eight countries across Europe, including Ukraine. In the EPPICC HIV/HCV co-infection study, children aged $\geq$ 18 months, adolescents and young adults aged $\leq$ 25 years were eligible for inclusion if they were infected with HIV and with chronic HCV acquired vertically or in childhood, irrespective of acquisition route. Anonymized individual-patient data were collected according to a standard protocol using a modified HIV Cohort Data Exchange Protocol [Turkova et al., 2015].

Children in either cohort who had fewer than two CD4 measurements were excluded from the analysis. Children with known spontaneous viral clearance of HCV (i.e. disappearance of HCV RNA in $\geq$ 2 consecutive serum samples taken 6 months apart) were also excluded.

3.3.2 Definitions

HCV-infected children were identified by detection of positive HCV antibody and/or ≥ 2 positive HCV RNA detected on two separate clinic visits at least 3 months apart. HIV infection in children was defined as detection of HIV antibody and/or positive HIV RNA or DNA PCR in a minimum of two samples obtained on separate visits. "Co-infected children" in this study refers to HIV/HCV co-infected children whilst "mono-infected children" refers to HIV mono-infected children. Our threshold for detection of HIV viral load was set to 50 copies/ml.

3.3.3 Ethical approval

Each participating cohort within EPPICC followed local ethical guidelines. The Ukraine Paediatric HIV Cohort has approval from the UCL Research Ethics Committee and local institutional review boards.

3.3.4 Age-adjustment of CD4 counts using z-scores

It is well known that healthy children usually demonstrate changes in their CD4 count as they approach adulthood [Hulstaert et al., 1994]. This is partly due to rapid changes in thymic activity. One difficulty posed by these age-associated changes is that a direct comparison of raw CD4 counts between children of different age groups cannot be made. One way to overcome this is by considering the proportion of T lymphocytes which are $CD4^+$ ("CD4 percentages"). However, despite the acceptance and wide usage of CD4 percentages across the literature [De Beaudrap et al., 2008, Kekitiinwa et al., 2008, Puthanakit et al., 2009] and in comparison to raw CD4 counts, they have been shown to be a weaker prognostic marker of disease progression in HIV-infected individuals [HIV Paediatric Prognostic Markers Collaborative Study et al., 2010].

Another way of overcoming age effects when modelling CD4 counts is by using CD4 z-scores. A child's z-score represents their CD4 rank relative to an age-matched population of healthy children and described using the standard normal distribution. The formula used for calculating the individual z-scores was derived by implementing the LMS method on data obtained from a population of HIV-negative children born to HIV-positive European mothers [Wade and Ades, 1994, Wade and Ades, 1998]. To facilitate model fitting to the dataset, I have chosen to normalise the raw CD4 count data into CD4 z-scores. A CD4 z-score of 0 implies that a child has a normal CD4 count for their age whilst scores of ± 2 indicate that a child is on the 97.7% and 2.3% centiles respectively of the expected CD4 lymphocyte count for their age.

One drawback of using CD4 z-scores is their relative instability at very low CD4 counts [HIV Paediatric Prognostic Markers Collaborative Study et al., 2010]. For instance, a very young child would not be able to achieve a very low z-score in keeping with a very low CD4 count. Furthermore, the lowest attainable z-score (i.e. CD4 count of 0) for any child is dependent on their age (-4.31 at 4 weeks, -9.34 at age 1yr and -77 at age 3 yrs). In addition to this, at very low CD4 counts, the z-score function becomes very sensitive to small changes in age and is less predictable. Due to the unpredictable behaviour of CD4 z-scores at low CD4 counts, a total of 28 CD4 z-scores below -12 from 24 children were excluded from this analysis. This threshold was taken from extensive work done by Joanna Lewis in her doctoral thesis where detailed sensitivity analysis was conducted on CD4 z-scores in a similar population of European children [Joanna Lewis, 2012].

3.3.5 Mixed-effects modelling

Lewis et al [Lewis et al., 2012] developed a mathematical model to describe the recovery of HIV-infected children's CD4 z-scores with time on ART. Within the mixed-effects framework, it is assumed that recovery of CD4 z-scores follows an asymptotic pattern as illustrated in Figure 3.2 [Lewis et al., 2012]. In this model, the CD4 z-scores start at a below-healthy initial value, int_i and following ART initiation, the z-score increases, tending in the long-term to a higher, stable level, asy_i. This pattern of CD4$^+$ T cell recovery is described by the following equation:

$$z_{ij} = asy_i - (asy_i - int_i)e^{-c_i t_{ij}} + \epsilon_{ij} \tag{3.1}$$

where z_{ij} represents the function f given in equation 2.1 and int_i, asy_i and c_i are the child-specific parameters, $\mathbf{\Phi_i}$ (chapter 2). The function z_{ij} denotes CD4 z-score for child i at time t_{ij} after commencing ART, asy_i represents the steady state value of CD4 z-score for child i, int_i represents CD4 z-score at start of ART for child i (pre-ART CD4 z score) and c_i is a parameter that describes the rate of increase in CD4 z-score for child i, $\frac{\ln 2}{c_i}$ being the time taken ($t_{1/2}$) to achieve half of the total recovery from int_i to asy_i. The term ϵ_{ij} is the "residual error" which represents measurement error, random variation and model mis-specification leading to differences between the recorded data and the form of the curve in Figure 3.2. In their work, this approach provided a good description of CD4$^+$ T cell dynamics in children starting ART and variations of this model have also been used elsewhere to predict immune recovery in HIV-infected individuals on ART [Li et al., 2011]. The asymptotic model is very useful for describing reconstitution because all three parameters are biological meaningful. Furthermore, the presence of an

exponential function in equation 3.1 suggests an underlying deterministic process which could be modelled mechanistically using differential equations.

For each child, the parameters are modelled as a combination of a population-average value ("fixed effect") and a child-specific deviation from this average ("random effect"):

$$int_i = int_i^f + int_i^r, \qquad asy_i = asy_i^f + asy_i^r, \qquad c_i = c_i^f e^{c_i^r} \tag{3.2}$$

where superscripts f and r represent fixed and random effects, respectively. In line with the notation used in section 2.1, the fixed effects are modelled as a linear combination of covariate effects:

$$\left(int_i^f, asy_i^f, \ln(c_i^f)\right) = \boldsymbol{A_i\theta} \tag{3.3}$$

where $\boldsymbol{\theta}$ is a vector of parameters to be inferred. $\boldsymbol{A_i}$ is a design matrix which links the observations to the fixed effects. The vector of random effects, $\boldsymbol{u_i} = (int_i^r, asy_i^r, c_i^r)$ has a multivariate normal distribution with a mean of 0 and independent of the residual errors (ϵ_{ij}).

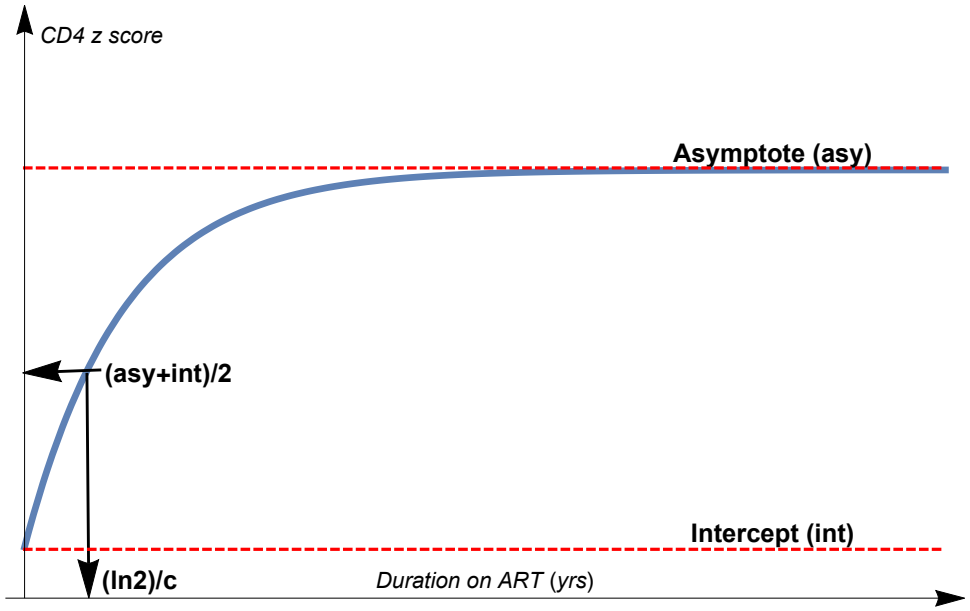


Figure 3.2: The above diagram illustrates a typical recovery profile of CD4 T cells following commencement of ART in HIV-infected children (adapted from [Lewis et al., 2012])

3.3.6 Software and algorithms

Initial attempts at fitting the asymptotic model using the "nlme" package in R proved to be very challenging. Convergence was often very difficult to achieve with random effects on all three parameters, hence, the implementation was done in the NONMEM software. All mixed-effects model fitting was by maximum likelihood using the FOCE algorithm [Beal et al., 2014]. Data analysis and manipulation was done in the R language [R Core Team, 2015] whilst predictions generated from the model were plotted in R and Wolfram Mathematica [Wolfram Research, Inc., 2015]. Stepwise covariate analysis and visual predictive checks were performed in PsN (Perl speaks NONMEM) [Lindbom et al., 2005].

3.3.7 Covariate analysis

The aim of the statistical analysis was to construct a multivariate model around the fixed effects components *int*, *asy* and *c*. I used a forward and backward stepwise selection with exit *p-values* of 0.05 and 0.01 respectively (likelihood ratio test) to investigate potential effects on CD4 recovery of: age and AIDS status at start of ART, gender, HCV status, pre-ART HIV viral load and EPPICC cohort. Random effects were added to all three parameters and a correlation structure was specified between asy_i and int_i.

3.4 Results

Patient characteristics

Characteristics of the study population are described in table 3.1. A total of 355 HIV mono-infected and 46 co-infected children were included for analysis. Median age at start of ART was 3.1 years (interquartile range (IQR): 1.31-5.65 years) for the co-infected group and 4.4 years (IQR: 1.73-7.07 years) for the mono-infected group. Median year of birth was 2004 (IQR: 2001-2006; range: 1989-2011).

The mono-infected children were all from Ukraine whilst the co-infected children were from eight countries across Europe including Ukraine. The median number of CD4 measurements available per child was five for the entire population (IQR: 3-7) as well as for the mono-infected children (IQR: 4-7) and four for the co-infected children (IQR: 3-6). Median follow up period on ART was 4.2 years for the entire population (IQR: 2.7-5.3 years), 4.1 years for the mono-infected children (IQR: 2.7-5.2 years) and 5.1 years for the co-infected children (IQR: 3.1-5.6 years). Figure 3.3 shows individual plots of CD4 concentration and HIV viral load for each child with an overall trend fitted through the data using a local regression line.

Table 3.1: Characteristics of study population for the EPPICC mono-infected and co-infected cohort

Characteristics	HIV^+	HIV^+/HCV^+
Number of children	355	46
Gender		
Male	171 (48.2%)	20 (43.5%)
Female	184 (51.8%)	25 (54.3%)
CD4 counts per child		
Mean (IQR)	6 (4-7)	6 (3-6)
Age at start of ART (yrs)		
Mean (IQR)	4.9 (1.7-7.1)	3.8 (1.3-5.7)
HIV transmission route		
Vertical	349 (98.3%)	45 (97.8%)
Unknown	3 (0.85%)	-
HCV transmission route		
Vertical	-	40 (87%)
Transfusion	-	6 (13%)
Unknown	3 (0.85%)	1 (2.2%)
AIDS status		
Yes	145 (42.4%)	6 (13%)
No	197 (57.6%)	39 (84.8%)
Pre-ART HIV viral load		
$<$ 50 copies/ml	11 (3.1%)	-
$\geq$ 50 copies/ml	203 (57.2%)	24 (52.2%)
Unknown	141 (39.7%)	22 (47.8%)
Median logvl	10.6 (7.49-12.93)	12.8 (11.65-13.61)
Latest HIV viral load		
$<$ 50 copies/ml	204 (57.5%)	39 (84.8%)
$\geq$ 50 copies/ml	118 (33.2%)	4 (8.7%)
Unknown	33 (9.3%)	3 (6.5%)
European centres		
Spain	-	2 (4.35%)
Germany	-	1 (2%)
Poland	-	1 (2%)
Russia	-	17 (37%)
Switzerland	-	3 (7%)
UK	-	2 (4.35%)
Ukraine	355 (100%)	18 (39%)
Italy	-	2 (4.35%)
HCV genotype		
Genotype 1	-	19 (41.3%)
Genotype 2	-	2 (4.4%)
Genotype 3	-	12 (26.1%)
Genotype 4	-	3 (6.5%)
Unknown	-	10 (21.7%)

Undetectable HIV viral load is defined as viral load $<$ 50 copies/ml. The values of log viral load are median values (interquartile range in parenthesis), logvl: log viral load, HIV^+: HIV mono-infected, HIV^+/HCV^+: HIV/HCV co-infected

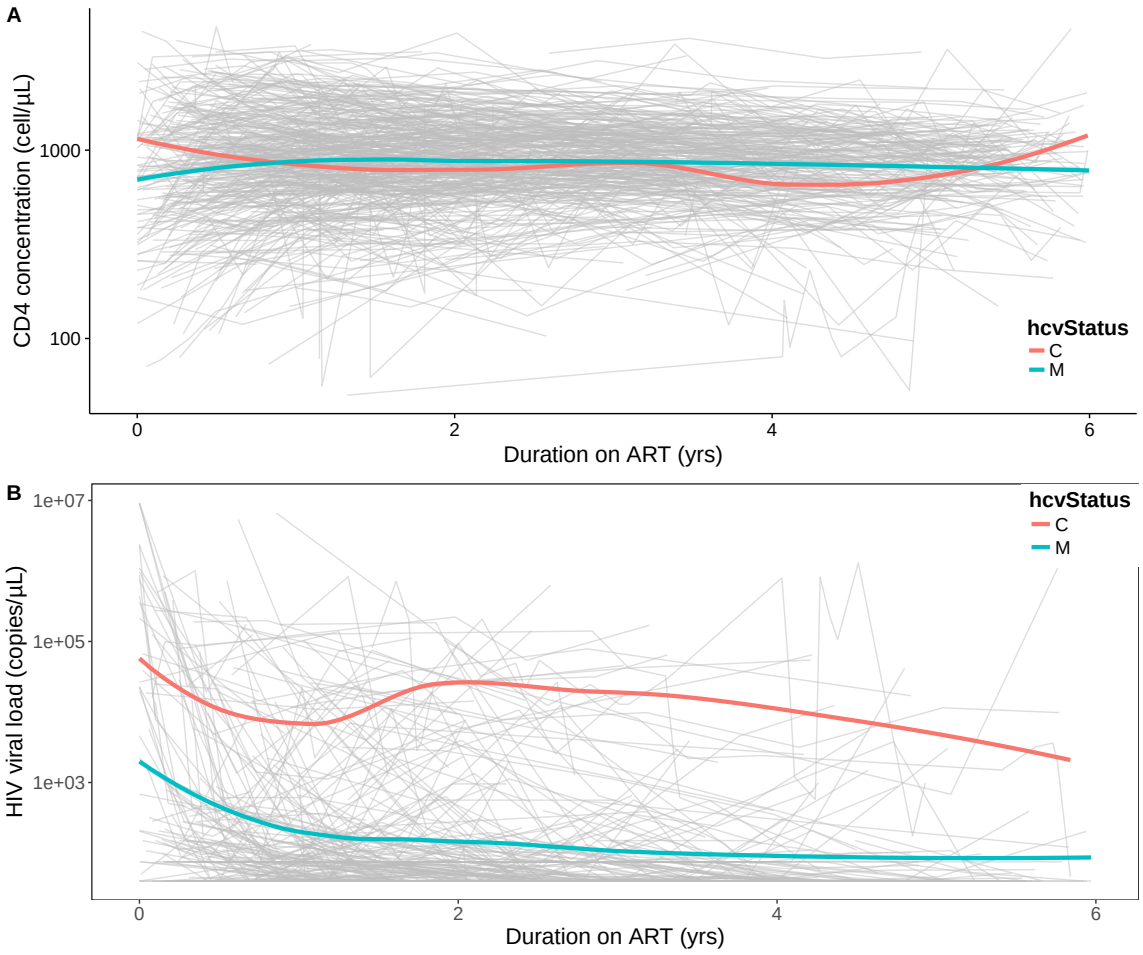


Figure 3.3: **A**: Plot of raw CD4 count (log scale) plotted against duration on ART, **B**: Plot of HIV viral load (log scale) over duration on ART. In both plots, local regression lines (loess) are fitted through the co-infected data (C) and mono-infected data (M), grey lines represent individual children in the cohorts

Variance-covariance matrix

The covariance matrix of the mixed-effects model which describes the variability within the population in pre-ART and long-term z-scores is given below:

$$\begin{array}{c} \begin{matrix} int_i & asy_i & c_i \end{matrix} \\ \begin{pmatrix} 4.50 & 0.506 & 0 \\ 0.506 & 1.78 & 0 \\ 0 & 0 & 0.347 \end{pmatrix} \begin{matrix} int_i \\ asy_i \\ c_i \end{matrix} \end{array} \tag{3.4}$$

In the above covariance matrix (3.4), the diagonal numbers are the variances whilst the off-diagonal numbers are the covariances. The correlation coefficient between int_i and asy_i can be calculated from

their covariances as given below:

$$\rho = \frac{\omega_{int,asy}}{\sqrt{\omega_{int}^2}\sqrt{\omega_{asy}^2}}$$

where $\omega_{int,asy}$ is the covariance between int and asy, ω_{int}^2 and ω_{asy}^2 represent the variances on int and asy respectively. This gives a weakly positive correlation of 0.18 between pre-ART and long-term CD4 z-scores.

HIV Viral Load

In the raw dataset, a total of 1,475 HIV viral load measurements were available in 365 children. The data provided here was based on the latest HIV viral load reported in each child. These viral load measurements were available for 43 of the 46 (94%) co-infected children, 39 (91%) of whom had undetectable HIV viraemia at their latest blood test. Of the 355 HIV mono-infected children, data on latest HIV viral load was available in 322 (91%) children, 204 (63%) of whom had undetectable viral load. The 118 mono-infected children with detectable HIV viral load had a median log viral load of 5.98 log copies/ml (IQR: 4.6-9.0). The four co-infected children with detectable HIV viral load had a median log viral load of 5.1 log copies/ml (IQR: 5-5.8). Median duration on ART at the time of latest HIV viral load was 3.1 years for all 365 children combined (IQR: 1.5-4.6 years), 2.9 years for the mono-infected children (IQR: 1.5-4.3 years) and 5.4 years for the co-infected children (IQR: 2.6-7.5 years).

HCV Viral Load

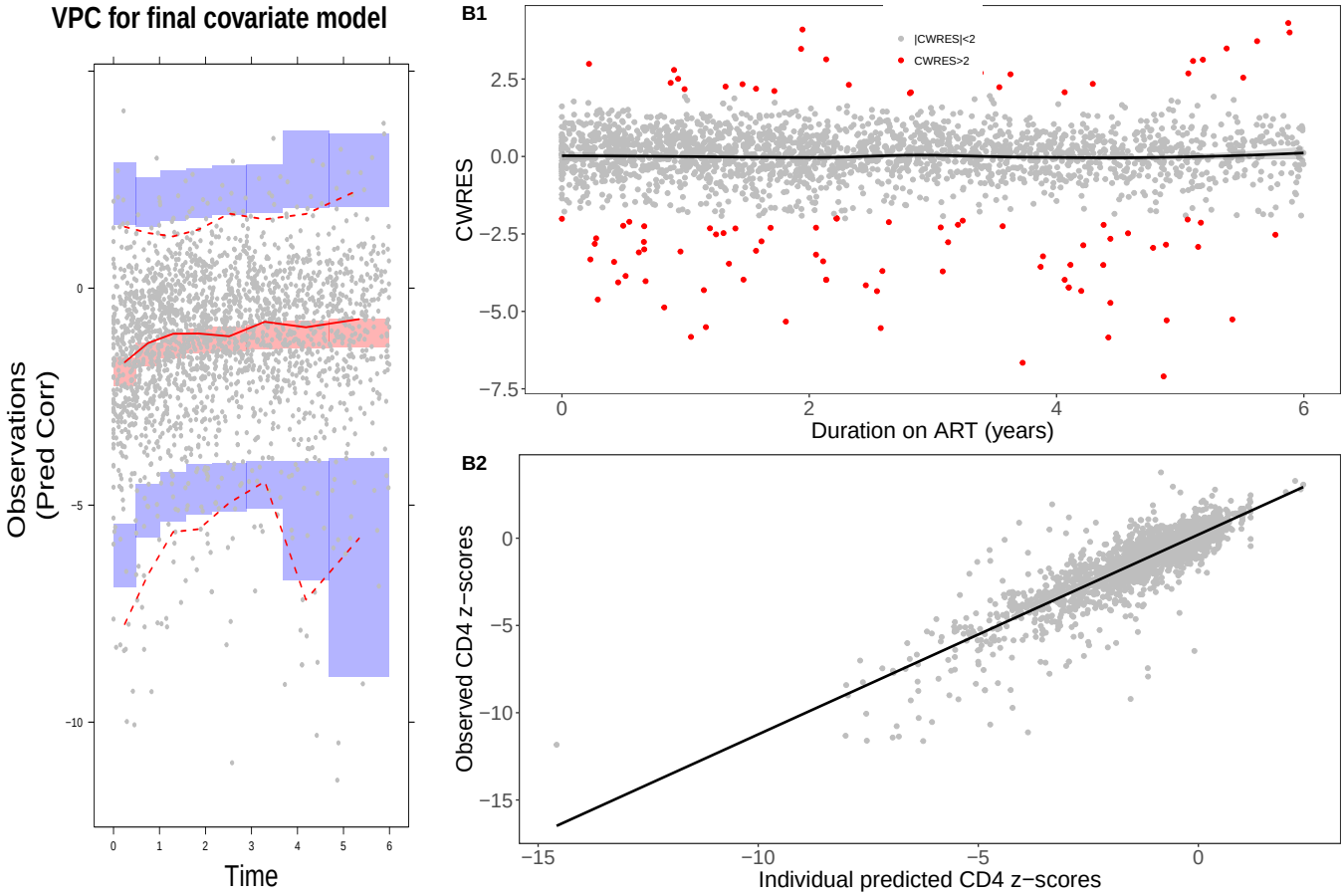
The data on HCV viral load was only available in 23 co-infected children. Of these 23 children, 5 had undetectable viral load at the most recent blood test. For the remaining 18 children with detectable viraemia, the latest blood test revealed a median HCV viral load of 4.46 log copies/ml (IQR: 3.28-5.64 log copies/ml)

Anti-HCV and anti-retroviral therapy

All ten co-infected children who received anti-HCV therapy had PEG-interferon and ribavirin. Of these, six failed treatment, three achieved spontaneous viral response and treatment outcome is unknown for one child. Data on ART regimen was available in 35 of 46 co-infected children and in all 355 mono-infected children. By far the most common ART drugs were Lamivudine, Zidovudine and Kaletra (Lopinavir and Ritonavir). A total of 33% of mono-infected children were on a three-ART regimen compared to 49% in the co-infected group.

Diagnostic plots

Diagnostic plots of final multivariate model can be found in Figure 3.4. It shows a good agreement between fitted and observed CD4 z-scores. The residual plot shows no particular trends over time as would be expected in an unbiased model. Figure 3.5 (A,C,D) shows that the assumption of normality in the random effects on all three parameters is reasonable. There's an outlier in the bottom left of Figure 3.4B2 which represents a co-infected child with very low CD4 measurement in the first 4 years of starting ART (CD4$^+$ count: 50 cells/μL). The visual predictive plot in 3.4A shows that the model overpredicts at low CD4 z-scores. However, in this case, the model underpredicted perhaps due to lack of information from children with very low CD4 counts in the data. This child also has very high viral load measurement in the first 4 years of therapy raising the possibility of poor compliance with medication or resistance to ART.



(A) (B)

Figure 3.4: (A): Visual predictive check for the final multivariate model generated from 5000 simulated datasets. The gray dots correspond to the observed dataset whilst the dashed and solid red lines correspond to 2.5th, 97.5th and 50th centiles respectively. The red and blue bands represent the 95% confidence intervals generated from the model-simulated centiles. (B1): Conditional weighted residual plot (CWRES) against time for the final multivariate model. Red points are cwres $>$ 2 or $<$ -2. (B2): Observed CD4 z-score plotted against fitted CD4 z-score.

Factors determining pre-ART and long-term CD4 z-score

On average (fixed effects), children started ART with a CD4 z-score of -2.39 (equivalent to $CD4^+$ count of 591 cells/μL in the average 4.3 yr old), corresponding to the 1st centile in uninfected children of the same age. The two predictors of pre-ART CD4 z-score were age at start of ART and EPPICC cohort. Pre-ART z-score was 0.31 $\pm$ 0.07 units lower per year older at ART initiation ($p < 0.0001$), and there was also an effect of cohort, with children from the UK and Spain starting ART with lower CD4 counts for their age while the children in the Italian cohort started with higher CD4 z-scores (Table 3.2 shows complete effect sizes). The only predictor of long-term age-adjusted CD4 count was the age at ART initiation: long-term z-score (asy_i) was 0.11 $\pm$ 0.0236 units lower per year older at ART initiation (Equation 3.6). The average long-term z-score of -1.06 from the final model corresponds to the 14th centile for healthy children (equivalent to 781 cells/μL in an average 4.3 yr old). Figure 3.6 shows model predictions (fixed effects) for three categories of HIV mono-infected children at 2 years, 4 years and 8 years. As expected, younger mono-infected children commenced ART at higher age-adjusted CD4 counts and also achieved higher long-term values of CD4 z-scores when compared to older children. According to the fixed effects estimates from the multivariate model, an average HIV mono-infected Ukrainian child of age 4.3 years is predicted to have a preART CD4 z-score of -2.39 $\pm$ 0.27 and a long-term z-score of -1.06 $\pm$ 0.08. These correspond to pre-ART and long-term CD4 concentrations of 591 and 781 cells/μL respectively. In the pre-ART HIV viral load data consisting of 365 children, an average Ukrainian child of age 5.2 years with viral load of 70,969 copies/ml was predicted to have a pre-ART and long-term CD4 z-score of -2.31 and -1.09 respectively (587 and 883 cells/μL). Equations 3.5 and 3.6 describe the relationships between parameter estimates and individual covariates retained at the end of the stepwise covariate selection.

$$int_i = -2.39 - 0.312(\text{age} - 4.3) + \begin{cases} 0 & \text{Ukraine} \\ 0.654 & \text{Russia} \\ -0.218 & \text{Switzerland} \\ -3.81 & \text{UK} \\ -18.2 & \text{Spain} \\ 2.60 & \text{Italy} \\ 0.293 & \text{Germany} \\ 0.397 & \text{Poland} \end{cases} \tag{3.5}$$

$$asy_i = -1.06 - 0.113(\text{age} - 4.3), \quad c_i = \begin{cases} 1.45\text{yr}^{-1} & \text{Mono-infected} \\ 0.36\text{yr}^{-1} & \text{Co-infected} \end{cases} \tag{3.6}$$

Hepatitis C co-infected children have a slower rate of increase in CD4 z-score than HIV mono-infected children

In the final multivariate model (equation 3.6), HCV co-infection was a significant predictor of the rate of CD4 z-score recovery, c. Co-infected children had a significantly reduced recovery rate of 0.36 per year compared to 1.45 per year in mono-infected children (Figure 3.6b, density plot in Figure 3.5B). This difference corresponds to a time to attain half the long-term recovery of 23 months in co-infected children, compared to 6 months in mono-infected children. Interestingly, HCV had no statistically significant effect on either pre-ART or long-term CD4 z-scores suggesting that although recovery was slower in co-infected children, they started ART with similar CD4 counts, and on long-term therapy do eventually achieve similar CD4 levels as their mono-infected peers. If co-infected and mono-infected children start ART at the same CD4$^+$ counts, it will take about four times longer for co-infected children to achieve similar long-term CD4$^+$ counts as mono-infected children (Figure 3.6).

Higher pre-ART HIV viral load is associated with a decreased pre-ART CD4 z-score

To investigate the effect of pre-ART viral load on CD4$^+$ T cell recovery, I analysed a subset of the original data consisting of 238 children who had pre-ART HIV viral load data. I compared the covariate model selection process with pre-ART HIV viral load excluded or included in addition to the other five

Table 3.2: Parameter estimates for the final multivariate model

Fixed effects	Estimate	SE	*p-values*	Ω
Intercept				
int	-2.39	0.27	-	4.50
int:age	-0.312	0.0685	< 0.0001	
int:Russia	0.654	0.557	-	
int:Switzerland	-0.218	0.816	-	
int:UK	-3.81	1.40	-	
int:Spain	-18.2	1.34	-	
int:Italy	2.60	0.739	-	
int:Germany	0.293	0.326	-	
Int:Poland	0.397	0.312	-	
c (yr^{-1})				
c	1.45	0.632	-	0.347
c:coinf	-0.816	0.119	0.0006	
Asymptote				
asy	-1.06	0.0791	-	1.78
asy:age	-0.113	0.0236	< 0.0001	
residual error	1.35	0.166	-	

The reference case is a Ukrainian child starting ART at 4.3 years (median age in dataset), HIV-monoinfected, with negative AIDS status. Hence the parameters *int*, *asy* and *c* are for this particular child. "age" represents age at start of ART, SE: standard errors of estimates, Coinf: HIV/HCV co-infected, 'int:age" represents the covariate interaction between age at start of ART and pre-ART CD4 z-score, Ω: Variances on random effects.

Table 3.3: Parameter estimates for the final multivariate model (with log viral load as covariate)

Fixed effects	Estimate	SE	*p-values*	Ω
Intercept				
int	-2.31	0.28	-	4.10
int:age	-0.392	0.10	< 0.0001	
int:logvl	-0.184	0.08	< 0.005	
c (yr^{-1})				
c	1.05	0.35	-	0.17
Asymptote				
asy	-1.06	0.079	-	2.08
asy:age	-0.136	0.032	< 0.0001	
residual error	1.26	0.153	-	

The reference case is a Ukrainian child starting ART at 5.2 years (median age in dataset). Hence the parameters *int*, *asy* and *c* are for this particular child. Age represents age at start of ART, SE: standard errors of estimates, 'int:age" represents the covariate interaction between age at start of ART and pre-ART CD4 z-score, logvl: log viral load, Ω: Variances on random effects.

covariates (age, AIDS status at start of ART, gender, HCV status and EPPICC cohort). Two models were explored: one model had correlation between asymptote and intercept whilst the other had no pre-specified

correlation structure. The model with the correlation structure was much more stable and revealed that higher pre-ART viral load was associated with a decreased pre-ART CD4 z-score. For every 1 unit increase in log viral load, the pre-ART CD4 z-score was reduced by 0.2 units. The covariate model retained the following relations: long-term CD4 z-score and age; pre-ART CD4 z-score with age and log viral load. Of note, HIV/HCV co-infection did not have any significant effect on any of the parameters obtained from this analysis (Table 3.3 shows full effect sizes).

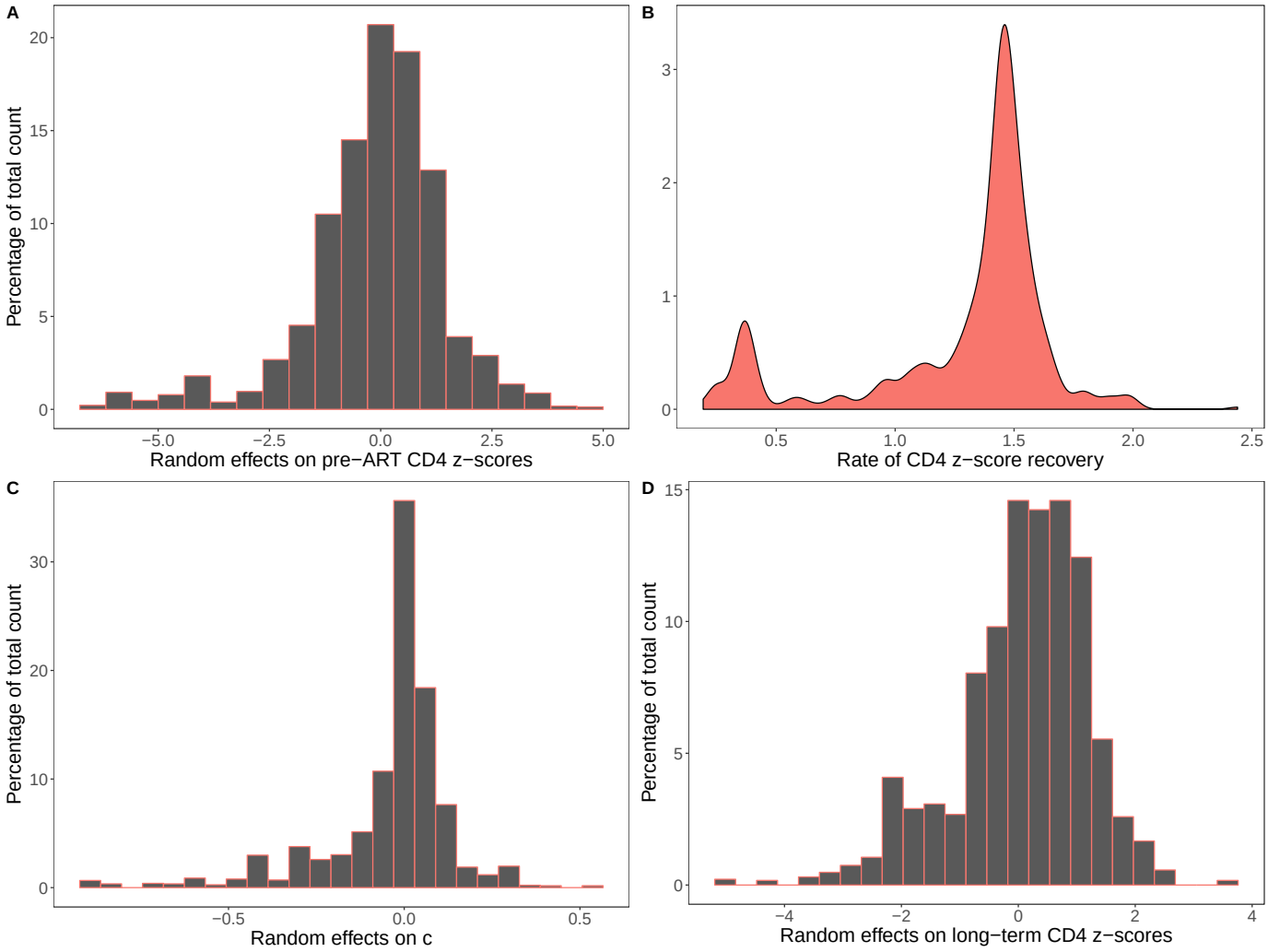


Figure 3.5: **A**. Histograms of random effects (ETA) on pre-ART CD4$^+$ z-score, **B**. Density plot of estimated rate of recovery in CD4$^+$ z-score, **C**. Histogram of random effects on rate of recovery in CD4$^+$ z-score, **D**. Histogram of estimated long-term CD4$^+$ z-scores.

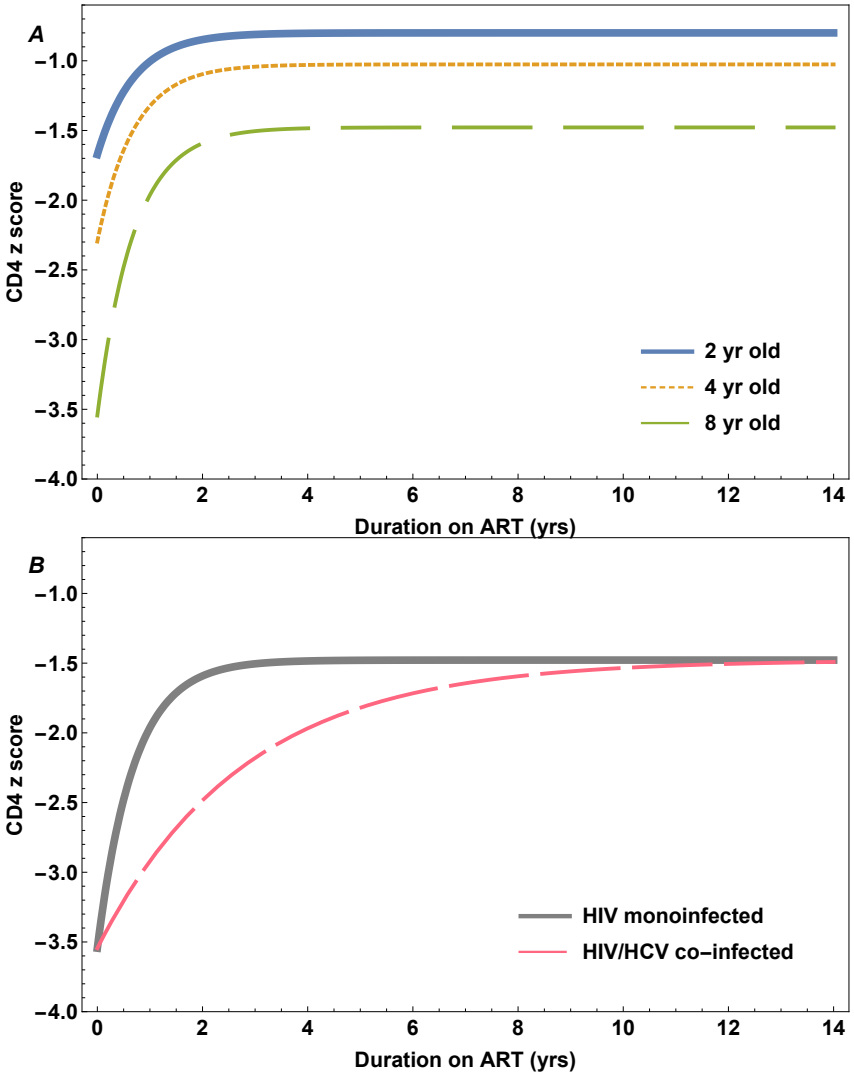


Figure 3.6: Model predictions for CD4 T cell recovery in co-infected and mono-infected children. **A**: Model recovery profiles (fixed effects) predicted for a 2 year old (solid blue line), 4 year old (short dashed lines-orange) and an 8 year old (long dashed lines-green) child (all HIV mono-infected). Younger mono-infected children start ART at higher intercept values and also achieve higher long-term values of age-adjusted CD4 counts when compared to older children. **B**: Fixed effects profiles predicted for a mono-infected (solid line) and a co-infected (long dashed line) 8-year old child. The HIV/HCV co-infected child is seen to have a significantly slower rate of increase in age-adjusted CD4 relative to the mono-infected child. However, both predictions have comparable intercept and asymptote values of their age-adjusted CD4 counts.

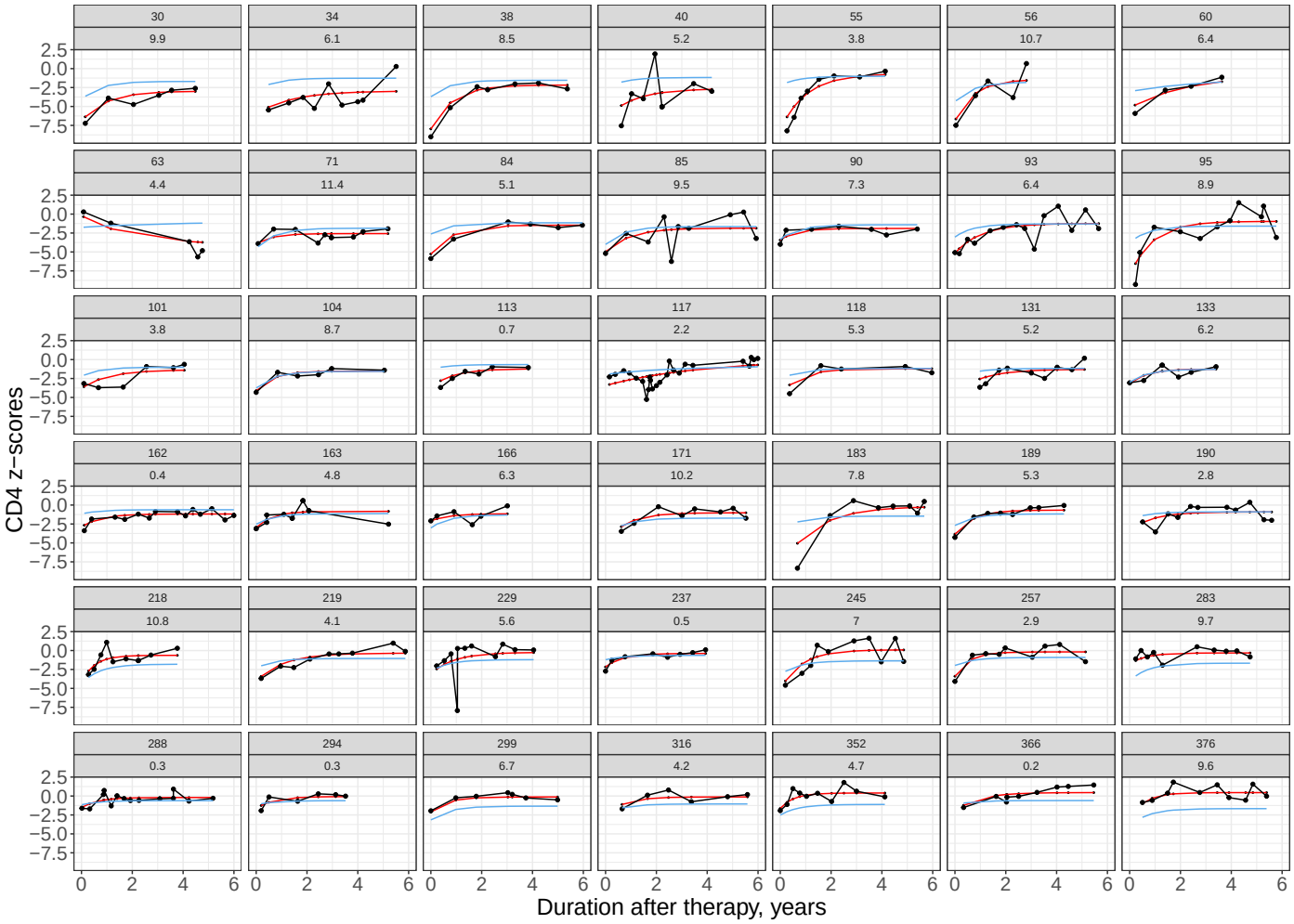


Figure 3.7: The above plot shows individual predictions for a randomly selected 42 children from the final multivariate model. Numbers in top panel shows IDs whilst second panel shows age at start of ART. Blue lines: population-level predictions, black line: observed data, red lines: individual predicted data

3.5 Discussion

In this chapter, I have investigated the impact of HCV co-infection on CD4$^+$ T cell reconstitution in HIV-infected children receiving ART. By fitting the asymptotic model to longitudinal data from 401 children, I have deduced that HIV/HCV co-infected children had significantly slower recovery of their age-adjusted CD4 counts than HIV mono-infected children. Despite this reduced rate of recovery, the co-infected children still managed to achieve long-term CD4$^+$ T cell levels comparable to HIV mono-infected children. The fixed effect estimates for c, pre-ART and long-term CD4 z-scores are consistent with others obtained in a different multi-centre European paediatric study [Lewis et al., 2012]. In a study by Marcus *et al*, it was shown that HIV/HCV co-infected adults have delayed CD4$^+$ T cell recovery on ART, relative

to their mono-infected counterparts [Marcus et al., 2015]. Similarly, a meta-analysis conducted by Tsiara *et al.* in 2013 involving 22,533 adult patients reported that even though HCV had a demonstrable adverse effect on immune reconstitution in the first 2 years of ART, this trend was not sustained in the long-term [Tsiara et al., 2013]. Our findings in children add to this by suggesting that this transient adverse effect of HCV co-infection in adults might be explained by a reduced rate of increase in CD4$^+$ T cell recovery. This is the first large-scale mathematical modelling analysis of the effect of HCV on CD4$^+$ T cell recovery in HIV/HCV co-infected children. The findings of this study are consistent with those of Micheloud *et al.*, whose smaller 2007 study found similar long-term CD4$^+$ T cell recovery in 19 co-infected and 25 mono-infected children [Micheloud et al., 2007].

A number of mechanisms could be responsible for driving slower rate of reconstitution in HIV/HCV co-infected children receiving ART. One possible explanation is reduced thymic output. Some studies have already shown that HCV mono-infection and HIV/HCV co-infection appears to negatively affect thymic output in adult patients [Shmagel et al., 2014, Hartling et al., 2013]. In children however, there is evidence that thymic output may recover on ART and this might explain why long-term CD4$^+$ counts were unaffected by HCV co-infection in this study [Sandgaard et al., 2014, Gibb et al., 2000]. Since younger children have higher thymic outputs and consequently, higher immunological reserve, it is likely that commencement of antiretroviral therapy may be sufficient to compensate for the potential adverse effect of HCV on immune reconstitution in the long term. Indeed, many studies that reported a difference in CD4$^+$ T cell reconstitution between co-infected and mono-infected patients only noted this in the first 1-2 yrs following therapy [Macías et al., 2003, Santin et al., 2008], and this difference has been observed to wane around 2 yrs into ART [Tsiara et al., 2013, Kaufmann et al., 2003]. Furthermore, in a very recent Russian study on 79 HIV-infected patients receiving HAART and with very low HIV viraemia (< 50 copies/ml), the co-infected subjects were found to have a reduced number of recent thymic emigrants (RTEs). This was not related to HCV viral load but it correlated with hepatic damage as indicated by deranged liver enzymes. However, there was no significant difference in the quantification of other subsets of CD4$^+$ T cells between co-infected patients and singly infected subjects [Korolevskaya et al., 2016].

Another possible mechanism that could give rise to slower CD4$^+$ T cell recovery in HIV/HCV co-infected individuals is increased T cell activation. Increased T cell activation is thought to impair CD4$^+$ T cell recovery through excessive cytokine production leading to immune exhaustion. In a study by Hunt *et al* consisting of 99 HIV-infected adults, increased T cell activation strongly correlated with HCV co-infection, lower nadir CD4$^+$ T cell count, shorter duration of sustained viral suppression and low level HIV viraemia

[Hunt et al., 2003]. On the other hand, effective treatment of HCV infection with interferon and ribavirin has been seen to reduce T cell activation [Gonzalez et al., 2009].

A third hypothesis is that poorer virological response in co-infected children means that CD4 concentration recovers more slowly. This is difficult to investigate in the cohorts included for analysis in this work for whom viral load data is sometimes patchy, but could be investigated in other children in future studies.

One further explanation for the reduced rate of CD4$^+$ T cell recovery in co-infected patients is increased apoptosis of CD4$^+$ T cells. Even though HIV infection is known to cause increased apoptosis of CD4 T cells, this loss of T lymphocytes is normally reduced with antiretroviral therapy. In a study by Korner *et al*, it was demonstrated that untreated HIV/HCV co-infected patients had an increased rate of CD4$^+$ T cell apoptosis. However, co-infected and mono-infected patients on ART had similar CD4$^+$ T cell apoptosis after 4 weeks of therapy [Körner et al., 2009].

One of the main strengths of this work is the number of HIV/HCV co-infected children included for analysis which is more than double the number included in the only previous analysis [Micheloud et al., 2007]. Furthermore, mixed-effects approach provides a statistically rigorous method appropriate for analysing longitudinal datasets where observations from an individual are often correlated [Jose C. Pinheiro, Douglas M. Bates, 2000].

One limitation of this analysis lies in the fact that all the HIV mono-infected children were from Ukraine whereas the HIV/HCV co-infected children were selected from eight countries across Europe (including Ukraine) [Majekodunmi et al., 2018]. Hence, laboratory testing of CD4 and viral load were conducted locally and to take account of this, we have added the effect of cohort as a covariate in informing model predictions. An additional limitation was that the effect of pre-ART HIV viral load could be investigated only in the subset of the children in whom it was available (238 children-59%). It is likely that the inability to detect a significant relationship between c and co-infection status is due to the reduced statistical power of this sub-analysis. Furthermore, not all the latest HIV viral loads were measured following the same duration of ART. This is likely to account for the greater proportion of co-infected children with viral suppression. The estimate obtained for the coefficient of the intercept value associated with the Spanish cohort was very low (-18.2) but with a proportionate standard error (1.34). This was due to the small number of children from the Spanish cohort (two children). However, this is unlikely to influence the main findings in this work which is that of differences between rate of CD4 z-score recovery between co-infected and mono-infected children. The different EPPICC cohorts had no significant covariate relation with rate of CD4 z-score recovery. In future work, the different EPPICC cohorts could be regrouped into larger

cohort prior to the covariate analysis.

Lastly, it has not been possible to explore the impact of HCV treatment on CD4$^+$ T cell reconstitution due to the small number of children treated within the cohort. The structural model used in fitting CD4$^+$ T cell data is empirical/semi-mechanistic in nature. Hence, mechanisms of CD4$^+$ T cell recovery cannot be understood fully with this model. Appropriate deterministic models involving differential equations would more suitable in this regards and provides a direction for future work. A fully mechanistic approach is explored further in chapter 5 of this report.

The delayed CD4$^+$ T cell recovery seen in HIV/HCV co-infected children suggests that early treatment with ART may be more critical in co-infected children. In order to minimize the time for which children are exposed to the increased risk of opportunistic infection associated with very low CD4$^+$ counts, it may be necessary to start therapy earlier in co-infected children because they will take longer to return to "safer" CD4$^+$ levels. As illustrated in Figure 3.6b, co-infected children take more than four times as long as mono-infected children to recover their CD4$^+$ levels. This adds to the case for early ART in co-infected children that has already been made: that good control of HIV viraemia is likely to lead to a better clearance of HCV [Falconer et al., 2008], and reinforces the latest PENTA guideline for treatment of paediatric HIV, which has diagnosis of HCV co-infection as a definite indication for commencing ART in children [Bamford et al., 2015]. Eradication of hepatitis C remains important in prevention of hepatocellular carcinoma as children progress into adulthood. A recent cross-sectional study by Thorne et al. revealed that a high percentage of HIV/HCV co-infected children and adolescents had progressive liver disease evidenced by hepatomegaly (42%) and raised alanine aminnotransferase enzyme (55%) [EPPICC in EuroCoord, 2016]. Of the 229 children included, only 11% had been successfully treated for hepatitis C. This emphasises the need for more effective HCV treatment in co-infected children.

A joint modelling approach used in a study by Del Fava et al investigating change of HCV and HIV infection prevalence and correlation between HCV and HIV infection between various regions in Italy amongst intravenous drug users [Del Fava et al., 2011]. They used a joint model with random effects models (generalised linear mixed effect and hierarchical Bayesian models) for the binomial prevalence data for hepatitis C and HIV infection. They did not assume a prior association between HCV and HIV but investigated different hypotheses for the random effect covariance matrices with the aim of identifying the one that best describes the data. In future work, a similar approach could be used in investigating relationship between HIV/HCV co-infection and CD4$^+$ T cell recovery in children.

Further studies of vertically HIV/HCV co-infected children are therefore needed to fully understand

the implications of HCV co-infection on immune recovery. Additional studies on efficacy of the new direct acting antivirals are needed to fully understand whether the findings in this work will persist following their introduction in the paediatric population. In future work, multiple imputation of missing variables could be considered to ensure that more subjects are included for analysis.

Chapter 4

Predictors of poor immune recovery in children receiving anti-retroviral therapy

4.1 Introduction

4.1.1 Poor immune recovery in HIV-infected adults on ART

Despite full suppression of viral replication on ART, about 1 in 5 of all HIV-infected adults are unable to reconstitute their $CD4^+$ T cell count above 500 cells/μL [Autran et al., 1999, Piketty et al., 1998, Autran et al., 1997]. Furthermore, an estimated 16% of HIV-infected patients may fail to achieve $CD4^+$ T cell count $>$ 200 cells/μL [Kaufmann et al., 2003, Moore and Keruly, 2007]. This group of patients are referred to as immunological non-responders (IMNRs). Whilst there is no standard definition for IMNRs, most are identified as having $CD4^+$ T cell counts below 200 cells/μL following a wide range of treatment duration. Another definition of IMNR in the literature is based on an increase from baseline CD4 of $<$ 20% following ART [Li et al., 2011, Marziali et al., 2006]. In comparison to those with optimal $CD4^+$ T cell reconstitution, patients unable to recover their immune system sufficiently despite virologic suppression are at significantly increased risk of death, AIDS-defining events, cardiovascular disease, malignancies and severe opportunistic infections [Lapadula et al., 2013, Tan et al., 2008]. It is generally accepted that a $CD4^+$ T cell count of $>$ 500 cells/μL is representative of an adequate immune response as these patients tend to have morbidity and mortality rates comparable to uninfected individuals [Lewden et al., 2007]. Immune nonresponders are therefore a subset of inadequate immune responders which also include the less defined category of patients with CD4 count of 200-500 cells/μL. Factors associated with immunological nonresponse include older age [Kaufmann et al., 2002], longer duration of HIV infection before starting ART, active hepatitis C co-infection [Greub et al., 2000, Almeida et al., 2007] and a low nadir CD4 count prior to ART initiation [Moore and Keruly, 2007, Engsig et al., 2010, Kaufmann et al., 2005]. Of these factors, the best predictor of immune recovery is a lower nadir CD4 T cell count prior to commencement of ART. However, it has been observed that these factors alone do not sufficiently explain the morbidity

observed in viral-suppressed HIV-infected individuals.

Consequently, one mechanism that has been proposed in poor immunological responders is altered thymopoiesis [Delobel et al., 2006]. $CD4^+$ T cells are usually maintained in adults via a combination of thymic production and peripheral proliferation of naive $CD4^+$ T cells. Thymic production of $CD4^+$ T cells decreases with age, shifting production of $CD4^+$ T cells to peripheral proliferation. Studies visualising the thymus of HIV-infected adults using computed tomography (CT) scans have revealed a positive correlation between thymic size and naive/total $CD4^+$ T cell counts [Kolte et al., 2002, McCune et al., 1998, Smith et al., 2000]. HIV infection results in a significant functional and quantitative disruption in naive $CD4^+$ T cells both in the blood and lymphatic tissues [Autran et al., 1997, Hazenberg et al., 2000a, Douek et al., 1998]. Once commenced, ART reverses this disruption by causing a suboptimal increase in thymic output and total naive T cells as well as a decrease in naive $CD4^+$ T cell proliferation and reduced killing of $CD4^+$ T cells [Autran et al., 1997, Douek et al., 1998, Zhang et al., 1999]. However, the degree of immune disruption found in immunological nonresponders is greater than in HIV-infected patients with normal $CD4^+$ T cell recovery. In support of this, one study found very low levels of RTEs in circulation which correlated with the finding of no residual thymic tissue on both CT and PET scans of IMNRs [Tanaskovic et al., 2011]. The suggestion of altered thymopoiesis is further corroborated by the finding of reduced levels of naive T cells and other T cells expressing CD28 in IMNRs [Erikstrup et al., 2010, Isgrò et al., 2008].

The second mechanism thought to be driving poor immune recovery in immune nonresponders is altered haematopoiesis secondary to dysfunctional bone marrow (BM) and haematopoietic progenitor cells (HPCs). In a few limited studies, HIV has been shown to impair production of thymocyte precursors in the bone marrow [Marandin et al., 1996, Moses et al., 1998, Jenkins et al., 1998]. In addition to this, a number of studies have shown that a percentage of HPCs are susceptible to HIV infection through the expression of CD4, CXCR4 and CCR5 HIV receptors [Alexaki and Wigdahl, 2008].

A third mechanism that could be driving poor immune recovery in immune nonresponders is impaired cytokine production. Cytokines implicated in the regulation of T cell division and homeostasis include IL-2 (Interleukin 2), IL-7 and IL-15. Interleukin 7 plays a key role in promoting survival and proliferation of naive T cells in the peripheral circulation without changing their naive status [Schluns et al., 2000]. HIV-infected patients have shown higher circulating levels of IL-7 which correlated with downregulation of IL-7 receptor (IL-7R) on T cells compared to healthy individuals [Llano et al., 2001, Rethi et al., 2005]. Similarly, immune nonresponders had higher levels of IL-7 and decreased expression of IL-7R [Marziali et al., 2006]. Furthermore, increased stromal production of IL-7 has been observed in IMNRs compared to

HIV-infected subjects with optimal immune recovery [Isgrò et al., 2008, Bellistrì et al., 2010]. Since IMNRs mount appropriate IL-7 response, it has been concluded that impaired source of CD4 T cell production rather than cytokine signalling is responsible for poor immune recovery in this group of patients. Unlike IL-7, production of IL-2 and IL-15 have been noted to be compromised in HIV-infected patients [Ahmad et al., 2005, Sirskyj et al., 2008]. Therefore, improving the signalling via IL-2 and IL-15 could be helpful in reversing poor immune recovery observed in these patients. The recovery of CD4 T cells in the first 2 years of ART seems to be facilitated mainly by T cells expressing the IL-7 α-chain receptor [Hodge et al., 2011].

It is well known that primary infection with HIV results in the loss of a large number of cells including 50% of $CD4^+$ T cells in the lymph nodes [Mattapallil et al., 2005, Guadalupe et al., 2003]. Hence, another proposed mechanism is damage to secondary lymphatic tissue responsible for maintaining $CD4^+$ T cell population and subsequent replacement of functional tissue with collagen. The greater the degree of collagen deposition, the lower the number of naive $CD4^+$ T cells and CD4 count [Schacker et al., 2002]. Also, the amount of collagen-deposition has been indicative of the extent of immune recovery in HIV-infected patients [Schacker et al., 2005].

Other mechanisms include increased $CD4^+$ T cell apoptosis in patients with CD4 count < 500 cells/μL [Gougeon et al., 1996, Meyaard et al., 1992, Meyaard et al., 1994, Bottarel et al., 2001, Piconi et al., 2010], higher levels of immune activation and imbalance between pro- and anti-inflammatory T cells.

4.1.2 Poor immune recovery in HIV-infected children on ART

As discussed in chapters 1 and 3, the introduction of anti-retroviral therapy (ART) in 1980s/90s dramatically changed the outlook of HIV treatment, resulting in a remarkable decline in HIV-related morbidity and AIDS-related mortality. On commencing ART, HIV viral replication is suppressed and most HIV-infected children experience some recovery of $CD4^+$ T cells and this usually correlates with a better clinical outcome. However, despite ART and satisfactory virological control, a small percentage of HIV-infected children fail to achieve a satisfactory increase in their $CD4^+$ T cell population. Since poor CD4 recovery in the HIV-infected population has been strongly associated with increased risk of morbidity and mortality [Engsig et al., 2010, Antiretroviral Therapy Cohort Collaboration, 2008, Piketty et al., 2001], the main focus of therapy is to achieve an adequate recovery of $CD4^+$ T cells comparable to healthy subjects of the same age. Whilst there is currently no agreed definition of adequate $CD4^+$ T cell response in children under 5 years, in children aged 5 years and above, immunological failure (also immunological nonresponse) is defined as inability to achieve or maintain $CD4^+$ T cell count above 200 cells/μL in spite

of HIV viral suppression [World Health Organization, 2016] (Table 4.6). However, this criterion excludes children with poor immune response and viral load just above the limit of detection. Furthermore, this cut off is of limited use in identifying children with poor immune response since higher thymic output in children means they will have higher CD4 counts than adults patients. A study by Krogstad et al in 2015 investigating incomplete immune reconstitution in HIV-infected children and adolescents on ART revealed that out of 933 subjects included, only four children had CD4 counts below 200 cells/μL 2 years into viral load suppression [Krogstad et al., 2015]. To further complicate the picture, there are also children who experience good $CD4^+$ T cell recovery despite incomplete virological suppression [Chiappini et al., 2003]. Besides the depletion of CD4 counts and despite treatment with ART, it is increasingly recognised that HIV infection has other significant pathological manifestations on the immune system likely contributing to inadequate immune response.

One of the factors contributing to suboptimal reconstitution in children is higher age at start of therapy, usually implying a lower thymic output [Prendergast et al., 2007, Cotugno et al., 2012, Hainaut et al., 2003, Vigano et al., 2000].

There are very limited studies on poor $CD4^+$ T cell recovery in HIV-infected children receiving ART.

In a paper by Zanoni et al, 674 South African children on ART were analysed in a study investigating predictors of poor CD4 and weight recovery. They demonstrated that presence of chronic diarrhoea at start of ART and virologic failure predicted poor CD4 recovery 6 and 12 months following ART [Zanoni et al., 2012]. Poor CD4 recovery was defined based on WHO criteria for immunological failure in children and adolescents as well as other published reports in the literature [Kekitiinwa et al., 2008, World Health Organisation, 2010].

4.2 Objectives

The main aim of this chapter is to use mathematical models of immune reconstitution to establish predictors of poor immune recovery in HIV-infected children receiving ART. In addition, the following objectives shall be explored:

1. To identify risk factors for immunological non-response both overall and in children with viral suppression (VL $<$ 400 copies/ml, $\geq$ 6 months of ART).

2. To use a predictive model of $CD4^+$ T cell recovery to explore definitions of immunological non-response in HIV-infected children on ART.

4.3 Methods

4.3.1 Dataset

The data presented in this chapter was obtained from the European Pregnancy and Paediatric HIV Cohort Collaboration (EPPICC). EPPICC represents a network of cohort studies across Europe whose aim is to conduct epidemiological research on HIV-infected expectant mothers and children exposed to HIV during pregnancy. It aims to answer scientific questions that require a large sample size of patients which usually cannot be addressed by the local participating cohorts. A total of 22 different cohort from 19 countries contributed to the dataset. The cohorts supplied anonymised data in keeping with a standard operating procedure and submitted using the HIV Cohorts Data exchange Protocol (www.hicdep.org). The data was merged on 30th September, 2014. Children aged < 18 years starting anti-retroviral therapy naive from 1996 and on a minimum of three ART were eligible for inclusion. Each participating cohort was responsible for compliance with local and national ethics requirements. HIV infection in children was defined as detection of HIV antibodies after 18 months and/or positive HIV RNA or DNA PCR in a minimum of two samples obtained on separate visits (cord blood excluded). HIV-uninfected children were defined as children with negative HIV antibodies and/or negative HIV RNA on at least two occasions. Only vertically infected children were included for analysis.

4.3.2 Age-adjustment of CD4 counts

In chapter 3, CD4 z-scores were used to adjust raw CD4 counts with age. Another alternative to using CD4 z-scores are "CD4 for age" ratios. These age-adjusted CD4 counts are calculated by dividing the raw CD4 counts by the CD4 counts expected in a healthy child of the same age [Huenecke et al., 2008]. The logarithm of this ratio, ln(CD4-for-age), can then be calculated for each individual child and fitted to a suitable model of choice. Unlike z-scores, CD4 ratios (equation 4.1) can only have positive values.

$$\text{CD4 Ratio}_{ij} = \frac{\text{CD4 count}_{ij}}{\text{CD4 count in a healthy child of the same age}_{ij}} \tag{4.1}$$

By log-transforming these ratios, we can generate positive as well as negative values similar to z-scores. This also ensures that the random errors are normally distributed. A "CD4 for age" ratio of 1 indicates that the child has the expected CD4 count for their age; a ratio < 1 suggests that the child has a lower CD4 count for their age when compared to healthy children of the same age born to HIV-infected mothers

[Picat et al., 2013]. Likewise, a negative log ratio indicates a child has a lower CD4 count for their age whilst a value of zero indicates that they have a normal CD4 count for their age. $CD4^+$ ratios are preferred over z-scores for normalising datasets with an abundance of low $CD4^+$ counts such as that used in this chapter. This way, the unpredictable behaviour of CD4 z-scores at low $CD4^+$ T cells could be avoided.

4.3.3 Data analysis and model fitting

To identify poor immune responders, we recall the same mathematical model used in the analysis described in chapter 3. Initially, all CD4 counts were adjusted for age by converting them into CD4 ratios followed by a log-transformation as discussed in section 4.3.2. ln(CD4-for-age) was chosen over CD4 z-scores due to the large number of children with very low CD4 counts included for analysis (Thailand and CHIPS cohorts). Hence the mathematical model used in fitting the data is given as:

$$\ln(\text{CD4-for-age}_{ij}) = asy_i - (asy_i - int_i)e^{-c_i t_{ij}} \tag{4.2}$$

In order to minimize sparse data, all children with less than 2 CD4 measurements were excluded from the analysis. Likewise, non-perinatally infected children were excluded from the analysis. The data was divided into subsets A and B; children were randomly sampled within three distinct age groups (<1 yr, 1 to <10 yrs and ≥10 yrs) and from each EPPICC cohort for inclusion into either subsets. In this chapter, datasets A and B are often referred to as "A" and "B". This approach was adopted to ensure that EPPICC cohorts who had a greater percentage of older or younger children were not sampled with bias. The mathematical model in equation 4.2 was initially fitted to both datasets. The parameters of the final covariate model developed from dataset A (model A) were then applied to dataset B and likewise the parameters of the final covariate model from dataset B (model B) were applied to dataset A. I then compared predictions made from models A and B to the observed values in the datasets with a view to identifying outliers within the distribution of residual values which could represent poor immune responders. Poor responders were identified by subtracting twice the standard deviation (σ) of the fitted values from the individual predicted values ("IPRED" in NONMEM notation). If an observed value ("DV" in NONMEM notation) was less than IPRED $- 2\sigma$, the child was then considered as having at least one outlying residual value (potential poor immune responder) and investigated further. To minimise the effect of long-term ART resistance, the asymptotic model was fitted to data in the first 6 years following start of ART. Further attempts were made to identify poor immune responders by comparing predicted ln(CD4-for-age) to the WHO criteria for HIV-associated immunodeficiency [World Health Organisation,

2016b].

4.3.4 Covariate analysis

The aim of the statistical analysis was to construct a multivariate model around the fixed effects components *int*, *asy* and *c* for both datasets A and B. As outlined in section 2.4 (page 36), I used a forward and backward stepwise selection with exit *p-values* of 0.01 and 0.005 respectively to investigate potential effects on CD4-for-age recovery of: pre-ART age, log10 HIV viral load and AIDS status, gender, individual EPPICC cohorts. Random effects were added to all three parameters to allow for between subject variability. In order to reduce the number of parameters in the covariate model (overfitting), the 22 individual EPPICC cohorts were regrouped into four different regions based on demographic similarities: UK/Ireland, Thailand, Russia/Ukraine and other European countries. Only children who had all five covariates were included for analysis.

4.3.5 Software

Non-linear mixed-effects modelling was by maximum likelihood using the FOCE, SAEM and IMP algorithms in NONMEM 7.3 [Beal et al., 2014]. The FOCE algorithm was the first choice and in instances where convergence could not be achieved, the SAEM and IMP algorithms were chosen instead. Further data analysis and manipulation was done in python [Van, Rossum, Guido et al, 2017] and R language [R Core Team, 2015] whilst predictions generated from the model were plotted in R and Wolfram Mathematica [Wolfram Research, Inc., 2015]. Stepwise covariate analysis and visual predictive checks were implemented in PsN (Perl speaks NONMEM) [Lindbom et al., 2005].

4.4 Results

4.4.1 Patient characteristics

The final datasets used in fitting the covariate models are described in Table 4.3 and plotted in Figure 4.1. There were a total of 29,956 HIV viral load entries and 29,668 CD4 measurements. A total of 2204 children from 22 different EPPICC cohorts were selected for analysis. Median age at start of therapy was 5.8 yrs (IQR: 1.4-10 yrs) and 5.4 yrs (IQR: 1.4-10 yrs) for datasets A and B respectively (Figure 4.2). At start of therapy, the median baseline CD4 count for datasets A and B were 421 (201-972 cells/μL) and 427 (198-985 cells/μL) respectively. Median base $\log_{10}$ HIV viral load at start of therapy was 5 $\log_{10}$ copies/μL

(4.4-5.6 $\log_{10}$ copies/μL) for dataset A and 5.1 $\log_{10}$ copies/μL (4.5-5.7 $\log_{10}$ copies/μL) for dataset B. Latest HIV viral load data after 6 months of ART was available in 2111 children and revealed a median of 46 copies/mL (40-103 copies/mL) for the entire dataset. Of the 2111 children identified, 1057 were from the dataset A and had a median of 50 copies/mL (40-115copies/mL) whilst 1054 were from the dataset B with a median of 40 copies/mL (40-100copies/mL). A total of 883 (80%) and 872 (79%) children from A and B respectively had complete suppression of their HIV viral load at the last measurement (< 400 copies/mL) after 6 months of therapy. A summary of the ART regime received by the children is given in Table 4.3. Median duration on ART was very similar for both A (5.3 yrs: 3.2-5.8 yrs) and B (5.4 yrs: 3.0-5.8 yrs). Final covariate models obtained for the two datasets are described in equations 4.5 to 4.10.

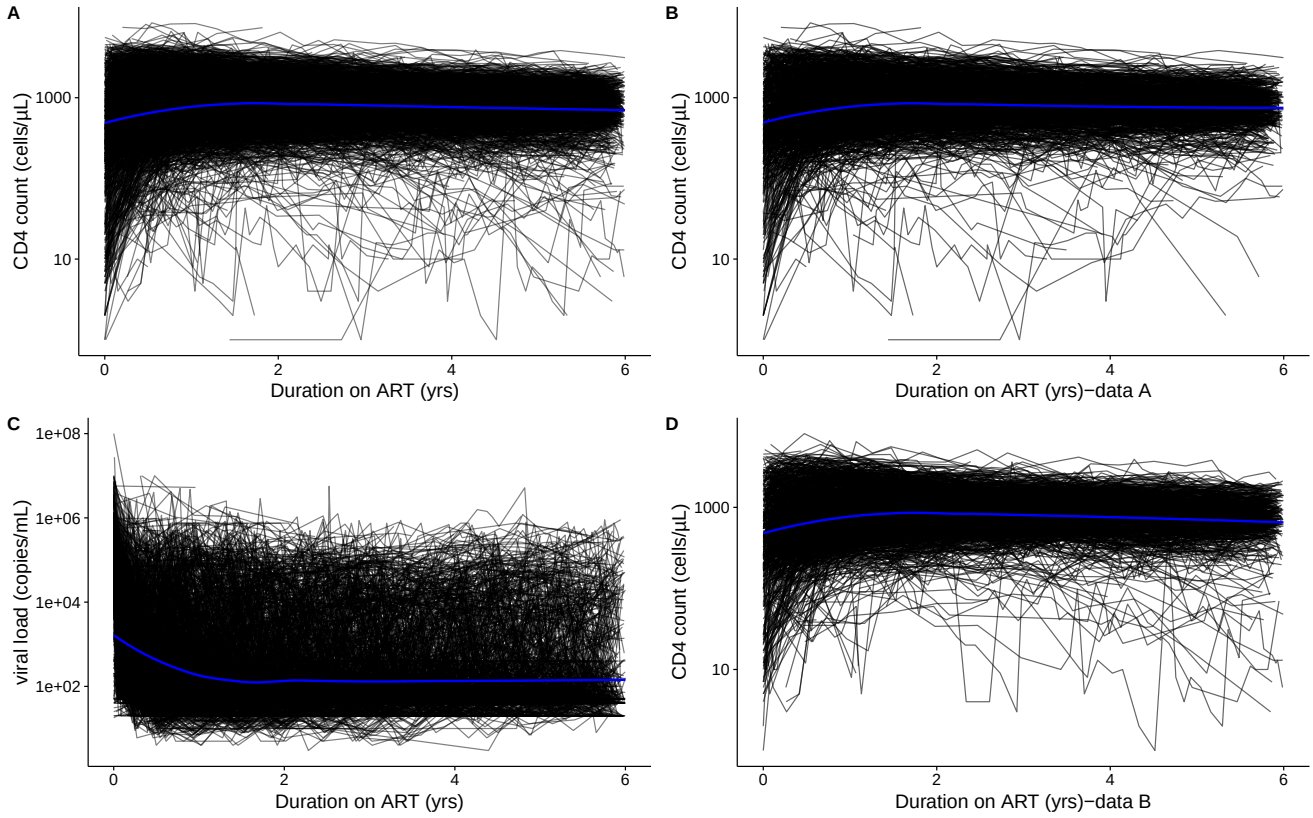


Figure 4.1: **A**: CD4 count (log scale) plotted against duration on ART therapy with local regression line fitted through data (data A and B). **B**: CD4 count (log scale) plotted against duration on ART therapy for data A. **C**: HIV viral load (log scale) plotted against duration on ART therapy with local regression line fitted through data. **D**: CD4 count (log scale) plotted against duration on ART therapy for data B.

4.4.2 Variance-covariance matrix

In building the covariate model, a full covariance matrix (on parameter random effects) was initially added to the base model. However, convergence could not be achieved and so covariances between random effects

on parameter values were removed. Full covariance was then added to the final covariate models for A and B. The addition of a full covariance structure to final covariate model A worsened the overall fit significantly (ΔOFV_A = -47) and only marginally improved the model fit in model B (ΔOFV_B = 12, 3 degrees of freedom). Note that ΔOFV indicates difference in objective function values between two fitted models (see chapter 2 for full explanation of OFV). The correlations are shown as the off diagonal numbers in the lower triangular matrices 4.3 (model A) and 4.4 (model B).

$$\begin{array}{c} \begin{array}{ccc} int_i & asy_i & c_i \end{array} \\ \begin{pmatrix} 1.16 & & \\ -0.02 & 0.47 & \\ -0.23 & 0.01 & 1.19 \end{pmatrix} \begin{array}{c} int_i \\ asy_i \\ c_i \end{array} \end{array} \tag{4.3}$$

$$\begin{array}{c} \begin{array}{ccc} int_i & asy_i & c_i \end{array} \\ \begin{pmatrix} 1.09 & & \\ -0.03 & 0.51 & \\ -0.11 & 0.11 & 0.75 \end{pmatrix} \begin{array}{c} int_i \\ asy_i \\ c_i \end{array} \end{array} \tag{4.4}$$

Final covariate model for dataset A

$$\begin{aligned} int_i = & \begin{cases} -1.32 - 0.16(\text{age} - 5.78) & \text{age} \leq 5.78 \\ -1.32 - 0.023(\text{age} - 5.78) & \text{age} > 5.78 \end{cases} + \begin{cases} 0 & \text{AIDS}^- \\ -0.725 & \text{AIDS}^+ \end{cases} \cdots \\ & \cdots + \begin{cases} 0 & \text{UK/Ireland} \\ -0.017 & \text{Rest of Europe} \\ -0.909 & \text{Thailand} \\ -0.305 & \text{Russia/Ukraine} \end{cases} \end{aligned} \tag{4.5}$$

$$c_i = 1.89 \times \begin{cases} 1 & \text{UK/Ireland} \\ 1.442 & \text{Rest of Europe} \\ 0.859 & \text{Thailand} \\ 1.407 & \text{Russia/Ukraine} \end{cases} \tag{4.6}$$

$$asy_i = -0.275 - 0.0433(\text{age} - 5.78) \tag{4.7}$$

Final covariate model for dataset B

$$\begin{aligned} int_i = &\begin{cases} -1.26 - 0.14(\text{age} - 5.4) & \text{age} \leq 5.4 \\ -1.26 - 0.043(\text{age} - 5.4) & \text{age} > 5.4 \end{cases} + \begin{cases} 0 & \text{AIDS}^- \\ -0.891 & \text{AIDS}^+ \end{cases} \cdots \\ &\cdots - 0.171(\text{blogvl} - 5.08) + \begin{cases} 0 & \text{UK/Ireland} \\ 0.148 & \text{Rest of Europe} \\ -0.885 & \text{Thailand} \\ 0.122 & \text{Russia/Ukraine} \end{cases} \end{aligned} \tag{4.8}$$

$$c_i = 1.52 \times \begin{cases} 1 & \text{AIDS}^- \\ 1.37 & \text{AIDS}^+ \end{cases} \times \begin{cases} 1 & \text{UK/Ireland} \\ 1.603 & \text{Rest of Europe} \\ 1.128 & \text{Thailand} \\ 0.937 & \text{Russia/Ukraine} \end{cases} \tag{4.9}$$

$$asy_i = \begin{cases} -0.219 - 0.00846(\text{age} - 5.4) & \text{age} \leq 5.4 \\ -0.219 - 0.0538(\text{age} - 5.4) & \text{age} > 5.4 \end{cases} \tag{4.10}$$

4.4.3 Factors affecting pre-ART and long-term age-adjusted CD4 counts

Final covariate model A

On average, children started ART with a ln(CD4-for-age) value of -1.32 (CD4 count of 323 cells/μL) corresponding to a CD4 count of 1210 cells/μL in a healthy HIV un-infected child of the same age. This estimate was for an average 5.8 yr old child with negative pre-ART AIDS status and from the CHIPs cohort (UK/Ireland). The covariate analysis demonstrated that a piecewise linear relationship was the best predictor of the effect of age at start of ART on pre-ART ln(CD4-for-age). For children $\leq$5.8 years, pre-ART ln(CD4-for-age) was 0.16 units lower for every year older at start of ART whilst in children >5.8 years, pre-ART decreased by 0.023 units for every yearly increase in age at start of therapy (Figure 4.4C

Table 4.1: Parameter estimates for the final multivariate model for the dataset A

Fixed effects	Estimate	SE	*p-values*	Ω
Intercept				
int	-1.32	0.098	< 0.0001	1.31
int:age ≤ 5.78	-0.16	0.024	-	
int:age > 5.78	-0.023	0.016	-	
int:AIDS$^+$	-0.725	0.106	-	
int:Europe	-0.017	0.103	-	
int:Thailand	-0.909	0.101	-	
int:Russia/Ukraine	-0.305	0.129	-	
c				
c	1.89	0.16	< 0.0001	1.32
c:Europe	0.442	0.2	-	
c:Thailand	-0.141	0.103	-	
c:Russia/Ukraine	0.407	0.304	-	
Asymptote				
asy	-0.275	0.016	< 0.0001	0.209
asy:age	-0.0433	0.003	-	
residual error	0.104	0.0003	-	

The reference case is an HIV-infected child from the UK/Ireland starting ART at 5.8 yrs (median age in dataset A) and with negative pre-ART AIDS status. Hence the estimated parameters *int*, *asy* and *c* are for this particular child. "Age" represents age at start of ART, SE: standard errors of estimates,'int:age" ("parameter:covariate") represents the covariate interaction between age at start of ART and pre-ART ln(CD4-for-age), Ω: Variances on random effects.

and equation 4.5). Other factors influencing pre-ART ln(CD4-for-age) include AIDS status at start of therapy and cohort effect. Notably, children from Thailand, Russia and Ukraine started ART with much lower CD4 counts compared to children from UK/Ireland. In the final covariate model, the only predictor of long-term ln(CD4-for-age) was age at start of therapy where long-term ln(CD4-for-age) was 0.04 units lower per year older at ART initiation. The average predicted long-term ln(CD4-for-age) for the median 5.8 yr old child was -0.275 (CD4 count of 920 cells/μL), equivalent to a CD4 count of 1211 cells/μL in a healthy HIV uninfected 5.8 yr old child. There was an effect of EPPICC cohort on rate of increase in ln(CD4-for-age) where children from Russia, Ukraine and other parts of Europe (excluding the UK) had a higher increase in rate of age-adjusted CD4$^+$ T cell recovery. Random effect on int_i (1.31) was greater than random effect on asy_i (0.21) reflecting greater variability in observed pre-ART values. Surprisingly, pre-ART base log HIV viral load was not a significant predictor of pre-ART ln(CD4-for-age) in this model. Full effect sizes from the final covariate model are given in Table 4.1. Plots generated from final covariate model A predictions are shown in Figure 4.3A and 4.3C.

Table 4.2: Parameter estimates for the final multivariate model for the dataset B

Fixed effects	Estimate	SE	*p-values*	Ω
Intercept				
int	-1.26	0.095	< 0.0001	1.17
int:age ≤ 5.4	-0.14	0.02	-	
int:age > 5.4	-0.043	0.016	-	
int:AIDS$^+$	-0.891	0.156	-	
int:Europe	0.148	0.089	-	
int:Thailand	-0.89	0.157	-	
int:Russia/Ukraine	0.122	0.117	-	
int:blogvl	-0.171	0.038	-	
c				
c	1.52	0.148	< 0.0001	0.58
c:Europe	0.603	0.24	-	
c:Thailand	0.128	0.16	-	
c:Russia/Ukraine	-0.0633	0.27	-	
c:AIDS$^+$	0.366	0.179	-	
Asymptote				
asy	-0.219	0.031	< 0.0001	0.26
asy:age ≤ 5.4	-0.0085	0.0092	-	
asy:age > 5.4	-0.0538	0.0068	-	
residual error	0.116	0.012	-	

The reference case is a child from the UK/Ireland starting ART at 5.4 yrs (median age in dataset B), base log viral load of 5.08 $\log_{10}$ copies/mL, HIV-monoinfected, with negative pre-ART AIDS status. Hence the estimated parameters *int*, *asy* and *c* are for this particular child. "Age" represents age at start of ART, SE: standard errors of estimates, int:age ('parameter:covariate") represents the covariate interaction between age at start of ART and pre-ART ln(CD4-for-age), Ω: Variances on random effects.

Final covariate model B

The final covariate model for dataset B as given in equations 4.8 to 4.10 shows that the average 5.4 yr old child from the UK/Ireland cohort with pre-ART base log viral load 5.08 $\log_{10}$ copies/mL and negative pre-ART AIDS status is predicted to have a pre-ART ln(CD4-for-age) of -1.26 (CD4 count of 356 cells/μL). Pre-ART ln(CD4-for-age) decreased by 0.2 units for every unit increase in HIV $\log_{10}$ viral load and similar to the final covariate model from dataset A, there was an EPPICC cohort effect on pre-ART with children from Thailand having the lowest int_i values. Both pre-ART and long-term ln(CD4-for-age) were strongly associated with pre-ART age via a piecewise linear relationship ($p < 0.0001$): compared to younger children ($\leq$5.4 yrs), older children (>5.4 yrs) experienced six times unit decrease in long-term ln(CD4-for-age) for every year older (Figure 4.4D). The rate of increase in age-adjusted CD4 counts (c_i) was strongly associated with pre-ART AIDS status and individual EPPICC cohort: children with AIDS diagnosis at start of therapy experienced 1.4 times higher recovery in their ln(CD4-for-age). Complete

EPPICC Data

	No.		No.
Total	2204	CD4 counts per child	14 (7-19)
Country		**Sex**	
Portugal		Female	1189 (54%)
Lisbon	5 (0.2%)	Male	1015 (46%)
Porto	12 (0.5%)	**AIDS diagnosis**	
Romania	18 (0.8%)	Yes	256 (12%)
Switzerland	20 (0.9%)	No	1948 (88%)
Belgium	25 (1.1%)	**Region**	
Sweden	45 (2.0%)	Europe (Other)	674 (31%)
Poland	47 (2.1%)	Russia/Ukraine	273 (12%)
Italy	98 (4.5%)	Thailand	393 (18%)
Russia	106 (4.8%)	UK/Ireland	864 (39%)
France	111 (5%)	**Age at HIV diagnosis**	
Holland	114 (5.2%)	Known	2023 (91.8%)
Spain		Unknown	181 (8.2%)
Barcelona	53 (2.4%)	Median (yrs)	3 (0.6-7.6)
Other	126 (5.7%)	**Age at sART**	
Ukraine	167 (7.6%)	Median (yrs)	5.6 (1.4-9.9)
Thailand	393 (17.8%)	**ART Regimen**	
UK	864 (39.2%)	3 NRTIs	52 (2.4%)
Baseline CD4		Other	54 (2.45%)
Median (cells/μL)	424 (200-975)	NNRTI + 3 NRTIs	123 (5.6%)
Baseline log10VL		NVP + 2 NRTIs	567 (25.7%)
Median	5.1 (2.4-5.6)	EFV + 2 NRTIs	669 (30.35%)
HIV Transmission		Boosted PI + NRTI	739 (33.5%)
Perinatal	2204	NVP +	695

Table 4.3: Median values are stated in units given in table with interquartile range in parenthesis, ART (antiretroviral therapy), Ab (antibody), sART (start of ART), VL (viral load), age at sART: age at start of therapy. The data was divided into Training and testing datasets each containing 1102 patients. NRTI: nucleoside reverse transcriptase inhibitor, NNRTI: non-nucleoside reverse transcriptase inhibitor, NVP: nevirapine, EFV: efavirenz, PI: protease inhibitor, NVP + represents number of children with NVP-based therapies

effect sizes are given in Table 4.2. Plots generated from final covariate model B predictions are shown in Figures 4.3B and 4.3D.

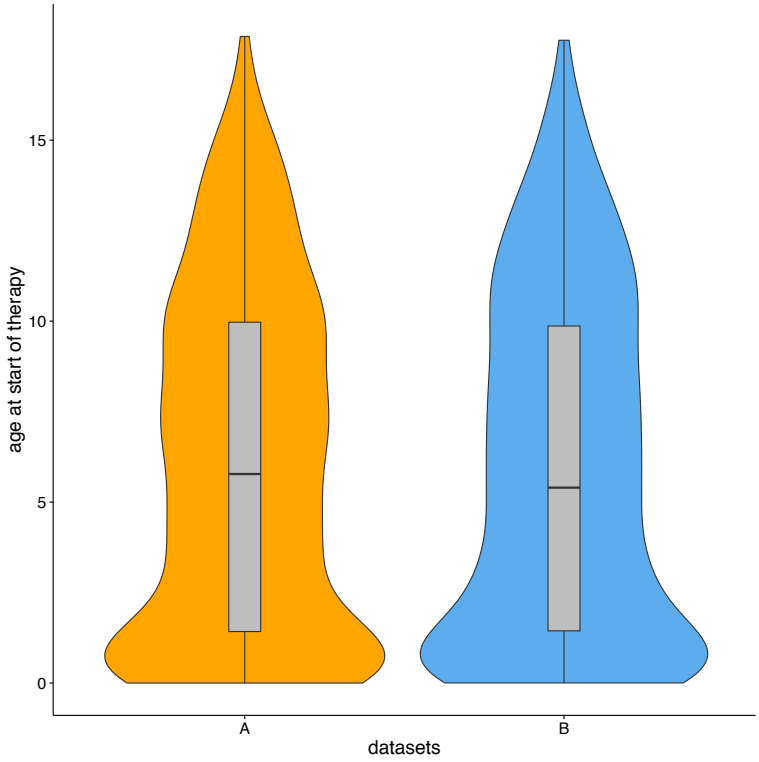


Figure 4.2: Violin plots of age at start of therapy for datasets A and B. The outer margin of the violin shows a rotated density plot demonstrating the high numbers of younger children in both cohorts. The inner plot is a box plot showing the respective quartiles [Hintze and Nelson, 1998, Wickham, 2009].

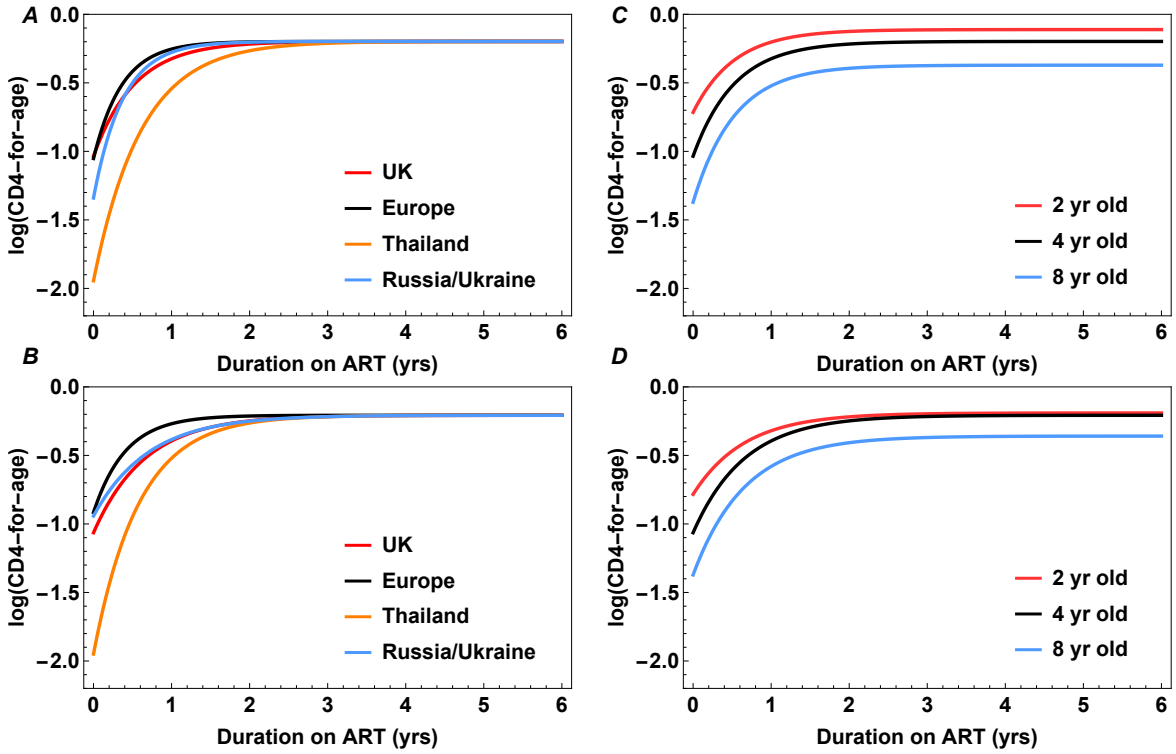


Figure 4.3: **A**: Effect of individual EPPICC Cohorts on overall ln(CD4-for-age) trajectories generated from the final covariate model A in a 4 yr old child with negative pre-ART AIDS status, **B**: Effect of individual EPPICC Cohorts on overall ln(CD4-for-age) trajectories for final covariate model B in a 4 yr old child with negative pre-ART AIDS status and a viral load of 5.08 log10 copies/mL, **C**: Plots of ln(CD4-for-age) trajectories in a 2 yr old, 4 yr old and an 8 yr old child all from the UK and with negative pre-ART AIDS status (final covariate model A), **D**: Plots of ln(CD4-for-age) trajectories in a 2 yr old, 4 yr old and an 8 yr old child all from the UK with negative pre-ART AIDS status and a viral load of 5.08 log10 copies/mL (final covariate model B)

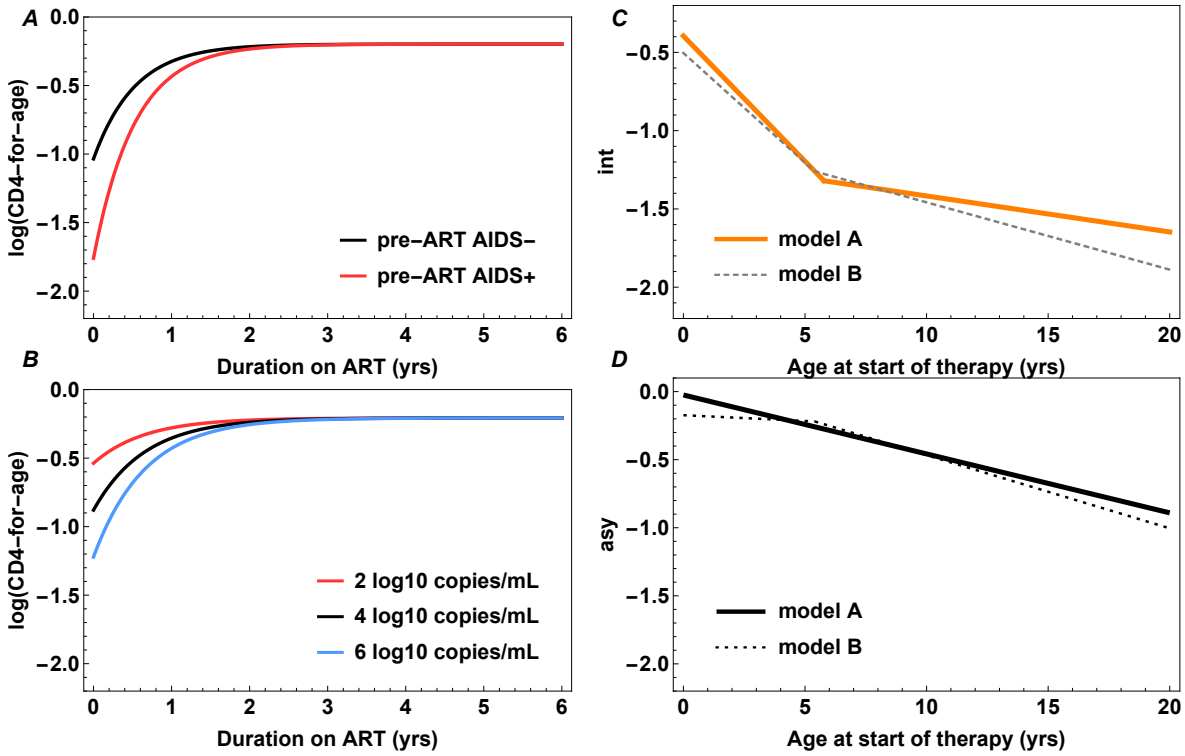


Figure 4.4: **A**: Effect of pre-ART AIDS status on ln(CD4-for-age) predictions in a 4 yr old old from the UK Cohort (predictions are from model A), **B**: Plots of ln(CD4-for-age) predictions in a 4 yr old child (UK) with negative pre-ART AIDS status and with viral loads of 2, 4 and 6 log10 copies/mL Effect of AIDS status on ln(CD4-for-age) predictions in a 4 yr old old from the UK Cohort, **C, D**: Effect of age at start of therapy on pre-ART and long-term ln(CD4-for-age) showing piece-wise linear relationships in three of the four effects. Continuous lines correspond to model A whilst dashed lines represent model B.

4.4.4 Diagnostic plots

Diagnostic plots for the final covariate models A and B are shown in Figures 4.6A-D and 4.5A-B. In addition they also show diagnostic plots generated from models A and B fitted to datasets B and A respectively (Figures 4.6E-H, 4.5C-D). The residuals are approximately normally distributed with a mean of zero and with the exception of a few very low outliers and some overpredictions after 4 years on ART (much more evident in plots from dataset B-Figures 4.6B,F). No other obvious trends/biases are evident with duration of therapy or population predictions indicating their independence with time and population predictions (unbiased heteroscedastic pattern). The low outlying residuals could be due to inaccuracies in predicting children with very low CD4 counts majority of whom are from the CHIPs and the Thailand cohort. Plots of observed ln(CD4-for-age) against predicted ln(CD4-for-age) in Figure 4.5 further confirm the absence of biases in the model predictions.

Visual predictive checks generated from both final covariate models A and B are shown in Figure

4.7. They show that the model simulated bands mostly match the observed data well especially at the lower (2.5th) and median (50th) centiles of dataset A (Figures 4.7a and 4.7b) with the exception of minor overpredictions notable after 4 years of ART. Similar trends hold for the simulations overlaid on dataset B (Figures 4.7c and 4.7d). Notably, overpredictions are seen after 3 yrs of ART. The reason for these overpredictions after 3 years on ART can be traced back to dataset B (Figure 4.1D) where more children with very low CD4 trajectories can be seen after 4 years of ART in comparison to dataset A. Overpredictions are noted at higher ln(CD4-for-age) values in both datasets A and B which could indicate that the assumption of log-normality especially at these values is inaccurate (Figures 4.8). The histograms in Figure 4.8 show that the observed ln(CD4-for-age) are skewed to the left and data is particularly sparse at ln(CD4-for-age)$>$ 0.

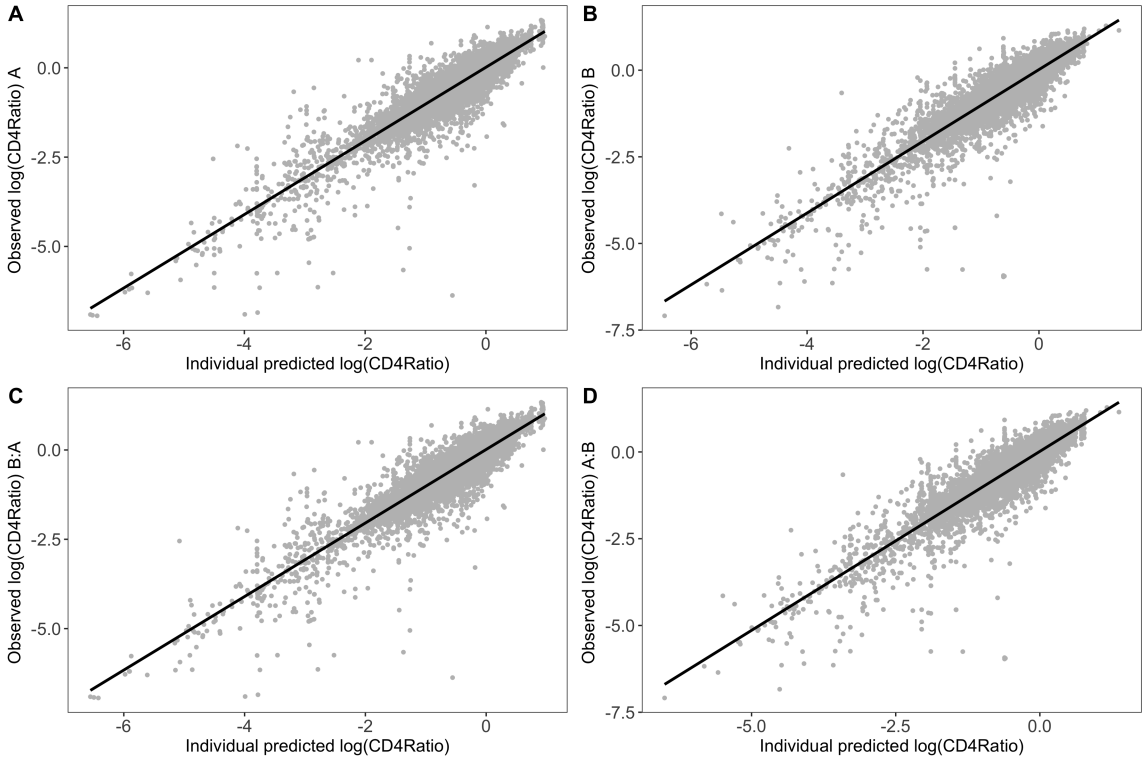


Figure 4.5: **A,C**: Observed ln(CD4-for-age) plotted against individual predicted ln(CD4-for-age) for the dataset A (Plot A-covariate model A, Plot C- covariate model B applied to dataset A). **B,D**: Observed ln(CD4-for-age) plotted against individual predicted ln(CD4-for-age) for dataset B (Plot A-final covariate model B, Plot B-final covariate model A applied to dataset B).

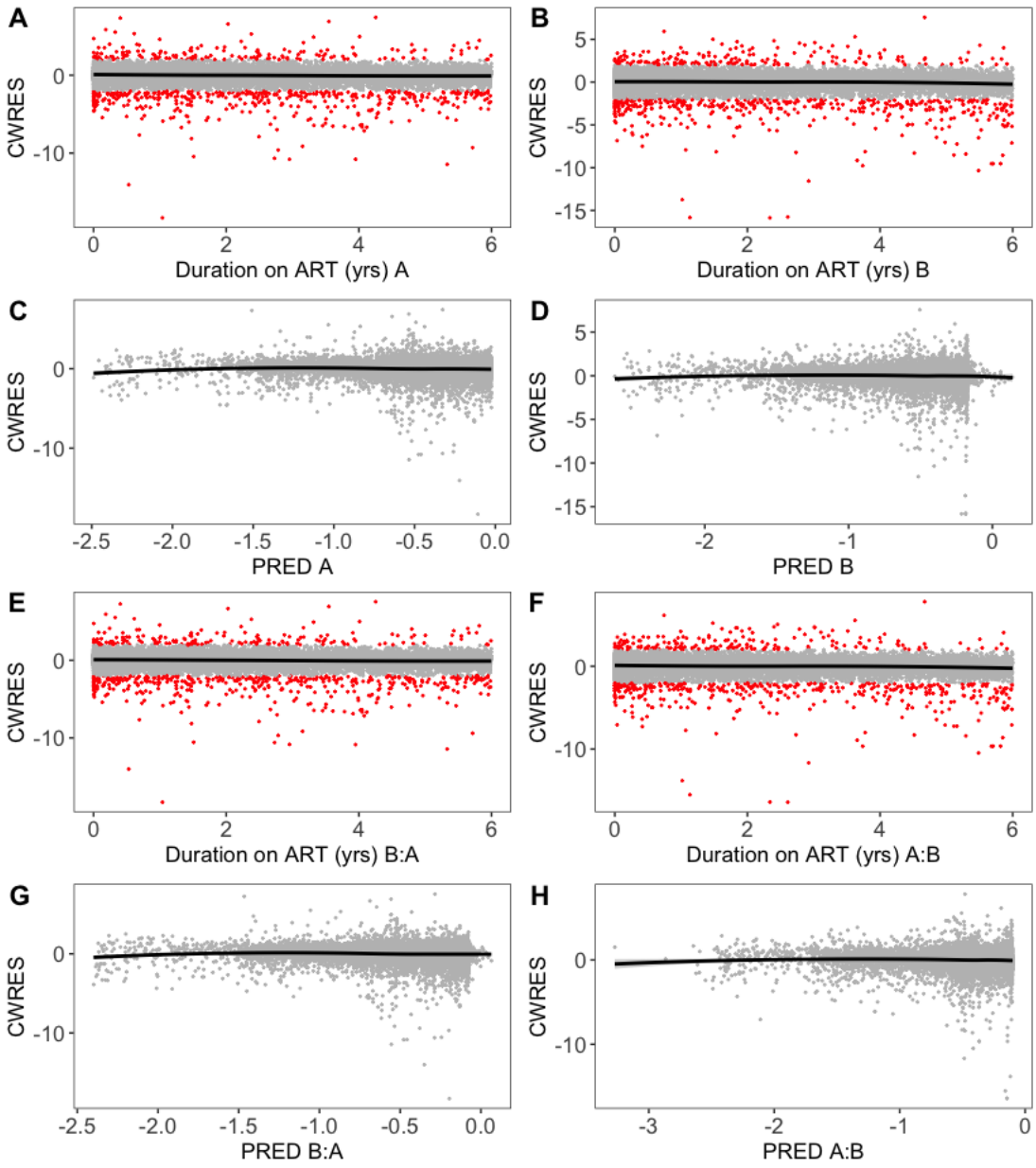
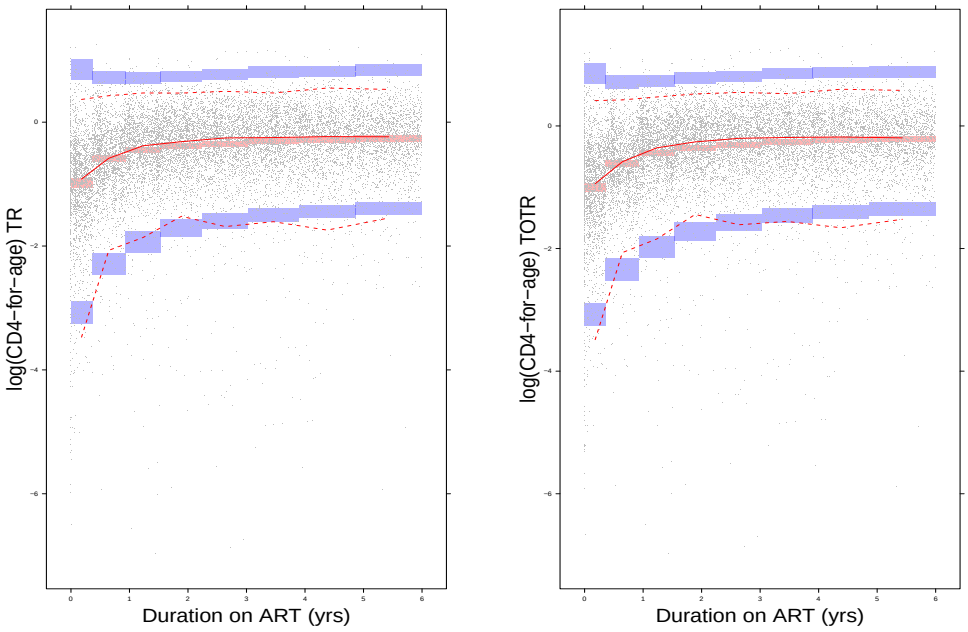
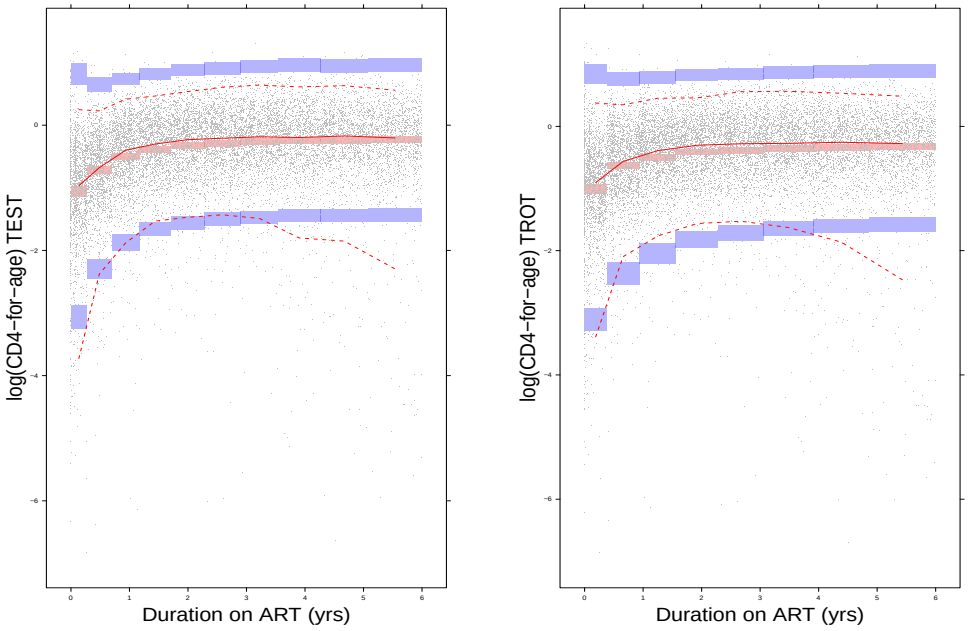


Figure 4.6: A,C: Plots of conditional weighted residuals for the final covariate model A plotted against duration on ART and population predictions respectively. G,E: Plots of conditional weighted residuals for the final covariate model B applied to the dataset B plotted against duration on ART and population predictions respectively. B,D: Plots of conditional weighted residuals for the final covariate model B plotted against duration on ART and population predictions respectively. F,H: Plots of conditional weighted residuals for the final covariate model A applied to dataset B plotted against duration on ART and population predictions respectively. Red data points represent standardised residuals > 2 standard deviations from the mean (0). Plots A,C,E,G (left column) represent results generated from the dataset A whilst plots B, D, F and H (right column) are results generated from dataset B.



(A)

(B)



(C)

(D)

Figure 4.7: (a): Visual predictive check (VPC) for the final multivariate model A overlaid with dataset A. (b): VPC for model B applied to the dataset A. (c): Final covariate model B overlaid with dataset B. (d): final covariate model A applied to the dataset B. The gray dots correspond to the observed dataset whilst the dashed and solid red lines correspond to 2.5th, 50th and 97.5th centiles respectively from the observed data. The red and blue bands represent the 95% confidence intervals generated from the model-simulated centiles. The x-axis shows duration on ART (first 6yrs of therapy).

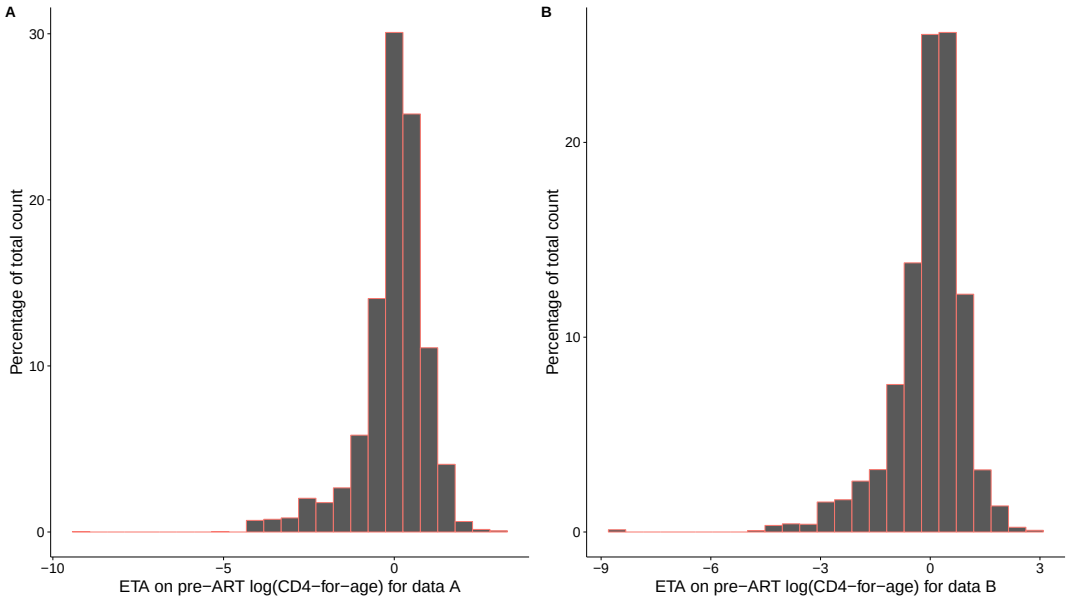


Figure 4.8: Histograms of random effects (ETA) on pre-ART log(CD4-for-age) for datasets A (**A**) and B (**B**).

4.4.5 Predictors of poor immune response

To identify poor immune responders, further analysis of the distribution of residuals was undertaken. In the non-linear mixed-effects framework, each parameter value is obtained by first estimating the average population parameter value and then adding a random effect to it. In addition to this, any other variability that cannot be explained by the fixed effect and the random effect is attributed to a residual error which could be due to model misspecification or laboratory errors. As such, children whose CD4 counts cannot be sufficiently explained by the estimated population mean with additional random variability will end up having high residual errors which may be identified as outliers on a diagnostic plot. As explained in section 4.3.3, children whose individual predicted ln(CD4-for-age) are greater than the observed ln(CD4-for-age) by more than 2σ from the mean were analysed further as potential outliers. Under the assumptions of normality, we would expect 95% of the residuals to be within 2σ of the mean. It is therefore expected that the presence of poor responders in the data would increase the number of residuals outside the 2σ threshold and beyond the anticipated 5%.

Analysis of residuals from final covariate model A and model B applied to dataset A

Further analysis of the intersection of children with outlying residuals between the results generated from the final covariate model B fitted to dataset A and the final covariate model A revealed a total of 219 children with at least one "outlying residual" value $> 2\sigma$ from the mean. Five percent of the residuals

from both models were outside the 2σ window ($\sigma = 0.3225$). The total number of data points in each child plotted against the outlying residuals is shown in Figure 4.9A and illustrates that majority of the outlying residuals (75%) were from children with one outlying data point. As expected, the number of outlying residuals are seen to increase proportionately with total number of data points.

Analysis of residuals from final covariate model B and model A fitted to dataset B

Analysis of the residuals generated from final covariate model B and final covariate model A fitted to dataset B gives a very similar picture to the findings in dataset A. The intersection of all outlying residuals from both results (model A applied to B and model B) showed a total of 210 HIV-infected children with residuals $> 2\sigma$ from the mean ($\sigma = 0.3362$). As shown in Figure 4.9B, 69% of the children have one outlying residual.

The analysis of residuals from both datasets described above shows that poor immune responders cannot be identified purely from the distribution of the residuals given that these findings are to be expected under the assumptions of normality. Hence an alternative approach is needed to identify poor immune responders in the EPPICC cohort.

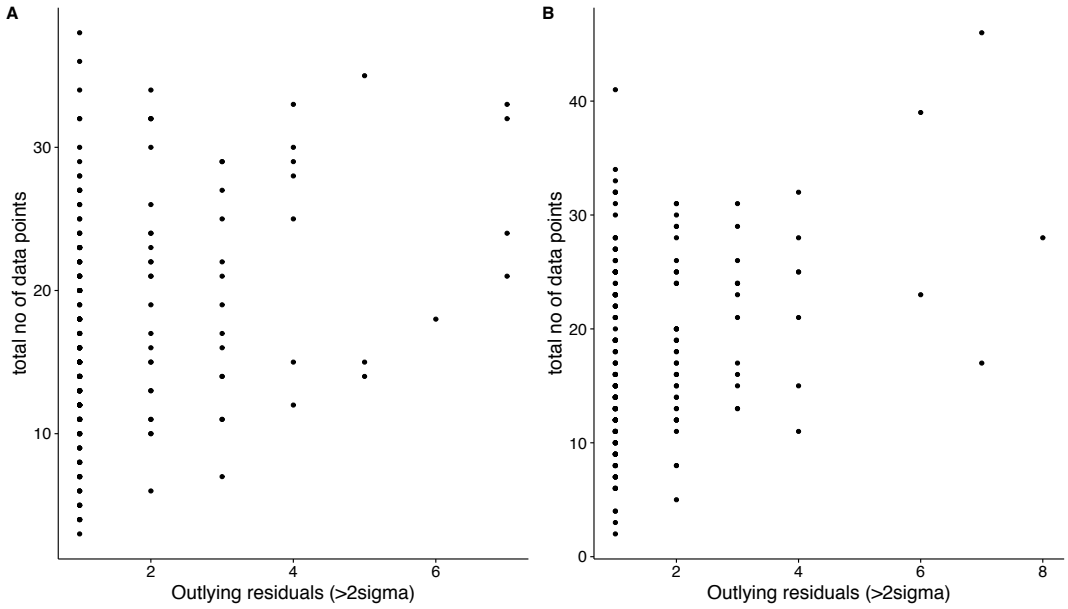


Figure 4.9: Plots of total no of data points against outlying residuals in both dataset A (**A**) and dataset B (**B**). Each dot represents an individual child.

4.4.6 Individual CD4$^+$ T cell predictions

In this section, the WHO criteria for HIV-associated immune deficiency outlined in Table 4.6 [World Health Organisation, 2016b] were used in identifying poor immune responders on the basis of their predicted long-term CD4$^+$ counts. To facilitate this, all predicted ln(CD4-for-age) were converted back to CD4 counts using the function described in appendix B.2. According to the WHO guidelines, advanced and severe HIV-associated immunodeficiency are defined immunologically as CD4$^+$ concentrations within the range 200-349 cells/μL and $<$ 200 cells/μL respectively for children $\geq$ 5 yrs. For the purpose of this study, poor responders are hereby defined as children with individual predicted final CD4 counts $<$ 350 cell/μL. Using the above WHO thresholds, the asymptotic model was able to identify two distinct groups according to long-term predicted ln(CD4-for-age): mild to normal immune responders and advanced to severe poor immune responders. A total of 80 children (7.3%) from the dataset A and 82 children (7.4%) from dataset B fell into the category of advanced to severe immunodeficiency. However, when considering only virally suppressed children (HIV viral load $<$ 400 copies/mL), 38 children (4.3%) from dataset A and 34 children (3.9%) from dataset B were found to have advanced or severe immunodeficiency. Table 4.4 outlines summary of poor responders in both datasets.

In dataset A, children within the advanced HIV immunodeficiency group had a median age at start of therapy of 11 yrs (IQR: 8.1-13.2 yrs) which is almost double the median age of 5.8 yrs (IQR: 1.4-10 yrs) in the entire dataset. Of the 54 children in the advanced immunodeficiency group, only three were diagnosed with AIDS at start of therapy. A plot of all advanced poor responders with suppressed viral load for dataset A is shown in Figure 4.11 and illustrates a marked variability between the population-level estimates (blue line) and individual estimates (red line) which could not be explained by either the covariates or the random effects. Of note, the vast majority of the individual level estimates are seen to be well below the population fixed effect which is mostly above the 500 cells/μL threshold for mild-normal immunodeficiency category outlined in Table 4.6. There are instances of over and under-estimation of the final CD4 predictions but the overall trend within the observations were captured by the model estimates. As shown in Figure 4.12, very similar trends are observed in the cohort of advanced poor immune responders from dataset B. Median age at start of therapy of 11.5 yrs (IQR: 9.5-14.2 yrs) was much higher compared to the median age of 5.4 yrs for the rest of the cohort. Six of the 51 children were diagnosed with AIDS at start of therapy and there were more children in dataset B with incomplete CD4 profiles over the duration of ART.

In datasets A and B, the largest number of virally suppressed poor responders (advanced and severe)

was from the UK/Ireland (50% and 56%) followed by Thailand (18%). Individual plots of severe poor immune responders for virally suppressed children in both datasets is shown in Figure 4.13. A summary of all children in the poor immune responders category can be found in Table 4.5. Median predicted final CD4 count in virally suppressed poor responders was 282 cells/μL (241-307) for cohorts in dataset A. These values are very comparable to the corresponding median observed final CD4 count of 271 cells/μL (224-404). In comparison, normal responders from the same dataset had a median predicted final CD4 count of 843 cells/μL (637-1146).

For children in dataset B, median predicted final CD4 count in virally suppressed poor responders was 297 cells/μL (213-319) whilst their corresponding median observed final CD4 count was 271 cells/μL (224-362). The normal responders from the same cohort had a median predicted final CD4 count of 850 cells/μL (645-1137). Two deceased children from the Thailand cohort were identified amongst the poor immune responders where cause of death listed include cardiovascular and HIV-related deaths. Amongst the virally suppressed poor immune responders in dataset A, a total of 8 children (21%) were on a nevirapine-containing regimen. In comparison, 265 children (31%) received nevirapine amongst the mild-normal responders. Similar Figures are noted in dataset B (Table 4.5). The violin plots in Figure 4.10 show a marked difference in age at start of therapy between the three immunological groups with the advanced to severe responders having a much higher age at start of therapy (Also table 4.5).

Summary of poor responders using WHO criteria for HIV-associated immunodeficiency

Category	Dataset A	%	Dataset B	%
All patients				
Mild-Normal	1022	92.7	1020	92.6
Advanced	54	4.9	51	4.6
Severe	26	2.4	31	2.8
vl $\leq$ 400 copies/mL				
Yes	883	80.1	872	79.1
No	219	19.9	230	20.9
Mild-Normal	845	95.7	838	96.1
Advanced	32	3.6	26	3.0
Severe	6	0.7	8	0.9

Table 4.4: Table showing numbers of poor responders by comparing predicted long-term CD4 counts with WHO guidelines for HIV-associated immunodeficiency. vl: HIV viral load. Normal, Mild, Advanced and Severe represents the four categories of immunodeficiency. "Severe" represents children with predicted long-term CD4 concentration <200 cells/μL, "Advanced" represents children with long-term predicted CD4 count <350 but >200, "Mild-Normal" represents children with CD4 concentration $\geq$350 cells/μL. Predicted CD4 counts considered here are the final CD4 count over the 6 years of therapy.

WHO summary table for advanced to severe category

	A	%	B	%
Cohort				
UK	19	50	19	55.9
Thailand	7	18	6	17.6
Spain	2	5.3	5	14.7
Italy	3	7.9	-	-
Ukraine	3	7.9	2	5.9
France	1	2.6	-	-
Holland	1	2.6	-	-
Sweden	1	2.6	-	-
Belgium	1	2.6	-	-
Russia	-	-	1	2.9
Poland	-	-	1	2.9
Total	38		34	-
Age at sART	-	-	-	
Median (yrs)	11	-	11.5	-
Median (IQR)	8-13	-	10-15	-
pre-ART AIDS				
Yes	2	5.3	5	14.7
No	36	94.7	29	85.3
Nevirapine				
Yes (AS)	8	21	8	24
Yes (MN)	265	31	264	32

Table 4.5: Summary table for children with advanced to severe immunodeficiency with complete suppression of viral load ($\leq$ 400 copies/mL). Under Nevirapine, AS: advanced to severe responders, MN: mild to normal responders.

WHO $CD4^+$ T cell criteria for HIV immunodeficiency

Immunological marker	$\leq$ 11 months	12-35 months	35-59 months	$\geq$ 5 yrs
None	-	-	-	$>$ 500
Mild	-	-	-	350-499
Advanced	-	-	-	200-349
Severe				
%CD4+	$<$ 25%	$<$ 20%	$<$ 15%	$<$ 15%
CD4 count (cells/μL)	$<$ 1500	$<$ 750	$<$ 350	$<$ 200

Table 4.6: WHO criteria for severe immunodeficiency. None, Mild, Advanced are all in CD4 counts for the $\geq$ 5yrs category [World Health Organisation, 2016b].

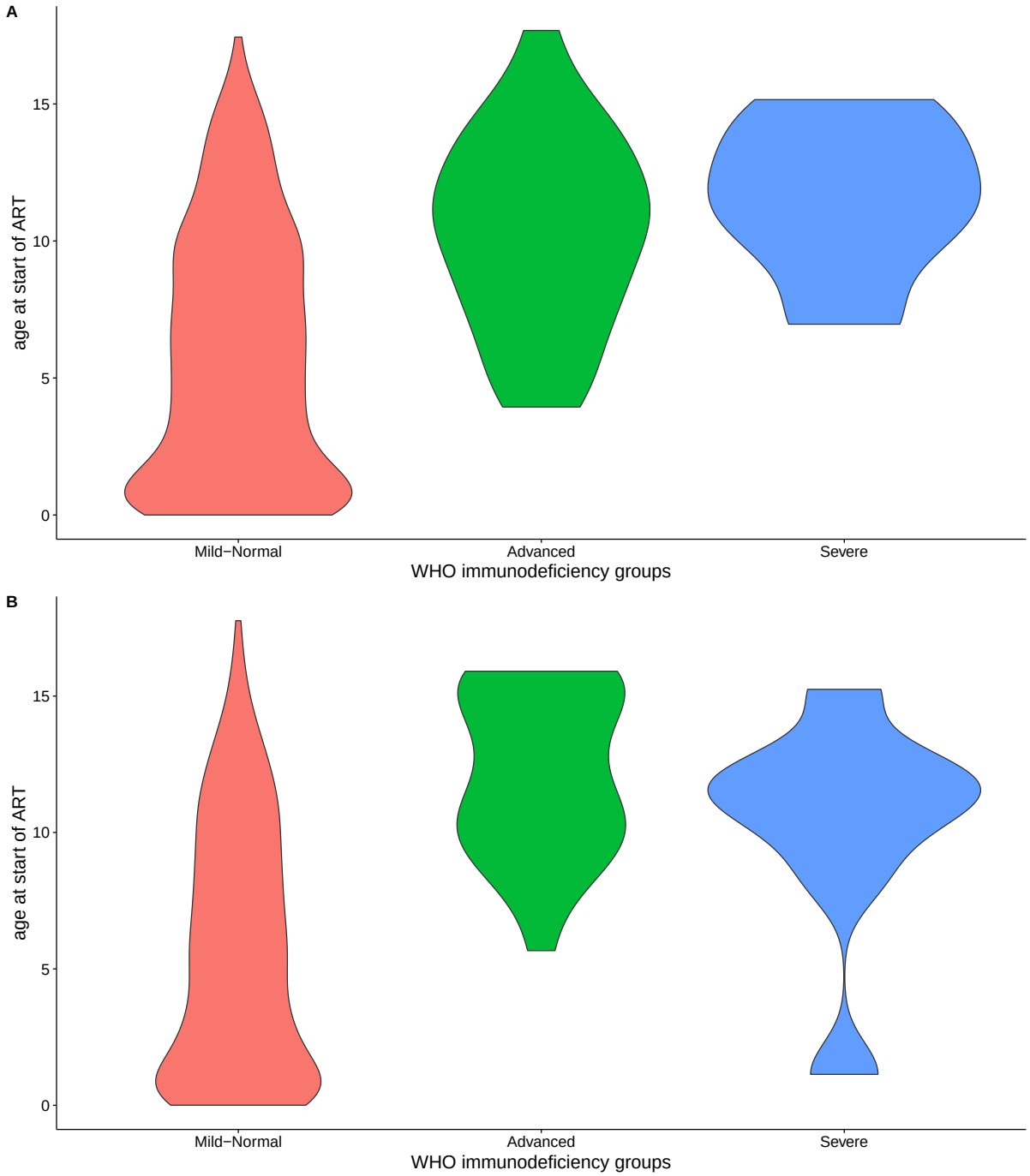


Figure 4.10: Violin plots of age at start of therapy in all three categories of immune responders based on WHO criteria for immunodeficiency (dataset A-above and dataset B-below) [Hintze and Nelson, 1998]

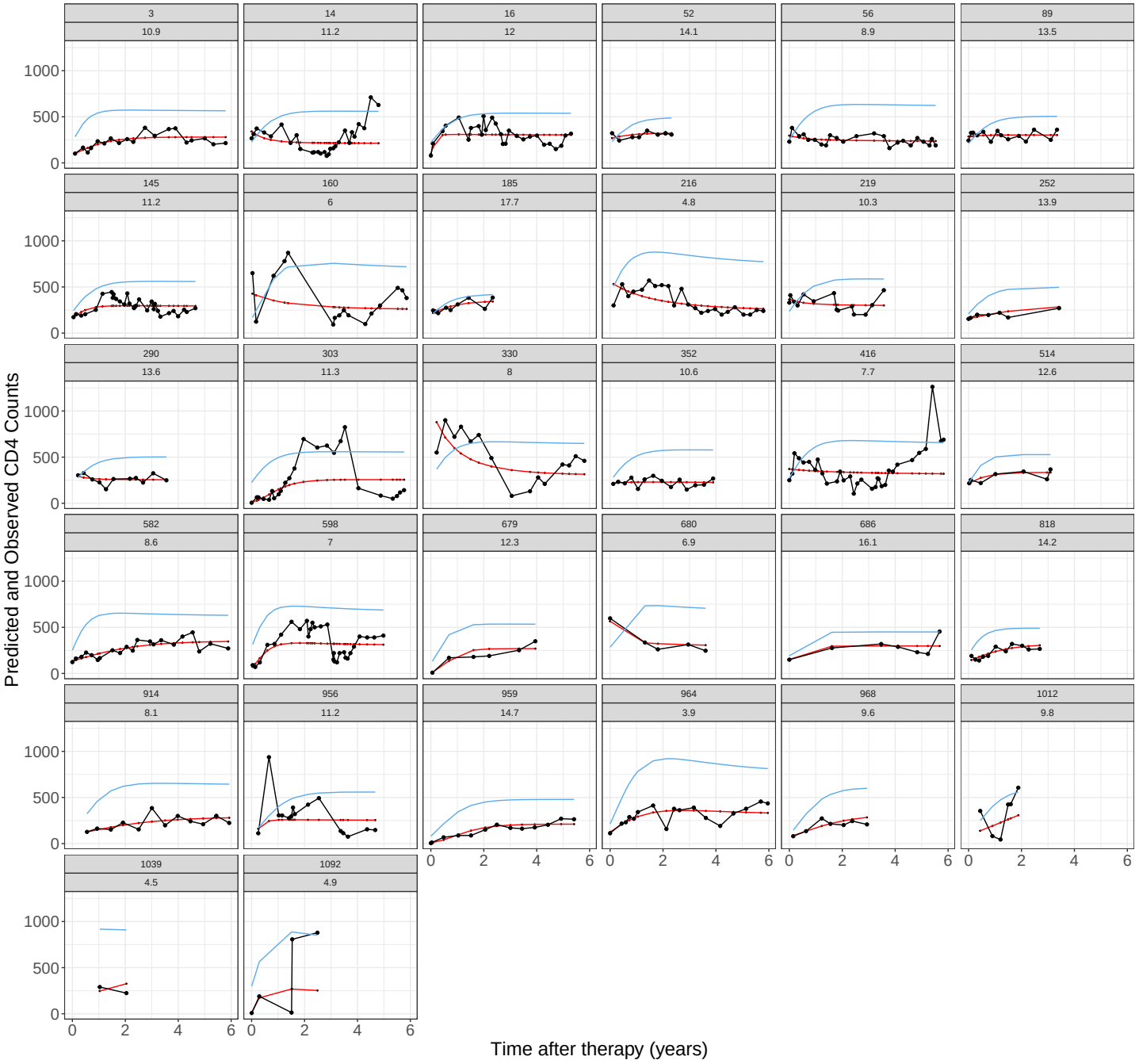


Figure 4.11: Plots of virally suppressed children within the advanced HIV immune deficiency category for dataset A. Blue line: population prediction; Red line: individual predictions; Black line: observed CD4 counts. Top panel of each box shows the ID and age at start of ART for each child.

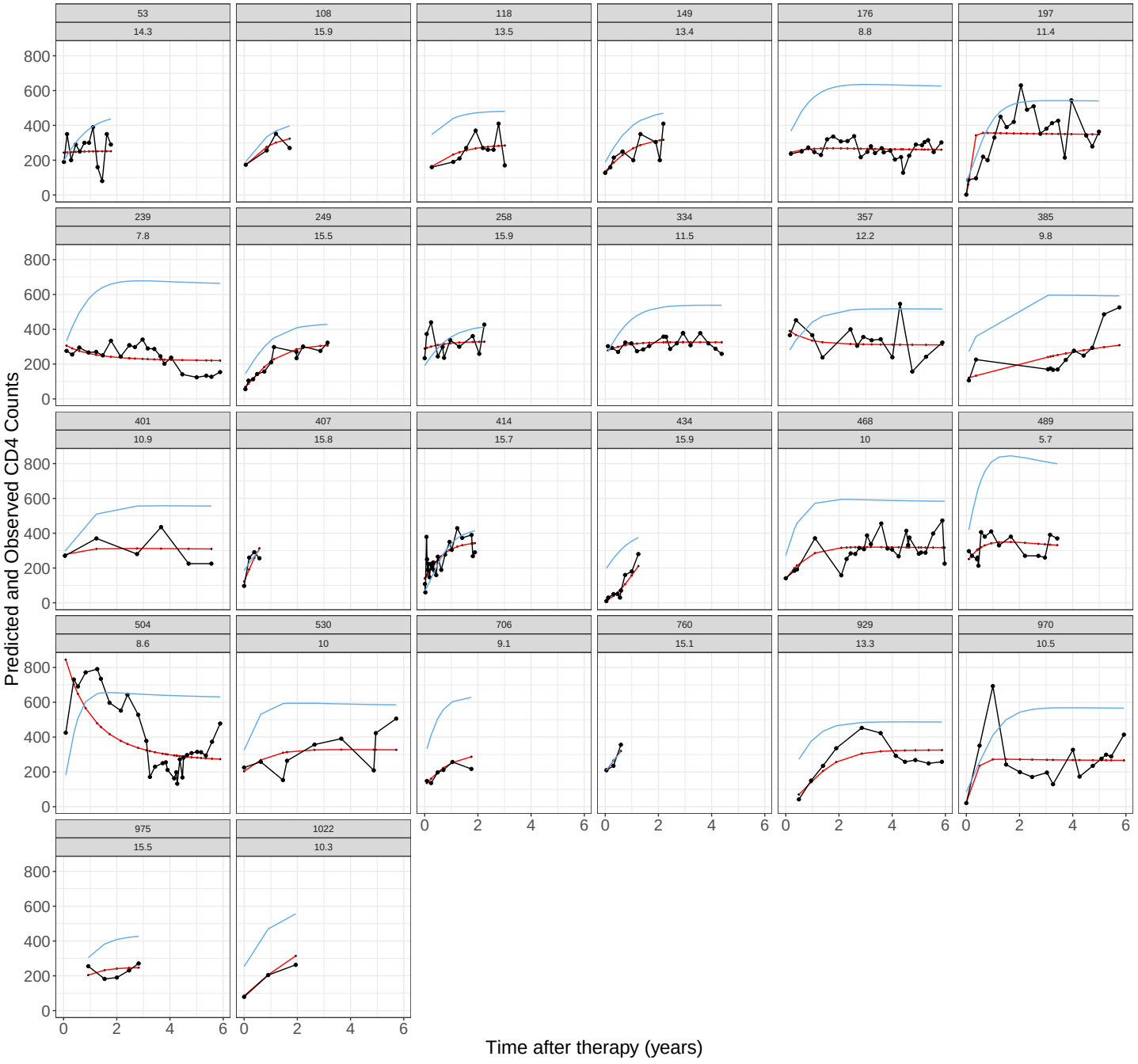


Figure 4.12: Plots of virally suppressed children within the advanced HIV immune deficiency category for dataset B. Blue line: population prediction; Red line: individual predictions; Black line: observed CD4 counts. Top panel of each box shows the ID and age at start of ART for each child.

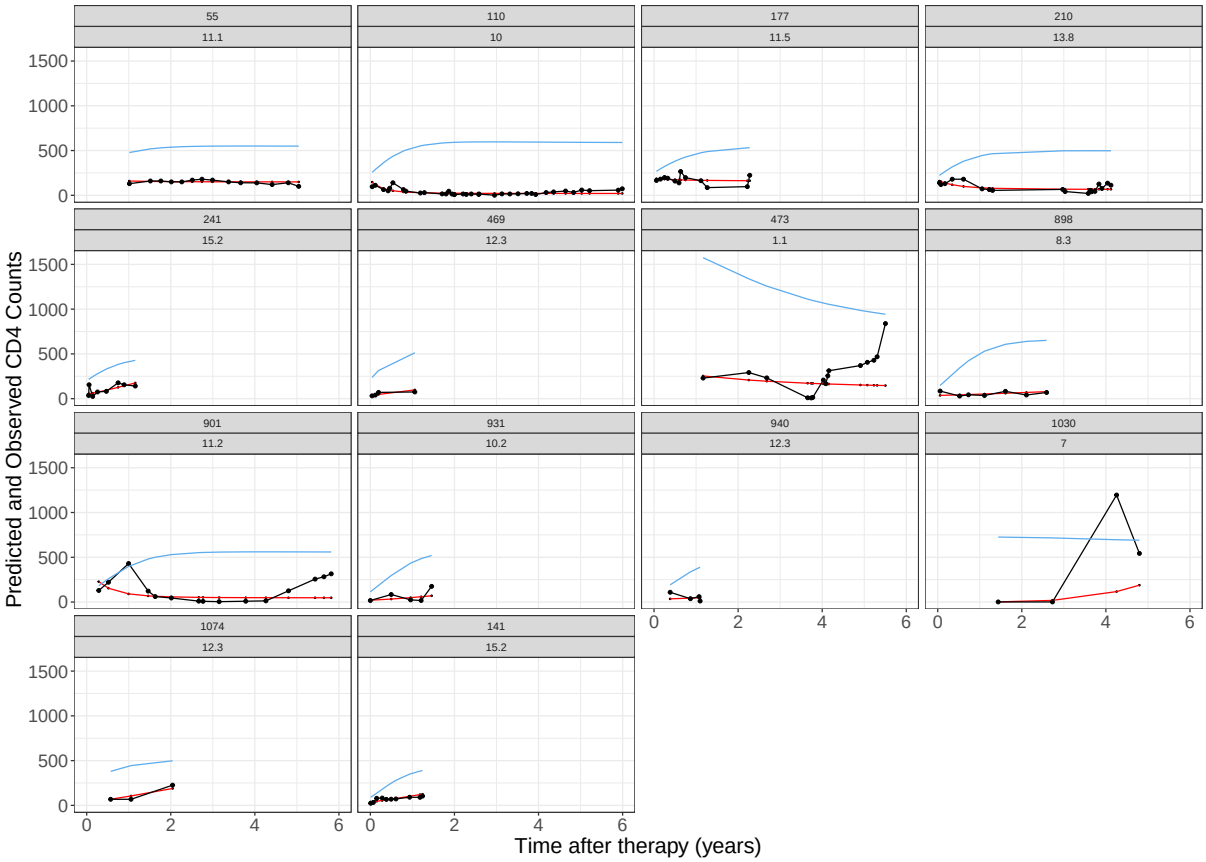


Figure 4.13: Plots of virally suppressed children within the severe HIV immune deficiency category for datasets A and B. Blue line: population prediction; Red line: individual predictions; Black line: observed CD4 counts. Top panel of each box shows the ID and age at start of ART for each child.

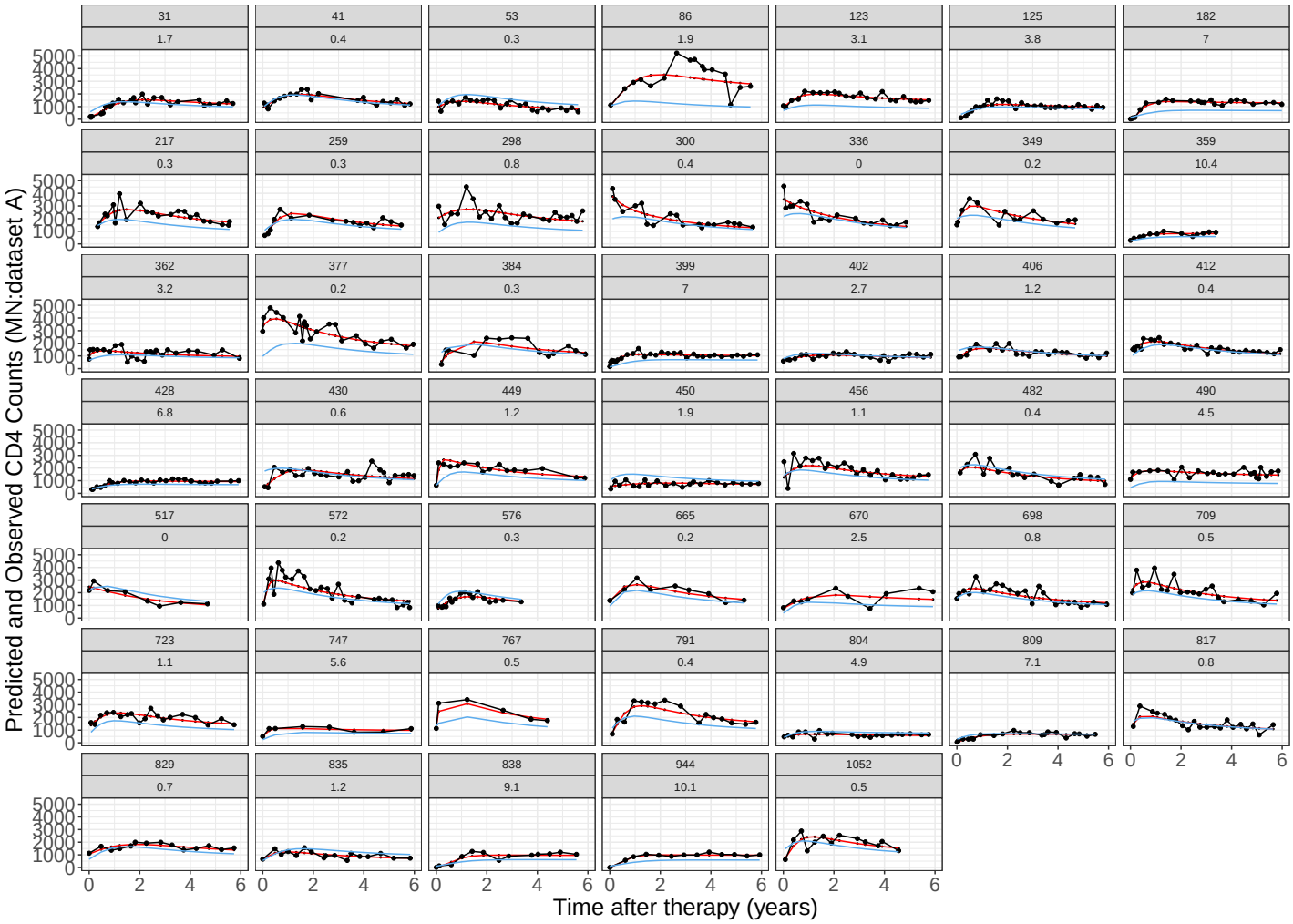


Figure 4.14: Plots of mild to normal immune responders in dataset A. Blue line: population prediction; Red line: individual predictions; Black line: observed CD4 counts. Top panel of each box shows the ID and age at start of ART for each child (yrs). Age of child with value of zero is a 3 day old child.

4.5 Discussion

In this chapter, I have investigated predictors of poor immune response in a large EPPICC cohort of 2204 HIV-infected children started on ART. The entire dataset was divided into two equal halves and a covariate model was fitted to each using the mathematical model described in equation 4.2. The two covariate models were then applied to the other halves of the datasets from which the model had not been originally developed (model B on dataset A, model A on dataset B). Fitted residuals from all four analyses were explored looking for a higher proportion of outlying data points outside the 2σ criterion. However, this approach did not reveal any further outlying residuals beyond what is expected under the assumptions of log-normality (Figure 4.6), perhaps due to the fact that σ was based on residual values.

Hence poor immune responders could not be identified through this approach. Predicted CD4 counts were compared to the World Health Organisation criteria for HIV-associated immunodeficiency as outlined in Table 4.6. These thresholds used in identifying poor immune responders have been used extensively in other studies [Zanoni et al., 2012, Lederman et al., 2011]. Analysing the two datasets separately provided a way to compare results from one to the other. Although the covariates retained in the two models are not precisely the same, the analysis shows that the WHO criteria were robust in consistently identifying poor immune responders across the two datasets. Dividing the dataset into two further illustrated that different outcomes can be obtained from a covariate analysis on subsets of data drawn from the same pool and care has to be taken when deciding which findings are of clinical importance.

The results of the covariate analysis showed that the only variable associated with long-term ln(CD4-for-age) was age at start of ART. This is in agreement with previous work [Majekodunmi et al., 2017, Lewis et al., 2012, Picat et al., 2013]. The analysis here further revealed that a piecewise linear relationship suitably described the effect of age at start of ART on int_i and asy_i. Figure 4.4D shows that rate of decrease in long-term ln(CD4-for-age) is about six times less for children under 5.4 yrs compared to children over 5.4 yrs (covariate model B). Similarly, for both datasets, a piecewise linear model with a break point at 5.8 yrs was the best fit between int_i and age at start of therapy. The choice of a piecewise linear function is reflective of the rapid changes in CD4 concentration in younger children due to volume contraction of the thymic epithelial space with increasing age [Steinmann et al., 1985]. This is further supported in the WHO guidelines where children above 5 yrs have the same immunodeficiency threshold as adults whilst children under 5 yrs have variable thresholds dependent on age. Factors affecting rate of $CD4^+$ T cell recovery, c_i include pre-ART AIDS status and individual cohort effect. Children diagnosed with AIDS pre-ART had lower int_i values and therefore experienced higher rates of recovery of their $CD4^+$ T cells possibly via IL-7 induced proliferation of $CD4^+$ T cells. In addition to IL-7, elevated IL-1A has been observed in IMNRs when compared to HIV-infected controls with optimal immunological recovery [Marziali et al., 2006, Delobel et al., 2006, Erikstrup et al., 2010, Nakanjako et al., 2011]. However, it is not clear whether elevated IL-1A is a consequence of poor immune recovery or a reason for it. Elevated levels of IL-1A is normally reflective of a normal and healthy immune response following infection. However, in HIV-infection, IL-1A has been associated with disease progression and is a stronger prognostic indicator when compared to the traditional $CD4^+$ T cell count or viral load measurements [Giorgi et al., 1999, Hazenberg et al., 2003b, Liu et al., 1997, Sousa et al., 2002, Grossman et al., 2002]. This is further evidenced in elite controllers, a group of HIV-infected patients who do not progress to AIDS and maintain normal CD4

counts with undetectable viral loads without ART. These elite controllers have been noted to have reduced levels of IL-1A in comparison to normal HIV progressors [Owen et al., 2010]. The reason for IL-1A's strong predictive property is unclear but levels have been noted to decline with ART [Massanella et al., 2010].

The final covariate model from both datasets differed slightly in the variables retained on int_i and c_i. In addition to the covariates retained in model A, model B retained relationships between int_i and logvl as well as between c_i and pre-ART AIDS status. This could be due to some differences in the characteristics of the children included in the two datasets. As expected, pre-ART ln(CD4-for-age) decreased with increasing age at start of therapy and older children were predicted to have lower long-term ln(CD4-for-age). Gender was not a significant covariate in overall $CD4^+$ T cell recovery.

According to Table 4.4, a total of 38 (4.3%) virally suppressed children were in the advanced to severe category of poor immune responders in dataset A. There were a further 42 children who were not fully virally suppressed in the advanced to severe poor immune responders group. In dataset B, a total of 34 (3.9%) virally suppressed children were in the advanced to severe category and a further 48 children were not fully virally suppressed. In both datasets, the number of children in the advanced and severe categories is much reduced in the virally suppressed category compared to the mixed group. This suggests an association between poor immune response in these children and lack of viral suppression perhaps due to HIV resistance mutations, poor compliance with medication or co-infection. This finding is well supported by a number of paediatric studies where good viral load suppression has been linked to optimal $CD4^+$ T cell recovery [Zanoni et al., 2012, Machado et al., 2007, Kovacs et al., 2005].

The overall percentage of advanced to severe poor immune responders (7.35% $\pm$0.05) in this study was lower than expected in the adult population perhaps due to the large number of younger children under 2 years at start of therapy included for analysis. The density component of the violin plot in Figure 4.2 shows high peaks in both datasets amongst children under the age of 2 years.

A larger proportion of children who received a nevirapine-based ART regimen were found amongst the virally suppressed mild-normal immune responders (31%) in both datasets A and B. In comparison, only 21% of children used a nevirapine-based regimen amongst the poor immune responders. Literature findings on $CD4^+$ T cell recovery in nevirapine-based regimen is inconclusive. Whilst some evidence suggest that there is no difference in $CD4^+$ T cell recovery between nevirapine based and PI-based regimen [The AVANTI and INCAS Study Groups, 2000, Podzamczer et al., 2002], other studies have found smaller increases in $CD4^+$ T cells with nevirapine based-regimen compared to PI-based regimen [Van Leth et al., 2004]. A study comparing nevirapine-based treatment to protease inhibitor-based regimen amongst children

who have been previously exposed to nevirapine found reduced viral suppression in the nevirapine-based treatment group [Fifth IAS Conference on HIV Pathogenesis, 2009].

Amongst the virally suppressed poor immune responders, median age at start of therapy (11-12 yrs) was almost double that of the entire cohort (5.4-5.8 yrs) indicating that a later age at start of therapy is indeed a major risk factor for poor immune response. It is therefore likely to be the case that the longer the duration of HIV infection before therapy, the greater the extent of damage to the immune system and lymphatic tissues [Schacker et al., 2002]. Hence, children who start therapy at an older age and with low CD4 counts are at greater risk of incomplete immune recovery. This has been reflected in the 2016 WHO guidelines which now recommends commencement of ART for all HIV-infected individuals irrespective of age or CD4 count [World Health Organisation, 2016a].

Whilst it is appreciated that some of the poor immune responders may go on to increase their CD4 counts beyond the first six years of therapy, there is no doubt about the clear difference between the CD4 trajectories of normal responders and poor responders. This difference is both qualitative and quantitative. Majority of the poor responders have flat-looking CD4 trajectories indicating a lack of robust immune response. In dataset A, median fixed effect for rate of recovery (c_i) was 1.3/yr (IQR: 0.7-2.0/yr) for the poor responders and 2.1/yr (1.4-2.9/yr) for the normal responders (virally suppressed). This gives a time to achieve half of the maximal increase in CD4 count of 120 days and 192 days respectively. Median fixed effect for rate of recovery decreased to 0.8/yr (0.2-1.1/yr) in the severe poor immune responders category giving a time to achieve half of the maximum increase in CD4 count of 10 months.

In this work, a large scale analysis has been conducted using statistical techniques that are appropriate for analysing longitudinal datasets. By including a large cohort of children, statistical power of the analysis is increased and effects that could not be seen in smaller analyses were captured. The approach of dividing the dataset into two provides an avenue to validate results and compare findings.

Limitations of the study include inaccurate assumptions of log-normality giving rise to less accurate predictions at higher ln(CD4-for-age) as illustrated on the visual predictive check (Figure 4.7). Although the covariate analysis identified factors associated with poor immune response, it does not reveal mechanisms driving the recovery process. In older children, it would be useful to examine other factors that may give a fuller picture of CD4 recovery dynamics including naive $CD4^+$ T cell homeostasis, cytokine profiles and other T cell subsets (CD4/CD8 naive and memory cells). There was no data available on these covariates and this could provide an angle for further future work. Pre-ART HIV viral load was investigated as a covariate alongside other variables. Since the dynamics of HIV viral load is closely linked to that of CD4

recovery, it would be better to model this mechanistically in future work.

Identifying factors responsible for poor immune recovery is crucial in effective management of HIV-infected children. Furthermore, it allows high risk patients to be targeted early and monitored closely. The analysis conducted here has identified some factors responsible for poor immune response using WHO thresholds for HIV-associated immune deficiency. These include higher age at start of therapy and poor HIV viral load control. Higher proportion of poor immune responders amongst the UK cohort warrants further investigation of other clinical correlates.

Chapter 5

Modelling homeostasis of naive $CD4^+$ T cells in HIV-infected children on ART

5.1 Introduction

HIV infection is characterised by a gradual loss of $CD4^+$ T cells eventually resulting in increased susceptibility to opportunistic infections. Through effective combination therapy (ART), viral replication can be suppressed allowing $CD4^+$ T cell population to recover slowly back to healthy levels. This manner of $CD4^+$ T cell recovery takes place via redistribution of T cells from the lymphoid tissues, peripheral expansion as well as production of naive T cells from the thymus. Since HIV-infected children on ART tend to rely more on naive T cells for immune recovery, understanding the dynamics of naive $CD4^+$ T cell recovery in children is critical to effective management of paediatric HIV infection.

5.1.1 Naive $CD4^+$ T cell homeostasis

As introduced in section 1.3.3, T cell receptor excision circles (TRECs) are circular pieces of DNA which are by-products of thymic TCR gene re-arrangement [Livak and Schatz, 1996, Douek et al., 1998, Kong et al., 1998]. TRECs have the unique property in that they are stable and do not divide during mitosis, they are predominantly of thymic origin and are consistent in 70% of $\alpha\beta$ T cells [Douek et al., 1998, Kong et al., 1999, Verschuren et al., 1997]. As a result, average cellular TREC content decreases with each cycle of cell division. This property has distinguished TRECs as surrogate markers for thymic output and can be used to determine the proliferative history of naive $CD4^+$ T cells at the population-level [Douek et al., 1998, Kong et al., 1998]. Whilst a reduction in mean cellular TREC content is strongly associated with decreasing thymic output, it could also be an indication of increased peripheral cellular division or death characteristic of HIV infection [Hazenberg et al., 2000a, Ye and Kirschner, 2002, Douek, 2004]. Therefore, mean TREC content per cell provides an incomplete measure of thymic function. For these reasons, TRECs have been used to identify T cells produced by the thymus. CD31 molecules are a member of the Ig family expressed on a number of cells including T cells, mast cells, NK cells, monocytes,

granulocytes, endothelial cells and platelets [Newman, 1997]. Naive CD4$^+$ T cells ("thymic naive CD4$^+$ T cells") which co-express CD31$^+$ are high in TRECs (about eight times higher) whilst the central naive CD4$^+$ T cells which are CD31$^-$ are low in TRECs [Kohler et al., 2005, Tanaskovic et al., 2010]. However, TREC$^+$ T cells remain in the peripheral lymphoid tissues many years after thymectomy suggesting that not all TREC$^+$ CD4 T cells are recent thymic emigrants [Fink, 2013]. The non-replicative nature of TRECs ensures that only one daughter T cell receives the episomal DNA giving further evidence that not all TREC$^+$ naive cells are RTEs [Hazenberg et al., 2003a].

In keeping with thymic output, the number of CD31$^+$ CD4$^+$ T cells decreases drastically with age. Evidence from cord blood samples shows that 90-95% of naive CD4$^+$ T cells express CD31 in contrast to just ∼30% in a 60 year old adult [Kilpatrick et al., 2008, Kimmig et al., 2002, Gomez et al., 2003, Duszczyszyn et al., 2006]. CD31$^+$ T cells represent a subset of T cells that are yet to undergo significant homeostatic T cell proliferation and are influenced by thymic output, cell division and loss through cell death or loss of CD31 expression [Wightman et al., 2010]. There is further evidence in the literature for IL-7-induced CD31$^+$ naive T cell expansion without loss of the CD31 cell surface marker [Kilpatrick et al., 2008, Azevedo et al., 2009]. Therefore, CD31 expression is not a clear marker for RTEs but instead represents a sub-population of naive CD31$^+$ CD4$^+$ T cells which include RTEs and other CD31$^+$ T cells that have not yet undergone significant cell division following thymic egress [Kohler and Thiel, 2009]. Unlike CD31$^+$ T cells, the population of CD31$^-$ T cells remains stable throughout adulthood. This results in a cumulative increase in the percentage of CD31$^-$ T cells resident within the naive CD4$^+$ T cell pool with increasing age [Kimmig et al., 2002, Kilpatrick et al., 2008].

In healthy individuals, the average estimated lifespan of naive T cells is 6-9 years [Vrisekoop et al., 2008] and from 20 years of age, the thymus is known to contribute 10% of daily naive T cell production [den Braber et al., 2012]. Using deuterium labeling study in chronically infected HIV$^+$ patients, it was discovered that the lifespans of naive CD4$^+$ T cells decreased to 1.7 yrs [Vrisekoop et al., 2015]. In 20-25 year olds, an estimated $1.7 - 17 \times 10^7$ RTE cells are produced daily [Vrisekoop et al., 2008, den Braber et al., 2012].

5.1.2 Naive CD4$^+$ T cell homeostasis and TREC dynamics in HIV infection

A number of studies have shown that TREC content decreases in HIV-infected individuals. In addition, HIV$^+$ individuals who have sub-optimal TREC frequency and are receiving ART have been observed to significantly increase their average TREC content [Douek et al., 1998, Zhang et al., 1999]. This is in

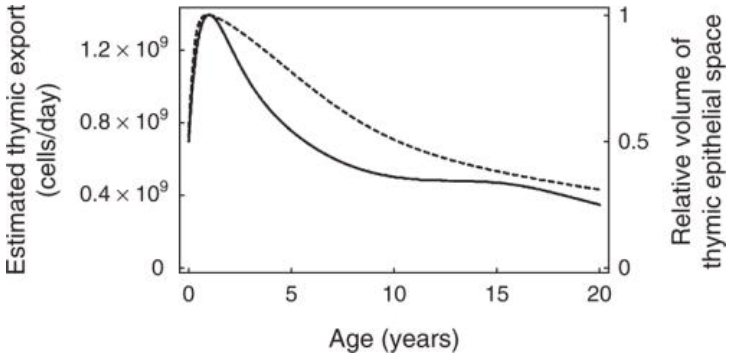


Figure 5.1: Plot of expected Thymic output of naive CD4 T cells as a function of age. Solid line: mean thymic export, Dashed line: relative volume of thymic epithelial space. Diagram taken from Bains *et al.* [Bains et al., 2009b]

keeping with findings in children suggesting that antiretroviral therapy increased thymic output in HIV$^+$ children [Sandgaard et al., 2014]. An early longitudinal study by Hatzakis *et al.* involving 131 HIV$^+$ participants showed that in addition to HIV viral load and CD4$^+$ T cell count, peripheral T cell pool TREC concentration was an important predictor of AIDS and death in HIV-infected persons [Hatzakis et al., 2000]. A more recent prospective longitudinal study which investigated 69 HIV-infected patients discovered that low baseline RTE percentage ($< 50\%$) predicted poor HIV disease progression [Zakhour et al., 2016]. In a mathematical model of TREC data, it was shown that alongside thymic output, dynamic processes including T cell division or death could affect TREC frequency [Hazenberg et al., 2000a]. Furthermore, it is known that HIV infection increases T cell division which in turn leads to a dilution in the average TREC content per cell, emphasising the need for careful interpretation of TREC data [Castro et al., 2016, Douek et al., 1998, Zhang et al., 1999]. However, increased naive T cell loss in HIV-infection is expected to increase average TREC content through renewal of the T cell pool [Hazenberg et al., 2000a, Dutilh and de Boer, 2003]. Hence increased naive T cell death counteracts the TREC-depleting effect of increased T cell division.

Mathematical models of naive T cell homeostasis could therefore give further insights into the mechanisms behind these disturbed homeostases.

5.1.3 Mathematical models of naive CD4 T cell homeostasis

In this section, two deterministic modelling approaches for naive T cell homeostasis are examined.

TREC-naive T cell dynamics model (Hazenberg model)

One of the earliest mathematical models of naive T cell homeostasis was published in Hazenberg et al [Hazenberg et al., 2000a]. In this study, they developed a mathematical model to describe the dynamics of TREC content of CD45RA$^+$CD27$^+$ T cells in both HIV-infected and uninfected individuals. They formulated two differential equations models; one for naive CD4$^+$ or CD8$^+$ T cells (N) and the other to model total amount of TRECs (T) within the naive T cell pool.

$$\frac{dN}{dt} = \sigma(a) + \alpha N - \delta N \tag{5.1}$$

$$\frac{dT}{dt} = c\sigma(a) - \delta T - \delta_i T \tag{5.2}$$

where $\sigma(a)$ is the thymic output as a function of age, α and δ are proliferation and death rates of naive T cells, c is the TREC content in recent thymic emigrants, δ_i is the decrease in TRECs due to T cell priming or intracellular degradation. By expressing the average TREC content of each naive T cell (A) as a ratio of TRECs to naive T cells, $A = \frac{T}{N}$, they were able to establish that in healthy subjects, decreasing thymic output only affected TREC concentration in the presence of naive T cell division. Furthermore, the fast decrease in TRECs seen in HIV-1 infection and the corresponding increase following ART are better explained by peripheral T cell division rather than thymic dysfunction.

Resting-dividing naive T cell model

A mechanistic model of CD4 T cell homeostasis for healthy adults and children was published by Hapuarachchi et al [Hapuarachchi et al., 2013]. They modelled naive CD4 T cells in two compartments; the resting (X) and the dividing (Y) T cell compartments (Figure 5.2). The model was expressed with the following coupled differential equations:

$$\frac{dX}{dt} = \theta + 2rY - \left(\lambda_{(X+Y)} + \delta_{(X+Y)}\right) X \tag{5.3}$$

$$\frac{dY}{dt} = \lambda_{(X+Y)} X - rY - \mu_{(Y)} \tag{5.4}$$

where λ, r, δ and μ are first order rate constants and θ is a zero order rate constant. Rate of entry into cells is given as $\lambda = \lambda_0 e^{\frac{-N(t)}{\epsilon}}$ where $N(t) = X(t) + Y(t)$.

The death rates of the resting and activated CD4 T cells were modelled as density dependent: $\delta = \delta_0 e^{\frac{N(t)}{\rho}}$, $\mu = \mu' Y$. The above model was an improvement of a previous model which did not consider the effect of

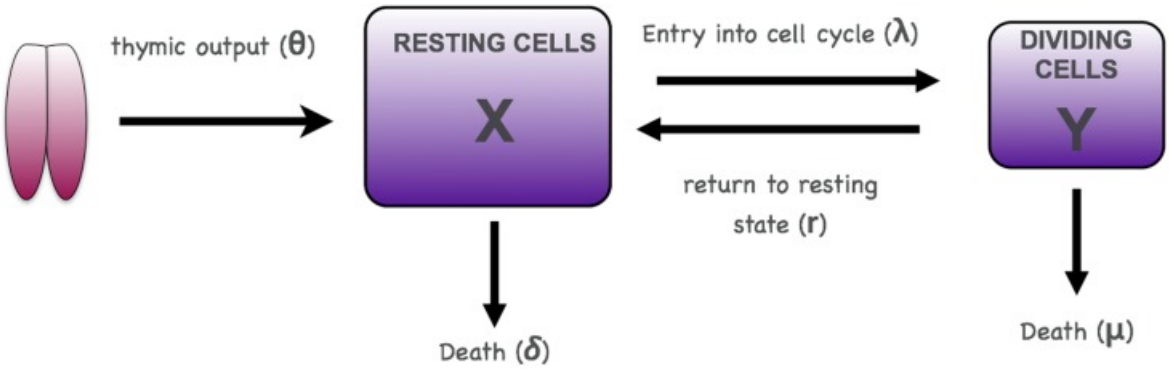


Figure 5.2: T compartment model of resting and dividing CD4 T cells. Diagram taken from [Hapuarachchi et al., 2013]

thymic output in naive CD4 T cell homeostasis [Yates et al., 2007, Callard et al., 2003]. They found that a decrease in the resource parameters resulted in a corresponding reduction in the number of CD4 T cells at homeostatic equilibrium. The model accurately predicted the concentrations of T cells during the first 3 years of life.

5.2 Objectives

1. The main aim of this chapter is to understand how HIV and ART affect the dynamics of the naive CD4$^+$ T cell compartments. Naive T cell compartments to be considered include recent thymic emigrants and central naive T cells.

2. To model the dynamics of CD31$^+$ T cells in HIV-infected children receiving ART.

Initially, a model is fitted to total naive CD4$^+$ T cell homeostasis using two different approaches and comparing parameter estimates from these. I then go on to formulate three different models of naive CD4$^+$CD31$^+$CD45RA$^+$ T cells.

5.3 Methods

5.3.1 The ARROW immunophenotype substudy

The ARROW trial (Anti-Retroviral Research fOr Watoto) was a multi-centre study conducted across four sites (three in Uganda and one in Zimbabwe) which enrolled ART-naive HIV-infected children aged between 3 months and 17 yrs over an 18 month period (2006-2008) [ARROW Trial Team, 2016]. A total of 1206 ART-naive children were enrolled and followed up over a period of 3.5 years. The aims of the study were two-fold: firstly to investigate the efficacy and safety of ART in the absence of regular haematological

monitoring; secondly to investigate the effect of an initial four-drug regimen ART (three NRTIs and one NNRTI) continued with a maintenance dose of three-drug regimen compared to the standard three-drug regimen. The children were randomised into either clinical disease monitoring or laboratory and clinical disease monitoring groups. Independent of monitoring protocol, these children were further allocated into one of three treatment arms. In accordance with the WHO guidelines, all three treatment groups received abacavir, lamivudine and one of nevirapine (< 3 years) or efavirenz (> 3 years). Children in group A were administered this three-drug regimen throughout the study. Groups B and C both received an additional fourth drug (zidovudine) initially for 36 weeks and then discontinued either zidovudine or one of efavirenz and nevirapine based on what therapy was given.

Over a 3.5 year period, routine full blood count, lymphocyte counts and liver function tests were collected every 3 months. It was standard practice for clinicians to follow up HIV-infected patients with routine blood tests including CD4$^+$ count and viral load at 3 monthly intervals. However, in resource-limited settings, these tests are both very expensive and not always available. In the ARROW trial, all enrolled children had laboratory tests done but only a randomly selected half of them had their results returned to their clinicians (laboratory and clinically driven monitory-LCM vs clincially driven monitory-CDM). In these children, clinical decisions were made based on physical examination and medical history. In cases where there were serious concerns about the wellbeing of these children, their clinicians were able to promptly request laboratory tests to facilitate medical care. Irrespective of study groups, abnormal blood tests were reported back to the corresponding clinicians for further review.

In terms of disease prognosis and mortality, the study showed that laboratory driven monitoring was not superior to clinically driven monitoring. Similarly, there was also no significant difference between the four-drug regimen group and the three-drug regimen cohort. This led the authors to suggest that limited resources normally spent on laboratory monitoring should be channeled into the administration of ART drugs [ARROW Trial team et al., 2013].

A percentage of children recruited from two Ugandan centres were selected for an immunophenotying substudy. Samples were analysed by flow cytometry on two separate panels: RA/CD31/Ki67 and RA/CD31/HLA-DR and recorded at weeks 4, 12, 24, 36, 48 and then every 24 weeks. Median follow up period was 3.5yrs and a total of 210 children were included for analysis in this work. Total CD4$^+$ and CD8$^+$ T cells were measured in real time. All children had whole-blood immunophenotyping, using anti-CD4-PerCP (Becton Dickinson [BD]), anti-CD45RA-APC (Caltag Medsytems), anti-CD31-PE (eBioscience), and either anti-Ki67- FITC (BD; after nuclear membrane permeabilization) or anti- HLA-DR-FITC (BD),

with data acquired on a BD FACSCalibur flow cytometer. Analysis was undertaken using Cellquest (BD). Viral loads were assayed at 1, 6, 9, and 12 months after starting ART and every 6 months thereafter [Prendergast et al., 2016].

5.3.2 Age-adjustment of CD4$^+$ T cell counts

In chapter 4, changes in CD4$^+$ T cell with age were scaled by log-transforming the CD4 ratios. Another way of dealing with age differences in CD4$^+$ counts is by incorporating it as part of the model building process as published by Bains et al [Bains et al., 2009b]. The advantage of this method is that age-adjustment is modelled with the raw data and therefore eliminating the need to transform the data prior to the model fitting. Furthermore, since the data is transformed within the model, raw CD4$^+$ counts are produced from the model fitting which makes interpretation easier. A total of three CD4 counts from three children greater than 7000 were excluded from analysis as were children with < 2 CD4 measurements (7442, 8561, 11227 cells/μL). The high CD4 counts above 7000 were likely due to laboratory errors.

In this chapter, I shall be examining age-adjustment within the model using a one compartment model of total naive CD4$^+$ T cells and making comparison of parameter estimates to investigate the effect of competition for resources.

5.3.3 Modelling naive CD4 T cell homeostasis

One-compartment model without competition for resources

To investigate naive CD4$^+$ T cell homeostasis, a single compartment turnover model illustrated in Figure 5.3 is initially considered:

$$\frac{dN}{dt} = \theta + N(\mu - \delta) \tag{5.5}$$

where the central compartment represents total naive CD4$^+$ T cell count $N(t)$ and cells from the thymus enter the compartment at a zero order rate constant θ and subsequently undergoing mitosis to either become two daughter cells or die at first order rate constants μ and δ respectively.

At steady state, $\frac{dN}{dt} = 0$, hence $\hat{N} = \frac{\theta}{\delta - \mu}$

$\delta - \mu > 0$ is a necessary condition for this steady state to exist. This implies $\delta > \mu$.

The analytical solution to the first order differential equation given above in equation 5.5 is:

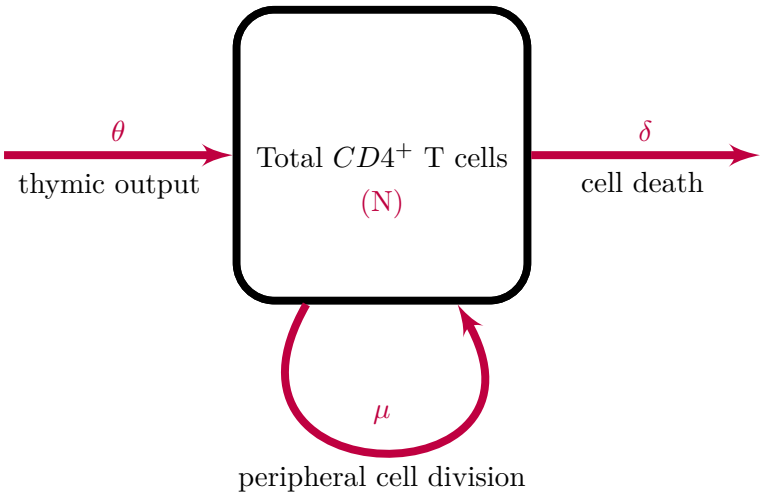


Figure 5.3: The above is a schematic of a single compartment model representing total CD4 T cell pool homeostasis. θ represents thymic contribution to the CD4 T cell pool, μ represents peripheral T cell division, δ represents cell death.

$$N(t) = \frac{\theta}{\delta - \mu} + \left(N_0 - \frac{\theta}{\delta - \mu}\right) e^{(\mu - \delta)t} \tag{5.6}$$

Modelling the effect of age on thymic output

In their work, Bains et al have described changes in thymic output with age mathematically where they used the stability of TRECs in estimating thymic production of T cells with age [Bains et al., 2009b] (Figure 5.1). They eliminated the dilution in TREC concentration subsequent to cell proliferation by using age-related Ki67 levels as a surrogate for cell division. Ki67 is a nuclear protein that is expressed in all phases of the cell cycle (G1, S, G2 and mitosis) and gradually lost by non-dividing cells over a time course of a few days. As such, Ki67 can be used to identify both actively dividing and recently divided cells [Gossel et al., 2017]. The function below from Bains et al. describes thymic output with respect to age a:

$$\theta(a) = \theta_0 \times \frac{y(a)V(a)\Gamma}{0.02\eta(c - \Gamma)} \tag{5.7}$$

where a is the age at CD4 measurement; η is the duration of Ki67 expression ($\eta = 0.52$days); c (0.25) and Γ (0.08) are values derived from TREC concentrations of CD4 T cells as they exit the thymus; V(a) is the expected CD4 T cell concentration with age (days); $y(a)$ represents the fraction of CD4 T cells expressing Ki67 with age (in days),

$$y(a) = 0.02e^{-\lambda a} \tag{5.8}$$

$$V(a) = 496 + 2074e^{-0.000857a} \tag{5.9}$$

as shown in Figure 5.4. In equation 5.7, θ_0 is added in such a way that the fraction of the expected age-dependent thymic output can be calculated whilst age-related changes are preserved.

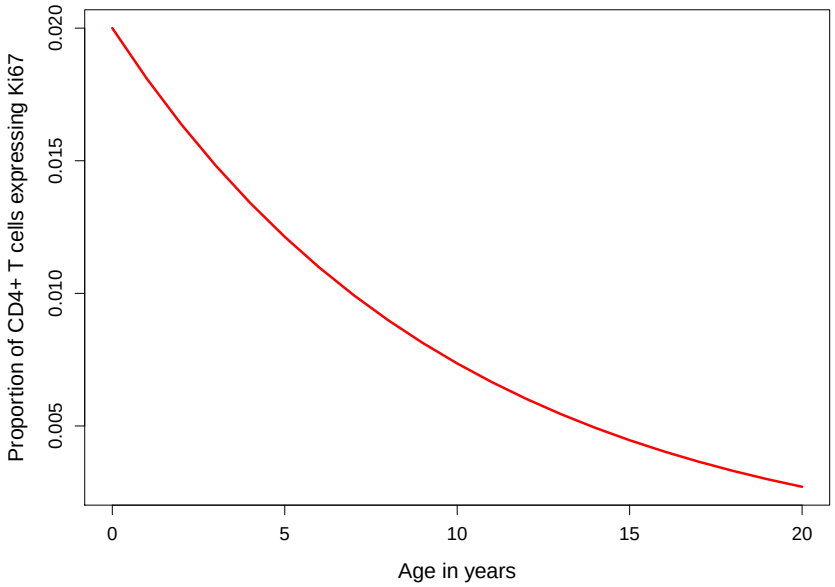


Figure 5.4: Proportion of CD4 T cells expressing Ki67. Function used in generating plot was taken from [Bains et al., 2009b] using exponential decay constant, $\lambda = 0.00027$.

Modelling competition and age effect on cell proliferation and loss

There is evidence from the literature to suggest that naive CD4 T cells compete for signals such as cytokines and sp-MHC which enable them to survive and undergo cell division. Accordingly, as the number of T cells increases, proliferation rate decreases and death rate increases and vice versa. This observation was made in thymectomised children where reduced T cell concentrations corresponded to increased levels of proliferation marker Ki67 [van Gent et al., 2011]. In section 5.1.3, a 2-compartment model incorporating competition for resources has already been proposed [Hapuarachchi et al., 2013]. The effect of competition for resources is therefore added to the one-compartment model for naive CD4 T cell homeostasis using the exponential functions described in equations 5.10 and 5.11. Since Ki67 is a marker for cell division, the

effect of age on fluctuating Ki67 concentration can be used to model duration of decline in cell proliferation and loss in childhood [Bains et al., 2009a]. Therefore, a function describing Ki67 expression with age ($y(a)$) has been added to the proliferation and loss parameters as shown in the following equations:

$$\mu = y(a)\mu_0 e^{c_\mu\left(1-\frac{N(t)}{V(a)}\right)} \tag{5.10}$$

$$\delta = y(a)\delta_0 e^{c_\delta\left(\frac{N(t)}{V(a)}-1\right)} \tag{5.11}$$

where c_δ and c_μ are magnitudes of the effect of competition for signalling resources on cell death and proliferation respectively. It follows that μ_0 and δ_0 are division and death rates when a child's naive CD4 concentration is equal to the expected CD4 concentration in a healthy child of the same age (i.e. $\frac{N(t)}{V(a)} = 1$). In order to account for age-related fluctuation in T cell count, the naive CD4 T cell concentrations have been adjusted with total CD4 T cell concentrations [Rollo Hoare, 2015, Hoare et al., 2017]. I have achieved this by using a function developed by Bains et al based on fitted regression estimated in a separate study from HIV uninfected European children [Huenecke et al., 2008] :

$$V(a) = 496 + 2074e^{-0.000857a} \tag{5.12}$$

where $V(a)$ is the expected CD4 concentration for age (days).

Full structural model for total naive CD4 T cells

The full structural model is given as:

$$\frac{dN(a,t)}{dt} = \theta(a) + N(a,t)(\mu(a,t) - \delta(a,t)),$$

where

$$y(a) = 0.02\ e^{-\lambda a} \tag{5.13}$$

$$V(a) = 496 + 2074e^{-0.000857a} \tag{5.14}$$

$$\theta(a) = \theta_0 \times \frac{y(a)V(a)\Gamma}{0.02\eta(c-\Gamma)} \tag{5.15}$$

$$\mu(a,t) = y(a)\mu_0 e^{c_\mu\left(1-\frac{N(t)}{V(a)}\right)} \tag{5.16}$$

$$\delta(a,t) = y(a)\delta_0 e^{c_\delta\left(\frac{N(t)}{V(a)}-1\right)} \tag{5.17}$$

In equations 5.13 to 5.17, $y(a)$ is the fraction of the expected CD4 T cell concentration with age (age in days), λ is the exponential decay constant of the plot shown in Figure 5.4 (day^{-1}), $V(a)$ is the expected CD4 T cell concentration with age (age in days), θ_0 is the proportion of theoretical or expected thymic output with age (cells/day), c (0.25) and Γ (0.08) are values derived from TREC concentrations of CD4 T cells as they exit the thymus, division and proliferations rates are in day^{-1} [Bains et al., 2009b]. As described in table 5.2, N_0 is the initial total naive CD4 T cell concentration.

Since there is no closed form solution for the above differential equation 5.3.3, it was solved numerically in NONMEM 7.3 using the ADVAN13 general non-linear kinetics subroutine and fitted to data using a combination of the SAEM and importance sampling EM algorithms. Further details of the codes can be found in appendix A.3. Using non-linear mixed-effects modelling, we shall be estimating the following parameters: θ_0, N_0, μ_0, δ_0, c_μ, c_δ and λ.

5.3.4 Modelling naive CD31$^+$CD4$^+$ naive T cells

Evidence from one study suggests that naive CD31$^+$ T cells subjected to repeated in vitro T cell receptor stimulation eventually become CD31$^-$ cells indicating that CD31$^+$ cells produced by the thymus may lose CD31 on further peripheral division [Demeure et al., 1996]. Further experimental evidence suggests that CD4$^+$ T cells are able to retain CD31 expression for an indeterminate number of divisions leading to an accumulation of cells which have either not proliferated or have undergone one or more divisions [den Braber et al., 2012]. In this chapter, models of CD31$^+$ T cells are explored and fitted to longitudinal data obtained from the ARROW immunophenotyping study. There are currently no published mathematical models of CD31$^+$ T cells in the literature. The model described below was taken from work done by Joanna Lewis as part of her doctoral thesis [Joanna Lewis, 2012].

1. Firstly, consider a situation where CD31 expression is lost after one T cell division (Figure 5.5a). The rate of change of CD4$^+$CD31$^+$CD45RA$^+$ T cells is given as:

$$\frac{dT_0}{dt} = \theta - (\mu + \delta)T_0 \tag{5.18}$$

At steady state, the rate of change is zero and the total number of CD4$^+$CD31$^+$CD45RA$^+$ T cells is:

$$\hat{T_0} = \frac{\theta}{\mu + \delta}$$

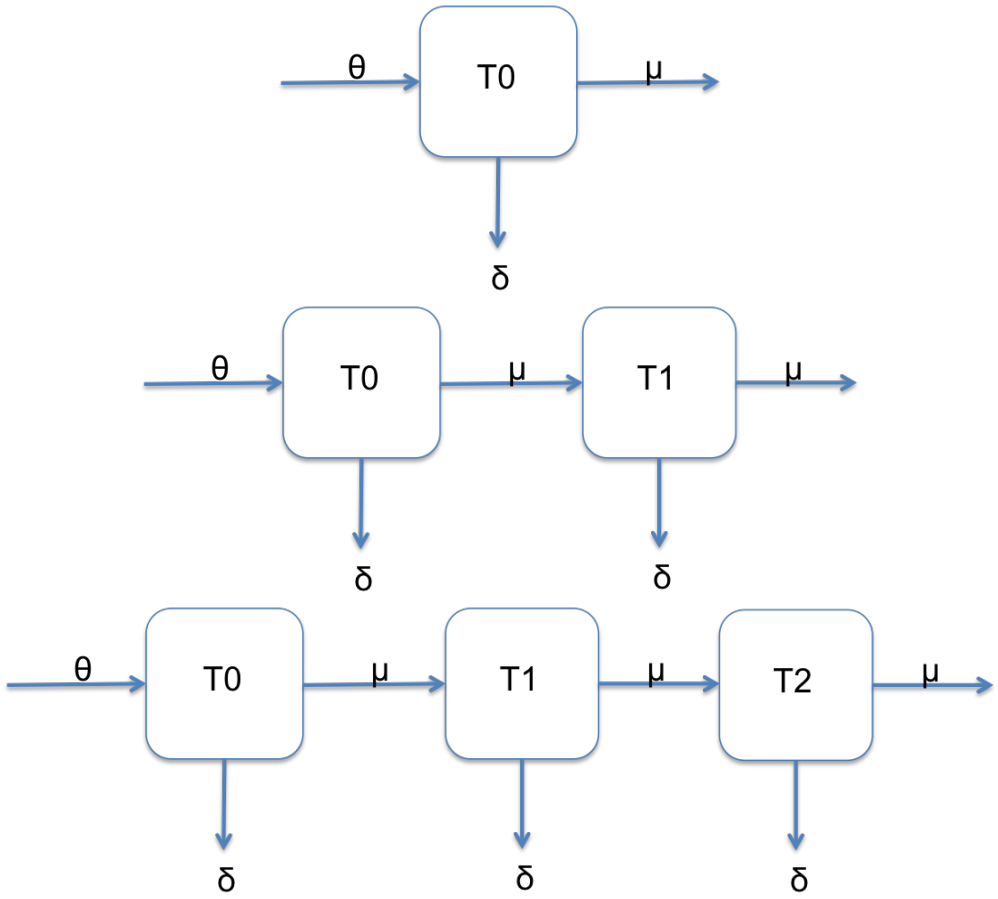


Figure 5.5: Model illustrating CD31 retention for zero (top-a), one (middle-b) and two (bottom-c) divisions following thymic egress of naive T cells. T cells are produced at rate θ, they proliferate at rate μ and die at rate δ. μ and δ are not dependent on the number of cellular divisions. Figure reproduced from [Joanna Lewis, 2012]

2. Similarly, a scenario where CD31 expression is retained after the first division but lost thereafter can be modelled with the following equation:

$$\frac{dT_1}{dt} = 2\mu T_0 - (\mu + \delta)T_1 \tag{5.19}$$

At steady state (assuming equal division and death rates in both compartments), total number of CD31 T cells is given as:

$$\hat{T_1} = \frac{2\mu \hat{T_0}}{\mu + \delta}$$

3. Lastly, the idea can be extended to model a situation where CD31 expression is retained after two cellular divisions:

$$\frac{dT_2}{dt} = 2\mu T_1 - (\mu + \delta)T_2 \tag{5.20}$$

At steady state,

$$\hat{T_2} = \frac{2\mu \hat{T_1}}{\mu + \delta}$$

From the above equations, both $\hat{T_1}$ and $\hat{T_2}$ can be expressed in terms of $\hat{T_0}$ as follows:

$$\hat{T_0} = \frac{\theta}{\mu + \delta}$$

$$\hat{T_1} = \frac{2\mu \hat{T_0}}{\mu + \delta} = \frac{2\mu\theta}{(\mu + \delta)^2}$$

$$\hat{T_2} = \frac{2\mu \hat{T_1}}{\mu + \delta} = \frac{4\mu^2\theta}{(\mu + \delta)^3}$$

Hence,

$$\hat{T_0} = \frac{\theta}{\mu + \delta},\ \hat{T_1} = \frac{2\theta\mu}{(\mu + \delta)^2},\ \hat{T_2} = \frac{4\theta\mu^2}{(\mu + \delta)^3}$$

Total CD31$^+$ T cells after three cellular divisions:

$$\hat{T_0} + \hat{T_1} + \hat{T_2} = \frac{\theta}{\mu + \delta} + \frac{2\mu\theta}{(\mu + \delta)^2} + \frac{4\mu^2\theta}{(\mu + \delta)^3} = \frac{\theta}{\mu + \delta} \sum_{k=0}^{2} \left(\frac{2\mu}{\mu + \delta}\right)^k$$

The above expression can be easily generalised to obtain steady state total CD31$^+$ naive T cells for n-compartments:

$$\hat{T_{\text{CD31total}}} = \frac{\theta}{\mu + \delta} \sum_{k=0}^{n} \left(\frac{2\mu}{\mu + \delta}\right)^k = \frac{\theta}{\mu + \delta} \left(\frac{1 - \left(\frac{2\mu}{\mu+\delta}\right)^{n+1}}{1 - \left(\frac{2\mu}{\mu+\delta}\right)}\right) \tag{5.21}$$

where

$$\frac{2\mu}{\mu + \delta} < 1$$

The total number of naive CD4$^+$ T cells (both CD31$^+$ and CD31$^-$) can be found by taking the limits of $\hat{T_{\text{CD31total}}}$ as $n \to \infty$

$$\lim_{n\to\infty} \hat{T_{\text{total}}} = \frac{\theta}{\mu+\delta}\left(\frac{1}{1-\frac{2\mu}{\mu+\delta}}\right) = \frac{\theta}{\delta-\mu}$$

This means in the long term, dynamics are only influenced mostly by proliferation, death and thymic output. i.e.

$$\frac{dT}{dt} = \theta - (\delta-\mu)T$$

which gives a steady state value of:

$$\hat{T} = \frac{\theta}{\delta-\mu}$$

Proportion of CD31$^+$ T cells at steady state (ratio of CD31$^+$ T cells to total CD4$^+$ naive T cells):

$$\frac{\hat{T_{CD31}}}{\hat{T_{total}}} = \frac{\theta}{\mu+\delta}\left(\frac{1-\left(\frac{2\mu}{\mu+\delta}\right)^{n+1}}{1-\left(\frac{2\mu}{\mu+\delta}\right)}\right) \times \frac{\delta-\mu}{\theta} = \frac{\delta-\mu}{\delta+\mu}\left(\frac{1-\left(\frac{2\mu}{\mu+\delta}\right)^{n+1}}{1-\left(\frac{2\mu}{\mu+\delta}\right)}\right) \tag{5.22}$$

Notably, equation 5.22 illustrates that the steady state value of CD31 ratio is independent of thymic output and only determined by cell death (δ), proliferation (μ) and number of divisions (n). Hence, proportion of CD31$^+$ T cells is increased by slower proliferation, faster death rate and larger k.

Analytical solutions to ODEs

In this section, analytical solutions to the differential equations given in equations 5.18, 5.19 and 5.20 are obtained.

- Consider equation 5.18:

$$\frac{dT_0(t)}{dt} + (\mu+\delta)T_0(t) = \theta$$

where $T_0(0) = A$. Using an integrating factor of $e^{(\mu+\delta)t}$,

$$T_0(t) = \frac{\theta}{\mu+\delta} + \left(A - \frac{\theta}{\mu+\delta}\right)e^{-(\mu+\delta)t}$$

If $\mu+\delta = \alpha$,

$$T_0(t) = \frac{\theta}{\alpha} + \left(A - \frac{\theta}{\alpha}\right)e^{-\alpha t}$$

- Consider equation 5.19:

$$\frac{dT_1(t)}{dt} + (\mu + \delta)T_1(t) = 2\mu T_0(t)$$

where $T_1(0) = B$.

Similarly,

$$T_1(t) = \frac{2\mu\theta}{\alpha^2} + e^{-\alpha t}\left(2\mu t\left(A - \frac{\theta}{\alpha}\right) - \frac{2\mu\theta}{\alpha^2} + B\right),$$

where $\mu + \delta = \alpha$.

- Consider equation 5.20:

$$\frac{dT_2(t)}{dt} + (\mu + \delta)T_2(t) = 2\mu T_1(t)$$

where $T_2(0) = C$ and $\mu + \delta = \alpha$. The solution is given as:

$$T_2(t) = \frac{4\mu^2\theta}{\alpha^3} + e^{-\alpha t}\left(2A\mu^2t^2 - \frac{2\mu^2\theta t^2}{\alpha} + 2\mu Bt - \frac{4\mu^2\theta t}{\alpha^2} - \frac{4\mu^2\theta}{\alpha^3} + C\right)$$

Expressions for T_0, T_1 and T_2:

$$T_0(t) = \frac{\theta}{\alpha} + \left(A - \frac{\theta}{\alpha}\right)e^{-\alpha t}$$

$$T_1(t) = \frac{2\mu\theta}{\alpha^2} + \left(2\mu t\left(A - \frac{\theta}{\alpha}\right) - \frac{2\mu\theta}{\alpha^2} + B\right)e^{-\alpha t}$$

$$T_2(t) = \frac{4\mu^2\theta}{\alpha^3} + \left(2A\mu^2t^2 - \frac{2\mu^2\theta t^2}{\alpha} + 2\mu Bt - \frac{4\mu^2\theta t}{\alpha^2} - \frac{4\mu^2\theta}{\alpha^3} + C\right)e^{-\alpha t}$$

Total sum of CD4$^+$CD45RA$^+$CD31$^+$ T cells:

$$\begin{aligned}
&\sum_{k=0}^{2} T_k(t) \\
&\quad = \left(\frac{\theta}{\alpha} + \frac{2\mu\theta}{\alpha^2} + \frac{4\mu^2\theta}{\alpha^3}\right) + \left(A + B + C - \frac{\theta}{\alpha} + 2\mu At + 2A\mu^2t^2 - \frac{2\mu\theta t}{\alpha} - \frac{2\mu\theta}{\alpha^2}\right. \\
&\quad \left. - \frac{2\mu^2\theta t^2}{\alpha} + 2\mu Bt - \frac{4\mu^2\theta t}{\alpha^2} - \frac{4\mu^2\theta}{\alpha^3}\right)e^{-\alpha t} \\
&\quad = P + Q(\theta, \mu, \delta, t)e^{-\alpha t}
\end{aligned} \tag{5.23}$$

where $P = \frac{\theta}{\alpha} + \frac{2\mu\theta}{\alpha^2} + \frac{4\mu^2\theta}{\alpha^3}$ and

$$Q(\theta, \mu, \delta, t) = T_0(0) + T_1(0) + T_2(0) - \frac{\theta}{\alpha} + 2\mu At + 2A\mu^2 t^2 - \frac{2\mu\theta t}{\alpha} - \frac{2\mu\theta}{\alpha^2} - \frac{2\mu^2\theta t^2}{\alpha} + 2\mu Bt - \frac{4\mu^2\theta t}{\alpha^2} - \frac{4\mu^2\theta}{\alpha^3}$$

As $t \to \infty$, $e^{-\alpha t}$ tends to zero and hence the sum tends to P (where $\alpha \neq 0$). Since $\delta > 0$ and $\mu > 0$, $\alpha = \mu + \delta > 0$.

Full structural model for CD31$^+$ naive T cells

$$y(a) = 0.02\ e^{-\lambda a} \tag{5.24}$$

$$V(a) = 496 + 2074e^{-0.000857a} \tag{5.25}$$

$$\theta(a) = \theta_0 \times \frac{y(a)V(a)\Gamma}{0.02\eta(c - \Gamma)} \tag{5.26}$$

$$\mu(a,t) = y(a)\mu_0 e^{c_\mu\left(1 - \frac{N(t)}{V(a)}\right)} \tag{5.27}$$

$$\delta(a,t) = y(a)\delta_0 e^{c_\delta\left(\frac{N(t)}{V(a)} - 1\right)} \tag{5.28}$$

$$\frac{dT_0(a,t)}{dt} = \theta(a) - T_0(a,t)(\mu(a,t) + \delta(a,t)) \tag{5.29}$$

$$\frac{dT_1(a,t)}{dt} = 2\mu(a,t)T_0(a,t) - T_1(a,t)(\mu(a,t) + \delta(a,t)) \tag{5.30}$$

$$\frac{dT_2(a,t)}{dt} = 2\mu(a,t)T_1(a,t) - T_2(a,t)(\mu(a,t) + \delta(a,t)) \tag{5.31}$$

$$\frac{dN(a,t)}{dt} = \theta(a) + N(a,t)(\mu(a,t) - \delta(a,t)) \tag{5.32}$$

Model fitting

All three models generated from the three scenarios described in section 5.3.4 were fitted to longitudinal data from the ARROW immunophenotyping substudy. The ODE for the total naive CD4 count was added to further inform parameter estimates whilst the effect of competition for resources is added onto the death and proliferation parameters as described in equations 5.10 and 5.11.

5.4 Results

5.4.1 Patient characteristics

In total, 223 children from the ARROW immunophenotype study were included for analysis. A total of 1,769 CD4 measurements were available in the final dataset and the children started ART at a median age of 5 yrs (IQR: 2-9 yrs) with a median duration on ART of 5 yrs. All participants were recruited from

two sites: Entebbe (22%) and MRC/JCRC Uganda (78%). A summary of patient characteristics is given in Table 5.1. Figure 5.6 shows plots of all four CD4$^+$ T cell subsets for 16 children. It shows that the evolution of CD31$^+$ T cells (red line) closely follows that of the total CD4$^+$ T cells (black line) indicating that CD31$^+$ T cells are an important source of naive CD4$^+$ T cells.

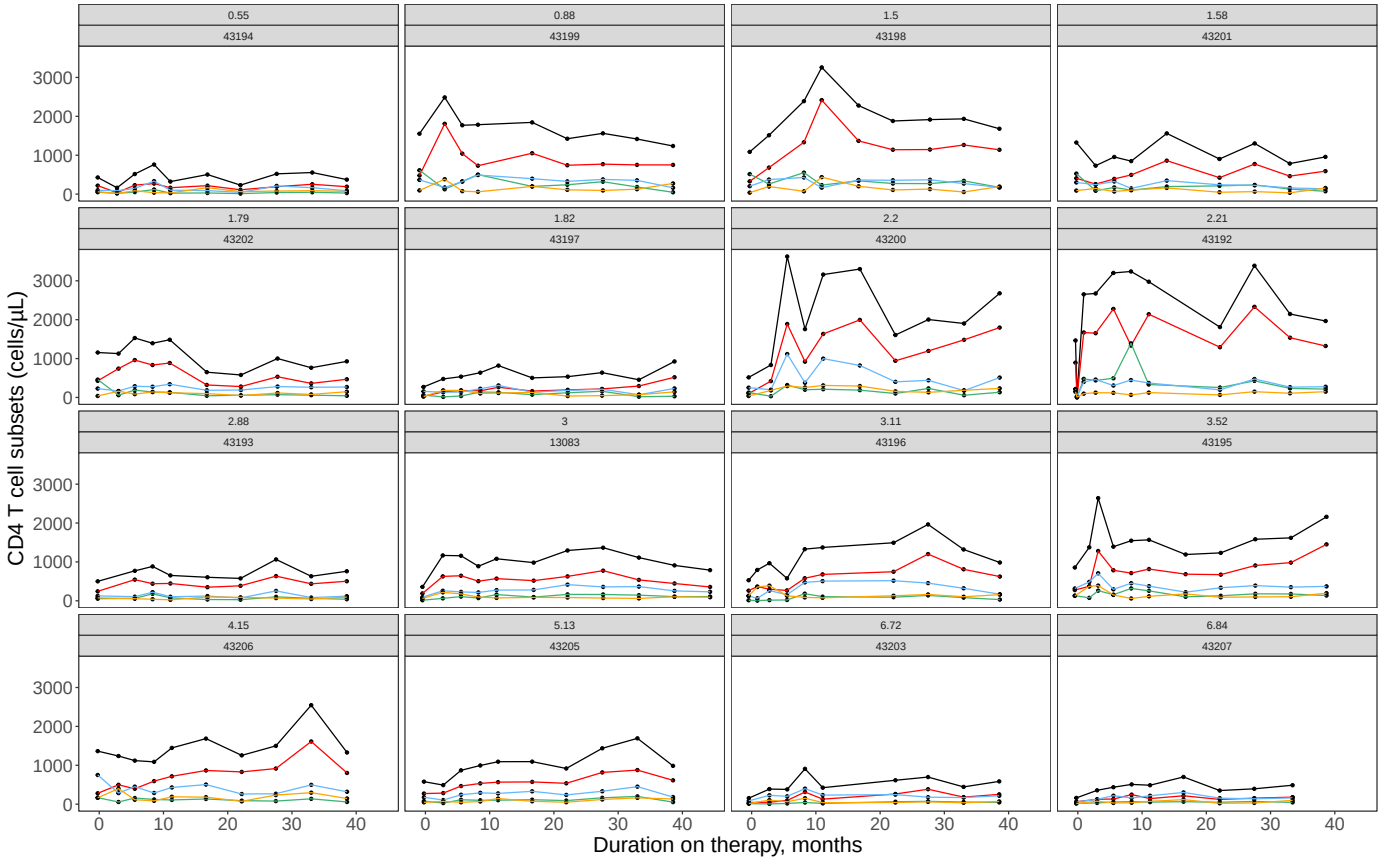


Figure 5.6: Plot of CD4 T cell subsets for 16 children in the ARROW immunophenotype substudy. Each line represents one of four CD4$^+$ T cell subsets whilst the black line represents total CD4 count. Red line: CD31$^+$CD45RA$^+$ cells (recent thymic emigrants), Blue line: CD31$^-$CD45RA$^-$cells (memory cells), Orange line: CD31$^+$CD45RA$^-$ cells (atypical memory cells), Green line: CD31$^-$CD45RA$^+$ cells (central naive cells). Age at start of therapy is shown in top panels along with individual IDs.

Summary of characteristics of the children from the ARROW substudy data

ARROW Data

	No.		No.
Total	223	Mean CD4 count per child	8 (7-9)
Centres		**Gender**	
Entebbe	50 (22%)	Female	118 (53%)
JCRC	173 (78%)	Male	105 (47%)
Group		**Drug Arms**	
CDM	113 (51%)	A	70 (31%)
LCM	110 (49%)	B	80 (36%)
Duration on ART		C	73 (33%)
Median (yrs)	5.2 (5.2-5.3)	**Age at sART**	
last VL	98 (4.5%)	Median (yrs)	5 (2.1-9.2)
total	209	**Trial EP**	
<400 copies/mL	154 (74%)	2006-2008	
≥400 copies/mL	55 (26%)	**last WAZ**	
Median (copies/mL)	80 (80-552)	Median	-1.35 (-2.03, -0.53)

Table 5.1: JCRC: Joint Clinical Research Centre, Kampala, Uganda. Entebbe: MRC/Uganda Virus Research Institute Programme on AIDS, Entebbe, Uganda. Viral load of 80copies/mL was the lower threshold of detection, trial EP: trial enrolment period, VL-viral load, sART: start of ART, WAZ: weight-for-age z-score (available in all 223 children). Interquartile range provided in parentheses

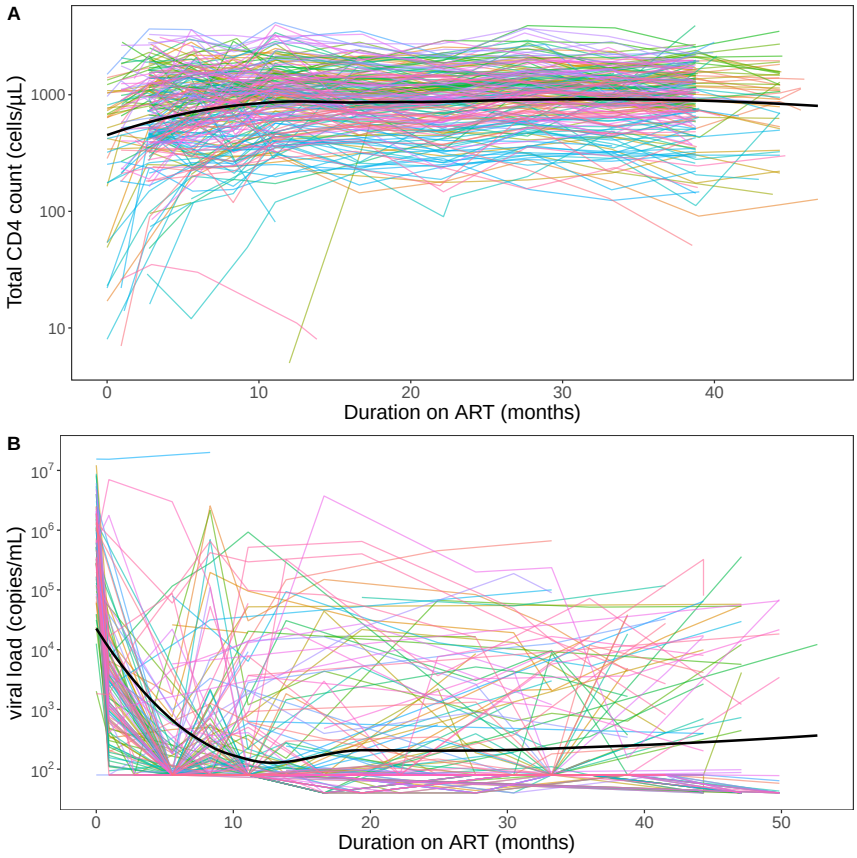


Figure 5.7: Plot of CD4 T cell (A) and HIV viral load (B) against duration on ART for the ARROW children (log scale). Black line represents local regression fitted to the data.

5.4.2 One-compartment model of total naive T cell homeostasis

Initially the one-compartment model given in equation 5.5 was fitted to total naive CD4$^+$ T cells in all 223 children. Two scenarios with and without age and competition effects on the proliferation and loss parameters were considered (equation 5.13). The parameters obtained from the two models are given in Table 5.3. In both models, the Ki67 decay constant (λ in Table 5.3) was estimated directly as a fixed effect from the data. As illustrated by Figure 5.9 and table 5.3, overall, TN model B ("Total naive" with competition effect) produced a better fit compared to TN model A which had no competition effect (ΔOFV = 137, two degrees of freedom-see section 2.2.1 for more details). Visual predictive plots of the two models illustrated in Figure 5.9 show that the model with competition for resources is better at predicting higher values of total naive CD4$^+$ T cell counts at the upper centiles. The estimate obtained for λ in models A and B respectively ($\lambda_A = \lambda_B = 0.0004$) was very close to the estimate used in Bains *et al.* which was sourced from experimental data ($\lambda = 0.00027$) [Bains et al., 2009b]. Estimates for proportion

of proliferating cells (μ_0) were 0.0009/day for model A and 0.09/day for model B. This gives an average between-division time for total naive CD4$^+$ T cells of 1111 days and 11 days respectively. This indicates that 0.1% (model A) and 9% (model B) of the total naive CD4$^+$ T cells are dividing at any given day. The estimate from model B is more reflective of findings from deuterium-labeled studies showing increased CD4$^+$ T cell turnover in HIV-infected children receiving ART [Vrisekoop et al., 2015]. In a study by Bains et al., mean inter-division time for naive T cells in healthy children between 1 and 5 yrs was found to be 125 days [Bains et al., 2009a]. Model B gave an estimated residence time ($\frac{1}{\delta_0}$) for total naive CD4$^+$ T cells of 8 days indicating a loss rate of 13% per day. This is a much higher loss rate compared to an estimate of 1.4% in healthy children [Bains et al., 2009a]. Estimated thymic outputs for models A and B were 3.5% and 1% of the predicted thymic output from Bains et al. [Bains et al., 2009b].

Table 5.2: Parameter definitions table

Parameter	Definition
θ_0	Proportion of theoretical thymic output [Bains et al., 2009b](cells/day)
μ_0	Proportion of expected naive CD4$^+$ T cell division (/day)
δ_0	Proportion of expected naive CD4$^+$ T cell loss (/day)
c_μ	Magnitude of competition proliferation
c_δ	Magnitude of competition loss
N_0	Initial total naive CD4$^+$ T cell concentration (cells/μL)

Table 5.3: Comparison of parameter estimates from the total naive CD4 T cell models with and without competition effect

	TN Model A		TN Model B	
Parameter	value	Ω	value	Ω
N_0 (cells/μL)	275	2.17	275	2.04
θ_0 (/day)	0.035	3.45	0.01	6.64
μ_0 (/day)	0.0009	4.35	0.09	2.25
δ_0 (/day)	0.00075	4.18	0.134	9.96
c_μ	-	-	0.571	1.41
c_δ	-	-	2.27	0.62
λ	0.0004	-	0.0004	-
OFV	21038	-	20901	-
σ	0.12	-	0.104	-
n_p	5	-	7	-

n_p: total no of fixed effects parameters, Ω: Variance on random effects, TN: Total naive model. TN Model A: model without competition or age effect on net loss. TN Model B: model with competition and age effect added to the proliferation and loss parameters. For TN Model A, the parameters μ_0 and δ_0 are just μ and δ respectively

CD31$^+$ T cell models

Parameter estimates from the three CD31$^+$ T cell models described in Figure 5.5 are given in Table 5.4. Random effects were added to all parameters with the exception of the Ki67 exponential constant which was estimated as a fixed effect parameter. As expected, the total number of CD31$^+$ T cells decreased with increasing cycles of proliferation (212 $\rightarrow$ 77 cells/μL). Estimated thymic outputs for models 1-3 were 3.1%, 2.4% and 2% of the predicted output from Bains et al. respectively [Bains et al., 2009b]. As the cells undergo further proliferation, the number of CD31$^+$ T cells drops with a corresponding increase in division rate to restore homeostasis. In terms of drop in OFV, the model describing loss of CD31 expression after the third division is the best of the three models (ΔOFV = 41). Diagnostic plots for the fitted models are shown in Figure 5.11 which illustrates a normal distribution of residuals over time and with population predictions with no obvious bias. Estimates of Ki67 decay constant, λ, were between 0.0005 to 0.001 for the three models. Proportion of cells proliferating per day was greatest in model 3 with an estimate of 10% (3.4% and 8.5% in models 1 and 2) giving a mean inter-division time of 10 days. Proportion of cells dying per day was lowest in model 2 which had an estimate of 31% (52% and 51% in models 1 and 3) giving an average residence time of 3 days. Both proliferation and loss parameters are much higher in comparison to estimates obtained from total naive CD4 T cells in healthy children (loss: 0.0143/day, division: 0.008/day [Bains et al., 2009a]).

Visual predictive checks for all three models are given in Figure 5.10 which shows a good overlap between model confidence bands and observed data. In Figure 5.8, overall, proportion of thymic output, proliferation and loss parameters are seen to decrease firstly between age 1 and 5 yrs followed by a transient increase between 5 and 7 yrs and then a gradual decrease with increasing age at start of therapy. This pattern is more prominent in the proportion of theoretical thymic output and loss estimates. Furthermore, the resultant effect of competition for resources on the loss rate is greater than that of proliferation. Individual CD31$^+$ T cell predictions of 30 children from model 3 are given in Figure 5.12.

A number of parameters had high shrinkage and these include $\theta_0, \mu_0, \delta_0,$ and c_μ (32%-60%), where a shrinkage 20-30% is considered high [Savic and Karlsson, 2009]. This indicates that information needed to estimate these parameters is lacking from the data.

Table 5.4: Comparison of parameter estimates from all three CD31$^+$ models

Parameter	Model 1		Model 2		Model 3	
	value	Ω	value	Ω	value	Ω
N_0 (cells/μL)	212	2.49	117	2.66	77	2.54
θ_0 (/day)	0.031	1.44	0.024	1.61	0.02	1.76
μ_0 (/day)	0.034	1.25	0.085	1.77	0.1	1.19
δ_0 (/day)	0.52	1.39	0.31	1.50	0.51	1.07
c_μ	0.243	0.46	0.27	0.39	0.214	0.7
c_δ	5.38	0.55	5.55	0.56	4.12	0.9
λ	0.0014	-	0.001	-	0.0005	-
OFV	20498	-	20484	-	20457	-
σ	0.12	-	0.12	-	0.12	-
n_p	7	-	8	-	9	-

n_p: total no of fixed effects parameters, Ω: Variances on random effects.

5.4.3 Poor CD31$^+$ T cell recovery is associated with reduced proportion of theoretical thymic output, slower death and faster proliferation rates in HIV-infected children on ART

In the fitted model 3, children with final predicted CD31$^+$ T cell count $<$ 200 cells/μL were considered poor immune responders. Reference range for normal CD31$^+$ T cells in healthy children up to 9 years of age is 460-1684 cells/μL [Shearer et al., 2003]. Of the 223 children analysed, a total of 39 children (17.5%) were poor immune responders. Median proportion of theoretical thymic output for poor responders was 0.012 cells/day (IQR: 0.0063-0.021 cells/day) in comparison to the median value of 0.026 cells/day (0.016-0.04 cells/day) in the remaining 184 children. Median proportion of daily cell loss was 28% (19-40%) for the poor responders and 57% (42-76%) for the remaining children. Median proportion of cells proliferating per day was 20% (7.5-26%, mean: 21%) for the poor responders and 11% (7.3-16%) for the mild-normal responders. These findings indicate that on average, poor responders are proliferating more quickly, die more slowly and have a poorer thymic output compared to the rest of the cohort. This is in agreement with the steady state analysis illustrated in equation 5.22 which suggests that proportion of CD31$^+$ T cells increased in the long-term with slower proliferation, faster death rate and CD31 expression over a larger number of cellular divisions.

Correlation matrix for random effects

The correlation matrix for the random effects obtained for parameters in CD31$^+$ T cell model 3 is given below in equation 5.33. Diagonal values are the standard deviations of the random effects on parameter

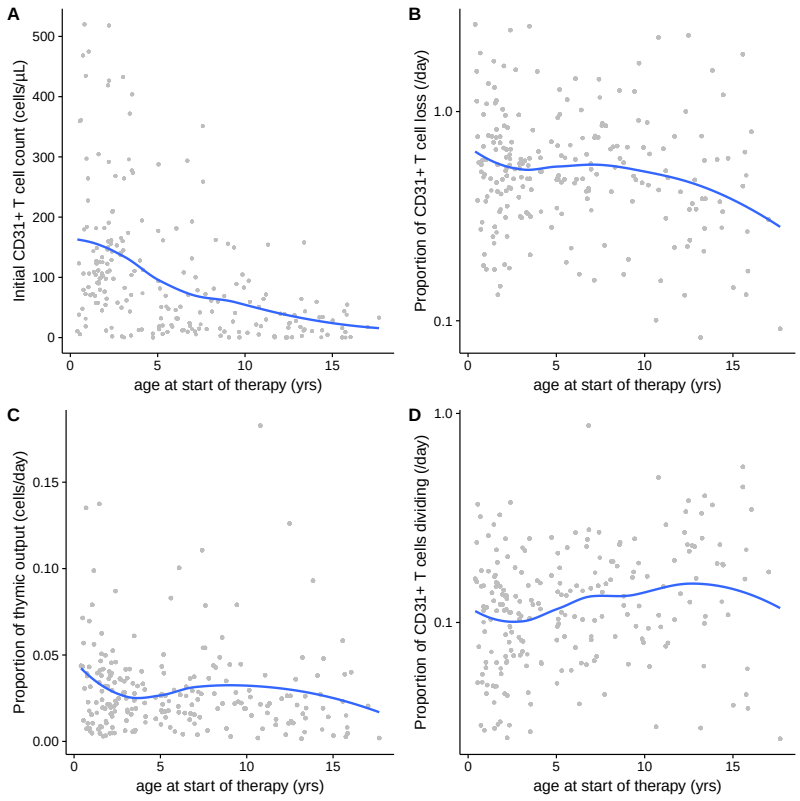


Figure 5.8: **A**: Plot of estimated initial CD31$^+$ T cell count against pre-ART age. **C**: Proportion of thymic output plotted against pre-ART age, **B**: Proportion of CD31$^+$ T cell loss plotted against pre-ART age ($\log_{10}$ scale), **D**: Proportion of CD31$^+$ T cell proliferation plotted against pre-ART age ($\log_{10}$ scale). The blue lines represent local regression through the data.

values whilst the off-diagonal variables are the correlation coefficients. It highlights a number of parameters with strong correlations-for example the initial CD31$^+$ T cell count in model 1, T_0 is strongly correlated with initial CD31$^+$ T cell count in models 2 (T_1) and 3 (T_2). Likewise, there is a strong correlation between T_1 and T_2 as well as between T_0, μ_0 and δ_0. There is a strong positive correlation between proportion of theoretical thymic output, θ_0 and proportion of CD4 T cell loss, δ_0.

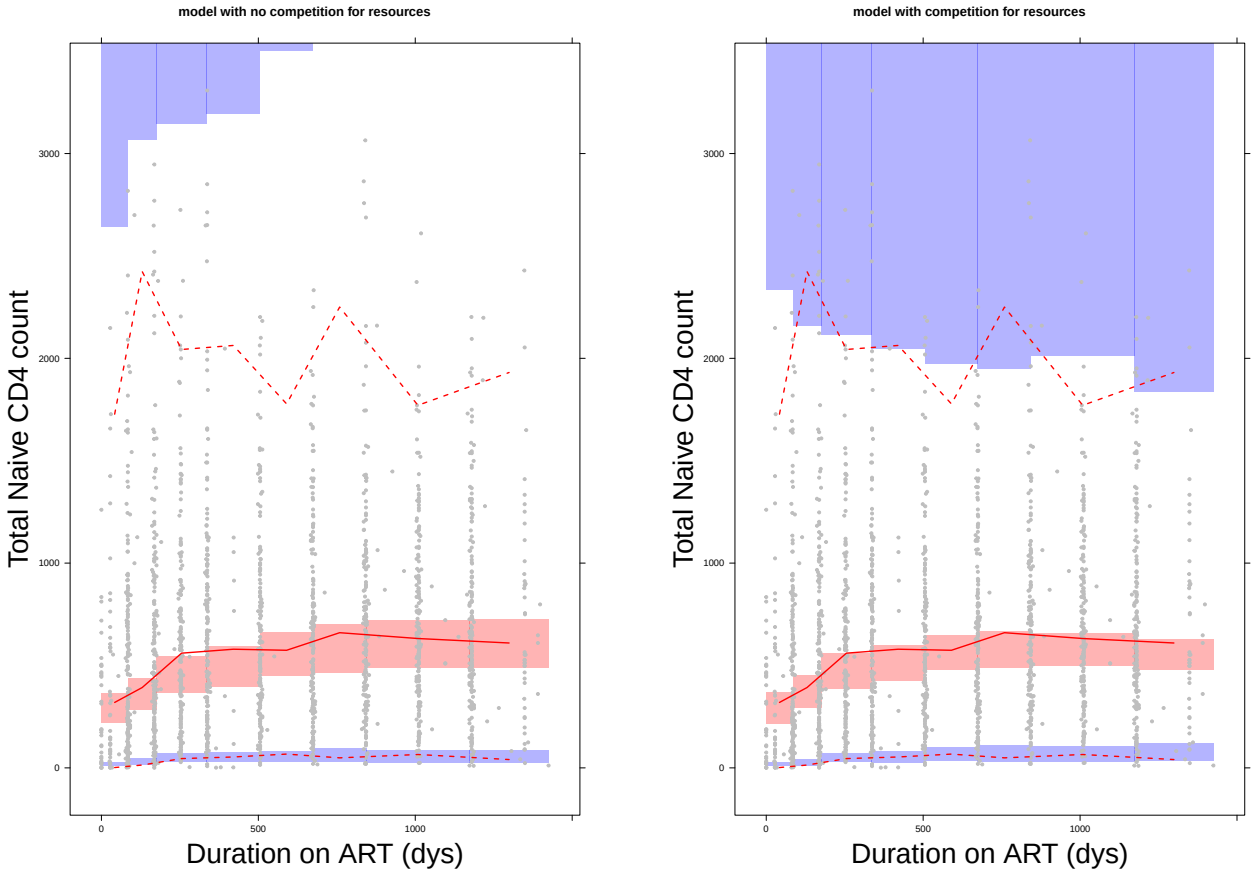
$$
\begin{array}{c}
\begin{array}{cccccccc} T_0 & \theta_0 & \mu_0 & \delta_0 & c_\mu & c_\delta & T_1 & T_2 \end{array} \\
\begin{pmatrix}
1.59 & & & & & & & \\
0.15 & 1.33 & & & & & & \\
-0.05 & 0.63 & 1.09 & & & & & \\
0.06 & 0.87 & 0.53 & 1.03 & & & & \\
-0.31 & 0.29 & 0.14 & 0.052 & 0.84 & & & \\
0.084 & -0.66 & -0.49 & -0.26 & -0.71 & 0.96 & & \\
0.90 & -0.015 & -0.35 & -0.07 & -0.35 & 0.42 & 1.67 & \\
0.90 & -0.09 & -0.45 & -0.18 & -0.33 & 0.18 & 0.88 & 1.74
\end{pmatrix}
\begin{array}{c} T_0 \\ \theta_0 \\ \mu_0 \\ \delta_0 \\ c_\mu \\ c_\delta \\ T_1 \\ T_2 \end{array}
\end{array}
\tag{5.33}
$$

Variance-covariance matrix

The variance covariance matrix for the random effects obtained from CD31$^+$ T cell model 3 is given in equation 5.34 below.

$$
\begin{array}{c}
\begin{array}{cccccccc} T_0 & \theta_0 & \mu_0 & \delta_0 & c_\mu & c_\delta & T_1 & T_2 \end{array} \\
\begin{pmatrix}
2.54 & & & & & & & \\
0.31 & 1.76 & & & & & & \\
-0.08 & 0.92 & 1.19 & & & & & \\
0.09 & 1.19 & 0.59 & 1.07 & & & & \\
-0.42 & 0.32 & 0.13 & 0.05 & 0.7 & & & \\
0.13 & -0.84 & -0.51 & -0.26 & -0.57 & 0.91 & & \\
2.39 & -0.32 & -0.65 & -0.13 & -0.48 & 0.67 & 2.8 & \\
2.50 & -0.2 & -0.85 & -0.33 & -0.48 & 0.30 & 2.57 & 3.04
\end{pmatrix}
\begin{array}{c} T_0 \\ \theta_0 \\ \mu_0 \\ \delta_0 \\ c_\mu \\ c_\delta \\ T_1 \\ T_2 \end{array}
\end{array}
\tag{5.34}
$$

5.4.4 Diagnostic plots



(A) (B)

Figure 5.9: (A): Visual predictive check (VPC) for total naive model A overlaid with data. (B): VPC for total naive model B applied to the data A. The gray dots correspond to the observed dataset whilst the dashed and solid red lines correspond to 2.5th, 50th and 97.5th centiles respectively from the observed data. The red and blue bands represent the 95% confidence intervals generated from the model-simulated centiles.

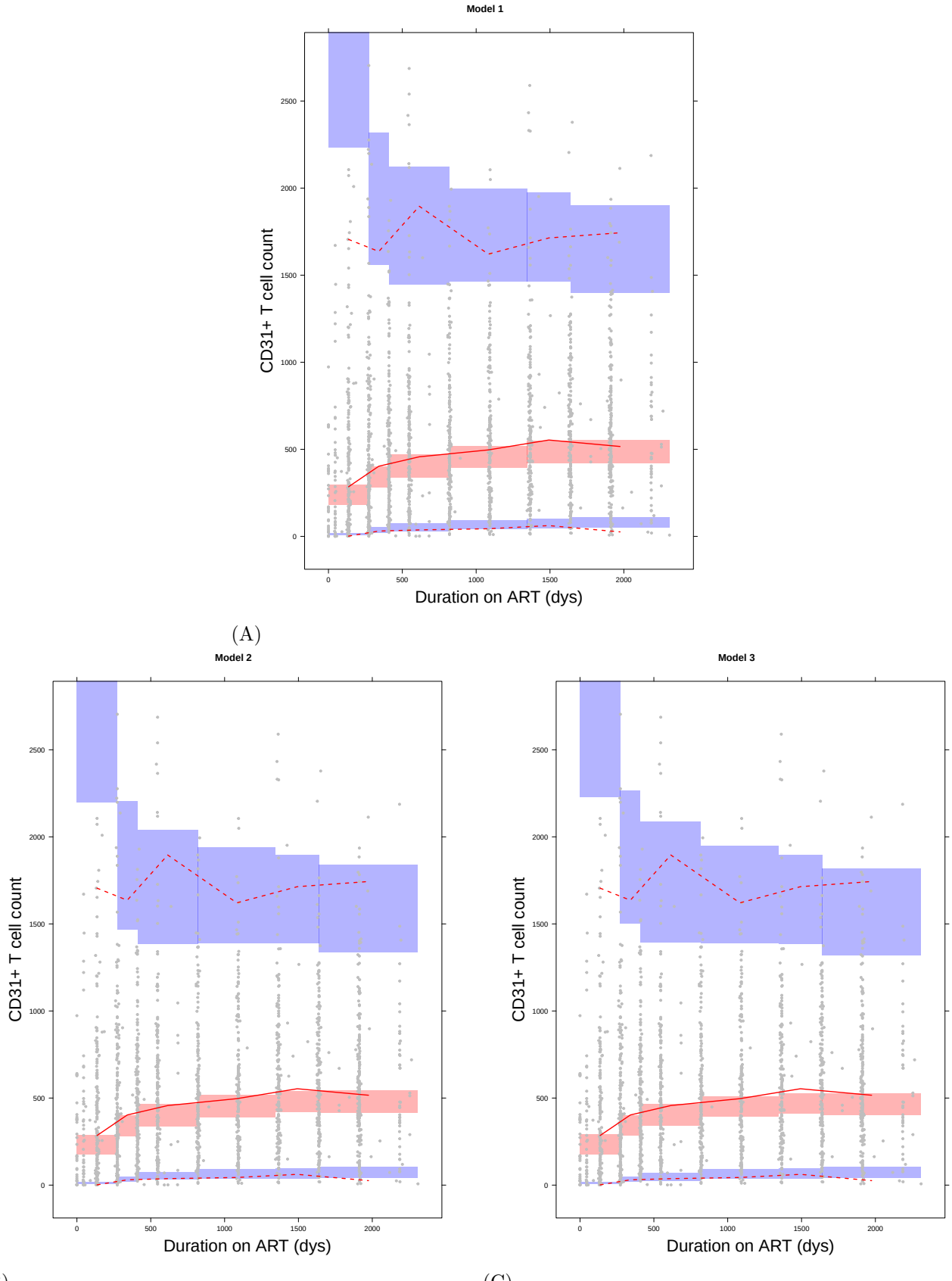


Figure 5.10: (A) to (C): Visual predictive checks (VPCs) for $CD31^+$ T cells models 1, 2 and 3 overlaid with original data. The gray dots correspond to the observed dataset whilst the dashed and solid red lines correspond to 2.5th, 50th and 97.5th centiles respectively from the observed data. The red and blue bands represent the 95% confidence intervals generated from the model-simulated centiles. Each VPC was generated from 2000 simulated datasets.

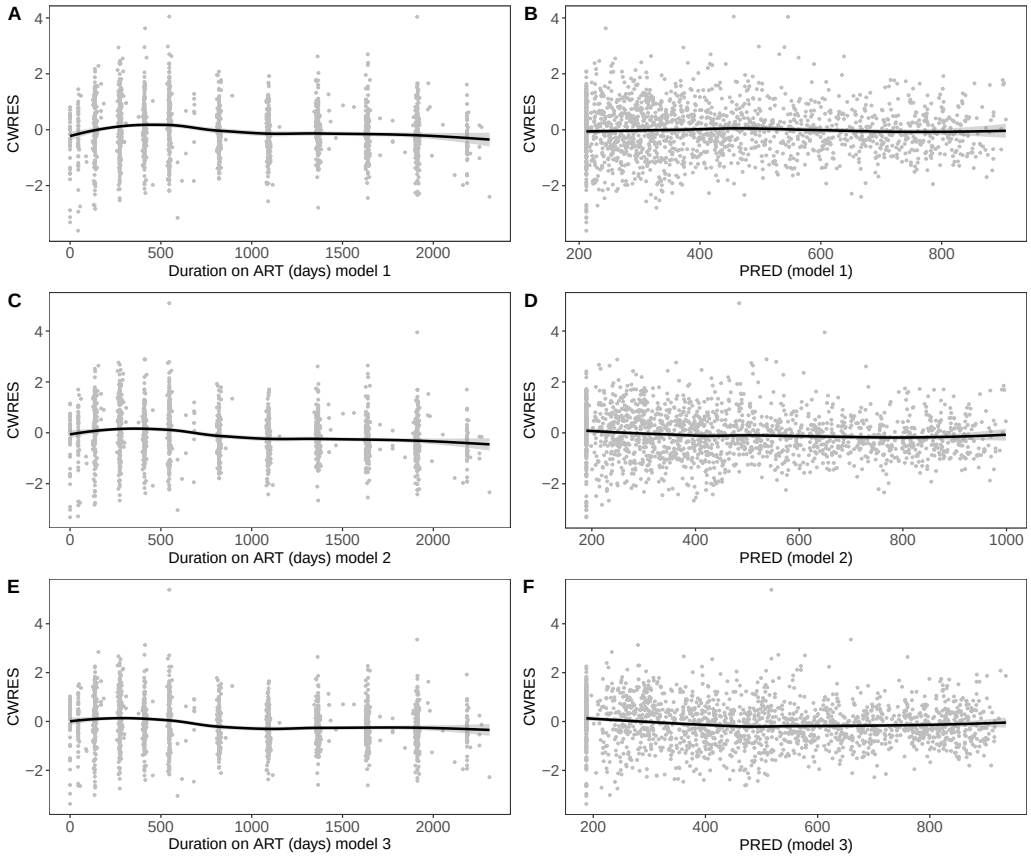


Figure 5.11: Residual plots for all three CD31$^+$ T cell models. Residuals on the left (A,C,E) are plotted against duration on ART (days) whilst residuals on the right are plotted against population predicted CD31$^+$ T cell counts (cells/μL).

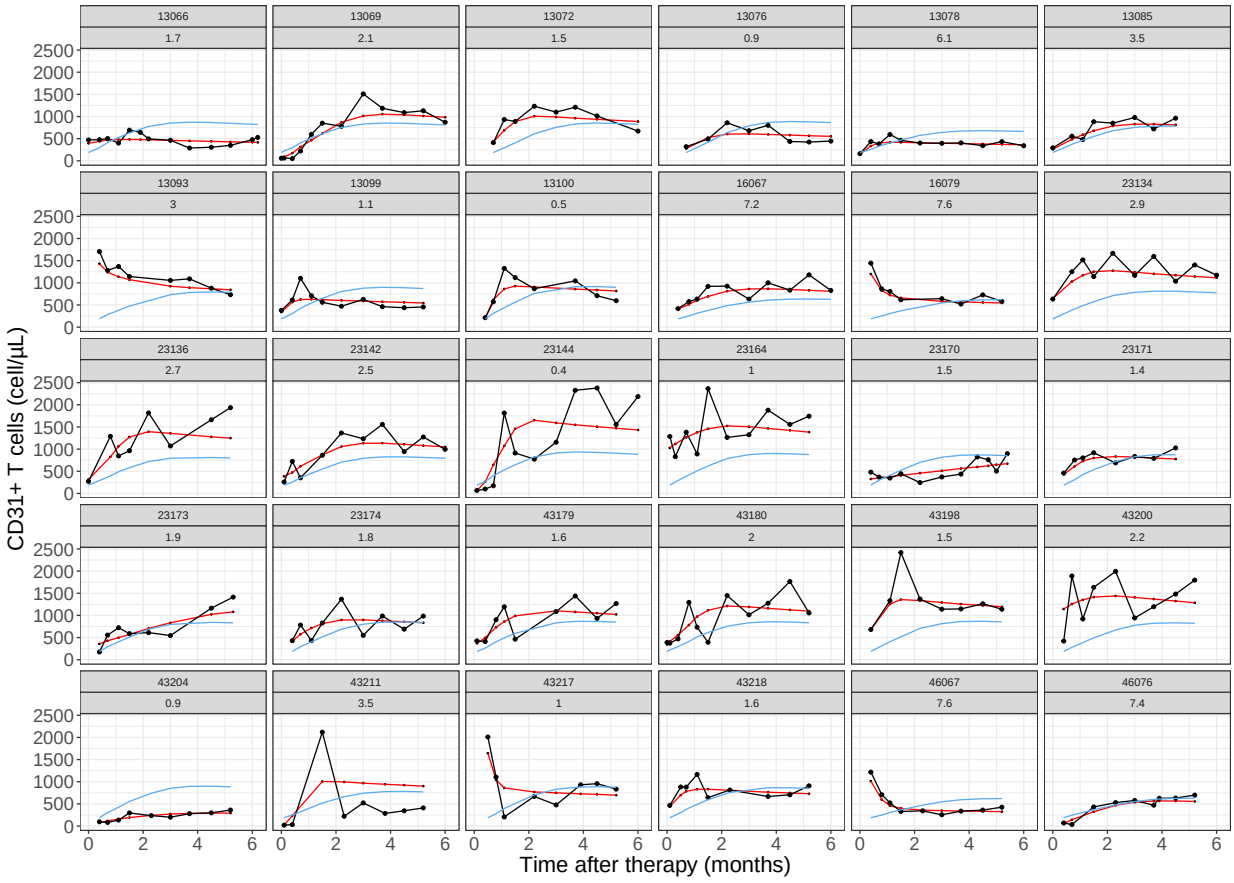


Figure 5.12: Individual predictions from the CD31$^+$ T cell model 3. Blue line: population prediction; Red line: individual predictions; Black line: observed CD4 counts. Top panel of each box shows the ID and age at start of ART for each child (yrs).

5.5 Discussion

In this chapter, models of total naive CD4$^+$ and CD31$^+$ T cells homeostasis were developed and fitted to longitudinal data obtained from 223 HIV-infected African children receiving ART. Unlike the approaches outlined in chapters 3 and 4, this was done without the need to transform naive CD31$^+$ T cells prior to fitting the model.

The model diagnostics and parameter estimates showed that the total naive CD4$^+$ T cell model with competition effect produced a better fit and gave more biologically plausible parameters and predictions especially at higher CD4$^+$ T cell concentrations compared to the model without competition effects. The addition of competition effects to the total naive CD4$^+$ T cell model is more representative of *in vivo* dynamics where resources are limited. In Figure 5.9a, the model without competition (model B) over-predicts total naive CD4$^+$ T cells at higher concentrations. Although the model with competition

effect estimated a larger proportion of T cells dividing ($\mu_0 = 0.09$/day), it also estimated a greater percentage of them dying ($\delta_0 = 0.134$/day). The magnitude of cell death ($c_\delta = 2.27$) is four times greater than that of cell proliferation ($c_\mu = 0.571$). Exponential decay constant from the competition-based model ($\lambda = 0.0004$) was faster compared to estimates from healthy children ($\lambda = 0.00027$). This is likely a reflection of faster turnover of total naive CD4$^+$ T cells in HIV-infected children. Hence, the model with competition effects was chosen as a basis for further exploration of CD31$^+$ T cell homeostasis.

In order to build models of naive CD4$^+$ T cell homeostasis, there is a need to understand how T cells proliferate and differentiate within the human body and to what degree they can divide without losing their CD31 phenotype. There are currently no published models of CD31$^+$ T cells in the literature and the CD31$^+$ T cell models developed in this work were based on a very limited understanding of *in vivo* dynamics. Of the three CD31$^+$ T cell models fitted to data, evidence from diagnostic plots and model analysis showed that model 3 produced the overall best fit. Hence further discussion of model results will be based on this model. The proportion of theoretical thymic output estimated was 2% of the predicted output from Bains et al. [Bains et al., 2009b]. The estimate for the Ki67 exponential decay constant was 0.0005 which is twice the value from Bains et al. (0.00027). The estimate from Bains et al was obtained from healthy children data and may therefore not mirror dynamics in HIV-infected children. Proportion of CD31$^+$ T cells proliferating in model 3 was 10% which is a larger estimate when compared to healthy children (0.8%). Likewise, proportion of cell loss (51%) was higher than estimates from healthy children (1.4%). Although equivalent estimates are unavailable in the literature, it agrees with the observation that HIV increases both T cell proliferation and loss [Hazenberg et al., 2000a, Zhang et al., 1999]. In an adult study by Wightman *et al*, it was shown that patients started on ART had a higher proportion of CD31$^+$ CD4$^+$ T cells when compared to ART naive HIV-infected and HIV-uninfected subjects [Wightman et al., 2010]. In addition, they found that HIV infected both CD31$^+$ and CD31$^-$ naive T cells indicating that both cell types are potential reservoirs in HIV-infected individuals. A lower estimate for proportion of thymic output could be due to the absence of viral load dynamics in the model.

Mechanistic models such as the one developed in this work give further insights into the recovery mechanisms of CD4$^+$ T cells in HIV-infected children. Furthermore, they can be used to examine mechanisms contributing to poor immune response in HIV-infected children receiving ART which was the subject of the preceding chapter (chapter 4). According to model 3, children with lower long-term CD31$^+$ T cell estimates (< 200 cells/μL) had much lower proportion of theoretical thymic outputs. This could indicate that poor thymic output stemming from a physiological or HIV-induced aetiology is a

reason for poor CD31$^+$ T cell recovery. This has implications for total CD4$^+$ T cell recovery in children since CD31$^+$ cells are an important source of naive CD4$^+$ T cells in HIV-infected children. In Figure 5.6, the CD31$^+$ T cell subset are observed to be the largest source of CD4$^+$ T cells and children who had poor immune recovery also showed a poor CD31$^+$ T cell recovery. In a study by Li et al., CD31% was measured in both immunological responders and nonresponders following successful ART. Thirteen virally suppressed immunological nonresponders were found to have lower proportion of recent thymic emigrants as evidenced by measured CD31%. This led them to conclude that thymic exhaustion could be responsible for the poor immune recovery observed in these patients [Li et al., 2011]. In a separate analysis of the ARROW data, children who died before the end of the trial were found to have lower thymic output and poorly reconstituted their CD4$^+$ T cell pool [Picat et al., 2013, Rollo Hoare, 2015]. In a recent study by Zakhour *et al*, 69 perinatally infected HIV-positive children and adolescents on therapy were followed up prospectively between January 2010 and 2012. They found that a low baseline RTE % of <50% independently predicted the clinical course of HIV disease progression in patients. As such, patients with lower RTE % (<50%) had worse CD4$^+$ T cell recovery as well as poor control of their viral load [Zakhour et al., 2016].

Proportion of theoretical thymic output, proliferation, loss and initial CD31$^+$ T cell count are seen to decrease with increasing age at start of ART. In particular, the local regression fit through the data in Figure 5.8 shows a biphasic decline: between birth and 5 years followed by a gradual increase and then a final decline from about 7 years. The reason for this could be physiological changes in thymic output with increasing age. There could also be a survival bias in that higher thymic output could be required in children who survived up to 7 years. Consequently, the increase between 5-7 yrs could be a reflection of children surviving longer with higher thymic outputs.

The high shrinkage observed in proliferation and loss parameters indicate that the two variables could be replaced by a single net loss parameter. Attempts at fitting the CD31$^+$ T cell models together with the total naive CD4$^+$ T cells models so far failed to converge for models 2 and 3. The reason for the fitting was to allow parameter estimates to draw additional information from the total naive CD4$^+$ T cell data.

This study has established some foundational work for mechanistic modelling of CD31$^+$ T cell homeostasis in HIV-infected children on ART. A number of simplifying assumptions were made in the model-building process (Figure 5.5). Firstly, the model assumes all naive T cells divide and die equally. However, in mice, there is evidence that recent thymic emigrants have a shorter lifespan compared to longer-lived naive T cells [Van Hoeven V., 2012]. Nonetheless, if this were to be the case, the expression

for the proportion of CD31$^+$ T cells in the long-term (equation 5.22) would remain independent of thymic output. Secondly, in order to keep the model simple, the effect of viral load was not added to the model. This could be responsible for the lower estimates obtained for proportion of thymic outputs in the CD31$^+$ model 3. Prior to considering models with more than 3 cycles of proliferation, viral load should be incorporated to make the basic model more realistic. It is not clear at this stage the appropriate number of divisions suitable for modeling CD31 T cells-this will require further modelling work with due consideration of viral load dynamics.

Other future work could include modifying covariance structure to reduce co-linearities and investigating potential effect of covariates on parameter estimates.

Contributions

Joanna Lewis [Joanna Lewis, 2012] proposed the model described in Figure 5.5 and formulated the equations described in section 5.3.4 (up to equation 5.22). I went further to find analytical solutions to the ODE models and computed their steady state behaviours (section 5.3.4, equations 5.3.4 to 5.23). The equations described in 5.10 and 5.11 were proposed by Rollo Hoare [Rollo Hoare, 2015, Hoare et al., 2017]. I produced all analytical and numerical implementations for ODEs in NONMEM 7.3.

Chapter 6

Conclusions

6.1 How can modelling CD4$^+$ T cell reconstitution improve understanding of HIV in children?

Understanding the processes underlying T cell recovery and homeostasis is fundamental in predicting outcomes in HIV-infected individuals started on anti-retroviral therapy. Despite progressive research efforts, the dynamics of CD4$^+$ T cells still poses many unanswered questions which are of primary interest both to the scientific community and clinicians who are involved in the day-to-day management of infected individuals.

The overall aim of this thesis was to investigate factors affecting CD4$^+$ T cell recovery in HIV-infected children starting ART as well as understand changes in naive T cell homeostasis over the course of immune recovery. An improved overview of immune recovery would facilitate a more efficient and individualised clinical management of HIV-infected children and could also inform global public health policies. In addition, it will help to identify early on patients who are at risk of poor immune recovery from start of ART who may benefit from targeted interventions with a view to improving their overall clinical outcome.

The technique of non-linear mixed-effects regression has been widely used in modelling repeated datasets such as longitudinal CD4$^+$ T cell measurements. This technique enables us to answer questions regarding longitudinal changes in the evolution of T cell recovery and factors that may be contributing to this. Both mechanistic and empirical models have been used to investigate changes in CD4$^+$ T cells over time in children [De Beaudrap et al., 2008, Hapuarachchi et al., 2013] and in adults [Thiébaut et al., 2006, De Beaudrap et al., 2009] receiving ART. Whilst the vast majority of the mechanistic models emphasised direct anti-viral killing of infected CD4$^+$ T cells [Guedj et al., 2007, Lavielle et al., 2011], other evidence suggested that there were other as yet unidentified factors contributing to T cell loss in HIV infection [Yates et al., 2007].

6.2 Age at start of ART and thymic output are key determinants of immune recovery in children

The first section of this study investigated the impact of hepatitis C co-infection on $CD4^+$ T cell recovery in 402 children receiving ART. The model chosen was an empirical asymptotic reconstitution model which has been widely published across the literature in modelling $CD4^+$ T cell recovery in children [De Beaudrap et al., 2008, Lewis et al., 2012, Li et al., 2011]. Fitting the model to data from the EPPICC and Ukraine Cohort showed that co-infected children had a much reduced rate of increase in $CD4^+$ T cell recovery. In addition to this, there were age and cohort effects on long-term and pre-ART CD4 z-scores respectively. This is the first time an empirical model describing $CD4^+$ T cell recovery has been used to investigate the effect of hepatitis C co-infection on T cell recovery in a large cohort of European children.

The findings in this study support previous studies [Lewis et al., 2012] demonstrating the importance of age on $CD4^+$ T cell reconstitution. While this study did not find evidence of any long-term effect of HIV/HCV co-infection on $CD4^+$ T cell recovery, the slower rate of recovery on ART does suggest that HIV/HCV co-infection may have other, as yet unidentified, immunological consequences. As other studies have shown, HIV-infected individuals with very small decreases in CD4 count are deemed to have increased risk of cardiovascular pathology and malignancies [Lichtenstein et al., 2010, Guiguet et al., 2009, Triant et al., 2010]. Early treatment of co-infected children with ART could therefore reduce overall morbidity and mortality [Lapadula et al., 2013, Tan et al., 2008].

There are various barriers to successful treatment of HCV in co-infected patients. Some of these include increased drug interactions and contraindication with ART drugs, hepatotoxicity and poor compliance to anti-HCV therapy. In children, there is a very limited number of antiretroviral drugs available to the clinician and interactions may prevent the use of certain drugs in combination (zidovudine, diadanosine and ribavirin) [England et al., 2009]. For these reasons, HCV clearance is very difficult to achieve in co-infected patients. In practice, both infections are not usually treated simultaneously and HIV viral control with ART usually precedes HCV therapy [Taylor et al., 2012]. Evidence from the literature and other cohort studies suggests that HIV/HCV co-infected patients have an increased risk of developing cirrhotic liver disease at a much early stage when compared to HCV mono-infected patients [Hernandez and Sherman, 2011]. Whilst these complications are not common in childhood, co-infected children are inevitably exposed to similar risks as they progress into adulthood. Therefore, more effective hepatitis C drugs such as the direct acting antivirals may need to be approved in children to minimize these risks.

Having established the negative impact of hepatitis C co-infection on $CD4^+$ T cell recovery, the next section of this report investigated other factors that could be responsible for poor immune recovery in HIV-infected children. The same empirical monophasic asymptotic model was fitted to a larger EPPICC dataset consisting of 2,204 children using a ln(CD4-for-age) transformation. Poor immune responders were identified by comparing the final predicted $CD4^+$ T cell count (after at least 6 months on ART) with the WHO guidelines for advanced to severe immunodeficiency. The analysis showed that higher age at start of ART and poor viral load control were both associated with poor immune recovery. There were a total of 7.3% and 7.4% advanced-severe poor immune responders in datasets A and B respectively. When considering only virally suppressed children ($\leq$400 copies/mL), the incidence of advanced to severe poor immune responders reduced to 4.3% and 3.9% in datasets A and B respectively. The percentage of poor immune responders in these cohorts is much lower when compared to an estimate of 20% in adults [Piketty et al., 1998, Autran et al., 1997]. This is not too surprising since younger children up to the age of 5 yrs have a much more active thymus resulting in higher circulating $CD4^+$ T cells compared to adults [Shearer et al., 2003, Schatorjé et al., 2012]. In addition, there is also a qualitative difference in the overall strength and magnitude of immune response in children compared to adults where children typically have a less effective immune response [Siegrist and Aspinall, 2009, Sandberg et al., 2003]. Hence, beyond the actual numbers of $CD4^+$ T cells, it is likely that functional immunological factors such as chronic T cell activation, T cell differentiation and cytokine profiles are just as important in understanding mechanisms contributing to poor immunological response.

Identifying predictors of poor immune recovery sets the stage for further investigation of the possible mechanisms responsible for incomplete $CD4^+$ T cell recovery using mechanistic models. A one-compartment model predicting $CD4^+$ T cell recovery in HIV-infected children has recently been developed [Rollo Hoare, 2015]. For the first time, this model was modified to predict $CD31^+$ T cell recovery in HIV-infected children from the ARROW trial. Age-related changes in $CD4^+$ T cell dynamics were incorporated directly on thymic output, proliferation and loss parameters. Although there were a number of strong colinearities between parameter values, this model accurately described evolution of $CD31^+$ T cells following commencement of ART. Addition of competition for resources made the model more reflective of *in vivo* dynamics. Poor immune responders were shown to have significantly reduced proportion of theoretical thymic outputs, had faster proliferation and slower death rates.

These findings illustrate the potential value of $CD31^+$ T cells as surrogate markers for identifying poor immune responders in HIV-infected children early on in therapy. It is therefore foreseeable in the near

future that $CD31^+$ count could be used alongside total $CD4^+$ count and viral load in facilitating effective management of HIV-infected children.

6.3 Future Work

In chapter 3, an empirical model was applied to investigate impact of hepatitis C co-infection on $CD4^+$ T cell recovery in HIV-infected children on ART. Initially, the model fitting was attempted using the "nlme" package in R. However, convergence was very difficult to achieve especially with random effect on the rate parameter (c_i) and outcome was very sensitive to small changes in starting estimates. The models did not converge despite tuning parameters and increasing number of iterations. In cases where the model did converge, results were not always reliable when visualised with diagnostic plots. The difficulties encountered with the "nlme" package is possibly due to the very sparse nature of the co-infection dataset. NONMEM software was able to handle random effects on all parameters and has the added benefit of a variety of EM algorithms to implement this. In addition to this, NONMEM has the advantage of having very robust ODE solving routines able to solve differential equations numerically (this is used in chapter 5). However, there were difficulties with convergence on addition of a full covariance structure on random effects of parameter estimates. In most cases, this was due to rounding errors in the parameter estimates. At other times, it was due to a nonpositive semi-definite Hessian matrix (matrix of the second derivatives of the objective function with respect to the parameters) indicating that a global minimum has not been achieved at the estimated parameters. These errors persisted despite tuning parameters, increasing number of iterations and lowering the number of significant digits required in the estimates. Hence the model could only accommodate covariance on two of the three parameters. This could be an indication that the data is uninformative but also emphasises the need for better fitting algorithms which can minimize convergence problems and produce robust parameter estimates. The dataset available for analysis in chapter 3 had a disproportionately large number of HIV mono-infected children which were all from Ukraine (89%). In the future, this model should be applied to a much more balanced dataset of co-infected and mono-infected European children. Effect of viral load on $CD4^+$ T cell recovery was investigated by adding this as a covariate. However, there is evidence to suggest that the dynamics of CD4 T cells and viral load is better modelled jointly using a mechanistic approach [Thiébaut et al., 2003].

The overall aim of treating patients with ART is not only to increase the $CD4^+$ T cell count but to improve the overall efficacy of immune response. In identifying poor immune responders, future work should include data on functional immunology of $CD4^+$ T cells such as cytokine profiles and T cell

activation markers. These can be incorporated into the model building process to give a fuller picture of T cell reconstitution. Where possible, the model should be applied to a different population of HIV-infected children (e.g. African children) to further investigate the generalisability of the results. The work on poor immune responders was restricted to the first 6 years of data on ART. It is not known how far into ART the asymptotic model can be applied and this could be explored much further in other studies.

The mechanistic model of $CD31^+$ T cells developed in chapter 5 is a very basic model that doesn't take into consideration the viral load dynamics. This is the first time a mechanistic model has been developed from $CD31^+$ T cells in HIV-infected children. Nonetheless the simulations and diagnostic plots show that the model describes the ARROW data well. In future work, viral load should be modelled jointly with $CD31^+$ T cells and age should be incorporated in a time-varying fashion to allow thymic output to reflect developmental and physiological changes in the immune system. Addition of viral load will reveal further insights as to any obvious patterns of viral dynamics that could be contributing to longitudinal trends in CD4 T cell recovery.

Throughout this work, models were fitted using maximum likelihood estimation methods which are widely implemented in many software packages. In the future, it would be worth applying Bayesian methods to fit the models. Bayesian methods allow the addition of priors such as variance covariance matrices which can be useful when applying a model to new datasets and distributions around parameter estimates can be obtained. A potential disadvantage of this is the introduction of errors in the event of wrong priors. One of the main challenges encountered in fitting ODE models in NOMEM is that of excessive run times especially with larger datasets. In the future, parallelization may be necessary in order to obtain results promptly and other techniques for validating covariate models such as the bootstrap could be achieved in a reasonable time frame. Such bootstrap procedures will have to be run using dedicated computer clusters with the necessary softwares pre-installed.

This doctoral thesis aimed to describe factors affecting $CD4^+$ T cell recovery in HIV-infected children receiving ART and to identify mechanisms driving this process with a view to shedding more light on poor immunological recovery. By using an empirical model, I have been able to gain further insights into the dynamics of hepatitis C co-infection and other factors affecting $CD4^+$ T cell reconstitution. In the latter part of this work, a mechanistic model has been developed which facilitates further understanding of naive

$CD4^+$ T cell homeostasis and therefore has great potential for future work which is envisaged to have applications in paediatric HIV, T cell homeostasis and clinical management of HIV-infected patients.

Bibliography

[Abdel-Hady et al., 2014] Abdel-Hady, M., Bansal, S., Davison, S. M., Brown, M., Tizzard, S. A., Mulla, S., Barnes, E., Davies, P., Mieli-Vergani, G., and Kelly, D. A. (2014). Treatment of chronic viral hepatitis C in children and adolescents: UK experience. *Arch Dis Child*, 99(6):505–10.

[Abram et al., 2010] Abram, M. E., Ferris, A. L., Shao, W., Alvord, W. G., and Hughes, S. H. (2010). Nature, position, and frequency of mutations made in a single cycle of HIV-1 replication. *J Virol*, 84(19):9864–78.

[Ahmad et al., 2005] Ahmad, A., Ahmad, R., Iannello, A., Toma, E., Morisset, R., and Sindhu, S. T. A. K. (2005). IL-15 and HIV infection: lessons for immunotherapy and vaccination. *Curr HIV Res*, 3(3):261–70.

[Alexaki and Wigdahl, 2008] Alexaki, A. and Wigdahl, B. (2008). Hiv-1 infection of bone marrow hematopoietic progenitor cells and their role in trafficking and viral dissemination. *PLoS Pathog*, 4(12):e1000215.

[Almeida et al., 2005] Almeida, A. R. M., Rocha, B., Freitas, A. A., and Tanchot, C. (2005). Homeostasis of T cell numbers: from thymus production to peripheral compartmentalization and the indexation of regulatory T cells. *Semin Immunol*, 17(3):239–49.

[Almeida et al., 2007] Almeida, M., Cordero, M., Almeida, J., and Orfao, A. (2007). Abnormal cytokine production by circulating monocytes and dendritic cells of myeloid origin in ART-treated HIV-1+ patients relates to CD4+ T-cell recovery and HCV co-infection. *Curr HIV Res*, 5(3):325–36.

[Anderson et al., 1998] Anderson, R. W., Ascher, M. S., and Sheppard, H. W. (1998). Direct HIV cytopathicity cannot account for CD4 decline in AIDS in the presence of homeostasis: a worst-case dynamic analysis. *J Acquir Immune Defic Syndr Hum Retrovirol*, 17(3):245–52.

[Antiretroviral Therapy Cohort Collaboration, 2008] Antiretroviral Therapy Cohort Collaboration (2008). Life expectancy of individuals on combination antiretroviral therapy in high-income countries: a collaborative analysis of 14 cohort studies. *Lancet*, 372(9635):293–9.

[ARROW Trial Team, 2016] ARROW Trial Team (2016). Anti-Retroviral Research for Watoto Trial. www.arrowtrial.org.

[ARROW Trial team et al., 2013] ARROW Trial team, Kekitiinwa, A., Cook, A., Nathoo, K., Mugyenyi, P., Nahirya-Ntege, P., Bakeera-Kitaka, S., Thomason, M., Bwakura-Dangarembizi, M., Musiime, V., Munderi, P., Naidoo-James, B., Vhembo, T., Tumusiime, C., Katuramu, R., Crawley, J., Prendergast, A. J., Musoke, P., Walker, A. S., and Gibb, D. M. (2013). Routine versus clinically driven laboratory monitoring and first-line antiretroviral therapy strategies in African children with HIV (ARROW): a 5-year open-label randomised factorial trial. *Lancet*, 381(9875):1391–403.

[Autran et al., 1997] Autran, B., Carcelain, G., Li, T. S., Blanc, C., Mathez, D., Tubiana, R., Katlama, C., Debré, P., and Leibowitch, J. (1997). Positive effects of combined antiretroviral therapy on CD4+ T cell homeostasis and function in advanced HIV disease. *Science*, 277(5322):112–6.

[Autran et al., 1999] Autran, B., Carcelaint, G., Li, T., Gorochov, G., Blanc, C., Renaud, M., Durali, M., Mathez, D., Calvez, V., Leibowitch, J., Katlama, C., and Debré, P. (1999). Restoration of the immune system with antiretroviral therapy. *Immunology letters*, 66(1-3):207–211.

[Azevedo et al., 2009] Azevedo, R. I., Soares, M. V. D., Barata, J. T., Tendeiro, R., Serra-Caetano, A., Victorino, R. M. M., and Sousa, A. E. (2009). IL-7 sustains CD31 expression in human naive CD4+ T cells and preferentially expands the CD31+ subset in a PI3K-dependent manner. *Blood*, 113(13):2999–3007.

[Bains et al., 2009a] Bains, I., Antia, R., Callard, R., and Yates, A. J. (2009a). Quantifying the development of the peripheral naive CD4+ T-cell pool in humans. *Blood*, 113(22):5480–7.

[Bains et al., 2009b] Bains, I., Thiébaut, R., Yates, A. J., and Callard, R. (2009b). Quantifying thymic export: combining models of naive T cell proliferation and TCR excision circle dynamics gives an explicit measure of thymic output. *J Immunol*, 183(7):4329–36.

[Bamford et al., 2015] Bamford, A., Turkova, A., Lyall, H., Foster, C., Klein, N., Bastiaans, D., Burger, D., Bernadi, S., Butler, K., Chiappini, E., Clayden, P., Della Negra, M., Giacomet, V., Giaquinto, C.,

Gibb, D., Galli, L., Hainaut, M., Koros, M., Marques, L., Nastouli, E., Niehues, T., Noguera-Julian, A., Rojo, P., Rudin, C., Scherpbier, H., Tudor-Williams, G., Welch, S., and (PENTA Steering Committee) (2015). Paediatric European Network for Treatment of AIDS (PENTA) guidelines for treatment of paediatric HIV-1 infection 2015: optimizing health in preparation for adult life. *HIV Med*.

[Bauer, 2013] Bauer, R. (2013). *Technical Guide on the Expectation-Maximization Population Analysis Methods in the NONMEM 7 Program*. ICON Development Solutions, 820 W. Diamond Avenue Suite 100 Gaithersburg, MD 20878.

[Beal and Sheiner, 1980] Beal, S. and Sheiner, L. (1980). The NONMEM system. *The American Statistician*, 34(2):118–119.

[Beal et al., 2014] Beal, S., Sheiner, L., Boeckmann, A., and ICON plc (2014). *Non-linear Mixed Effects Modelling*. ICON Development Solutions, 7.3.0 edition.

[Beal and Sheiner, 1988] Beal, S. L. and Sheiner, L. B. (1988). *NONMEM Users Guide–Part II Users Supplemental Guide*. The Regents of the University of California, San Francisco, USA, electronic copy produced march 2008 and august 2011 edition.

[Bellistrì et al., 2010] Bellistrì, G. M., Casabianca, A., Merlini, E., Orlandi, C., Ferrario, G., Meroni, L., Galli, M., Magnani, M., Monforte, A. d., and Marchetti, G. (2010). Increased bone marrow interleukin-7 (IL-7)/IL-7R levels but reduced IL-7 responsiveness in HIV-positive patients lacking CD4+ gain on antiviral therapy. *PLoS One*, 5(12):e15663.

[Benova et al., 2014] Benova, L., Mohamoud, Y. A., Calvert, C., and Abu-Raddad, L. J. (2014). Vertical transmission of hepatitis C virus: systematic review and meta-analysis. *Clin Infect Dis*, 59(6):765–73.

[Berger et al., 1999] Berger, E. A., Murphy, P. M., and Farber, J. M. (1999). Chemokine receptors as hiv-1 coreceptors: roles in viral entry, tropism, and disease. *Annu Rev Immunol*, 17:657–700.

[Berzins et al., 2002] Berzins, S. P., Uldrich, A. P., Sutherland, J. S., Gill, J., Miller, J. F. A. P., Godfrey, D. I., and Boyd, R. L. (2002). Thymic regeneration: teaching an old immune system new tricks. *Trends Mol Med*, 8(10):469–76.

[Birger et al., 2015] Birger, R., Kouyos, R., Dushoff, J., and Grenfell, B. (2015). Modeling the effect of HIV coinfection on clearance and sustained virologic response during treatment for hepatitis C virus. *Epidemics*, 12:1–10.

[Bortolotti et al., 1998] Bortolotti, F., Resti, M., Giacchino, R., Crivellaro, C., Zancan, L., Azzari, C., Gussetti, N., Tasso, L., and Faggion, S. (1998). Changing epidemiologic pattern of chronic hepatitis C virus infection in Italian children. *J Pediatr*, 133(3):378–81.

[Bortolotti et al., 2008] Bortolotti, F., Verucchi, G., Cammà, C., Cabibbo, G., Zancan, L., Indolfi, G., Giacchino, R., Marcellini, M., Marazzi, M. G., Barbera, C., Maggiore, G., Vajro, P., Bartolacci, S., Balli, F., Maccabruni, A., Guido, M., and Italian Observatory for HCV Infection and Hepatitis C in Children (2008). Long-term course of chronic hepatitis C in children: from viral clearance to end-stage liver disease. *Gastroenterology*, 134(7):1900–7.

[Bottarel et al., 2001] Bottarel, F., Bonissoni, S., Lucia, M. B., Bragardo, M., Bensi, T., Buonfiglio, D., Mezzatesta, C., DiFranco, D., Balotta, C., Capobianchi, M. R., Dianzani, I., Cauda, R., and Dianzani, U. (2001). Decreased function of Fas in patients displaying delayed progression of HIV-induced immune deficiency. *Hematol J*, 2(4):220–7.

[Boyman et al., 2009] Boyman, O., Létourneau, S., Krieg, C., and Sprent, J. (2009). Homeostatic proliferation and survival of naïve and memory T cells. *Eur J Immunol*, 39(8):2088–94.

[Boyman et al., 2007] Boyman, O., Purton, J. F., Surh, C. D., and Sprent, J. (2007). Cytokines and T-cell homeostasis. *Curr Opin Immunol*, 19(3):320–6.

[Bujdoso et al., 1989] Bujdoso, R., Young, P., Hopkins, J., Allen, D., and McConnell, I. (1989). Non-random migration of CD4 and CD8 T cells: changes in the CD4: CD8 ratio and interleukin 2 responsiveness of efferent lymph cells following in vivo antigen challenge. *Eur J Immunol*, 19(10):1779–84.

[Callard et al., 2003] Callard, R. E., Stark, J., and Yates, A. J. (2003). Fratricide: a mechanism for T memory-cell homeostasis. *Trends Immunol*, 24(7):370–5.

[Castro et al., 2016] Castro, M., Lythe, G., Molina-París, C., and Ribeiro, R. M. (2016). Mathematics in modern immunology. *Interface Focus*, 6(2):20150093.

[Chiappini et al., 2003] Chiappini, E., Galli, L., Zazzi, M., and de Martino, M. (2003). Immunological recovery despite virological failure is independent of human immunodeficiency virus-type 1 resistant mutants in children receiving highly active antiretroviral therapy. *J Med Virol*, 70(4):506–12.

[Cihlar and Fordyce, 2016] Cihlar, T. and Fordyce, M. (2016). Current status and prospects of HIV treatment. *Curr Opin Virol*, 18:50–56.

[Ciofani and Zúñiga-Pflücker, 2010] Ciofani, M. and Zúñiga-Pflücker, J. C. (2010). Determining $\gamma\delta$ versus $\alpha\beta$ T cell development. *Nat Rev Immunol*, 10(9):657–63.

[Claassen et al., 2012] Claassen, M. A. A., de Knegt, R. J., Turgut, D., Groothuismink, Z. M. A., Janssen, H. L. A., and Boonstra, A. (2012). Negative regulation of hepatitis C virus specific immunity is highly heterogeneous and modulated by pegylated interferon-alpha/ribavirin therapy. *PLoS One*, 7(11):e49389.

[Claassen et al., 2013] Claassen, M. A. A., Janssen, H. L. A., and Boonstra, A. (2013). Role of T cell immunity in hepatitis C virus infections. *Curr Opin Virol*, 3(4):461–7.

[Cleveland, 1979] Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American statistical association*, 74(368):829–836.

[Cleveland and Devlin, 1988] Cleveland, W. S. and Devlin, S. J. (1988). Locally weighted regression: an approach to regression analysis by local fitting. *Journal of the American statistical association*, 83(403):596–610.

[Conte et al., 2000] Conte, D., Fraquelli, M., Prati, D., Colucci, A., and Minola, E. (2000). Prevalence and clinical course of chronic hepatitis C virus (HCV) infection and rate of HCV vertical transmission in a cohort of 15,250 pregnant women. *Hepatology*, 31(3):751–5.

[Cotugno et al., 2012] Cotugno, N., Douagi, I., Rossi, P., and Palma, P. (2012). Suboptimal immune reconstitution in vertically HIV infected children: a view on how HIV replication and timing of HAART initiation can impact on T and B-cell compartment. *Clin Dev Immunol*, 2012:805151.

[Dabis et al., 2001] Dabis, F., Elenga, N., Meda, N., Leroy, V., Viho, I., Manigart, O., Dequae-Merchadou, L., Msellati, P., Sombie, I., and DITRAME Study Group (2001). 18-Month mortality and perinatal exposure to zidovudine in West Africa. *AIDS*, 15(6):771–9.

[Dalyon et al., 1999] Dalyon, B., Lavielle, M., and Moulines, E. (1999). Convergence of a stochastic approximation version of the expectation maximization algorithm. *Annals of Statistics*, 27(1):94–128.

[De Beaudrap et al., 2009] De Beaudrap, P., Etard, J.-F., Diouf, A., Ndiaye, I., Guèye, N. F., Guèye, P. M., Sow, P. S., Mboup, S., Ndoye, I., Ecochard, R., Eric, D., and ANRS 1215/90 Study Group (2009).

Modeling CD4+ cell count increase over a six-year period in HIV-1-infected patients on highly active antiretroviral therapy in Senegal. *Am J Trop Med Hyg*, 80(6):1047–53.

[De Beaudrap et al., 2008] De Beaudrap, P., Rouet, F., Fassinou, P., Kouakoussui, A., Mercier, S., Ecochard, R., and Msellati, P. (2008). CD4 cell response before and after HAART initiation according to viral load and growth indicators in HIV-1-infected children in Abidjan, Côte d'Ivoire. *J Acquir Immune Defic Syndr*, 49(1):70–6.

[De Rossi et al., 2002] De Rossi, A., Walker, A. S., Klein, N., De Forni, D., King, D., and Gibb, D. M. (2002). Increased thymic output after initiation of antiretroviral therapy in human immunodeficiency virus type 1-infected children in the Paediatric European Network for Treatment of AIDS (PENTA) 5 Trial. *J Infect Dis*, 186(3):312–20.

[Del Fava et al., 2011] Del Fava, E., Shkedy, Z., Hens, N., Aerts, M., Suligoi, B., Camoni, L., Vallejo, F., Wiessing, L., and Kretzschmar, M. (2011). Joint modeling of hcv and hiv co-infection among injecting drug users in italy and spain using individual cross-sectional data. *Statistical Communications in Infectious Diseases*, 3(1).

[Delobel et al., 2006] Delobel, P., Nugeyre, M.-T., Cazabat, M., Sandres-Sauné, K., Pasquier, C., Cuzin, L., Marchou, B., Massip, P., Cheynier, R., Barré-Sinoussi, F., Izopet, J., and Israël, N. (2006). Naive T-cell depletion related to infection by X4 human immunodeficiency virus type 1 in poor immunological responders to highly active antiretroviral therapy. *J Virol*, 80(20):10229–36.

[Demeure et al., 1996] Demeure, C. E., Byun, D. G., Yang, L. P., Vezzio, N., and Delespesse, G. (1996). CD31 (PECAM-1) is a differentiation antigen lost during human CD4 T-cell maturation into Th1 or Th2 effector cells. *Immunology*, 88(1):110–5.

[Dempster et al., 1977] Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the royal statistical society. Series B (methodological)*, pages 1–38.

[den Braber et al., 2012] den Braber, I., Mugwagwa, T., Vrisekoop, N., Westera, L., Mögling, R., de Boer, A. B., Willems, N., Schrijver, E. H. R., Spierenburg, G., Gaiser, K., Mul, E., Otto, S. A., Ruiter, A. F. C., Ackermans, M. T., Miedema, F., Borghans, J. A. M., de Boer, R. J., and Tesselaar, K. (2012). Maintenance of peripheral naive T cells is sustained by thymus output in mice but not humans. *Immunity*, 36(2):288–97.

[Douek, 2004] Douek, D. (2004). Thymic output and HIV infection: on the right TREC. *Immunity*, 21(6):744–5.

[Douek et al., 2000] Douek, D. C., Koup, R. A., McFarland, R. D., Sullivan, J. L., and Luzuriaga, K. (2000). Effect of HIV on thymic function before and after antiretroviral therapy in children. *J Infect Dis*, 181(4):1479–82.

[Douek et al., 1998] Douek, D. C., McFarland, R. D., Keiser, P. H., Gage, E. A., Massey, J. M., Haynes, B. F., Polis, M. A., Haase, A. T., Feinberg, M. B., Sullivan, J. L., Jamieson, B. D., Zack, J. A., Picker, L. J., and Koup, R. A. (1998). Changes in thymic function with age and during the treatment of HIV infection. *Nature*, 396(6712):690–5.

[Dowd et al., 2009] Dowd, K. A., Netski, D. M., Wang, X.-H., Cox, A. L., and Ray, S. C. (2009). Selection pressure from neutralizing antibodies drives sequence evolution during acute infection with hepatitis C virus. *Gastroenterology*, 136(7):2377–86.

[Durmaz, 2012] Durmaz, O. (2012). Hepatitis C infection in childhood. *Clin Res Hepatol Gastroenterol*, 36(3):294–6.

[Dustin and Rice, 2007] Dustin, L. B. and Rice, C. M. (2007). Flying under the radar: the immunobiology of hepatitis C. *Annu Rev Immunol*, 25:71–99.

[Duszczyszyn et al., 2006] Duszczyszyn, D. A., Beck, J. D., Antel, J., Bar-Or, A., Lapierre, Y., Gadag, V., and Haegert, D. G. (2006). Altered naive CD4 and CD8 T cell homeostasis in patients with relapsing-remitting multiple sclerosis: thymic versus peripheral (non-thymic) mechanisms. *Clin Exp Immunol*, 143(2):305–13.

[Dutilh and de Boer, 2003] Dutilh, B. E. and de Boer, R. J. (2003). Decline in excision circles requires homeostatic renewal or homeostatic death of naive T cells. *J Theor Biol*, 224(3):351–8.

[El-Shabrawi and Kamal, 2013] El-Shabrawi, M. H. and Kamal, N. M. (2013). Burden of pediatric hepatitis C. *World J Gastroenterol*, 19(44):7880–8.

[England et al., 2009] England, K., Thorne, C., Castelli-Gattinara, G., Vigano, A., El Mehabresh, M. I., and Newell, M.-L. (2009). HIV and HCV progression in parenterally coinfected children. *Curr HIV Res*, 7(3):346–53.

[England et al., 2006] England, K., Thorne, C., and Newell, M.-L. (2006). Vertically acquired paediatric coinfection with HIV and hepatitis C virus. *Lancet Infect Dis*, 6(2):83–90.

[Engsig et al., 2010] Engsig, F. N., Gerstoft, J., Kronborg, G., Larsen, C. S., Pedersen, G., Røge, B., Jensen, J., Nielsen, L. N., and Obel, N. (2010). Long-term mortality in HIV patients virally suppressed for more than three years with incomplete CD4 recovery: a cohort study. *BMC Infect Dis*, 10:318.

[EPPICC in EuroCoord, 2016] EPPICC in EuroCoord (2016). The European Paediatric HIV/HCV Co-infection Study Group in the European Pregnancy and Paediatric HIV Cohort Collaboration. Co-infection with HIV and HCV in 229 children and young adults living in Europe. *AIDS*.

[Erikstrup et al., 2010] Erikstrup, C., Kronborg, G., Lohse, N., Ostrowski, S. R., Gerstoft, J., and Ullum, H. (2010). T-cell dysfunction in HIV-1-infected patients with impaired recovery of CD4 cells despite suppression of viral replication. *J Acquir Immune Defic Syndr*, 53(3):303–10.

[Esteban et al., 2008] Esteban, J. I., Sauleda, S., and Quer, J. (2008). The changing epidemiology of hepatitis C virus infection in Europe. *J Hepatol*, 48(1):148–62.

[European Paediatric Hepatitis C Virus Network, 2005] European Paediatric Hepatitis C Virus Network (2005). Three broad modalities in the natural history of vertically acquired hepatitis C virus infection. *Clin Infect Dis*, 41(1):45–51.

[Falconer et al., 2008] Falconer, K., Gonzalez, V. D., Reichard, O., Sandberg, J. K., and Alaeus, A. (2008). Spontaneous HCV clearance in HCV/HIV-1 coinfection associated with normalized CD4 counts, low level of chronic immune activation and high level of T cell function. *J Clin Virol*, 41(2):160–3.

[Fifth IAS Conference on HIV Pathogenesis, 2009] Fifth IAS Conference on HIV Pathogenesis, editor (2009). *Nevirapine (NVP) vs lopinavir-ritonavir (LPV/r)-based antiretroviral therapy (ART) in single dose nevirapine (sdNVP)-exposed HIV-infected infants: preliminary results from the IMPAACT P1060 trial*, Cape Town, South Africa. Fifth IAS Conference on HIV Pathogenesis, Treatment and Prevention.

[Fink, 2013] Fink, P. J. (2013). The biology of recent thymic emigrants. *Annu Rev Immunol*, 31:31–50.

[Floreani, 2013] Floreani, A. (2013). Hepatitis C and pregnancy. *World J Gastroenterol*, 19(40):6714–20.

[Freitas and Rocha, 2000] Freitas, A. A. and Rocha, B. (2000). Population biology of lymphocytes: the flight for survival. *Annu Rev Immunol*, 18:83–111.

[Gale and Foy, 2005] Gale, Jr, M. and Foy, E. M. (2005). Evasion of intracellular host defence by hepatitis C virus. *Nature*, 436(7053):939–45.

[Geraci, 2014] Geraci, M. (2014). Linear quantile mixed models: The lqmm package for laplace quantile regression. *Journal of Statistical Software*, 57(13).

[Gerlach et al., 2003] Gerlach, J. T., Diepolder, H. M., Zachoval, R., Gruener, N. H., Jung, M.-C., Ulsenheimer, A., Schraut, W. W., Schirren, C. A., Waechtler, M., Backmund, M., and Pape, G. R. (2003). Acute hepatitis C: high rate of both spontaneous and treatment-induced viral clearance. *Gastroenterology*, 125(1):80–8.

[Gibb et al., 2000] Gibb, D. M., Newberry, A., Klein, N., de Rossi, A., Grosch-Woerner, I., and Babiker, A. (2000). Immune repopulation after HAART in previously untreated HIV-1-infected children. Paediatric European Network for Treatment of AIDS (PENTA) Steering Committee. *Lancet*, 355(9212):1331–2.

[Giorgi et al., 1999] Giorgi, J. V., Hultin, L. E., McKeating, J. A., Johnson, T. D., Owens, B., Jacobson, L. P., Shih, R., Lewis, J., Wiley, D. J., Phair, J. P., Wolinsky, S. M., and Detels, R. (1999). Shorter survival in advanced human immunodeficiency virus type 1 infection is more closely associated with T lymphocyte activation than with plasma virus burden or virus chemokine coreceptor usage. *J Infect Dis*, 179(4):859–70.

[Golubovskaya and Wu, 2016] Golubovskaya, V. and Wu, L. (2016). Different subsets of t cells, memory, effector functions, and car-t immunotherapy. *Cancers (Basel)*, 8(3).

[Gomez et al., 2003] Gomez, I., Hainz, U., Jenewein, B., Schwaiger, S., Wolf, A. M., and Grubeck-Loebenstein, B. (2003). Changes in the expression of CD31 and CXCR3 in CD4+ naïve T cells in elderly persons. *Mech Ageing Dev*, 124(4):395–402.

[Gonzalez et al., 2009] Gonzalez, V. D., Falconer, K., Blom, K. G., Reichard, O., Mørn, B., Laursen, A. L., Weis, N., Alaeus, A., and Sandberg, J. K. (2009). High levels of chronic immune activation in the T-cell compartments of patients coinfected with hepatitis C virus and human immunodeficiency virus type 1 and on highly active antiretroviral therapy are reverted by alpha interferon and ribavirin treatment. *J Virol*, 83(21):11407–11.

[Gossel et al., 2017] Gossel, G., Hogan, T., Cownden, D., Seddon, B., and Yates, A. J. (2017). Memory CD4 T cell subsets are kinetically heterogeneous and replenished from naive T cells at high levels. *Elife*, 6.

[Gougeon et al., 1996] Gougeon, M. L., Lecoeur, H., Dulioust, A., Enouf, M. G., Crouvoiser, M., Goujard, C., Debord, T., and Montagnier, L. (1996). Programmed cell death in peripheral lymphocytes from HIV-infected persons: increased susceptibility to apoptosis of CD4 and CD8 T cells correlates with lymphocyte activation and with disease progression. *J Immunol*, 156(9):3509–20.

[Greub et al., 2000] Greub, G., Ledergerber, B., Battegay, M., Grob, P., Perrin, L., Furrer, H., Burgisser, P., Erb, P., Boggian, K., Piffaretti, J. C., Hirschel, B., Janin, P., Francioli, P., Flepp, M., and Telenti, A. (2000). Clinical progression, survival, and immune recovery during antiretroviral therapy in patients with HIV-1 and hepatitis C virus coinfection: the Swiss HIV Cohort Study. *Lancet*, 356(9244):1800–5.

[Grossman et al., 2002] Grossman, Z., Meier-Schellersheim, M., Sousa, A. E., Victorino, R. M. M., and Paul, W. E. (2002). CD4+ T-cell depletion in HIV infection: are we closer to understanding the cause? *Nat Med*, 8(4):319–23.

[Guadalupe et al., 2003] Guadalupe, M., Reay, E., Sankaran, S., Prindiville, T., Flamm, J., McNeil, A., and Dandekar, S. (2003). Severe CD4+ T-cell depletion in gut lymphoid tissue during primary human immunodeficiency virus type 1 infection and substantial delay in restoration following highly active antiretroviral therapy. *J Virol*, 77(21):11708–17.

[Guedj et al., 2007] Guedj, J., Thiébaut, R., and Commenges, D. (2007). Maximum likelihood estimation in dynamical models of HIV. *Biometrics*, 63(4):1198–206.

[Guiguet et al., 2009] Guiguet, M., Boué, F., Cadranel, J., Lang, J.-M., Rosenthal, E., Costagliola, D., and Clinical Epidemiology Group of the FHDH-ANRS CO4 cohort (2009). Effect of immunodeficiency, HIV viral load, and antiretroviral therapy on the risk of individual malignancies (FHDH-ANRS CO4): a prospective cohort study. *Lancet Oncol*, 10(12):1152–9.

[Haase, 1999] Haase, A. T. (1999). Population biology of HIV-1 infection: viral and CD4+ T cell demographics and dynamics in lymphatic tissues. *Annu Rev Immunol*, 17:625–56.

[Hainaut et al., 2003] Hainaut, M., Ducarme, M., Schandene, L., Peltier, C. A., Marissens, D., Zissis, G., Mascart, F., and Levy, J. (2003). Age-related immune reconstitution during highly active antiretroviral therapy in human immunodeficiency virus type 1-infected children. *Pediatr Infect Dis J*, 22(1):62–9.

[Hapuarachchi et al., 2013] Hapuarachchi, T., Lewis, J., and Callard, R. E. (2013). A mechanistic model for naive CD4 T cell homeostasis in healthy adults and children. *Front Immunol*, 4:366.

[Hartling et al., 2013] Hartling, H. J., Gaardbo, J. C., Ronit, A., Salem, M., Laye, M., Clausen, M. R., Skogstrand, K., Gerstoft, J., Ullum, H., and Nielsen, S. D. (2013). Impaired thymic output in patients with chronic hepatitis C virus infection. *Scand J Immunol*, 78(4):378–86.

[Hasse et al., 2011] Hasse, B., Ledergerber, B., Furrer, H., Battegay, M., Hirschel, B., Cavassini, M., Bertisch, B., Bernasconi, E., Weber, R., and Swiss HIV Cohort Study (2011). Morbidity and aging in HIV-infected persons: the Swiss HIV cohort study. *Clin Infect Dis*, 53(11):1130–9.

[Hatzakis et al., 2000] Hatzakis, A., Touloumi, G., Karanicolas, R., Karafoulidou, A., Mandalaki, T., Anastassopoulou, C., Zhang, L., Goedert, J. J., Ho, D. D., and Kostrikis, L. G. (2000). Effect of recent thymic emigrants on progression of HIV-1 disease. *Lancet*, 355(9204):599–604.

[Haynes et al., 2000] Haynes, B. F., Markert, M. L., Sempowski, G. D., Patel, D. D., and Hale, L. P. (2000). The role of the thymus in immune reconstitution in aging, bone marrow transplantation, and HIV-1 infection. *Annu Rev Immunol*, 18:529–60.

[Hazenberg et al., 2003a] Hazenberg, M. D., Borghans, J. A. M., de Boer, R. J., and Miedema, F. (2003a). Thymic output: a bad TREC record. *Nat Immunol*, 4(2):97–9.

[Hazenberg et al., 2000a] Hazenberg, M. D., Otto, S. A., Cohen Stuart, J. W., Verschuren, M. C., Borleffs, J. C., Boucher, C. A., Coutinho, R. A., Lange, J. M., Rinke de Wit, T. F., Tsegaye, A., van Dongen, J. J., Hamann, D., de Boer, R. J., and Miedema, F. (2000a). Increased cell division but not thymic dysfunction rapidly affects the T-cell receptor excision circle content of the naive T cell population in HIV-1 infection. *Nat Med*, 6(9):1036–42.

[Hazenberg et al., 2003b] Hazenberg, M. D., Otto, S. A., van Benthem, B. H. B., Roos, M. T. L., Coutinho, R. A., Lange, J. M. A., Hamann, D., Prins, M., and Miedema, F. (2003b). Persistent immune activation in HIV-1 infection is associated with progression to AIDS. *AIDS*, 17(13):1881–8.

[Hazenberg et al., 2004] Hazenberg, M. D., Otto, S. A., van Rossum, A. M. C., Scherpbier, H. J., de Groot, R., Kuijpers, T. W., Lange, J. M. A., Hamann, D., de Boer, R. J., Borghans, J. A. M., and Miedema, F. (2004). Establishment of the CD4+ T-cell pool in healthy children and untreated children infected with HIV-1. *Blood*, 104(12):3513–9.

[Hazenberg et al., 2000b] Hazenberg, M. D., Stuart, J. W., Otto, S. A., Borleffs, J. C., Boucher, C. A., de Boer, R. J., Miedema, F., and Hamann, D. (2000b). T-cell division in human immunodeficiency virus (HIV)-1 infection is mainly due to immune activation: a longitudinal analysis in patients before and during highly active antiretroviral therapy (HAART). *Blood*, 95(1):249–55.

[Hernandez and Sherman, 2011] Hernandez, M. D. and Sherman, K. E. (2011). HIV/hepatitis C coinfection natural history and disease progression. *Curr Opin HIV AIDS*, 6(6):478–82.

[Hintze and Nelson, 1998] Hintze, J. L. and Nelson, R. D. (1998). Violin plots: A box plot-density trace synergism. *The American Statistician*, 52(2):181–184.

[HIV Paediatric Prognostic Markers Collaborative Study et al., 2010] HIV Paediatric Prognostic Markers Collaborative Study, Boyd, K., Dunn, D. T., Castro, H., Gibb, D. M., Duong, T., Aboulker, J. P., Bulterys, M., Cortina-Borja, M., Gabiano, C., Galli, L., Giaquinto, C., Harris, D. R., HugheS, M., McKinney, R., Mofenson, L., Moye, J., Newell, M. L., Pahwa, S., Palumbo, P., Rudin, C., Sharland, M., Shearer, W., Thompson, B., and Tookey, P. (2010). Discordance between CD4 cell count and CD4 cell percentage: implications for when to start antiretroviral therapy in HIV-1 infected children. *AIDS*, 24(8):1213–7.

[Hoare et al., 2017] Hoare, R. L., Veys, P., Klein, N., Callard, R., and Standing, J. F. (2017). Predicting CD4 T-cell reconstitution following paediatric haematopoietic stem cell transplantation. *Clin Pharmacol Ther*.

[Hochman and Balistreri, 2007] Hochman, J. A. M. and Balistreri, W. F. M. (2007). Acute and Chronic Viral Hepatitis. In Suchy, F. J., Sokol, R. J., and Balistreri, W. F., editors, *Liver Disease in Children*, pages 369–446. Cambridge University Press, third edition.

[Hodge et al., 2011] Hodge, J. N., Srinivasula, S., Hu, Z., Read, S. W., Porter, B. O., Kim, I., Mican, J. M., Paik, C., Degrange, P., Di Mascio, M., and Sereti, I. (2011). Decreases in IL-7 levels during antiretroviral treatment of HIV infection suggest a primary mechanism of receptor-mediated clearance. *Blood*, 118(12):3244–53.

[Holford, 2005] Holford, N. (2005). The visual predictive check-superiority to standard diagnostic (Rorshach) plots. *14th meeting of the Population Approach Group in Europe, Pamplona, Spain, PAGE Abstract*, 738.

[Hooker et al., 2007] Hooker, A. C., Staatz, C. E., and Karlsson, M. O. (2007). Conditional weighted residuals (CWRES): a model diagnostic for the FOCE method. *Pharm Res*, 24(12):2187–97.

[Huenecke et al., 2008] Huenecke, S., Behl, M., Fadler, C., Zimmermann, S. Y., Bochennek, K., Tramsen, L., Esser, R., Klarmann, D., Kamper, M., Sattler, A., von Laer, D., Klingebiel, T., Lehrnbecher, T., and Koehl, U. (2008). Age-matched lymphocyte subpopulation reference values in childhood and adolescence: application of exponential regression analysis. *Eur J Haematol*, 80(6):532–9.

[Hulstaert et al., 1994] Hulstaert, F., Hannet, I., Deneys, V., Munhyeshuli, V., Reichert, T., De Bruyere, M., and Strauss, K. (1994). Age-related changes in human blood lymphocyte subpopulations. II. Varying kinetics of percentage and absolute count measurements. *Clin Immunol Immunopathol*, 70(2):152–8.

[Hunt et al., 2003] Hunt, P. W., Martin, J. N., Sinclair, E., Bredt, B., Hagos, E., Lampiris, H., and Deeks, S. G. (2003). T cell activation is associated with lower CD4+ T cell gains in human immunodeficiency virus-infected patients with sustained viral suppression during antiretroviral therapy. *J Infect Dis*, 187(10):1534–43.

[Indolfi et al., 2013] Indolfi, G., Azzari, C., and Resti, M. (2013). Perinatal transmission of hepatitis C virus. *J Pediatr*, 163(6):1549–1552.e1.

[Indolfi et al., 2015] Indolfi, G., Bartolini, E., Serranti, D., Azzari, C., and Resti, M. (2015). Hepatitis C in Children Co-infected With Human Immunodeficiency Virus. *J Pediatr Gastroenterol Nutr*, 61(4):393–9.

[International Genetically Modified Machine (iGEM) Foundation, 2016] International Genetically Modified Machine (iGEM) Foundation (2016). Acquired Immune Deficiency Syndrome. http://2015.igem.org/Team:Freiburg/Project/Diseases.

[Isgrò et al., 2008] Isgrò, A., Leti, W., De Santis, W., Marziali, M., Esposito, A., Fimiani, C., Luzi, G., Pinti, M., Cossarizza, A., Aiuti, F., and Mezzaroma, I. (2008). Altered clonogenic capability and stromal cell function characterize bone marrow of HIV-infected subjects with low CD4+ T cell counts despite viral suppression during HAART. *Clin Infect Dis*, 46(12):1902–10.

[Jameson, 2005] Jameson, S. C. (2005). T cell homeostasis: keeping useful T cells alive and live T cells useful. *Semin Immunol*, 17(3):231–7.

[Jang et al., 2013] Jang, Y.-S., Kang, W., Chang, D.-Y., Sung, P. S., Park, B.-C., Yoo, S. H., Park, Y. W., and Shin, E.-C. (2013). CD27 engagement by a soluble CD70 protein enhances non-cytolytic antiviral activity of CD56bright natural killer cells by IFN-γ secretion. *Clin Immunol*, 149(3):379–87.

[Jenkins et al., 1998] Jenkins, M., Hanley, M. B., Moreno, M. B., Wieder, E., and McCune, J. M. (1998). Human immunodeficiency virus-1 infection interrupts thymopoiesis and multilineage hematopoiesis in vivo. *Blood*, 91(8):2672–8.

[Joanna Lewis, 2012] Joanna Lewis (2012). *Mathematical and Statistical Modelling of CD4 T cell reconstitution in HIV-infected children starting antiretroviral therapy.* PhD thesis, University College London, Gower St, London WC1E 6BT.

[Jonsson and Karlsson, 1998] Jonsson, E. N. and Karlsson, M. O. (1998). Automated covariate model building within NONMEM. *Pharm Res*, 15(9):1463–8.

[Jonsson et al., 2007] Jonsson, F., Jonsson, E. N., Bois, F. Y., and Marshall, S. (2007). The application of a bayesian approach to the analysis of a complex, mechanistically based model. *J Biopharm Stat*, 17(1):65–92.

[Jose C. Pinheiro, Douglas M. Bates, 2000] Jose C. Pinheiro, Douglas M. Bates (2000). *Mixed-Effects Models in S and S-PLUS.* Springer Verlag New York, LLC.

[Karlsson and Holford, 2008] Karlsson, M. O. and Holford, N. (2008). A tutorial on visual predictive checks. *17th meeting of the Population Approach Group in Europe, Marsielle, France, PAGE Abstract*, 1434.

[Karlsson and Savic, 2007] Karlsson, M. O. and Savic, R. M. (2007). Diagnosing model diagnostics. *Clin Pharmacol Ther*, 82(1):17–20.

[Karnsakul et al., 2009] Karnsakul, W., Alford, M. K., and Schwarz, K. B. (2009). Managing pediatric hepatitis C: current and emerging treatment options. *Ther Clin Risk Manag*, 5(3):651–60.

[Kaufmann et al., 2002] Kaufmann, G. R., Bloch, M., Finlayson, R., Zaunders, J., Smith, D., and Cooper, D. A. (2002). The extent of HIV-1-related immunodeficiency and age predict the long-term CD4 T lymphocyte response to potent antiretroviral therapy. *AIDS*, 16(3):359–67.

[Kaufmann et al., 2005] Kaufmann, G. R., Furrer, H., Ledergerber, B., Perrin, L., Opravil, M., Vernazza, P., Cavassini, M., Bernasconi, E., Rickenbach, M., Hirschel, B., Battegay, M., and Swiss HIV Cohort Study (2005). Characteristics, determinants, and clinical relevance of CD4 T cell recovery to <500 cells/microL in HIV type 1-infected individuals receiving potent antiretroviral therapy. *Clin Infect Dis*, 41(3):361–72.

[Kaufmann et al., 2003] Kaufmann, G. R., Perrin, L., Pantaleo, G., Opravil, M., Furrer, H., Telenti, A., Hirschel, B., Ledergerber, B., Vernazza, P., Bernasconi, E., Rickenbach, M., Egger, M., Battegay, M., and Swiss HIV Cohort Study Group (2003). CD4 T-lymphocyte recovery in individuals with advanced HIV-1 infection receiving potent antiretroviral therapy for 4 years: the Swiss HIV Cohort Study. *Arch Intern Med*, 163(18):2187–95.

[Ke and Wang, 2001] Ke, C. and Wang, Y. (2001). Semiparametric nonlinear mixed-effects models and their applications. *Journal of the American Statistical Association*, 96(456):1272–1281.

[Kekitiinwa et al., 2008] Kekitiinwa, A., Lee, K. J., Walker, A. S., Maganda, A., Doerholt, K., Kitaka, S. B., Asiimwe, A., Judd, A., Musoke, P., Gibb, D. M., Collaborative HIV Paediatric Study (CHIPS) Steering Committe, and Mulago Cohort Team (2008). Differences in factors associated with initial growth, CD4, and viral load responses to ART in HIV-infected children in Kampala, Uganda, and the United Kingdom/Ireland. *J Acquir Immune Defic Syndr*, 49(4):384–92.

[Kenneth Todar, 2006] Kenneth Todar (2006). *Todar's Online Textbook of Bacteriology*. Madison, Wisconsin: University of Wisconsin-Madison Department of Bacteriology.

[Keonker, 2004] Keonker, R. (2004). Quantile regression for longitudinal data. *Journal of Multivariate Analysis*, 91(1):74 – 89. Special Issue on Semiparametric and Nonparametric Mixed Models.

[Kieper et al., 2004] Kieper, W. C., Burghardt, J. T., and Surh, C. D. (2004). A role for TCR affinity in regulating naive T cell homeostasis. *J Immunol*, 172(1):40–4.

[Kilpatrick et al., 2008] Kilpatrick, R. D., Rickabaugh, T., Hultin, L. E., Hultin, P., Hausner, M. A., Detels, R., Phair, J., and Jamieson, B. D. (2008). Homeostasis of the naive CD4+ T cell compartment during aging. *J Immunol*, 180(3):1499–507.

[Kim and Chung, 2009] Kim, A. Y. and Chung, R. T. (2009). Coinfection with HIV-1 and HCV–a one-two punch. *Gastroenterology*, 137(3):795–814.

[Kimmig et al., 2002] Kimmig, S., Przybylski, G. K., Schmidt, C. A., Laurisch, K., Möwes, B., Radbruch, A., and Thiel, A. (2002). Two subsets of naive T helper cells with distinct T cell receptor excision circle content in human adult peripheral blood. *J Exp Med*, 195(6):789–94.

[Klenerman and Semmo, 2006] Klenerman, P. and Semmo, N. (2006). Cellular immune responses against persistent hepatitis C virus: gone but not forgotten. *Gut*, 55(7):914–6.

[Kohler and Thiel, 2009] Kohler, S. and Thiel, A. (2009). Life after the thymus: CD31+ and CD31- human naive CD4+ T-cell subsets. *Blood*, 113(4):769–74.

[Kohler et al., 2005] Kohler, S., Wagner, U., Pierer, M., Kimmig, S., Oppmann, B., Möwes, B., Jülke, K., Romagnani, C., and Thiel, A. (2005). Post-thymic in vivo proliferation of naive CD4+ T cells constrains the TCR repertoire in healthy human adults. *Eur J Immunol*, 35(6):1987–94.

[Kokordelis et al., 2014] Kokordelis, P., Krämer, B., Körner, C., Boesecke, C., Voigt, E., Ingiliz, P., Glässner, A., Eisenhardt, M., Wolter, F., Kaczmarek, D., Nischalke, H. D., Rockstroh, J. K., Spengler, U., and Nattermann, J. (2014). An effective interferon-gamma-mediated inhibition of hepatitis C virus replication by natural killer cells is associated with spontaneous clearance of acute hepatitis C in human immunodeficiency virus-positive patients. *Hepatology*, 59(3):814–27.

[Kolte et al., 2002] Kolte, L., Dreves, A.-M., Ersbøll, A. K., Strandberg, C., Jeppesen, D. L., Nielsen, J. O., Ryder, L. P., and Nielsen, S. D. (2002). Association between larger thymic size and higher thymic output in human immunodeficiency virus-infected patients receiving highly active antiretroviral therapy. *J Infect Dis*, 185(11):1578–85.

[Kong et al., 1998] Kong, F., Chen, C. H., and Cooper, M. D. (1998). Thymic function can be accurately monitored by the level of recent T cell emigrants in the circulation. *Immunity*, 8(1):97–104.

[Kong et al., 1999] Kong, F. K., Chen, C. L., Six, A., Hockett, R. D., and Cooper, M. D. (1999). T cell receptor gene deletion circles identify recent thymic emigrants in the peripheral T cell pool. *Proc Natl Acad Sci U S A*, 96(4):1536–40.

[Körner et al., 2009] Körner, C., Krämer, B., Schulte, D., Coenen, M., Mauss, S., Fätkenheuer, G., Oldenburg, J., Nattermann, J., Rockstroh, J. K., and Spengler, U. (2009). Effects of HCV co-infection on apoptosis of CD4+ T-cells in HIV-positive patients. *Clin Sci (Lond)*, 116(12):861–70.

[Korolevskaya et al., 2016] Korolevskaya, L. B., Shmagel, K. V., Saidakova, E. V., Shmagel, N. G., Slobodchikova, S. V., and Chereshnev, V. A. (2016). Effect of Hepatitis C Virus Coinfection on the Content of CD4(+) and CD8(+) T Cell Subpopulations in HIV-Infected Patients Receiving Antiretroviral Therapy. *Bull Exp Biol Med*, 161(2):281–3.

[Kovacs et al., 2005] Kovacs, A., Montepiedra, G., Carey, V., Pahwa, S., Weinberg, A., Frenkel, L., Capparelli, E., Mofenson, L., Smith, E., McIntosh, K., Burchett, S. K., and Pediatric AIDS Clinical Trials Group 366 Study Team (2005). Immune reconstitution after receipt of highly active antiretroviral therapy in children with advanced or progressive hiv disease and complete or partial viral load response. *J Infect Dis*, 192(2):296–302.

[Krogstad et al., 2015] Krogstad, P., Patel, K., Karalius, B., Hazra, R., Abzug, M. J., Oleske, J., Seage, 3rd, G. R., Williams, P. L., Borkowsky, W., Wiznia, A., Pinto, J., Van Dyke, R. B., and Pediatric HIVAIDS Cohort Study, IMPAACT 219C, and NICHD International Site Development Initiative (NISDI) Investigators (2015). Incomplete immune reconstitution despite virologic suppression in HIV-1 infected children and adolescents. *AIDS*, 29(6):683–93.

[Lapadula et al., 2013] Lapadula, G., Cozzi-Lepri, A., Marchetti, G., Antinori, A., Chiodera, A., Nicastri, E., Parruti, G., Galli, M., Gori, A., Monforte, A. d., and ICONA Foundation Study (2013). Risk of clinical progression among patients with immunological nonresponse despite virological suppression after combination antiretroviral treatment. *AIDS*, 27(5):769–79.

[Lavielle et al., 2011] Lavielle, M., Samson, A., Karina Fermin, A., and Mentré, F. (2011). Maximum likelihood estimation of long-term HIV dynamic models and antiviral response. *Biometrics*, 67(1):250–9.

[Lechner et al., 2000] Lechner, F., Wong, D. K., Dunbar, P. R., Chapman, R., Chung, R. T., Dohrenwend, P., Robbins, G., Phillips, R., Klenerman, P., and Walker, B. D. (2000). Analysis of successful immune responses in persons infected with hepatitis C virus. *J Exp Med*, 191(9):1499–512.

[Lederman et al., 2011] Lederman, M. M., Calabrese, L., Funderburg, N. T., Clagett, B., Medvik, K., Bonilla, H., Gripshover, B., Salata, R. A., Taege, A., Lisgaris, M., McComsey, G. A., Kirchner, E., Baum, J., Shive, C., Asaad, R., Kalayjian, R. C., Sieg, S. F., and Rodriguez, B. (2011). Immunologic failure despite suppressive antiretroviral therapy is related to activation and turnover of memory CD4 cells. *J Infect Dis*, 204(8):1217–26.

[Lewden et al., 2007] Lewden, C., Chene, G., Morlat, P., Raffi, F., Dupon, M., Dellamonica, P., Pellegrin, J.-L., Katlama, C., Dabis, F., Leport, C., Agence Nationale de Recherches sur le Sida et les Hepatites Virales (ANRS) CO8 APROCO-COPILOTE Study Group, and Agence Nationale de Recherches sur le Sida et les Hepatites Virales (ANRS) CO3 AQUITAINE Study Group (2007). HIV-infected adults with a CD4 cell count greater than 500 cells/mm3 on long-term combination antiretroviral therapy reach same mortality rates as the general population. *J Acquir Immune Defic Syndr*, 46(1):72–7.

[Lewis et al., 2012] Lewis, J., Walker, A. S., Castro, H., De Rossi, A., Gibb, D. M., Giaquinto, C., Klein, N., and Callard, R. (2012). Age and CD4 count at initiation of antiretroviral therapy in HIV-infected children: effects on long-term T-cell reconstitution. *J Infect Dis*, 205(4):548–56.

[Li et al., 2011] Li, T., Wu, N., Dai, Y., Qiu, Z., Han, Y., Xie, J., Zhu, T., and Li, Y. (2011). Reduced thymic output is a major mechanism of immune reconstitution failure in HIV-infected patients after long-term antiretroviral therapy. *Clin Infect Dis*, 53(9):944–51.

[Lichtenstein et al., 2010] Lichtenstein, K. A., Armon, C., Buchacz, K., Chmiel, J. S., Buckner, K., Tedaldi, E. M., Wood, K., Holmberg, S. D., Brooks, J. T., and HIV Outpatient Study (HOPS) Investigators (2010). Low CD4+ T cell count is a risk factor for cardiovascular disease events in the HIV outpatient study. *Clin Infect Dis*, 51(4):435–47.

[Lindbom et al., 2005] Lindbom, L., Pihlgren, P., Jonsson, E. N., and Jonsson, N. (2005). PsN-Toolkit–a collection of computer intensive statistical methods for non-linear mixed effect modeling using NONMEM. *Comput Methods Programs Biomed*, 79(3):241–57.

[Liu et al., 1997] Liu, Z., Cumberland, W. G., Hultin, L. E., Prince, H. E., Detels, R., and Giorgi, J. V. (1997). Elevated CD38 antigen expression on CD8+ T cells is a stronger marker for the risk of chronic HIV disease progression to AIDS and death in the Multicenter AIDS Cohort Study than CD4+ cell count, soluble immune activation markers, or combinations of HLA-DR and CD38 expression. *J Acquir Immune Defic Syndr Hum Retrovirol*, 16(2):83–92.

[Livak and Schatz, 1996] Livak, F. and Schatz, D. G. (1996). T-cell receptor alpha locus V(D)J recombination by-products are abundant in thymocytes and mature T cells. *Mol Cell Biol*, 16(2):609–18.

[Llano et al., 2001] Llano, A., Barretina, J., Gutiérrez, A., Blanco, J., Cabrera, C., Clotet, B., and Esté, J. A. (2001). Interleukin-7 in plasma correlates with CD4 T-cell depletion and may be associated with

emergence of syncytium-inducing variants in human immunodeficiency virus type 1-positive individuals. *J Virol*, 75(21):10319–25.

[Lu and Huang, 2014] Lu, X. and Huang, Y. (2014). Bayesian analysis of nonlinear mixed-effects mixture models for longitudinal data with heterogeneity and skewness. *Stat Med*, 33(16):2830–49.

[Maartens et al., 2014] Maartens, G., Celum, C., and Lewin, S. R. (2014). Hiv infection: epidemiology, pathogenesis, treatment, and prevention. *Lancet*, 384(9939):258–71.

[Machado et al., 2007] Machado, D. M., Gouvêa, A. d. F. B., Cardoso, M. R., Beltrão, S. V., Cunegundes, K. S., Bononi, F., Almeida, F., Cavalheiro, K., Angelis, D. S. A. d., and Succi, R. C. d. M. (2007). Factors associated with clinical, immunological and virological responses in protease-inhibitor-experienced brazilian children receiving highly active antiretroviral therapy containing lopinavir-ritonavir. *Braz J Infect Dis*, 11(1):16–9.

[Macías et al., 2003] Macías, J., Pineda, J. A., Lozano, F., Corzo, J. E., Ramos, A., León, E., García-García, J. A., Fernández-Rivera, J., Mira, J. A., and Gómez-Mateos, J. (2003). Impaired recovery of CD4+ cell counts following highly active antiretroviral therapy in drug-naïve patients coinfected with human immunodeficiency virus and hepatitis C virus. *Eur J Clin Microbiol Infect Dis*, 22(11):675–80.

[Majekodunmi et al., 2018] Majekodunmi, A. O., Callard, R., Klein, N., and Lewis, J. (2018). Re: Recovery of CD4 T Cells in HIV/HCV Coinfected Children: Is it Really Impaired? *Pediatr Infect Dis J*, 37(3):278–279.

[Majekodunmi et al., 2017] Majekodunmi, A. O., Thorne, C., Malyuta, R., Volokha, A., Callard, R. E., Klein, N. J., Lewis, J., and European Paediatric HIV/HCV Co-infection Study group in the European Pregnancy and Paediatric HIV Cohort Collaboration and the Ukraine Paediatric HIV Cohort Study in EuroCoord (2017). Modelling CD4 T Cell Recovery in Hepatitis C and HIV Co-infected Children Receiving Antiretroviral Therapy. *Pediatr Infect Dis J*, 36(5):e123–e129.

[Marandin et al., 1996] Marandin, A., Katz, A., Oksenhendler, E., Tulliez, M., Picard, F., Vainchenker, W., and Louache, F. (1996). Loss of primitive hematopoietic progenitors in patients with human immunodeficiency virus infection. *Blood*, 88(12):4568–78.

[Marcellin et al., 1997] Marcellin, P., Boyer, N., Gervais, A., Martinot, M., Pouteau, M., Castelnau, C., Kilani, A., Areias, J., Auperin, A., Benhamou, J. P., Degott, C., and Erlinger, S. (1997). Long-term

histologic improvement and loss of detectable intrahepatic HCV RNA in patients with chronic hepatitis C and sustained response to interferon-alpha therapy. *Ann Intern Med*, 127(10):875–81.

[Marcus et al., 2015] Marcus, J. L., Leyden, W. A., Chao, C. R., Xu, L., Quesenberry, Jr, C. P., Tien, P. C., Klein, D. B., Towner, W. J., Horberg, M. A., and Silverberg, M. J. (2015). Differences in response to antiretroviral therapy by sex and hepatitis c infection status. *AIDS Patient Care STDS*, 29(7):370–8.

[Marie Davidian, David M. Giltinan, 1995] Marie Davidian, David M. Giltinan (1995). *Non-linear models for repeated measurement data*. Chapman and Hall.

[Marinho et al., 2014] Marinho, R. T., Vitor, S., and Velosa, J. (2014). Benefits of curing hepatitis C infection. *J Gastrointestin Liver Dis*, 23(1):85–90.

[Marziali et al., 2006] Marziali, M., De Santis, W., Carello, R., Leti, W., Esposito, A., Isgrò, A., Fimiani, C., Sirianni, M. C., Mezzaroma, I., and Aiuti, F. (2006). T-cell homeostasis alteration in HIV-1 infected subjects with low CD4 T-cell count despite undetectable virus load during HAART. *AIDS*, 20(16):2033–41.

[Massanella et al., 2010] Massanella, M., Negredo, E., Pérez-Alvarez, N., Puig, J., Ruiz-Hernández, R., Bofill, M., Clotet, B., and Blanco, J. (2010). CD4 T-cell hyperactivation and susceptibility to cell death determine poor CD4 T-cell recovery during suppressive HAART. *AIDS*, 24(7):959–68.

[Mast et al., 2005] Mast, E. E., Hwang, L.-Y., Seto, D. S. Y., Nolte, F. S., Nainan, O. V., Wurtzel, H., and Alter, M. J. (2005). Risk factors for perinatal transmission of hepatitis C virus (HCV) and the natural history of HCV infection acquired in infancy. *J Infect Dis*, 192(11):1880–9.

[Mattapallil et al., 2005] Mattapallil, J. J., Douek, D. C., Hill, B., Nishimura, Y., Martin, M., and Roederer, M. (2005). Massive infection and loss of memory CD4+ T cells in multiple tissues during acute SIV infection. *Nature*, 434(7037):1093–7.

[McCune, 2001] McCune, J. M. (2001). The dynamics of CD4+ T-cell depletion in HIV disease. *Nature*, 410(6831):974–9.

[McCune et al., 1998] McCune, J. M., Loftus, R., Schmidt, D. K., Carroll, P., Webster, D., Swor-Yim, L. B., Francis, I. R., Gross, B. H., and Grant, R. M. (1998). High prevalence of thymic tissue in adults with human immunodeficiency virus-1 infection. *J Clin Invest*, 101(11):2301–8.

[Mehta et al., 2002] Mehta, S. H., Cox, A., Hoover, D. R., Wang, X.-H., Mao, Q., Ray, S., Strathdee, S. A., Vlahov, D., and Thomas, D. L. (2002). Protection against persistence of hepatitis C. *Lancet*, 359(9316):1478–83.

[Meyaard et al., 1992] Meyaard, L., Otto, S., Jonker, R., Mijnster, M., Keet, R., and Miedema, F. (1992). Programmed death of T cells in HIV-1 infection. *Science*, 257(5067):217–219.

[Meyaard et al., 1994] Meyaard, L., Otto, S. A., Keet, I. P., Roos, M. T., and Miedema, F. (1994). Programmed death of T cells in human immunodeficiency virus infection. No correlation with progression to disease. *J Clin Invest*, 93(3):982–8.

[Micheloud et al., 2007] Micheloud, D., Jensen, J., Bellón, J. M., Gonzalez, R., Mellado, M. J., Navarro, M. L., Muñoz-Fernández, M. A., Resino, S., and Spanish Group of Pediatric HIV Infection (2007). Long-term response to highly active antiretroviral therapy in human immunodeficiency virus and hepatitis C virus coinfected children: 6 years of follow-up. *Pediatr Infect Dis J*, 26(11):1061–4.

[Mohri et al., 2001] Mohri, H., Perelson, A. S., Tung, K., Ribeiro, R. M., Ramratnam, B., Markowitz, M., Kost, R., Hurley, A., Weinberger, L., Cesar, D., Hellerstein, M. K., and Ho, D. D. (2001). Increased turnover of T lymphocytes in HIV-1 infection and its reduction by antiretroviral therapy. *J Exp Med*, 194(9):1277–87.

[Moir et al., 2011] Moir, S., Chun, T.-W., and Fauci, A. S. (2011). Pathogenic mechanisms of hiv disease. *Annu Rev Pathol*, 6:223–48.

[Monga et al., 2001] Monga, H. K., Rodriguez-Barradas, M. C., Breaux, K., Khattak, K., Troisi, C. L., Velez, M., and Yoffe, B. (2001). Hepatitis C virus infection-related morbidity and mortality among patients with human immunodeficiency virus infection. *Clin Infect Dis*, 33(2):240–7.

[Moore and Keruly, 2007] Moore, R. D. and Keruly, J. C. (2007). CD4+ cell count 6 years after commencement of highly active antiretroviral therapy in persons with sustained virologic suppression. *Clin Infect Dis*, 44(3):441–6.

[Moses et al., 1998] Moses, A., Nelson, J., and Bagby, Jr, G. C. (1998). The influence of human immunodeficiency virus-1 on hematopoiesis. *Blood*, 91(5):1479–95.

[Murali-Krishna and Ahmed, 2000] Murali-Krishna, K. and Ahmed, R. (2000). Cutting edge: naive T cells masquerading as memory cells. *J Immunol*, 165(4):1733–7.

[Murphy and Weaver, 2017] Murphy, K. and Weaver, C. (2017). *Janeway's immunobiology*. Garland Science, Taylor and Francis Group, LLC, 9 edition.

[Nakanjako et al., 2011] Nakanjako, D., Ssewanyana, I., Mayanja-Kizza, H., Kiragga, A., Colebunders, R., Manabe, Y. C., Nabatanzi, R., Kamya, M. R., and Cao, H. (2011). High T-cell immune activation and immune exhaustion among individuals with suboptimal CD4 recovery after 4 years of antiretroviral therapy in an African cohort. *BMC Infect Dis*, 11:43.

[Newman, 1997] Newman, P. J. (1997). The Biology of PECAM-1. *Journal of Clinical Investigation*, 99(1):3–8.

[Oehen and Brduscha-Riem, 1999] Oehen, S. and Brduscha-Riem, K. (1999). Naïve cytotoxic T lymphocytes spontaneously acquire effector function in lymphocytopenic recipients: A pitfall for T cell memory studies? *Eur J Immunol*, 29(2):608–14.

[Ometto et al., 2002] Ometto, L., De Forni, D., Patiri, F., Trouplin, V., Mammano, F., Giacomet, V., Giaquinto, C., Douek, D., Koup, R., and De Rossi, A. (2002). Immune reconstitution in HIV-1-infected children on antiretroviral therapy: role of thymic output and viral fitness. *AIDS*, 16(6):839–49.

[Owen et al., 2010] Owen, R. E., Heitman, J. W., Hirschkorn, D. F., Lanteri, M. C., Biswas, H. H., Martin, J. N., Krone, M. R., Deeks, S. G., Norris, P. J., and NIAID Center for HIV/AIDS Vaccine Immunology (2010). HIV+ elite controllers have low HIV-specific T-cell activation yet maintain strong, polyfunctional T-cell responses. *AIDS*, 24(8):1095–105.

[Paediatric European Network for Treatment of AIDS (PENTA), 2002] Paediatric European Network for Treatment of AIDS (PENTA) (2002). Comparison of dual nucleoside-analogue reverse-transcriptase inhibitor regimens with and without nelfinavir in children with HIV-1 who have not previously been treated: the PENTA 5 randomised trial. *Lancet*, 359(9308):733–40.

[Pantaleo et al., 1993] Pantaleo, G., Graziosi, C., and Fauci, A. S. (1993). New concepts in the immunopathogenesis of human immunodeficiency virus infection. *N Engl J Med*, 328(5):327–35.

[Penna et al., 2007] Penna, A., Pilli, M., Zerbini, A., Orlandini, A., Mezzadri, S., Sacchelli, L., Missale, G., and Ferrari, C. (2007). Dysfunction and functional restoration of HCV-specific CD8 responses in chronic hepatitis C virus infection. *Hepatology*, 45(3):588–601.

[Perz et al., 2006] Perz, J. F., Armstrong, G. L., Farrington, L. A., Hutin, Y. J. F., and Bell, B. P. (2006). The contributions of hepatitis B virus and hepatitis C virus infections to cirrhosis and primary liver cancer worldwide. *J Hepatol*, 45(4):529–38.

[Phoenix WinNonlin Certara, 2016] Phoenix WinNonlin Certara (2016). *Phoenix WinNonlin version 7.0.* Certara USA, Inc., 103 Carnegie Center, Suite 300, Princeton, NJ 08540 USA.

[Picat et al., 2013] Picat, M.-Q., Lewis, J., Musiime, V., Prendergast, A., Nathoo, K., Kekitiinwa, A., Nahirya Ntege, P., Gibb, D. M., Thiebaut, R., Walker, A. S., Klein, N., Callard, R., and ARROW Trial Team (2013). Predicting patterns of long-term CD4 reconstitution in HIV-infected children starting antiretroviral therapy in sub-Saharan Africa: a cohort-based modelling study. *PLoS Med*, 10(10):e1001542.

[Piconi et al., 2010] Piconi, S., Trabattoni, D., Gori, A., Parisotto, S., Magni, C., Meraviglia, P., Bandera, A., Capetti, A., Rizzardini, G., and Clerici, M. (2010). Immune activation, apoptosis, and Treg activity are associated with persistently reduced CD4+ T-cell counts during antiretroviral therapy. *AIDS*, 24(13):1991–2000.

[Piketty et al., 1998] Piketty, C., Castiel, P., Belec, L., Batisse, D., Si Mohamed, A., Gilquin, J., Gonzalez-Canali, G., Jayle, D., Karmochkine, M., Weiss, L., Aboulker, J. P., and Kazatchkine, M. D. (1998). Discrepant responses to triple combination antiretroviral therapy in advanced HIV disease. *AIDS*, 12(7):745–50.

[Piketty et al., 2001] Piketty, C., Weiss, L., Thomas, F., Mohamed, A. S., Belec, L., and Kazatchkine, M. D. (2001). Long-term clinical outcome of human immunodeficiency virus-infected patients with discordant immunologic and virologic responses to a protease inhibitor-containing regimen. *J Infect Dis*, 183(9):1328–35.

[Pinheiro et al., 2013] Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., and R Core Team (2013). *nlme: Linear and Nonlinear Mixed Effects Models.* http://CRAN.R-project.org/package=nlme, r package version 3.1-118 edition.

[Platt et al., 2016] Platt, L., Easterbrook, P., Gower, E., McDonald, B., Sabin, K., McGowan, C., Yanny, I., Razavi, H., and Vickerman, P. (2016). Prevalence and burden of HCV co-infection in people living with HIV: a global systematic review and meta-analysis. *Lancet Infect Dis*, 16(7):797–808.

[Podzamczer et al., 2002] Podzamczer, D., Ferrer, E., Consiglio, E., Gatell, J. M., Perez, P., Perez, J. L., Luna, E., González, A., Pedrol, E., Lozano, L., Ocaña, I., Llibre, J. M., Casiró, A., Aranda, M., Barrufet, P., Martínez-Lacasa, J., Miró, J. M., Badía, X., Casado, A., Lupo, S., Cahn, P., Maños, M., and Estela, J. (2002). A randomized clinical trial comparing nelfinavir or nevirapine associated to zidovudine/lamivudine in hiv-infected naive patients (the combine study). *Antivir Ther*, 7(2):81–90.

[Pollock and Welsh, 2002] Pollock, J. M. and Welsh, M. D. (2002). The WC1(+) gamma-delta T-cell population in cattle: a possible role in resistance to intracellular infection. *Vet Immunol Immunopathol*, 89(3-4):105–14.

[Prendergast et al., 2007] Prendergast, A., Tudor-Williams, G., Jeena, P., Burchett, S., and Goulder, P. (2007). International perspectives, progress, and future challenges of paediatric HIV infection. *Lancet*, 370(9581):68–80.

[Prendergast et al., 2016] Prendergast, A. J., Szubert, A. J., Berejena, C., Pimundu, G., Pala, P., Shonhai, A., Musiime, V., Bwakura-Dangarembizi, M., Poulsom, H., Hunter, P., Musoke, P., Kihembo, M., Munderi, P., Gibb, D. M., Spyer, M., Walker, A. S., Klein, N., and ARROW Trial Team (2016). Baseline Inflammatory Biomarkers Identify Subgroups of HIV-Infected African Children With Differing Responses to Antiretroviral Therapy. *J Infect Dis*, 214(2):226–36.

[Puthanakit et al., 2009] Puthanakit, T., Kerr, S., Ananworanich, J., Bunupuradah, T., Boonrak, P., and Sirisanthana, V. (2009). Pattern and predictors of immunologic recovery in human immunodeficiency virus-infected children receiving non-nucleoside reverse transcriptase inhibitor-based highly active antiretroviral therapy. *Pediatr Infect Dis J*, 28(6):488–92.

[R Core Team, 2015] R Core Team (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria (https://www.R-project.org).

[Rehermann, 2009] Rehermann, B. (2009). Hepatitis C virus versus innate and adaptive immune responses: a tale of coevolution and coexistence. *J Clin Invest*, 119(7):1745–54.

[Rehermann and Nascimbeni, 2005] Rehermann, B. and Nascimbeni, M. (2005). Immunology of hepatitis B virus and hepatitis C virus infection. *Nat Rev Immunol*, 5(3):215–29.

[Reiberger et al., 2011] Reiberger, T., Payer, B. A., Kosi, L., Heil, P. M., Rieger, A., Peck-Radosavljevic, M., and Vienna HIV Coinfection Study Group (2011). Concomitant highly active antiretroviral therapy

leads to smaller decline and faster recovery of CD4+ cell counts during and after pegylated interferon plus ribavirin therapy in HIV-hepatitis C virus coinfected patients. *J Infect Dis*, 203(12):1802–6.

[Rethi et al., 2005] Rethi, B., Fluur, C., Atlas, A., Krzyzowska, M., Mowafi, F., Grützmeier, S., De Milito, A., Bellocco, R., Falk, K. I., Rajnavölgyi, E., and Chiodi, F. (2005). Loss of IL-7Ralpha is associated with CD4 T-cell depletion, high interleukin-7 levels and CD28 down-regulation in HIV infected patients. *AIDS*, 19(18):2077–86.

[Rezzani et al., 2014] Rezzani, R., Nardo, L., Favero, G., Peroni, M., and Rodella, L. F. (2014). Thymus and aging: morphological, radiological, and functional overview. *Age (Dordr)*, 36(1):313–51.

[Ribeiro et al., 2002] Ribeiro, R. M., Mohri, H., Ho, D. D., and Perelson, A. S. (2002). In vivo dynamics of T cell activation, proliferation, and death in HIV-1 infection: why are CD4+ but not CD8+ T cells depleted? *Proc Natl Acad Sci U S A*, 99(24):15572–7.

[Robert J Bauer, Brian Sadler, 2015] Robert J Bauer, Brian Sadler (2015). Advanced Methods in NONMEM 7.3 Workshop Manual, Population Approach Group in Europe Conference 2015, Crete. In ICON plc, editor, *Advanced Methods in NONMEM 7.3*. ICON Development Solutions.

[Roberts et al., 1988] Roberts, J. D., Bebenek, K., and Kunkel, T. A. (1988). The accuracy of reverse transcriptase from HIV-1. *Science*, 242(4882):1171–3.

[Rollo Hoare, 2015] Rollo Hoare (2015). *Modelling Immune Reconstitution following Paediatric Haematopoietic Stem Cell Transplantation and in HIV-Infected Children*. PhD thesis, University College London, Gower St, London WC1E 6BT.

[Sandberg et al., 2003] Sandberg, J. K., Fast, N. M., Jordan, K. A., Furlan, S. N., Barbour, J. D., Fennelly, G., Dobroszycki, J., Spiegel, H. M. L., Wiznia, A., Rosenberg, M. G., and Nixon, D. F. (2003). HIV-specific CD8+ T cell function in children with vertically acquired HIV-1 infection is critically influenced by age and the state of the CD4+ T cell compartment. *J Immunol*, 170(8):4403–10.

[Sandgaard et al., 2014] Sandgaard, K. S., Lewis, J., Adams, S., Klein, N., and Callard, R. (2014). Antiretroviral therapy increases thymic output in children with HIV. *AIDS*, 28(2):209–14.

[Santin et al., 2008] Santin, M., Mestre, M., Shaw, E., Barbera, M. J., Casanova, A., Niubo, J., Bolao, F., Podzamczer, D., and Gudiol, F. (2008). Impact of hepatitis C virus coinfection on immune restoration

during successful antiretroviral therapy in chronic human immunodeficiency virus type 1 disease. *Eur J Clin Microbiol Infect Dis*, 27(1):65–73.

[SAS Institute, 2017] SAS Institute (2017). *SAS/STAT*. SAS Institute Inc., Cary, NC, USA.

[Savic and Karlsson, 2009] Savic, R. M. and Karlsson, M. O. (2009). Importance of shrinkage in empirical bayes estimates for diagnostics: problems and solutions. *AAPS J*, 11(3):558–69.

[Schacker et al., 2002] Schacker, T. W., Nguyen, P. L., Beilman, G. J., Wolinsky, S., Larson, M., Reilly, C., and Haase, A. T. (2002). Collagen deposition in HIV-1 infected lymphatic tissues and T cell homeostasis. *J Clin Invest*, 110(8):1133–9.

[Schacker et al., 2005] Schacker, T. W., Reilly, C., Beilman, G. J., Taylor, J., Skarda, D., Krason, D., Larson, M., and Haase, A. T. (2005). Amount of lymphatic tissue fibrosis in HIV infection predicts magnitude of HAART-associated change in peripheral CD4 cell count. *AIDS*, 19(18):2169–71.

[Schatorjé et al., 2012] Schatorjé, E. J. H., Gemen, E. F. A., Driessen, G. J. A., Leuvenink, J., van Hout, R. W. N. M., and de Vries, E. (2012). Paediatric reference values for the peripheral T cell compartment. *Scand J Immunol*, 75(4):436–44.

[Schluns et al., 2000] Schluns, K. S., Kieper, W. C., Jameson, S. C., and Lefrançois, L. (2000). Interleukin-7 mediates the homeostasis of naïve and memory CD8 T cells in vivo. *Nat Immunol*, 1(5):426–32.

[Schulze Zur Wiesch et al., 2012] Schulze Zur Wiesch, J., Ciuffreda, D., Lewis-Ximenez, L., Kasprowicz, V., Nolan, B. E., Streeck, H., Aneja, J., Reyor, L. L., Allen, T. M., Lohse, A. W., McGovern, B., Chung, R. T., Kwok, W. W., Kim, A. Y., and Lauer, G. M. (2012). Broadly directed virus-specific CD4+ T cell responses are primed during acute hepatitis C infection, but rapidly disappear from human blood with viral persistence. *J Exp Med*, 209(1):61–75.

[Selvapatt et al., 2015] Selvapatt, N., Ward, T., Bailey, H., Bennett, H., Thorne, C., See, L.-M., Tudor-Williams, G., Thursz, M., McEwan, P., and Brown, A. (2015). Is antenatal screening for hepatitis C virus cost-effective? A decade's experience at a London centre. *J Hepatol*, 63(4):797–804.

[Semmo et al., 2005] Semmo, N., Day, C. L., Ward, S. M., Lucas, M., Harcourt, G., Loughry, A., and Klenerman, P. (2005). Preferential loss of IL-2-secreting CD4+ T helper cells in chronic HCV infection. *Hepatology*, 41(5):1019–28.

[Shearer et al., 2003] Shearer, W. T., Rosenblatt, H. M., Gelman, R. S., Oyomopito, R., Plaeger, S., Stiehm, E. R., Wara, D. W., Douglas, S. D., Luzuriaga, K., McFarland, E. J., Yogev, R., Rathore, M. H., Levy, W., Graham, B. L., Spector, S. A., and Pediatric AIDS Clinical Trials Group (2003). Lymphocyte subsets in healthy children from birth through 18 years of age: the pediatric aids clinical trials group p1009 study. *J Allergy Clin Immunol*, 112(5):973–80.

[Sheiner and Beal, 1980] Sheiner, L. B. and Beal, S. L. (1980). Evaluation of methods for estimating population pharmacokinetics parameters. I. Michaelis-Menten model: routine clinical pharmacokinetic data. *J Pharmacokinet Biopharm*, 8(6):553–71.

[Sheiner et al., 1972] Sheiner, L. B., Rosenberg, B., and Melmon, K. L. (1972). Modelling of individual pharmacokinetics for computer-aided drug dosage. *Comput Biomed Res*, 5(5):411–59.

[Shin et al., 2013] Shin, E.-C., Park, S.-H., Nascimbeni, M., Major, M., Caggiari, L., de Re, V., Feinstone, S. M., Rice, C. M., and Rehermann, B. (2013). The frequency of CD127(+) hepatitis C virus (HCV)-specific T cells but not the expression of exhaustion markers predicts the outcome of acute HCV infection. *J Virol*, 87(8):4772–7.

[Shin et al., 2016] Shin, E.-C., Sung, P. S., and Park, S.-H. (2016). Immune responses and immunopathology in acute and chronic viral hepatitis. *Nat Rev Immunol*, 16(8):509–23.

[Shmagel et al., 2014] Shmagel, K. V., Saidakova, E. V., Korolevskaya, L. B., Shmagel, N. G., Chereshnev, V. A., Anthony, D. D., and Lederman, M. M. (2014). Influence of hepatitis C virus coinfection on CD4+ T cells of HIV-infected patients receiving HAART. *AIDS*, 28(16):2381–8.

[Siegrist and Aspinall, 2009] Siegrist, C.-A. and Aspinall, R. (2009). B-cell responses to vaccination at the extremes of age. *Nat Rev Immunol*, 9(3):185–94.

[Sirskyj et al., 2008] Sirskyj, D., Thèze, J., Kumar, A., and Kryworuchko, M. (2008). Disruption of the gamma c cytokine network in T cells during HIV infection. *Cytokine*, 43(1):1–14.

[Smith et al., 2000] Smith, K. Y., Valdez, H., Landay, A., Spritzler, J., Kessler, H. A., Connick, E., Kuritzkes, D., Gross, B., Francis, I., McCune, J. M., and Lederman, M. M. (2000). Thymic size and lymphocyte restoration in patients with human immunodeficiency virus infection after 48 weeks of zidovudine, lamivudine, and ritonavir therapy. *J Infect Dis*, 181(1):141–7.

[Snijdewind et al., 2015] Snijdewind, I. J. M., Smit, C., Schutten, M., Nellen, F. J. B., Kroon, F. P., Reiss, P., and van der Ende, M. E. (2015). Low mother-to-child-transmission rate of Hepatitis C virus in cART treated HIV-1 infected mothers. *J Clin Virol*, 68:11–5.

[Sousa et al., 2002] Sousa, A. E., Carneiro, J., Meier-Schellersheim, M., Grossman, Z., and Victorino, R. M. M. (2002). CD4 T cell depletion is linked directly to immune activation in the pathogenesis of HIV-1 and HIV-2 but only indirectly to the viral load. *J Immunol*, 169(6):3400–6.

[Spiegelhalter et al., 1999] Spiegelhalter, D., Thomas, A., and Best, N. (1999). *WinBUGS Version 1.2 User Manual*. MRC Biostatistics Unit.

[Sprent et al., 2008] Sprent, J., Cho, J.-H., Boyman, O., and Surh, C. D. (2008). T cell homeostasis. *Immunol Cell Biol*, 86(4):312–9.

[Sprent and Surh, 2002] Sprent, J. and Surh, C. D. (2002). T cell memory. *Annu Rev Immunol*, 20:551–79.

[Starr et al., 2003] Starr, T. K., Jameson, S. C., and Hogquist, K. A. (2003). Positive and negative selection of T cells. *Annu Rev Immunol*, 21:139–76.

[Steinmann et al., 1985] Steinmann, G. G., Klaus, B., and Müller-Hermelink, H. K. (1985). The involution of the ageing human thymic epithelium is independent of puberty. A morphometric study. *Scand J Immunol*, 22(5):563–75.

[Sulkowski et al., 2007] Sulkowski, M. S., Mehta, S. H., Torbenson, M. S., Higgins, Y., Brinkley, S. C., de Oca, R. M., Moore, R. D., Afdhal, N. H., and Thomas, D. L. (2007). Rapid fibrosis progression among HIV/hepatitis C virus-co-infected adults. *AIDS*, 21(16):2209–16.

[Surh and Sprent, 2000] Surh, C. D. and Sprent, J. (2000). Homeostatic T cell proliferation: how far can T cells be activated to self-ligands? *J Exp Med*, 192(4):F9–F14.

[Surh and Sprent, 2008] Surh, C. D. and Sprent, J. (2008). Homeostasis of naive and memory T cells. *Immunity*, 29(6):848–62.

[Tan et al., 2008] Tan, R., Westfall, A. O., Willig, J. H., Mugavero, M. J., Saag, M. S., Kaslow, R. A., and Kempf, M. C. (2008). Clinical outcome of HIV-infected antiretroviral-naive patients with discordant immunologic and virologic responses to highly active antiretroviral therapy. *J Acquir Immune Defic Syndr*, 47(5):553–8.

[Tanaskovic et al., 2011] Tanaskovic, S., Fernandez, S., French, M. A., Price, R. I., Song, S., Robins, P. D., and Price, P. (2011). Thymic tissue is not evident on high-resolution computed tomography and $[18-F]$fluoro-deoxy-glucose positron emission tomography scans of aviraemic HIV patients with poor recovery of CD4+ T cells. *AIDS*, 25(9):1235–7.

[Tanaskovic et al., 2010] Tanaskovic, S., Fernandez, S., Price, P., Lee, S., and French, M. A. (2010). CD31 (PECAM-1) is a marker of recent thymic emigrants among CD4+ T-cells, but not CD8+ T-cells or gammadelta T-cells, in HIV patients responding to ART. *Immunol Cell Biol*, 88(3):321–7.

[Taylor et al., 2012] Taylor, L. E., Swan, T., and Mayer, K. H. (2012). HIV coinfection with hepatitis C virus: evolving epidemiology and treatment paradigms. *Clin Infect Dis*, 55 Suppl 1:S33–42.

[Terilli and Cox, 2013] Terilli, R. R. and Cox, A. L. (2013). Immunity and hepatitis C: a review. *Curr HIV/AIDS Rep*, 10(1):51–8.

[The AVANTI and INCAS Study Groups, 2000] The AVANTI and INCAS Study Groups (2000). Highly active antiretroviral therapy including protease inhibitors does not confer a unique CD4 cell benefit. The AVANTI and INCAS Study Groups. *AIDS*, 14(10):1383–8.

[Thein et al., 2008] Thein, H.-H., Yi, Q., Dore, G. J., and Krahn, M. D. (2008). Natural history of hepatitis C virus infection in HIV-infected individuals and the impact of HIV in the era of highly active antiretroviral therapy: a meta-analysis. *AIDS*, 22(15):1979–91.

[Thiébaut et al., 2003] Thiébaut, R., Jacqmin-Gadda, H., Leport, C., Katlama, C., Costagliola, D., Le Moing, V., Morlat, P., Chêne, G., and APROCO Study Grooup (2003). Bivariate longitudinal model for the analysis of the evolution of HIV RNA and CD4 cell count in HIV infection taking into account left censoring of HIV RNA measures. *J Biopharm Stat*, 13(2):271–82.

[Thiébaut et al., 2006] Thiébaut, R., Jacqmin-Gadda, H., Walker, S., Sabin, C., Prins, M., Del Amo, J., Porter, K., Dabis, F., Chêne, G., and CASCADE Collaboration (2006). Determinants of response to first HAART regimen in antiretroviral-naïve patients with an estimated time since HIV seroconversion. *HIV Med*, 7(1):1–9.

[Thimme et al., 2002] Thimme, R., Bukh, J., Spangenberg, H. C., Wieland, S., Pemberton, J., Steiger, C., Govindarajan, S., Purcell, R. H., and Chisari, F. V. (2002). Viral and immunological determinants of hepatitis C virus clearance, persistence, and disease. *Proc Natl Acad Sci U S A*, 99(24):15661–8.

[Thimme et al., 2001] Thimme, R., Oldach, D., Chang, K. M., Steiger, C., Ray, S. C., and Chisari, F. V. (2001). Determinants of viral clearance and persistence during acute hepatitis C virus infection. *J Exp Med*, 194(10):1395–406.

[Thomas et al., 1998] Thomas, D. L., Villano, S. A., Riester, K. A., Hershow, R., Mofenson, L. M., Landesman, S. H., Hollinger, F. B., Davenny, K., Riley, L., Diaz, C., Tang, H. B., and Quinn, T. C. (1998). Perinatal transmission of hepatitis C virus from human immunodeficiency virus type 1-infected mothers. Women and Infants Transmission Study. *J Infect Dis*, 177(6):1480–8.

[Thomson et al., 2011] Thomson, E. C., Fleming, V. M., Main, J., Klenerman, P., Weber, J., Eliahoo, J., Smith, J., McClure, M. O., and Karayiannis, P. (2011). Predicting spontaneous clearance of acute hepatitis C virus in a large cohort of HIV-1-infected men. *Gut*, 60(6):837–45.

[Tosone et al., 2014] Tosone, G., Maraolo, A. E., Mascolo, S., Palmiero, G., Tambaro, O., and Orlando, R. (2014). Vertical hepatitis C virus transmission: Main questions and answers. *World J Hepatol*, 6(8):538–48.

[Townsend et al., 2014] Townsend, C. L., Byrne, L., Cortina-Borja, M., Thorne, C., de Ruiter, A., Lyall, H., Taylor, G. P., Peckham, C. S., and Tookey, P. A. (2014). Earlier initiation of art and further decline in mother-to-child hiv transmission rates, 2000-2011. *AIDS*, 28(7):1049–57.

[Townsend et al., 2008] Townsend, C. L., Cortina-Borja, M., Peckham, C. S., de Ruiter, A., Lyall, H., and Tookey, P. A. (2008). Low rates of mother-to-child transmission of hiv following effective pregnancy interventions in the united kingdom and ireland, 2000-2006. *AIDS*, 22(8):973–81.

[Triant et al., 2010] Triant, V. A., Regan, S., Lee, H., Sax, P. E., Meigs, J. B., and Grinspoon, S. K. (2010). Association of immunologic and virologic factors with myocardial infarction rates in a US healthcare system. *J Acquir Immune Defic Syndr*, 55(5):615–9.

[Tsiara et al., 2013] Tsiara, C. G., Nikolopoulos, G. K., Dimou, N. L., Bagos, P. G., Saroglou, G., Velonakis, E., and Hatzakis, A. (2013). Effect of hepatitis C virus on immunological and virological responses in HIV-infected patients initiating highly active antiretroviral therapy: a meta-analysis. *J Viral Hepat*, 20(10):715–24.

[Turkova et al., 2015] Turkova, A., Giacomet, V., Goetghebuer, T., Miloenko, M., Nicolini, L. A., Noguera-Julian, A., Rojo, P., Volokha, A., Indolfi, G., Giaquinto, C., and Thorne, C. (2015). HCV treatment in children and young adults with HIV/HCV co-infection in Europe. *J Virus Erad*, 1(3):179–84.

[Tuyama et al., 2010] Tuyama, A. C., Hong, F., Saiman, Y., Wang, C., Ozkok, D., Mosoian, A., Chen, P., Chen, B. K., Klotman, M. E., and Bansal, M. B. (2010). Human immunodeficiency virus (HIV)-1 infects human hepatic stellate cells and promotes collagen I and monocyte chemoattractant protein-1 expression: implications for the pathogenesis of HIV/hepatitis C virus-induced liver fibrosis. *Hepatology*, 52(2):612–22.

[van Gent et al., 2011] van Gent, R., Schadenberg, A. W. L., Otto, S. A., Nievelstein, R. A. J., Sieswerda, G. T., Haas, F., Miedema, F., Tesselaar, K., Jansen, N. J. G., and Borghans, J. A. M. (2011). Long-term restoration of the human T-cell compartment after thymectomy during infancy: a role for thymic regeneration? *Blood*, 118(3):627–34.

[Van Hoeven V., 2012] Van Hoeven V. (2012). Quantification of the lifespan and frequency of recent thymic emigrants. Presentation at the British Society for Immunology Mathematical Modelling Affinity Group Workshop.

[Van Leth et al., 2004] Van Leth, F., Wit, F. W. N. M., Reiss, P., Schattenkerk, J. K. M. E., Van Der Ende, M. E., Schneider, M. M. E., Mulder, J. W., Frissen, P. H. J., De Wolf, F., and Lange, J. M. A. (2004). Differential CD4 T-cell response in HIV-1-infected patients using protease inhibitor-based or nevirapine-based highly active antiretroviral therapy. *HIV Med*, 5(2):74–81.

[Van, Rossum, Guido et al, 2017] Van, Rossum, Guido et al (2017). *Python Language Reference, version 3.6*. Python Software Foundation, http://docs.python.org/py36/reference/index.html edition.

[Verschuren et al., 1997] Verschuren, M. C., Wolvers-Tettero, I. L., Breit, T. M., Noordzij, J., van Wering, E. R., and van Dongen, J. J. (1997). Preferential rearrangements of the T cell receptor-delta-deleting elements in human T cells. *J Immunol*, 158(3):1208–16.

[Vigano et al., 2000] Vigano, A., Vella, S., Saresella, M., Vanzulli, A., Bricalli, D., Di Fabio, S., Ferrante, P., Andreotti, M., Pirillo, M., Dally, L. G., Clerici, M., and Principi, N. (2000). Early immune reconstitution after potent antiretroviral therapy in HIV-infected children correlates with the increase in thymus volume. *AIDS*, 14(3):251–61.

[Villano et al., 1999] Villano, S. A., Vlahov, D., Nelson, K. E., Cohn, S., and Thomas, D. L. (1999). Persistence of viremia and the importance of long-term follow-up after acute hepatitis C infection. *Hepatology*, 29(3):908–14.

[Vrisekoop et al., 2008] Vrisekoop, N., den Braber, I., de Boer, A. B., Ruiter, A. F. C., Ackermans, M. T., van der Crabben, S. N., Schrijver, E. H. R., Spierenburg, G., Sauerwein, H. P., Hazenberg, M. D., de Boer, R. J., Miedema, F., Borghans, J. A. M., and Tesselaar, K. (2008). Sparse production but preferential incorporation of recently produced naive T cells in the human peripheral pool. *Proc Natl Acad Sci U S A*, 105(16):6115–20.

[Vrisekoop et al., 2015] Vrisekoop, N., Drylewicz, J., Van Gent, R., Mugwagwa, T., Van Lelyveld, S. F. L., Veel, E., Otto, S. A., Ackermans, M. T., Vermeulen, J. N., Huidekoper, H. H., Prins, J. M., Miedema, F., de Boer, R. J., Tesselaar, K., and Borghans, J. A. M. (2015). Quantification of naive and memory T-cell turnover during HIV-1 infection. *AIDS*, 29(16):2071–80.

[Wade and Ades, 1994] Wade, A. M. and Ades, A. E. (1994). Age-related reference ranges: significance tests for models and confidence intervals for centiles. *Stat Med*, 13(22):2359–67.

[Wade and Ades, 1998] Wade, A. M. and Ades, A. E. (1998). Incorporating correlations between measurements into the estimation of age-related reference ranges. *Stat Med*, 17(17):1989–2002.

[Wang, 2007] Wang, Y. (2007). Derivation of various NONMEM estimation methods. *J Pharmacokinet Pharmacodyn*, 34(5):575–93.

[Wickham, 2009] Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. 978-0-387-98140-6. Springer-Verlag New York.

[Wightman et al., 2010] Wightman, F., Solomon, A., Khoury, G., Green, J. A., Gray, L., Gorry, P. R., Ho, Y. S., Saksena, N. K., Hoy, J., Crowe, S. M., Cameron, P. U., and Lewin, S. R. (2010). Both CD31(+) and CD31(-) naive CD4(+) T cells are persistent HIV type 1-infected reservoirs in individuals receiving antiretroviral therapy. *J Infect Dis*, 202(11):1738–48.

[Wirth, 2012] Wirth, S. (2012). Current treatment options and response rates in children with chronic hepatitis C. *World J Gastroenterol*, 18(2):99–104.

[Wolfram Research, Inc., 2015] Wolfram Research, Inc. (2015). *Mathematica*. Wolfram Research, Inc., Champaign, Illinois, Version 10.3 edition.

[World Health Organisation, 2010] World Health Organisation (2010). Antiretroviral Therapy for HIV Infection in Infants and Children: Towards Universal Access: Recommendations for a Public Health Approach: 2010 Revision. Online.

[World Health Organisation, 2016a] World Health Organisation (2016a). Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection. http://www.who.int/hiv/pub/arv/arv-2016/en/.

[World Health Organisation, 2016b] World Health Organisation (2016b). WHO case definitions of HIV for surveillance and revised clinical and immunological classification of HIV-related disease in adults and children, 2007. http://www.who.int/hiv/pub/guidelines/HIVstaging150307.pdf.

[World Health Organisation, 2017a] World Health Organisation (2017a). Hepatitis C infection. http://www.who.int/mediacentre/factsheets/fs164/en/.

[World Health Organisation, 2017b] World Health Organisation (2017b). Mother-to-child transmission of hiv. http://www.who.int/hiv/topics/mtct/en/.

[World Health Organization, 2016] World Health Organization (2016). Anti-retroviral therapy of HIV infection in infants and children: towards universal access. Online.

[Wu and Zhang, 2006] Wu, H. and Zhang, J. (2006). *Nonparametric Regression Methods for Longitudinal Data Analysis: Mixed-Effects Modeling Approaches*. John Willey and Sons, New York.

[Wu and Zhang, 2002] Wu, H. and Zhang, J.-T. (2002). The study of long-term hiv dynamics using semi-parametric non-linear mixed-effects models. *Stat Med*, 21(23):3655–75.

[Xu et al., 2012] Xu, X. S., Yuan, M., Karlsson, M. O., Dunne, A., Nandy, P., and Vermeulen, A. (2012). Shrinkage in nonlinear mixed-effects population models: quantification, influencing factors, and impact. *AAPS J*, 14(4):927–36.

[Yates et al., 2007] Yates, A., Stark, J., Klein, N., Antia, R., and Callard, R. (2007). Understanding the slow depletion of memory CD4+ T cells in HIV infection. *PLoS Med*, 4(5):e177.

[Ye and Kirschner, 2002] Ye, P. and Kirschner, D. E. (2002). Reevaluation of T cell receptor excision circles as a measure of human recent thymic emigrants. *J Immunol*, 168(10):4968–79.

[Ye et al., 2003] Ye, P., Kourtis, A. P., and Kirschner, D. E. (2003). Reconstitution of thymic function in HIV-1 patients treated with highly active antiretroviral therapy. *Clin Immunol*, 106(2):95–105.

[Yeung et al., 2001] Yeung, L. T., King, S. M., and Roberts, E. A. (2001). Mother-to-infant transmission of hepatitis C virus. *Hepatology*, 34(2):223–9.

[Yeung et al., 2007] Yeung, L. T. F., To, T., King, S. M., and Roberts, E. A. (2007). Spontaneous clearance of childhood hepatitis C virus infection. *J Viral Hepat*, 14(11):797–805.

[Yuan and Johnson, 2012] Yuan, Y. and Johnson, V. E. (2012). Goodness-of-fit diagnostics for bayesian hierarchical models. *Biometrics*, 68(1):156–64.

[Zakhour et al., 2016] Zakhour, R., Tran, D. Q., Degaffe, G., Bell, C. S., Donnachie, E., Zhang, W., Pérez, N., Benjamins, L. J., Del Bianco, G., Rodriguez, G., Murphy, J. R., and Heresi, G. P. (2016). Recent Thymus Emigrant CD4+ T Cells Predict HIV Disease Progression in Patients With Perinatally Acquired HIV. *Clin Infect Dis*, 62(8):1029–35.

[Zanoni et al., 2012] Zanoni, B. C., Phungula, T., Zanoni, H. M., France, H., Cook, E. F., and Feeney, M. E. (2012). Predictors of poor CD4 and weight recovery in HIV-infected children initiating ART in South Africa. *PLoS One*, 7(3):e33611.

[Zeng et al., 2012] Zeng, M., Southern, P. J., Reilly, C. S., Beilman, G. J., Chipman, J. G., Schacker, T. W., and Haase, A. T. (2012). Lymphoid tissue damage in HIV-1 infection depletes naïve T cells and limits T cell reconstitution after antiretroviral therapy. *PLoS Pathog*, 8(1):e1002437.

[Zhang et al., 1999] Zhang, L., Lewin, S. R., Markowitz, M., Lin, H. H., Skulsky, E., Karanicolas, R., He, Y., Jin, X., Tuttleton, S., Vesanen, M., Spiegel, H., Kost, R., van Lunzen, J., Stellbrink, H. J., Wolinsky, S., Borkowsky, W., Palumbo, P., Kostrikis, L. G., and Ho, D. D. (1999). Measuring recent thymic emigrants in blood of normal and HIV-1-infected individuals before and after effective therapy. *J Exp Med*, 190(5):725–32.

Chapter A

NONMEM model scripts

A.1 CD4$^+$ T cell recovery in HIV/Hepatitis C co-infected children

```
$PROBLEM   Hepatitis C co-infection script
$INPUT      ID DV TIME AGEATARTSTART HCVSTATUS AGEATCD4 COUNTRY
            GENDER=DROP AIDS=DROP
$DATA       euroDataNM4.csv IGNORE=@
$PRED
;_____________________Covariate relations for C _____________________
;;; LRCHCVSTATUS-DEFINITION START
IF(HCVSTATUS.EQ.2) LRCHCVSTATUS = 1  ; Most common
IF(HCVSTATUS.EQ.1) LRCHCVSTATUS = ( 1 + THETA(13))
;;; LRCHCVSTATUS-DEFINITION END


;;; LRC-RELATION START
LRCCOV=LRCHCVSTATUS
;;; LRC-RELATION END
;_____________________Covariate relations for intercept _____________________
;;; INTCOUNTRY-DEFINITION START
IF(COUNTRY.EQ.8) INTCOUNTRY = 0  ; Most common
IF(COUNTRY.EQ.4) INTCOUNTRY = ( 0 + THETA(6))
IF(COUNTRY.EQ.6) INTCOUNTRY = ( 0 + THETA(7))
IF(COUNTRY.EQ.7) INTCOUNTRY = ( 0 + THETA(8))
IF(COUNTRY.EQ.2) INTCOUNTRY = ( 0 + THETA(9))
IF(COUNTRY.EQ.5) INTCOUNTRY = ( 0 + THETA(10))
```

```
IF(COUNTRY.EQ.3) INTCOUNTRY = ( 0 + THETA(11))
IF(COUNTRY.EQ.1) INTCOUNTRY = ( 0 + THETA(12))
;;; INTCOUNTRY-DEFINITION END

;;; INTAGEATARTSTART-DEFINITION START
INTAGEATARTSTART = ( 0 + THETA(5)*(AGEATARTSTART - 4.30))
;;; INTAGEATARTSTART-DEFINITION END

;;; INT-RELATION START
INTCOV=INTAGEATARTSTART+INTCOUNTRY
;;; INT-RELATION END
;_____________________Covariate relations for asymptote _____________________
;;; ASYAGEATARTSTART-DEFINITION START
ASYAGEATARTSTART = ( 0 + THETA(4)*(AGEATARTSTART - 4.30))
;;; ASYAGEATARTSTART-DEFINITION END

;;; ASY-RELATION START
ASYCOV=ASYAGEATARTSTART
;;; ASY-RELATION END

TVINT = THETA(1)
TVINT = INTCOV+TVINT
TVASY = THETA(2)
TVASY = ASYCOV+TVASY
TVLRC = THETA(3)
TVLRC = LRCCOV*TVLRC
INT = TVINT + ETA(1)
ASY = TVASY + ETA(2)
LRC = TVLRC*EXP(ETA(3))
IPRED = ASY-(ASY-INT)*EXP(-LRC*TIME)
Y= IPRED + EPS(1)
```

```
$THETA
-2.5181                       ; int
-1.06531                     ; asy
 1.44595                      ; c
$THETA
(-20,-0.113493,20)     ; ASYAGEATARTSTART1
$THETA
(-20,-0.319074,20)     ; INTAGEATARTSTART1
$THETA
(-20,0.684892,20)       ; INTCOUNTRY1
 (-20,0.0236298,20)    ; INTCOUNTRY2
 (-20,-17.4538,20)       ; INTCOUNTRY3
 (-20,-3.6333,20)         ; INTCOUNTRY4
 (-20,2.89443,20)         ; INTCOUNTRY5
 (-20,0.439309,20)       ; INTCOUNTRY6
 (-20,0.338835,20)       ; INTCOUNTRY7
$THETA
(-1,-0.773207,5)           ; LRCHCVSTATUS1
$OMEGA
4.78392
1.82605
0.3937
$SIGMA
1.33657
$ESTIMATION MAXEVAL=9999 METHOD=1 INTER SIGDIG=3 PRINT=1
$COVARIANCE PRINT=E
$TAB ID DV TIME PRED IPRED CWRES INT ASY LRC ETA1 ETA2 ETA3 AGEATARTSTART COUNTRY
NOPRINT NOAPPEND ONEHEADER FILE=sdtab1
$TAB ID DV TIME ASY INT LRC ETA1 ETA2 ETA3 ONEHEADER NOPRINT FILE=patab1
$TAB ID DV TIME AGEATARTSTART  ONEHEADER NOPRINT FILE=cotab1  ; continuous covariate
```

```
$TAB ID DV TIME COUNTRY ONEHEADER NOPRINT FILE=catab1  ; categorical covariate
$TAB ID DV TIME CWRES PRED IPRED ONEHEADER NOPRINT FILE=mytab1
```

A.2 Predictors of poor immune responders

A.2.1 Model A

```
;; Final covariate model A
$PROBLEM
$DATA      ../../finalThesisTrainData.csv IGNORE=@
$INPUT      ID DV TIME AGEATARTSTART COHORT GENDER=DROP REGION BLOGVL=DROP AIDS
$PRED
;_____________________ Covariate relations for C _____________________
;;; CREGION-DEFINITION START
IF(REGION.EQ.4) CREGION = 1  ; Most common
IF(REGION.EQ.1) CREGION = ( 1 + THETA(11))
IF(REGION.EQ.3) CREGION = ( 1 + THETA(12))
IF(REGION.EQ.2) CREGION = ( 1 + THETA(13))
;;; CREGION-DEFINITION END

;;; C-RELATION START
CCOV=CREGION
;;; C-RELATION END
;_____________________Covariate relations for intercept _____________________
;;; INTREGION-DEFINITION START
IF(REGION.EQ.4) INTREGION = 0  ; Most common
IF(REGION.EQ.1) INTREGION = ( 0 + THETA(8))
IF(REGION.EQ.3) INTREGION = ( 0 + THETA(9))
IF(REGION.EQ.2) INTREGION = ( 0 + THETA(10))
;;; INTREGION-DEFINITION END
```

```
;;; INTAIDS-DEFINITION START
IF(AIDS.EQ.0) INTAIDS = 0  ; Most common
IF(AIDS.EQ.1) INTAIDS = ( 0 + THETA(7))
;;; INTAIDS-DEFINITION END

;;; INTAGEATARTSTART-DEFINITION START
IF(AGEATARTSTART.LE.5.78) INTAGEATARTSTART = ( 0 + THETA(5)*(AGEATARTSTART - 5.78))
IF(AGEATARTSTART.GT.5.78) INTAGEATARTSTART = ( 0 + THETA(6)*(AGEATARTSTART - 5.78))
;;; INTAGEATARTSTART-DEFINITION END

;;; INT-RELATION START
INTCOV=INTAGEATARTSTART+INTAIDS+INTREGION
;;; INT-RELATION END
;_____________________Covariate relations for asymptote _____________________
;;; ASYAGEATARTSTART-DEFINITION START
ASYAGEATARTSTART = ( 0 + THETA(4)*(AGEATARTSTART - 5.78))
;;; ASYAGEATARTSTART-DEFINITION END

;;; ASY-RELATION START
ASYCOV=ASYAGEATARTSTART
;;; ASY-RELATION END

TVINT = THETA(1)
TVINT = INTCOV+TVINT
TVASY = THETA(2)
TVASY = ASYCOV+TVASY
TVC = THETA(3)
TVC = CCOV*TVC
INT = TVINT + ETA(1)
ASY = TVASY + ETA(2)
```

```
C = TVC*EXP(ETA(3))
IPRED = ASY-(ASY-INT)*EXP(-C*TIME)
Y = IPRED + EPS(1)

$THETA
-1.2959          ; int
-0.32091        ; asy
(0, 1.72343)    ; c
$THETA
-0.0430181     ; ASYAGEATARTSTART1
$THETA
-0.165713       ; INTAGEATARTSTART1
-0.00571004     ; INTAGEATARTSTART2
$THETA
-0.629958       ; INTAIDS1
$THETA
0.0595219       ; INTREGION1
-0.928997       ; INTREGION2
-0.225928       ; INTREGION3
$THETA
0.139205         ; CREGION1
-0.00876235   ; CREGION2
0.0838468      ; CREGION3
$OMEGA
 1.09431
 0.243181
 0.245
$SIGMA
0.12396
$ESTIMATION METHOD=SAEM INTERACTION NBURN=2000 NITER=80 CTYPE=3 PRINT=1
$ESTIMATION METHOD=IMP EONLY=1 INTERACTION ISAMPLE=1000 NITER=20 PRINT=1 CTYPE=1 NOABORT
```

```
$COVARIANCE PRINT=E
$TABLE  ID DV TIME PRED IPRED CWRES INT ASY C ETA1 ETA2 ETA3
            AGEATARTSTART REGION AIDS NOPRINT NOAPPEND ONEHEADER FILE=sdtab2
$TABLE  ID DV TIME ASY INT C ETA1 ETA2 ETA3 ONEHEADER NOPRINT FILE=patab2
$TABLE  ID DV TIME AGEATARTSTART ONEHEADER NOPRINT FILE=cotab2 ; continuous covariate
$TABLE  ID DV TIME REGION AIDS ONEHEADER NOPRINT FILE=catab2 ; categorical covariate
$TABLE  ID DV TIME CWRES PRED IPRED ONEHEADER NOPRINT FILE=mytab2
```

A.2.2 Model B

```
;; Final covariate model B
$PROBLEM
$DATA      ../../finalThesisTestData.csv IGNORE=@
$INPUT      ID DV TIME AGEATARTSTART COHORT GENDER=DROP REGION BLOGVL AIDS
$PRED
;_____________________Covariate relations for intercept ____________________
;;; INTAGEATARTSTART-DEFINITION START
IF(AGEATARTSTART.LE.5.40) INTAGEATARTSTART = ( 0 + THETA(15)*(AGEATARTSTART - 5.4))
IF(AGEATARTSTART.GT.5.40) INTAGEATARTSTART = ( 0 + THETA(16)*(AGEATARTSTART - 5.4))
;;; INTAGEATARTSTART-DEFINITION END

;;; INTREGION-DEFINITION START
IF(REGION.EQ.4) INTREGION = 0  ; Most common
IF(REGION.EQ.1) INTREGION = ( 0 + THETA(12))
IF(REGION.EQ.3) INTREGION = ( 0 + THETA(13))
IF(REGION.EQ.2) INTREGION = ( 0 + THETA(14))
;;; INTREGION-DEFINITION END
```

```
;;; INTBLOGVL-DEFINITION START
INTBLOGVL = ( 0 + THETA(11)*(BLOGVL - 5.08))
;;; INTBLOGVL-DEFINITION END

;;; INTAIDS-DEFINITION START
IF(AIDS.EQ.0) INTAIDS = 0  ; Most common
IF(AIDS.EQ.1) INTAIDS = ( 0 + THETA(10))
;;; INTAIDS-DEFINITION END

;;; INT-RELATION START
INTCOV=INTAIDS+INTBLOGVL+INTREGION+INTAGEATARTSTART
;;; INT-RELATION END
;_____________________Covariate relations for C _____________________
;;; CREGION-DEFINITION START
IF(REGION.EQ.4) CREGION = 1  ; Most common
IF(REGION.EQ.1) CREGION = ( 1 + THETA(7))
IF(REGION.EQ.3) CREGION = ( 1 + THETA(8))
IF(REGION.EQ.2) CREGION = ( 1 + THETA(9))
;;; CREGION-DEFINITION END

;;; CAIDS-DEFINITION START
IF(AIDS.EQ.0) CAIDS = 1  ; Most common
IF(AIDS.EQ.1) CAIDS = ( 1 + THETA(6))
;;; CAIDS-DEFINITION END

;;; C-RELATION START
CCOV=CAIDS*CREGION
;;; C-RELATION END
;_____________________Covariate relations for asymptote _____________________
;;; ASYAGEATARTSTART-DEFINITION START
IF(AGEATARTSTART.LE.5.40) ASYAGEATARTSTART = ( 0 + THETA(4)*(AGEATARTSTART - 5.4))
```

```
IF(AGEATARTSTART.GT.5.40) ASYAGEATARTSTART = ( 0 + THETA(5)*(AGEATARTSTART - 5.4))
;;; ASYAGEATARTSTART-DEFINITION END

;;; ASY-RELATION START
ASYCOV=ASYAGEATARTSTART
;;; ASY-RELATION END

TVINT = THETA(1)
TVINT = INTCOV+TVINT
TVASY = THETA(2)
TVASY = ASYCOV+TVASY
TVC = THETA(3)
TVC = CCOV*TVC
INT = TVINT + ETA(1)
ASY = TVASY + ETA(2)
C  = TVC*EXP(ETA(3))
IPRED = ASY-(ASY-INT)*EXP(-C*TIME)
Y = IPRED + EPS(1)

$THETA
-1.21417        ; int
-0.23364        ; asy
(0,1.37715)     ; c
$THETA
-0.0121368     ; ASYAGEATARTSTART1
-0.0519592     ; ASYAGEATARTSTART2
$THETA
0.54657           ; CAIDS1
$THETA
0.244146         ; CREGION1
0.150225         ; CREGION2
```

```
-0.274451        ; CREGION3
$THETA
-0.928506        ; INTAIDS1
$THETA
-0.145631         ; INTBLOGVL1
$THETA
0.177001          ; INTREGION1
-0.850708         ; INTREGION2
0.104971          ; INTREGION3
$THETA
-0.138126         ; INTAGEATARTSTART1
-0.0387801       ; INTAGEATARTSTART2
$OMEGA
1.05735
0.242787
0.5
$SIGMA
0.134575
$ESTIMATION MAXEVAL=9999 METHOD=1 INTER SIGDIG=3 PRINT=1
$COVARIANCE PRINT=E
$TABLE      ID DV TIME PRED IPRED CWRES INT ASY C ETA1 ETA2 ETA3
            AGEATARTSTART REGION AIDS BLOGVL NOPRINT NOAPPEND
            ONEHEADER FILE=sdtab3
$TABLE  ID DV TIME ASY INT C ETA1 ETA2 ETA3 ONEHEADER NOPRINT FILE=patab3
$TABLE  ID DV TIME AGEATARTSTART BLOGVL ONEHEADER NOPRINT FILE=cotab3 ; continuous covariate
$TABLE  ID DV TIME REGION AIDS ONEHEADER NOPRINT FILE=catab3 ; categorical covariate
$TABLE  ID DV TIME CWRES PRED IPRED ONEHEADER NOPRINT FILE=mytab3
```

A.3 Naive CD31$^+$ T cell homeostasis

```
;; Scenario 3: CD31 expression is retained after two T cell divisions and lost thereafter
```

```
$PROB CD31+ T cell model 3
$DATA cd31DataFull.csv IGNORE=@
$INPUT ID DV TIME HLARATIO AGEATCD4 AGEATCD4DAYS
$SUBROUTINE ADVAN13 TOL=9
$MODEL      COMP (CD4) COMP(CD4) COMP(CD4)
$PK
;------- Initiate population-level parameters for model ---------------------------------
  TVN = THETA(1)
 TVT0 = THETA(2)
 TVM0 = THETA(3)
 TVD0 = THETA(4)
 TVCM = THETA(5)
 TVCD = THETA(6)
 TVN2 = THETA(7)
 TVN3 = THETA(8)
 TVX  = THETA(9)
;------------ MU modelling to improve model run-times -------------------------------
MU_1 = LOG(TVN)
MU_2 = LOG(TVT0)
MU_3 = LOG(TVM0)
MU_4 = LOG(TVD0)
MU_5 = LOG(TVCM)
MU_6 = LOG(TVCD)
MU_7 = LOG(TVN2)
MU_8 = LOG(TVN3)
MU_9 = LOG(TVX)
;------- Naive (VA) T cell concentration with age ----------------------------------------
  VA = 496 + 2074*EXP(-0.000857*AGEATCD4DAYS)
;------- Include individual random effects on all parameters -------------
   N = EXP(MU_1 + ETA(1))
  TH = EXP(MU_2 + ETA(2))
```

```
   M = EXP(MU_3 + ETA(3))
   D = EXP(MU_4 + ETA(4))
  CM = EXP(MU_5 + ETA(5))
  CD = EXP(MU_6 + ETA(6))
  N2 = EXP(MU_7 + ETA(7))
  N3 = EXP(MU_8 + ETA(8))
  X  = EXP(MU_9 + ETA(9))
;----------- Include age effects on parameters -------------------------------
;----------- Estimate exponential constant -----------------------------------
TH_0 = TH * VA*0.02*0.905*EXP(-X*AGEATCD4DAYS)
PROL = M * 0.02*EXP(-X*AGEATCD4DAYS)
LOSS = D * 0.02*EXP(-X*AGEATCD4DAYS)
;------------------ Define initial concentration for each compartment---------------------
  A_0(1) = N
  A_0(2) = N2
  A_0(3) = N3
;-------------------------- Define the differential equations --------------------------
$DES
DADT(1) = TH_0 - A(1)*(LOSS * EXP(CD*((A(1)+A(2)+A(3))/VA-1)) +
PROL * EXP(CM*(-(A(1)+A(2)+A(3))/VA+1)))
DADT(2) = 2*PROL*A(1) - A(2)*(PROL * EXP(CM*(-(A(1)+A(2)+A(3))/VA+1)) +
LOSS * EXP(CD*((A(1)+A(2)+A(3))/VA-1)))
DADT(3) = 2*PROL*A(2) - A(3)*(PROL * EXP(CM*(-(A(1)+A(2)+A(3))/VA+1)) +
LOSS * EXP(CD*((A(1)+A(2)+A(3))/VA-1)))
;-------------------- Compare model output ipred to observed data ----------------------
$ERROR
IPRED = A(1) + A(2) + A(3)
Y = IPRED*(1 + EPS(1))
;------------ Initial estimates for the population-level parameters ---------------------
$THETA
(0, 300)    ;1. N, Initial CD31 T cell concentration prior to 1st division
```

```
(0, 0.007)  ;2. T0, proportion of thymic output
(0, 0.08)   ;3. M0, proliferation
(0, 0.40)   ;4. D0, loss
(0, 0.3)    ;5. CM, strength of proliferation
(0, 4.0)    ;6. CD, strength of loss
(0, 200)    ;7. N2_0, initial no of CD31 T cells prior to 2nd division
(0, 200)    ;7. N3_0, initial no of CD31 T cells prior to 3rd division
(0, 0.0005) ; exponential decay constant
;---------------Initial estimates for the random effect var-covariance matrix-----------------
$OMEGA BLOCK(8)
0.1
0.01 0.1
0.01 0.01 0.1
0.01 0.01 0.01 0.1
0.01 0.01 0.01 0.01 0.1
0.01 0.01 0.01 0.01 0.01 0.1
0.01 0.01 0.01 0.01 0.01 0.01 0.1
0.01 0.01 0.01 0.01 0.01 0.01 0.01 0.1
$OMEGA
0 FIX
;----------------Initial estimate for the residual error variance------------------------
$SIGMA
0.218
;-------------Estimation algorithm to be used (EONLY=1 Expectation Only)----------------
$ESTIMATION METHOD=SAEM INTERACTION NBURN=2000 NITER=1000 CTYPE=3 PRINT=1
$ESTIMATION METHOD=IMP EONLY=1 INTERACTION ISAMPLE=1000 NITER=20 PRINT=1 NSIG=3 MAPITER=0
$COVARIANCE PRINT=E
;----------------Output results of the model-fitting into a table-----------------------
$TABLE ID DV TIME IPRED PRED CWRES N N2 N3 TH M D TH_0 PROL LOSS CM CD ETA1 ETA2 ETA3 ETA4 ETA5
ETA6 ETA7 ETA8 NOPRINT ONEHEADER FILE=sdtab3
$TABLE ID N TH TH_0 M D CM CD PROL LOSS ETA1 ETA2 ETA3 ETA4 ETA5 ETA6 ETA7 ETA8 NOPRINT
```

```
ONEHEADER FILE=patab3
```

Chapter B

R Codes

B.1 CD4 z-score function

The below function written in R code was used to estimates CD4 z-scores. It accepts age in years and CD4 count (cell/μL). The function correlates well with the form described in [Wade and Ades, 1994, Wade and Ades, 1998]. Individual parameter values were provided by the MRC Clinical Trials Unit.

```
#####################################################################
########## age in years and CD4 count in cells/microL ######
#####################################################################
cd4.conc.zsc <- function (cd4,age){
a1 <-  0.62987
a2 <- -1.4017
a3 <-  0.83364
a4 <- -0.23954
a5 <-  0.57091
b1 <- -0.49740
b2 <-  0.75867
b3 <-  0.23400
c1 <-  0.26709
c2 <-  0.26079
c3 <-  0.43891

x <- (age^a5-1)/a5
```

```
m <- (a1*x-a2)*exp(-a3*x+a4)
l <- b1+b2*exp(-b3*x)
s <- c1+c2*exp(-c3*x)
(((((cd4/1000)^l)-1)/l)-m)/s
}
```

B.2 CD4 ratio function

```
#################################################################
################################## age in months #############
#################################################################


h.cd4.conc <- function(age){
  3.27863*10^3-2.35445873*10^3*(1-exp(-0.030371232*age))
}
```

B.3 Convertion of ln(CD4-for-age) to CD4 count

```
#################################################################
###########  Get CD4 count from log cd4 Ratio  ##############
#################################################################

logR_to_count <- function(logRatio, ageInYears){
  exp(logRatio)*(3.27863*10^3-2.35445873*10^3*(1-exp(-0.030371232*ageInYears*12)))
}
```