January 2012

Bayesian Inference on Mixed-effects Models with Skewed Distributions for HIV longitudinal Data

Ren Chen
University of South Florida, rchen@health.usf.edu

Bayesian Inference on Mixed-Effects Models with Skewed Distributions for HIV Longitudinal Data

by

Ren Chen

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Epidemiology and Biostatistics
College of Public Health
University of South Florida

Major Professor: Yangxin Huang, Ph.D.
Getachew A. Dagne, Ph.D.
Yiliang Zhu, Ph.D.
Patricia J. Emmanuel, M.D.
Henian Chen, M.D., Ph.D.

Date of Approval:
November 30, 2012

Keywords: mixed-effects model, HIV dynamics, Bayesian analysis, Markov chain Monte Carlo, skewed distribution

Dedication

This work is dedicated to my mother, Zhenhua Zhang, my husband Ralph and my lovely daughter Victoria, who have been giving me their selfless support for my study and research during all these years.

Acknowledgments

First of all, I would like to express sincere appreciation to my major advisor, Dr. Yangxin Huang, who has devoted himself on mentoring and nourishing me throughout my study and research. I could not have gone so far without his guidance and support. What I have learned from him and his research enables me to overcome many hurdles and accomplish what I have dreamed of. I believe his profound positive impact will continue to benefit my future professional work.

I would also like to thank Dr. Henian Chen, my committee member and director of biostatistics core in which I am working now, who provided great support and career guides. Special thanks are given to Drs. Yiliang Zhu, Getachew A. Dagne, and Patricia J. Emmanuel for positive input, helpful suggestions concerning the research presented in this dissertation and for taking valuable time out of their busy schedules to serve on my dissertation committee.

I also want to acknowledge and thank Dr. Jane Carver and all faculties and staff members at the Department of Epidemiology and Biostatistics, College of Public Health and Clinical and Transitional Science Institute, Morsani College of Medicine, University of South Florida.

Last but not the least, I would like to thank Dr. Michael Schell for being the Chairman of my defense committee.

Table of Contents

List of Tables

List of Figures

Abstract

Statistical models have greatly improved our understanding of the pathogenesis of HIV-1 infection and guided for the treatment of AIDS patients and evaluation of antiretroviral (ARV) therapies. Although various statistical modeling and analysis methods have been applied for estimating the parameters of HIV dynamics via mixed-effects models, a common assumption of distribution is normal for random errors and random-effects. This assumption may lack the robustness against departures from normality so may lead misleading or biased inference. Moreover, some covariates such as CD4 cell count may be often measured with substantial errors. Bivariate clustered (correlated) data are also commonly encountered in HIV dynamic studies, in which the data set particularly exhibits skewness and heavy tails. In the literature, there has been considerable interest in, via tangible computation methods, comparing different proposed models related to HIV dynamics, accommodating skewness (in univariate) and covariate measurement errors, or considering skewness in multivariate outcomes observed in longitudinal studies. However, there have been limited studies that address these issues simultaneously.

One way to incorporate skewness is to use a more general distribution family that can provide flexibility in distributional assumptions of random-effects and model random errors to produce robust parameter estimates. In this research, we developed Bayesian hierarchical models in which the skewness was incorporated by using skew-elliptical (SE) distribution and all of the inferences were carried out through Bayesian approach via Markov chain Monte Carlo (MCMC). Two real data set from HIV/AIDS clinical trial were used to illustrate the proposed models and methods.

This dissertation explored three topics. First, with an SE distribution assumption, we compared models with different time-varying viral decay rate functions. The effect of skewness on the model fitting was also evaluated. The associations between the estimated decay rates based on the best fitted model and clinical related variables such as baseline HIV viral load, CD4 cell count and long-term response status were also evaluated. Second, by jointly modeling via a Bayesian approach, we simultaneously addressed the issues of outcome with skewness and a covariate process with

measurement errors. We also investigated how estimated parameters were changed under linear, nonlinear and semiparametric mixed-effects models. Third, in order to accommodate individual clustering within subjects as well as the correlation between bivariate measurements such as CD4 and CD8 cell count measured during the ARV therapies, bivariate linear mixed-effects models with skewed distributions were investigated. Extended underlying normality assumption with SE distribution assumption was proposed. The impacts of different distributions in SE family on the model fit were also evaluated and compared.

Real data sets from AIDS clinical trial studies were used to illustrate the proposed methodologies based on the three topics and compare various potential models with different distribution specifications. The results may be important for HIV/AIDS studies in providing guidance to better understand the virologic responses to antiretroviral treatment. Although this research is motivated by HIV/AIDS studies, the basic concepts of the methods developed here can have generally broader applications in other fields as long as the relevant technical specifications are met. In addition, the proposed methods can be easily implemented by using the publicly available WinBUGS package, and this makes our approach quite accessible to practicing statisticians in the fields.

1 Introduction / Literature Review

1.1. Background

The history of human immunodeficiency virus (HIV) and acquired immunodeficiency syndrome (AIDS) can be traced back to 1981. In California and New York, various doctors reported that a small number of homosexual men had been diagnosed with rare forms of *Kaposi's* sarcoma and *Pneumocystis carinii* pneumonia, which are generally found in people with seriously compromised immune systems. By mid 1982, it was clear that they were more than isolated incidents and in September of that year, Centers for Disease Control and Prevention (CDC) used the term AIDS as an official diagnosis for this disease. Soon it was realized people could get HIV if they engaged in certain activities such as having unprotected sex, sharing needles, receiving a blood transfusion and if they were born to a mother with HIV infection.

HIV infection is considered as a pandemic by the World Health Organization (WHO). By the end of 2010 (UNAIDS 2011), an estimated 34 million people were living with HIV, up 17% from 2001. Approximately 16.8 million are women and 3.4 million are less than 15 years old. The estimated prevalence of HIV varies dramatically among regions: the most affected region is Sub-Saharan Africa and it accounts 68% HIV cases and 66% of HIV deaths; about 5% of the adult population in this area is infected. Prevalence is the lowest in Western and Central Europe (0.2%) and East Asia (0.1%). With the significant expansion of HIV prevention programs and access to antiretroviral therapy, the number of new infections and HIV/AIDS related deaths are decreasing. In 2010, there were 2.7 million new HIV infections, which was 15% less than in 2001 and 21% less than the number of new infections that occurred at the peak of the epidemic in 1997, and there were 1.8 million AIDS related deaths, which was 18% less than in 2001. In the United States, since the beginning of the HIV and AIDS epidemic, over half a million people have died of AIDS, and currently around 1.2 million people are living with HIV, however 20% of them are unaware of their infection.

HIV belongs to a class of viruses known as retroviruses. Retroviruses use ribonucleic acid

(RNA) to encode their genetic information and RNA is translated into deoxyribonucleic acid (DNA) during its life-cycle by a specific viral enzyme called reverse transcriptase. Viruses cannot grow or reproduce on their own so they must infect cells of a living organism in order to survive and make new copies. There are two types of HIV, HIV-1 and HIV-2, and both originated through the evolution of simian immunodeficiency virus (SIV). Although both types can be transmitted by sexual contact, blood, and from mother to child, compared with HIV-2, HIV-1 is more easily transmitted and patients with HIV-1 infection will more quickly progress to AIDS. Therefore, it is responsible for the majority of global HIV infections and AIDS cases.

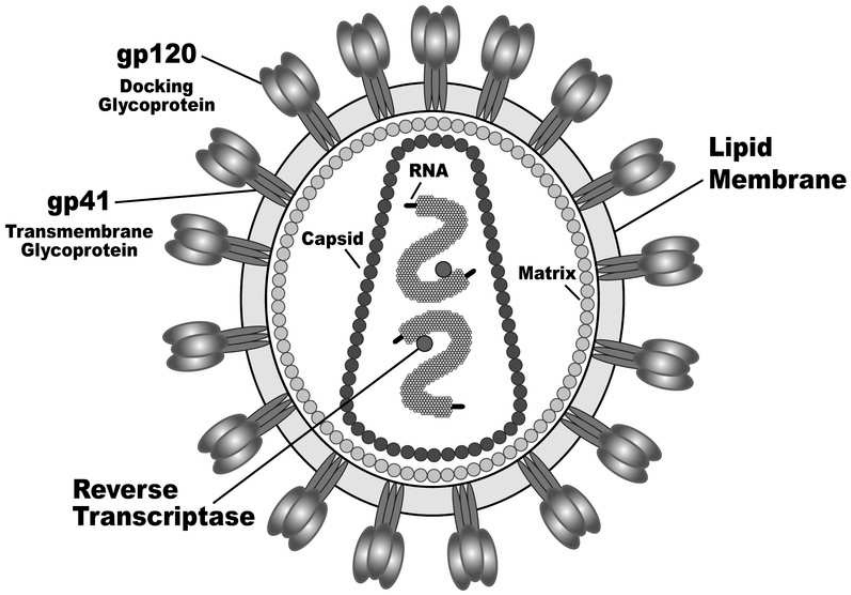

Figure 1.1: Diagram of HIV.

HIV virion is roughly spherical and has a diameter of about 1/10,000 *mm*, which is 60 times smaller than a red blood cell. As Figure 1.1 shows, the basic structure of HIV includes: (i) a lipid membrane. It is the outer envelope of the virus and consists of two layers of lipids. Different proteins are embedded in this viral envelope and form ”spikes” consisting of glycoprotein (gp) 120 and transmembrane gp41. Gp120 is needed to attach the virion to the host cell, and gp41 is critical for the cell fusion process; (ii) the HIV matrix proteins. They lie between the envelope and core; (iii) the viral core. It contains the viral capsule protein p24 which surrounds two single strands of RNA and the enzymes needed for HIV replication, such as reverse transcriptase, protease, ribonuclease,

and integrase. Among the nine virus genes coded on one long stand of RNA, three genes, gag, pol and env, contain information needed to make structural proteins for new virus particles.

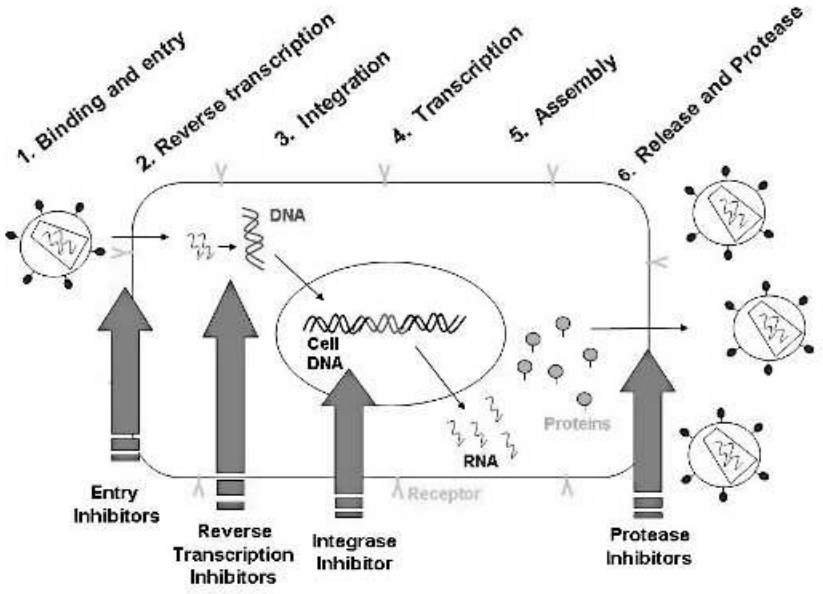

Figure 1.2: HIV replication.

There are six steps involved in HIV infection and replication (Figure 1.2). Step 1: binding and entry. By binding specific receptors on the surface of a target cell, such as CD4 positive T cells (i.e., CD4 cells), macrophages and microglial cells, HIV enters the host cells. The CD4 receptor is necessary but not sufficient to permit virus entry. The secondary receptors are "chemokine receptors" that bind to chemokines and are needed to facilitate the entering (Dragic et al., 1996); Step 2: reverse transcription. HIV uses an enzyme known as reverse transcriptase to convert its RNA into DNA; Step 3: integration. HIV DNA enters the nucleus of the target cell and inserts itself into the cell's DNA, where it may "hide" and stay inactive for years; Step 4: transcription. HIV DNA instructs the cell to make many copies of the original virus, along with some more specialized genetic materials for making longer proteins; Step 5: assembly. A special enzyme called protease cuts the longer HIV proteins into individual proteins. When these come together with the virus genetic material, a new virus is assembled; Step 6: release. The virus pushes itself out of the host cell and takes with it part of the cell membrane. This outer part covers the virus and contains all of the structures necessary for the virus to bind to a new CD4 cell and begin the virus life cycle process again. Knowing these

steps is critical in the development of medications that can interrupt the replication cycle. Current treatment strategy involves a combination of drugs that target different steps of HIV's life cycle such as entry inhibitors that prevent binding of HIV to the CD4 receptor, reverse transcriptase inhibitors that prevent the HIV RNA from being transcribed into DNA and protease inhibitors that prevent the assembly.

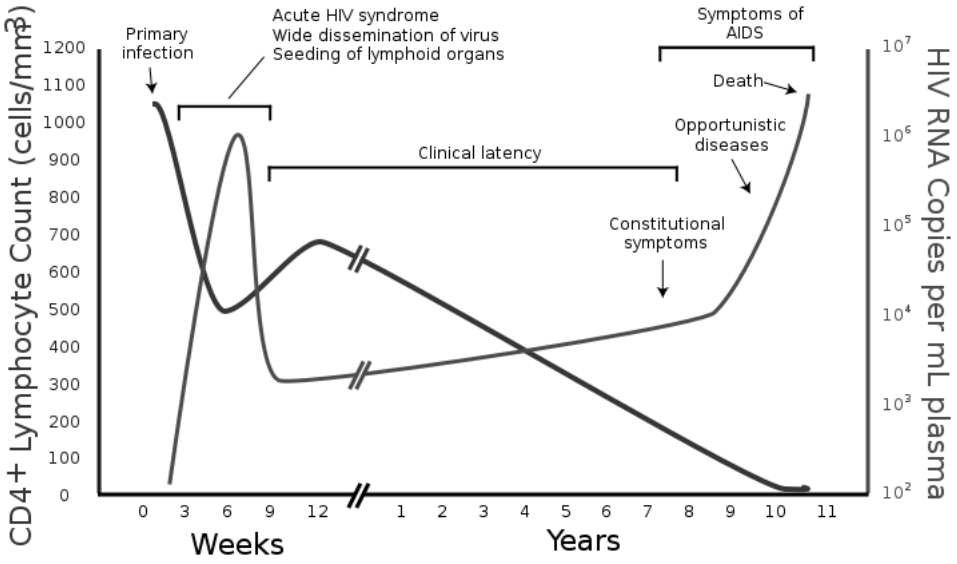

Figure 1.3: A generalized graph of the relationship between HIV copies and CD4 cell count over the average course of untreated HIV infection.

HIV infection generally can be broken into four stages: primary infection, clinical latency (asymptomatic) stage, symptomatic stage and AIDS (Figure 1.3). Stage 1: primary infection. This stage can last for a few weeks and patients are often accompanied by a short flu-like symptom, such as headache, nausea, sore throat or fever. During this stage, the amount of HIV in the peripheral blood increases sharply and the immune system starts to respond to the virus by producing HIV antibodies and cytotoxic lymphocytes. This process is known as "seroconversion". Since enzyme-linked immunosorbent assay (ELISA), which is the most commonly used method to test for HIV, uses blood, oral fluid or urine to detect HIV antibodies, the result may be negative if the ELISA test is done before seroconversion is complete. There is a corresponding decrease in the number of CD4 cells and an increase in CD8 cells. Patients are extremely infectious during this stage. Stage 2: clinically asymptomatic stage. This stage lasts for an average of ten years and patients are free

from major symptoms, although some may have swollen glands. During this stage, the immune system is able to mount an effective response, so the viral load starts to decrease and then stays at a constant low level. The number of CD4 cells rises and then slowly falls. People remain infectious and HIV antibodies are detectable in the blood so the antibody test will show a positive result. Although the viral load remains at a constant low level for years, virus replication is very active during this period. Stage 3: symptomatic HIV infection. Eventually, the immune system is severely damaged or "burned out" by years of activity. HIV mutates and becomes more pathogenic leading more immune cells destruction, while the body fails to keep up with replacing the lost cells. Symptomatic HIV infection is mainly caused by opportunistic infections that the normal immune system usually would prevent. This stage of HIV infection is often characterized by multi-system diseases and infections occurring in almost all body systems. Without any effective treatment, the immune suppression will continue to worsen. Stage 4: AIDS. Once the CD4 cell count is less than $200/mL$ or CD4 cell percentage is less than 15, AIDS will be diagnosed.

The CD4 cell, the major target cell for HIV, is a T lymphocytes. Under the microscope, lymphocytes can be divided into large and small lymphocytes. Large lymphocytes include natural killer (NK) cells, while small lymphocytes consist of T cells that mature from thymus and B cells that are bursa-derived. T cells are involved in cell-mediated immunity whereas B cells are primarily responsible for humoral immunity (relating to antibodies). The CD4 cell is a subset of T cells that express the cluster of differentiation 4 (CD4) and it is also known as T helper cell. These cells assist other white blood cells in immunologic processes. The normal CD4 cells account for 32% to 68% of total number of lymphocytes and range between 500 – $1600/mL$. Without effective HIV treatment, the hallmark decrease in CD4 cells that occurs during AIDS results in such a weakened immune system that the body can no longer fight infections or certain cancers, and eventually death ensues. The mechanisms of CD4 cell death in HIV infection are still not fully understood and are one of the most controversial issues in AIDS research. The mechanisms by which HIV can directly induce infected cell death include plasma membrane disruption or increased permeability due to continuous budding of the virion (Facui, 1988), increasing cellular toxicity due to build up of un-integrated liner viral DNA (Levy, 1993) and inactivation of anti-apoptotic genes (Nie et al., 2002). However, a longstanding question in HIV biology is how HIV viruses kill so many CD4 cells, despite the fact that most of them appear to be "bystander" cells that are not infected (Embretson et al., 1993). Re-

cent data demonstrate that the majority uninfected CD4 cells in peripheral blood and lymph nodes undergo three types of apoptosis (Varbanov et al., 2006), which is a tightly regulated programmed cell death (Evan et al., 1998). Several HIV proteins, such as Env and Vpr, have been found to be able to up-regulate Fas/FasL gene expression either on the infected cells or neighboring uninfected cells (Kaplan and Sieg, 1998), and these two genes will send signal of apoptosis to these cells.

CD8 cell is another type of T cell. It destroys virally infected cells and tumor cells so it is also known as cytotoxic T cell (T_c cells or CTLs). A healthy adult usually has 150 – 1,000$/mL$ CD8 cells and the normal ratio of CD4/CD8 is 1.0 – 3.7. In contrast to CD4 cells, CD8 cells often increase in people with HIV and the significance has not been well understood. Researches have revealed (Chevret et al., 1992; Krantz et al., 2011) that elevated total CD8 cell count was associated with greater risk of future virologic failure. The CD4/CD8 Ratio is used to help in diagnosing HIV, monitoring HIV progress, and making treatment decisions.

HIV diagnostic test is done by either detecting host antibodies made against different HIV proteins or by directly detecting the whole virus or components of virus (Iweala, 2004). Tests that detect host antibodies that are specific to the virus include ELISA, Western blot, the immunofluorescence assay (IFA), and the detuned assay. For screening purposes, ELISA is usually used first, and in order to minimize the risk of false positive results, a confirmatory test, such as Western blot or IFA, should be conducted before a patient is given the diagnosis of HIV infection. Detuned assay is used to distinguish recent HIV infection within the past 129 – 180 days from older HIV infections (Parekh et al., 2002). These tests may be negative during the acute infection or before seroconversion is completed. In contrast, three types of tests can directly detect the virus or parts of the virus as soon as people become infected with HIV. These tests include p24 antigen detection, peripheral blood mononuclear cell culture and RNA nucleic acid-based assays, such as reverse transcription followed by polymerase chain reaction (RT-PCR) and hybridization-based assays. Undetectable viral load is usually defined as less than 50 copies$/ml$. Until recently, this was the lowest detectable level for the commonly used tests in routine viral load monitoring. There are now some ultra-sensitive tests that can measure less than 20 copies$/ml$ and even 1 copy$/ml$ of plasma (Palmer et al., 2003).

It takes an average of 10 years after HIV infection to develop AIDS, and the viral load generally remains unchanged if measured repeatedly during those years. Originally, many people thought

the rate of HIV replication and disease process would be slow, which is not true. In 1995 and 1996, several important papers (Ho et al., 1995; Perelson et al., 1996; Wei et al., 1995) published in prestigious journals showed that HIV replication and the disease process are very vibrant. On average, plasma virions have a mean lifespan of 0.3 days (half-life = 0.24 days), and the average total HIV-1 production is 10.3×10^9 per day, the minimum duration of the HIV-1 life cycle in vivo is 1.2 days, and the average HIV-1 generation time is 2.6 days (generation time is defined as the time from release of a virion until it infects another cell and causes the release of a new generation of viral particles.) Because the high viral replication rate may result in a high mutation rate, Ho (1995) proposed the treatment strategy of "Hit Hard, Hit Early". "Hit Hard" requires simultaneously combining different medications in the treatment, while "Hit Early" means the treatment should start as early as HIV infection has been confirmed. Although the so called "cocktail" treatment approach proposed by Ho is still the most commonly used treatment strategy, "Hit Early" was abandoned quickly when clinicians realized the adverse effects outweighed the benefits. The treatment should be "Hit HIV-1 hard, but only when necessary" (Harrington et al., 2000). Based on 2012 U.S. Department of Health and Human Services Panel on Antiretroviral Guidelines for Adults and Adolescents (Guidelines, 2012), the initiation of antiretroviral therapy (ART) is optional if the CD4 cell count is $> 500\ /mL$, moderately recommended if the CD4 cell count is 350 to 500 $/mL$ and strongly recommended if the value is $< 350\ /mL$. Regardless of the CD4 cell count, ART is strongly recommended if patients have certain conditions such as pregnancy, history of an AIDS defining illness or hepatitis B (HBV) co-infection. The usual highly active antiretroviral therapy (HAART) combines three or more different medications such as two nucleoside reverse transcriptase inhibitors (NRTIs) and a protease inhibitor (PI), a non-nucleoside reverse transcriptase inhibitor (NNRTI) or other such combinations. These HAART regimens have been proven to be able to reduce the amount of active viruses and in some cases can lower the number of active viruses until it is undetectable by current blood testing techniques.

1.2. HIV dynamic models

The basic model for HIV infection includes three parts: target uninfected cell *T*, virus *V* and infected cell T^*. The equations that describe the basic model of viral dynamics before the treatment are:

$$\begin{aligned} \frac{dT}{dt} &= \rho - dT - kVT \\ \frac{dT^*}{dt} &= kVT - \delta T^* \\ \frac{dV}{dt} &= \eta T^* - cV \end{aligned} \tag{1.1}$$

where *T* is produced at a rate of ρ and dies at rate d, virus *V* is cleared from the body at rate c and infects the target cells *T* to T^* at rate of k, infected cell T^* dies at rate δ and produces new virus particles at a constant rate η. This is a system of nonlinear ordinary differential equations (ODE) without any closed form solution, however, we can derive various approximations and obtain an understanding of the system.

Before infection, $V = 0$, $T^* = 0$ and uninfected cells *T* are at equilibrium as $T = \rho/d$. Denote by $t = 0$ is the time when infection occurs. Suppose infection occurs with a certain amount of virus, so the initial conditions are $T_0 = \rho/d$, $T_0^* = 0$ and V_0. Similar as the condition that spread of an infectious disease in a population, whether or not the virus can grow and establish an infection depends on a crucial quantity called basic reproductive ratio R. R is defined as the number of newly infected cells arising from one infected cell when almost all cells are uninfected and $R = \frac{\rho k \eta}{d \delta c}$. If $R < 1$ then the virus will not spread, because every infected cell will on average produce less than one other infected cell. If starting with N infected cells, then on average, we expect roughly $\ln N / \ln(1 - R)$ rounds of replications before the virus population dies out. If on the other hand, $R > 1$, then, on average, every infected cell will produce more than one newly infected cell. The chain will generate an explosive multiplication of virus as $V(t) = V_0 \exp(rt)$, where r is the exponential growth rate of the virus population and it is given by the larger root of the equation r^2 +(δ + c)R +$\delta c(1 - r^2) = 0$, the approximation of $r = \delta(R - 1)$, which means each infected cell produces R newly infected cells before dying. Virus growth will not continue indefinitely because the supply of uninfected cells is limited.

During the short time since initiation of HAART treatment, the viral load decrease sharply. This change with time can be expressed by the differential equation as, $dV/dt = P - \lambda V$, where P is the viral production rate, λ is the decay rate of viral load, and V is the HIV viral load in plasma. If assuming a pretreatment steady state exists, $dV/dt = 0$, and a perfect treatment effect that no new infection or new virion produced, the HIV dynamics can be expressed as a simple one-exponential equation (Ho et al., 1995):

$$V(t) = V(0) \exp(-\lambda t) \tag{1.2}$$

where $V(t)$ is the viral load at time t and $V(0)$ is the viral load at the baseline. Equation (1.2) can only reasonably describe the behavior of the viral dynamics during 1–2 weeks after the initialization of treatment.

Assuming a perfect protease inhibitor treatment effect (Perelson et al., 1996), which means no new infectious virions (V_I) but some noninfectious virions (V_{NI}) will still be produced, the HIV dynamics can be expressed as the following system of ODE:

$$\begin{array}{rcl} \frac{dT^*}{dt} & = & kV_IT - \delta T^* \\ \frac{dV_I}{dt} & = & -cV_I \\ \frac{dV_{NI}}{dt} & = & N\delta T^* - cV_{NI} \end{array} \tag{1.3}$$

where N is the number of new virions produced per infected cell during its life time. Under the assumption of constant supply of target cell T and quasi-steady state before treatment ($dT^*/dt = 0$ and $dV/dt = 0$), a close form solution to the system of ODE (1.3) can be obtained:

$$V(t) \quad = \quad V_0 \exp(-\lambda t) + \tfrac{\lambda V_0}{\lambda - \delta} \times [\tfrac{\lambda V_0}{\lambda - \delta}\{\exp(-\delta t) - \exp(-\lambda t)\} - \delta t \exp(-\lambda t)] \tag{1.4}$$

where $V(t) = V_I(t) + V_{NI}(t)$, Perelson et al.(1996) applied equation (1.4) to more frequent measured HIV-1 RNA data during the first week of treatment. By nonlinear least-squares regression, the estimated half-life of free virions is about six hours and it is 1.6 days for productively infected cells.

Perelson et al.(1997) further extended the ODE (1.3) in order to include a longer period of treatment that a biphasical decay rate of plasma HIV-1 RNA was observed: an initial rapid exponential decline of nearly 2-logs (first phase), followed by a slower exponential decline (second phase). Two more target cells are added in the model: (i) long-lived infected cells, macrophages (M), will be infected into M^* with a rate of k_M, produce virions at rate of p and die with a rate of μ_M; (ii) latently infected lymphocytes (L) will be produced by a rate constant fk and die at a rate of μ_L. The HIV dynamics can be expressed as:

$$\begin{array}{rcl} \frac{dT^*}{dt} & = & kVT + \alpha L - \delta T^* \\ \frac{dL}{dt} & = & fkVT - \mu_L L \\ \frac{dM^*}{dt} & = & k_M VM - \mu_M M^* \\ \frac{dV}{dt} & = & N\delta T^* + pM^* - cV \end{array} \tag{1.5}$$

where latent infected cells L can become productively infected cells at rate of α. With the similar assumptions used for equation (1.4), a closed form solution to the system of ODE (1.5) is,

$$V(t) \quad = \quad V_0[A\exp(-\delta t) + B\exp(-\mu_L t) + C\exp(-\mu_M t) + (1 + A + B + C)] \qquad (1.6)$$

where A, B and C are functions of system parameters. Even with additional peripheral blood mononuclear cells information, this equation is too complicated to identify all parameters, therefore, some parameters are assumed to be known and replaced by the values from previous studies. The first six weeks since the treatment was used in equation (1.6) and the half-life of productively infected CD4 cells, long-lived infected cells and latently infected cells were estimated as 1.1 days, 14.1 days and 8.5 days, respectively.

Perfect treatment effect may not be a very reasonable assumption, especially after short period of treatment. Wu and Ding (1999) proposed a system of ODE that included a protease inhibitor efficacy parameter of γ, $0 \leq \gamma \leq 1$, while $\gamma = 0$ means the PI medications have no effect and $\gamma = 1$ means perfect effect. The original ODE they proposed included many parameters that either can be negligible if they are associated with the faster decays or can be approximated by constants if they are slow enough in the modeling time period or if they are impossible to be accurately estimated based on the HIV-1 viral load available. The simplified system of ODE Wu and Ding proposed is:

$$\begin{aligned} \tfrac{d}{dt}T^* &= kV_I T - \delta T^* \\ \tfrac{d}{dt}V_I &= (1-\gamma)P - cV_I \\ \tfrac{d}{dt}V_{NI} &= \gamma P + P^* + N\delta T^* - cV_{NI} \end{aligned} \qquad (1.7)$$

where P is the virus produced rate by productively infected cells, such as CD4 cell, P^* accounts for virus produced from "mysterious" infected cells such as Langerhans cells and microglial cells, or long-lived infected cells such as macrophages and latent infected cells, and $k, T^*, \delta, V_I, V_{NI}, N$ and c have the same meaning as ODE (1.3). A closed form solution to the system of ODE (1.7) is,

$$V(t) \quad = \quad \exp(P_1 - \lambda_1 t) + \exp(P_2 - \lambda_2 t) + (P_3 + P_4 t)\exp(-ct)$$

where $V(t) = V_I(t) + V_{NI}(t)$, $\lambda_1 = \delta$ and it is the first-phase viral decay rate that may represent the minimum turnover rate of productively infected cells, such as CD4, λ_2 is a possibly compound clearance rate of long-lived and latently infected cells and the value depends on the infection rate and destroyed rate by HIV virus. Because c has been estimated to be very rapid (less than 6 hours of half life), it can be negligible compared with other terms. Thus, the equation can be further simplified as a two-exponential equation:

$$V(t) \quad = \quad \exp(P_1 - \lambda_1 t) + \exp(P_2 - \lambda_2 t) \tag{1.8}$$

where P_1 and P_2 is initial viral production rate from productively infected cells, long-lived and latently infected cells, respectively. Nonlinear mixed-effects (NLME) modeling can be used in the estimation of the parameters in equation (1.8). NLME modeling will pool individual data together to estimate the population parameters first, then estimate the individual parameters by the empirical Bayesian method (Vonesh and Chinchili, 1996).

Although the "cocktail" HAART treatment can suppress HIV in 60 to 90% of cases, 30 to 60% of patients will end up as being considered treatment failure eventually because of the viral load rebound (Havlir et al., 2000). However, all of the equations introduced so far require the decay rate to be constant so they can't be applied to rebound values. Several extensions have been developed in order to catch up viral load response that include rebound data and three representatives are following:

(i) Extended from the ODE (1.1), Huang et al.(2003) proposed a viral dynamic model with a time varying treatment efficacy function $\gamma(t)$ as,

$$\begin{aligned} \frac{dT}{dt} &= \rho - dT - [1 - \gamma(t)]kTV \\ \frac{dT^*}{dt} &= [1 - \gamma(t)]kTV - \delta T^* \\ \frac{dV}{dt} &= N\delta T^* - cV \end{aligned}$$

where $\gamma(t)$ represents a time varying treatment efficacy and it can be modeled as a function of drug exposure and drug sensitivity.

(ii) Extended from one exponential equation (1.2) by replacing the constant decay rate with a time varying decay function (Wu, 2004):

$$V(t) = V(0)\exp(-\lambda(t)t)$$

(iii) Extended from two exponential (1.8) by replacing the second constant decay rate with a time varying decay rate function as (Wu and Zhang, 2002):

$$V(t) \quad = \quad \exp(P_1 - \lambda_1 t) + \exp(P_2 - \lambda_2(t)t)$$

Among these three extensions, the first one is a system of nonlinear ODE without a closed form, so compared with the other two, the computation is even more challengeable and the model may not converge, therefore, we will focus on either the one exponential or two exponential equation in the Chapter 2 and 3.

HIV progress status is usually measured via HIV viral load or CD4 cell count, which are both surrogate biomarkers. CD4 cell count is more often used as an endpoint for long follow-up trials or advanced patients population, but for trials with short follow-up periods, viral load is often used as a primary endpoint to quantify treatment effect, where CD4 cell count is viewed as a covariate to help predict virologic responses. However, we should be aware the possible issues of using either HIV viral load or CD4 cell count as the outcome. The possible troublesome aspects of using the viral load as the primary outcome include (i) if the viral load is measured by RT-PCR which is based on the viral fragments, the result may overestimate the number of infectious virus by an average factor of 60,000 (Nowak et al., 1991); the lack correlation between of viral load and infection was also noted in some publications (Perelson et al., 1993; 1999), where no evidence of virus by culture among the patients with detectable viral load; (ii) the lack of correlation between viral load and CD4 level such that the changes in viral load were only able to explain as little as 4% of change in the CD4 cell count (Rodriguez et al., 2006). Although CD4 cell count seems to be a better HIV progression indicator, especially for the study with a longer follow-up period, prediction may be risky since CD4 cell count models are often empirical (Wu and Ding, 1999; Wu, 2002). On the other hand, treating both viral load and CD4 cell count as a bivariate response (Sy et al., 2007) may be complicated, because the HIV dynamic model for viral load is nonlinear and CD4 cell count contains missing data.

1.3. Statistical inference in HIV dynamics

Various statistical inferences and analysis methods have been applied in HIV dynamics. Linear and nonlinear regression via least-squares (LS) estimation can be applied to very frequent measurements during the first 1 – 2 weeks after the treatment is initiated (Ho et al., 1995; Perelson et al., 1996; 1997; Wei et al., 1995). Because frequent viral load measurement is only achievable in small clinical studies and only subjects without any missing values can be included in LS, this method is considered to be less powerful than some other inferences.

Because viral load are measured repeatedly since the treatment, the values obtained from the same subject may be correlated but can be assumed to be independent if obtained from different subjects. One powerful tool to handle such longitudinal data is mixed-effects modeling, in which within-subject and between-subject variations are both considered (Laird and Ware, 1982). Linear mixed-effects (LME) and nonlinear mixed-effects (NLME) modeling approaches have been proposed in HIV dynamics (Wu et al., 1998; 2004; Wu and Ding, 1999). Semiparametric nonlinear mixed-effects (SNLME) modeling (Liu and Wu, 2007; Wu and Zhang, 2002; Wu et al., 2004) is proposed in order to allow the decay rate to vary with time so the rebound viral load can be included. Joint model approach via Monte Carlo EM algorithm can be applied to the NLME with covariate measurement errors and non-ignorable missing responses (Liu and Wu, 2007; Wu, 2002; 2004). Estimation of NLME is complex because usually the likelihood has no closed form solution, even for simple models. The Bayesian approach based on Markov chain Monte Carlo (MCMC) algorithm has been proposed for complex ODE and NLME (Huang et al., 2006; Huang and Dagne, 2011; 2012a; 2012b; Putter et al., 2002; Wu et al., 2005). To avoid the numerical computation of multiple integrals involved in the likelihood, likelihood approximation such as linearization, Laplace approximation, Stochastic approximation EM algorithm (SAEM) have been applied in HIV dynamics (Ding and Wu, 2000; Guedj et al., 2007; Kuhn and Lavielle, 2005; Wu, 2004).

Another complexity of viral load analysis is left censoring which occurs when viral loads are below a limit of qualification (LOQ), and if ignored, the censoring may induce biased parameter estimates. Different approaches have been proposed to address this problem (Fitzgerald et al., 2002; Hughes, 1999; Lavielle et al., 2011; Samson et al., 2006; Thiébaue et al., 2005).

The model random errors and random-effects in mixed-effect models are usually assumed to have a normal distribution and that assumption may not be satisfied in HIV viral load and CD4 cell count, so the estimation can be biased. Skewed distribution can be applied in order to consider this non-ignorable departure from normality (Huang et al., 2006; Huang and Dagne, 2012a; 2012b; Dagne and Huang, 2012).

CD4 and CD8 cell count can be used as surrogate biomarkers for HIV disease process. Shah et al.(1997) used an EM algorithm to fit a bivariate linear random-effects model. Sy et al.(1997) used the Fisher scoring method to fit a bivariate linear random-effects model including an integrated Orstein-Uhlenbeck process (IOU). IOU is a stochastic process that includes Brownian motion as a

special limiting case.

1.4. Skew-elliptical distributions

Linear and nonlinear mixed-effect models are powerful tools for analyzing repeated measures and clustered data. In these models, random-effects are included in order to account correlation. Usually either random-effects or model errors or both are assumed to follow a normal distribution. Although normality assumption may be reasonable for many situations, the skewness can still be obvious even after the variables have been transformed. Ignoring the departure from normality may cause biases or misleading results (Ghosh et al., 2007; Verbeke and Lesaffre, 1996). Ideally, we hope to use a more generalized distribution family that (i) has high flexibility in shapes and with a wide range of skewness and kurtosis; (ii) is mathematically tractable, which means it can retain nice properties of original family such that parameters can be directly linked to some aspects of known probability density function (pdf); (iii) allows us to easily apply the distributions in the existing software.

Skew-elliptical (SE) distribution is a parametric class of probability distributions that is extended from elliptical distribution by including an additional shape parameter for skewness. This class, which is usually obtained by using transformation and conditioning, contains many standard families such as multivariate skew-normal (SN), skew-*t* (ST), Student-*t* and normal distributions. Different versions of the multivariate SE distributions have been proposed. The version proposed by Azzalini et al.(1996; 1999) is based on conditioning one suitable random variable being greater than zero; SE distribution proposed by Jones and Faddy (2003) is scaled inverse $\boldsymbol{\chi}$ distribution; Fernandez and Steel (1998) developed a form that two Student-*t* distributions (with different scale parameters) in positive and negative domains are combined to form an SE distributions; We adopt a class of multivariate SE distributions proposed by Sahu et al.(2003), which is obtained by using transformation and conditioning, contains multivariate ST , SN, Student-*t* and normal distribution as special cases. A k-dimensional random vector $\boldsymbol{Y}$ follows a k-variate SE distribution if its pdf is given by

$$f(\boldsymbol{y}|\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta}; m_\nu^{(k)}) = 2^k f(\boldsymbol{y}|\boldsymbol{\mu}, \boldsymbol{A}; m_\nu^{(k)}) P(\boldsymbol{V} > \boldsymbol{0}) \tag{1.9}$$

where $\boldsymbol{A} = \boldsymbol{\Sigma} + \boldsymbol{\Delta}^2$, $\boldsymbol{\mu}$ is a location parameter vector, $\boldsymbol{\Sigma}$ is a covariance matrix, $\boldsymbol{\Delta}$ is a skewness diagonal matrix with the skewness parameter vector $\boldsymbol{\delta} = (\delta_1, \delta_2, \ldots, \delta_k)^T$, $\boldsymbol{V}$ follows the el-

liptical distribution $El\left(\boldsymbol{\Delta}A^{-1}(\boldsymbol{y}-\boldsymbol{\mu}), \boldsymbol{I}_k - \boldsymbol{\Delta}A^{-1}\boldsymbol{\Delta}; m_\nu^{(k)}\right)$ and the density generator function $m_\nu^{(k)}(u) = \frac{\Gamma(k/2)}{\pi^{k/2}}\frac{m_\nu(u)}{\int_0^\infty r^{k/2-1}m_\nu(u)dr}$, with $m_\nu(u)$ being a function such that $\int_0^\infty r^{k/2-1}m_\nu(u)dr$ exists. The function $m_\nu(u)$ provides the kernel of the original elliptical density and may depend on the parameter ν. We denote this SE distribution by $SE(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta}; m_\nu^{(k)})$. Two examples of $m_\nu(u)$, leading to important special cases used throughout the paper, are $m_\nu(u) = \exp(-u/2)$ and $m_\nu(u) = (u/\nu)^{-(\nu+k)/2}$, where $\nu > 0$. These two expressions lead to the multivariate SN and ST distributions, respectively. In the latter case, ν corresponds to the degree of freedom parameter.

1.4.1. Skew-t distribution

We briefly discuss a multivariate ST distribution introduced by Sahu et al.(2003) in this section. A k-dimensional random vector $\boldsymbol{Y}$ follows a k-variate ST distribution if its pdf is given by

$$f(\boldsymbol{y}|\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta}, \nu) = 2^k t_{k,\nu}(\boldsymbol{y}|\boldsymbol{\mu}, \boldsymbol{A})P(\boldsymbol{V} > \boldsymbol{0}) \tag{1.10}$$

we denote the k-variate t distribution with parameters $\boldsymbol{\mu}$, $\boldsymbol{A}$ and degrees of freedom ν by $t_{k,\nu}(\boldsymbol{\mu}, \boldsymbol{A})$ and the corresponding pdf by $t_{k,\nu}(\boldsymbol{y}|\boldsymbol{\mu}, \boldsymbol{A})$ henceforth, $\boldsymbol{V}$ follows the t distribution $t_{k,\nu+k}$. We denote this distribution by $ST_{k,\nu}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta})$. In particular, when $\boldsymbol{\Sigma} = \sigma^2\boldsymbol{I}_k$ and $\boldsymbol{\Delta} = \delta\boldsymbol{I}_k$, equation (1.10) simplifies to

$$f(\boldsymbol{y}|\boldsymbol{\mu}, \sigma^2, \delta, \nu) = 2^k(\sigma^2+\delta^2)^{-k/2}\frac{\Gamma((\nu+k)/2)}{\Gamma(\nu/2)(\nu\pi)^{k/2}}\left\{1 + \frac{(\boldsymbol{y}-\boldsymbol{\mu})^T(\boldsymbol{y}-\boldsymbol{\mu})}{\nu(\sigma^2+\delta^2)}\right\}^{-(\nu+k)/2}$$
$$\times T_{k,\nu+k}\left[\left\{\frac{\nu+(\sigma^2+\delta^2)^{-1}(\boldsymbol{y}-\boldsymbol{\mu})^T(\boldsymbol{y}-\boldsymbol{\mu})}{\nu+k}\right\}^{-1/2}\frac{\delta(\boldsymbol{y}-\boldsymbol{\mu})}{\sigma\sqrt{\sigma^2+\delta^2}}\right]$$

where $T_{k,\nu+k}(\cdot)$ denotes the cumulative distribution function (cdf) of $t_{k,\nu+k}(\boldsymbol{0}, \boldsymbol{I}_k)$. However, unlike in the SN distribution below, the ST density can not be written as the product of univariate ST densities. Here $\boldsymbol{Y}$ are dependent but uncorrelated.

The mean and covariance matrix of the ST distribution $ST_{k,\nu}(\boldsymbol{\mu}, \sigma^2\boldsymbol{I}_k, \boldsymbol{\Delta})$ are given by

$$E(\boldsymbol{Y}) = \boldsymbol{\mu} + (\nu/\pi)^{1/2}\frac{\Gamma((\nu-1)/2)}{\Gamma(\nu/2)}\boldsymbol{\delta}, \qquad cov(\boldsymbol{Y}) = \left[\sigma^2\boldsymbol{I}_k + \boldsymbol{\Delta}^2\right]\frac{\nu}{\nu-2} - \frac{\nu}{\pi}\left[\frac{\Gamma\{(\nu-1)/2\}}{\Gamma(\nu/2)}\right]^2\boldsymbol{\Delta}^2$$

The ST distribution of $\boldsymbol{Y}$ has two types of stochastic representation as follows, and each provides a convenience device for random number generation and implementation purpose.

(i). By the proposition of Sahu et al.(2003),

$$\boldsymbol{Y} = \boldsymbol{\mu} + \boldsymbol{\Delta}|\boldsymbol{X}_0| + \boldsymbol{\Sigma}^{1/2}\boldsymbol{X}_1 \tag{1.11}$$

where $\boldsymbol{X}_0$ and $\boldsymbol{X}_1$ are two independent random vectors following $t_{k,\nu}(\boldsymbol{0}, \boldsymbol{I}_k)$. Let $\boldsymbol{w} = |\boldsymbol{X}_0|$, then $\boldsymbol{w}$ follows a k-dimensional standard t distribution $t_{k,\nu}(\boldsymbol{0}, \boldsymbol{I}_k)$ truncated in the space $\boldsymbol{w} > \boldsymbol{0}$ (i.e., the standard half-t distribution). Thus, a hierarchical representation of (1.11) is given by

$$\boldsymbol{Y}|\boldsymbol{w} \sim t_{k,\nu+k}(\boldsymbol{\mu}+\boldsymbol{\Delta}\boldsymbol{w}, \omega\boldsymbol{\Sigma}), \quad \boldsymbol{w} \sim t_{k,\nu}(\boldsymbol{0}, \boldsymbol{I}_k)\boldsymbol{I}(\boldsymbol{w} > \boldsymbol{0}) \tag{1.12}$$

where $\omega = (\nu + \boldsymbol{w}^T\boldsymbol{w})/(\nu + k)$.

(ii) By Proposition 1 of Arellano-Valle et al.(2007), the ST of $\boldsymbol{Y}$ has another convenient stochastic representation as follows

$$\boldsymbol{Y} = \boldsymbol{\mu} + \boldsymbol{\Delta}|\boldsymbol{X}_0| + \xi^{-1/2}\boldsymbol{\Sigma}^{1/2}\boldsymbol{X}_1 \tag{1.13}$$

where $|\boldsymbol{X}_0|$ and $\boldsymbol{X}_1$ are two independent $N_k(\boldsymbol{0}, \boldsymbol{I}_k)$ random vectors. Let $\boldsymbol{w} = |\boldsymbol{X}_0|$, then $\boldsymbol{w}$ follows a k-dimensional standard normal distribution $N_k(\boldsymbol{0}, \boldsymbol{I}_k)$ truncated in the space $\boldsymbol{w} > \boldsymbol{0}$. Thus, following Sahu et al.(2003), a hierarchical representation of 1.13 is given by

$$\boldsymbol{Y}|\boldsymbol{w}, \xi \sim N_k(\boldsymbol{\mu}+\boldsymbol{\Delta}\boldsymbol{w}, \xi^{-1}\boldsymbol{\Sigma}), \quad \boldsymbol{w} \sim N_k(\boldsymbol{0}, \boldsymbol{I}_k)\boldsymbol{I}(\boldsymbol{w} > \boldsymbol{0}), \quad \xi \sim \boldsymbol{\Gamma}(\rho/\boldsymbol{2}, \rho/\boldsymbol{2}) \tag{1.14}$$

Note that the ST distribution presented in (1.12) or (1.14) can be reduced to the following three special distributions:

(a). An SN distribution $SN_k(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta})$ as $\nu \to \infty$ and $\xi \to 1$ with probability of 1 (based on equation of 1.14) or as $\nu \to \infty$ with probability of 1 (based on equation of 1.12);

(b). A Student-t distribution $t_{k,\nu}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ as $\boldsymbol{\Delta} = \boldsymbol{0}$;

(c). A normal distribution $N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ if both conditions of (a) and (b) are satisfied.

In order to better understand the shape of an ST distribution, plots of an ST density as a function of the skewness parameter with $\delta = -3, 0, 3$ are shown in Figure 1.4(a).

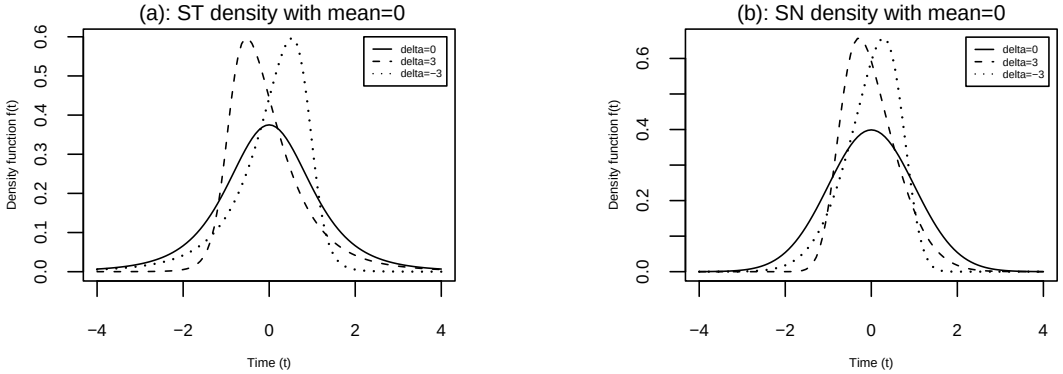

Figure 1.4: The univariate skew-t (df $\nu = 4$) and skew-normal density functions with precision $\sigma^2 = 1$ and skewness parameter δ = 0, -3 and 3, respectively.

1.4.2. Skew-normal distribution

We briefly discuss a multivariate SN distribution introduced by Sahu et al.(2003) in this section. A k-dimensional random vector $\boldsymbol{Y}$ follows a k-variate SN distribution, if its pdf is given by

$$f(\boldsymbol{y}|\boldsymbol{\mu},\boldsymbol{\Sigma},\boldsymbol{\Delta}) = 2^k|\boldsymbol{A}|^{-1/2}\phi_k\{\boldsymbol{A}^{-1/2}(\boldsymbol{y}-\boldsymbol{\mu})\}P(\boldsymbol{V}>\boldsymbol{0}), \tag{1.15}$$

where $\boldsymbol{V} \sim N_k\{\boldsymbol{\Delta}\boldsymbol{A}^{-1}(\boldsymbol{y}-\boldsymbol{\mu}), \boldsymbol{I}_k - \boldsymbol{\Delta}\boldsymbol{A}^{-1}\boldsymbol{\Delta}\}$, and $\phi_k(\cdot)$ is the pdf of $N_k(\boldsymbol{0},\boldsymbol{I}_k)$. We denote the above distribution by $SN_k(\boldsymbol{\mu},\boldsymbol{\Sigma},\boldsymbol{\Delta})$. An appealing feature of equation (1.15) is that it gives independent marginal when $\boldsymbol{\Sigma} = diag(\sigma_1^2,\sigma_2^2,\ldots,\sigma_k^2)$. The pdf (1.15) thus reduces to

$$f(\boldsymbol{Y}|\boldsymbol{\mu},\boldsymbol{\Sigma},\boldsymbol{\Delta}) = \prod_{i=1}^k \left[\frac{2}{\sqrt{\sigma_i^2+\delta_i^2}}\phi\left\{\frac{y_i-\mu_i}{\sqrt{\sigma_i^2+\delta_i^2}}\right\}\boldsymbol{\Phi}\left\{\frac{\delta_i}{\sigma_i}\frac{y_i-\mu_i}{\sqrt{\sigma_i^2+\delta_i^2}}\right\}\right],$$

where $\phi(\cdot)$ and $\boldsymbol{\Phi}(\cdot)$ are the pdf and cdf of the standard normal distribution, respectively.

The mean and covariance matrix are given by

$$E(\boldsymbol{Y}) = \boldsymbol{\mu} + \sqrt{2/\pi}\boldsymbol{\delta}, \qquad cov(\boldsymbol{Y}) = \boldsymbol{\Sigma} + (1-2/\pi)\boldsymbol{\Delta}^2$$

It is noted that when $\boldsymbol{\Delta} = \boldsymbol{0}$, the SN distribution reduces to usual normal distribution. In addition, the SN distribution is a special case of the ST distribution. That is, the ST distribution reduces to the SN distribution when the degree of freedom is large. In order to better understand the shape of an SN distribution, plots of an SN density as a function of the skewness parameter with $\delta = -3, 0$, and 3 are shown in Figure 1.4(b).

1.5. Specific aims

A common assumption in mixed-effect model for random errors and random-effects is normal distribution. This assumption may lack robustness against departure from normality and can be greatly affected by outliers too, therefore, the results may be somewhat misleading. In HIV/AIDS studies, the viral load, CD4 and CD8 cell count can exhibit obvious skewness, even after transformation. It will be valuable to explore whether a general skewed distribution such as ST or SN will bring a better model fitting. Also due to the nature of HIV dynamics, the related models can be very complicated and associated intensive computation burden in the inference. Non-convergence of the algorithms may exist under the framework of likelihood estimation. Besides these issues, there are at least three specific questions that have not been satisfactorily answered:

First, it is important to use entire HIV viral load data to have a better understand about the disease progress and to compare the effect of different medications. However, among all of those

models that can be applied to include the rebound data, it is unclear which one is preferred, or whether different distributions will affect the model fit, or whether the estimated parameters can be good predictors for some long-term result as as treatment failure.

Second, in order to explain individual difference in HIV dynamics, covariates, such as CD4, are often used in the model. However, CD4 values may be measured with substantial errors or at a different schedule as the viral load measurement. Also, LME, NLME and SNLME can be used for short, middle and long term of HIV dynamics data, respectively. Although they have some of the same parameters such as the first decay rate which is the minimal turn over of the productively infected cells, it is unclear whether this estimation obtained from different models is constant, and if not, which model will yield more reasonable estimations.

Third, using HIV viral load as a surrogate to predict the disease progress might be problematic. For example, the amount of infectious virus may be overestimated, therefore, the CD4 cell count seems to be a better indicator. However, the mechanism by which the CD4 cell count change during the HIV progress is not clear. Although using bivariate outcomes of CD4 and CD8 cell count appear to be superior to any of these cell count alone or their ratio (Ir et al., 1990), the distribution of CD4 and CD8 cell count shows skewness with heavy tails, and no model has been proposed to consider CD4 and CD8 as outcome simultaneously with skewed distribution assumption.

Via the Bayesian approach and assuming an SE distribution, this dissertation research is organized as follow:

Aim 1. Related to the first question of multiple models for entire HIV viral load follow-up, in Chapter 2, we explored different models with time-varying decay rate function in order to find which one has the best fit. We also assumed different distributions in each model to check the effect of skewness on the model fit. After finding the best fitted model, we explored the applications of the estimated decay rate, such as their association with decay rate, CD4 cell count and viral load rebound status. To the best of our knowledge, no time-varying decay rate function was checked or had been found to have any significant association with the long-term outcome such as viral load rebound, although some research found the constant decay rate may reflect

the potency of antiviral therapies in the short term. For the purpose of model comparisons, we used one AIDS clinical trial study data to do the model comparisons and then checked the validity of the conclusions based on another AIDS clinical trial study.

Aim 2. Related to the second question of covariate with measurement errors and skewness, in Chapter 3, we compared the three most commonly used models for short, middle and long term HIV dynamics. CD4 was included as an important covariate in the models. A critical question is whether these models produce coherent estimates of viral decay rates, and if not, which model is appropriate and should be used in practice. In addition, one common assumption is that model random errors is normally distributed, but the normality assumption may be unrealistic, particularly, if the data exhibit skewness. Moreover, some covariates, such as CD4 cell count, may often be measured with substantial errors. We addressed these issues simultaneously by jointly modeling the response variable with skewness and a covariate process with measurement errors. A real data set from an AIDS clinical trial study was used to present the proposed models.

Aim 3. Related to the third question of CD4 and CD8 as being biomarkers during ARV and considering their dependence on common predictors, in Chapter 4, we applied a joint bivariate linear mixed-effects (BLME) model that can include CD4 and CD8 cell count simultaneously as the outcomes, while the observed skewness in the data was considered by applying an SE distribution. The baseline viral load, patients' age, time-varying drug efficacy and the group of treatments were also included as covariates in the BLME.

2 Mixed-effects models with skewed distributions for time-varying viral decay rate in HIV dynamics

2.1. Introduction

Mathematical modeling is an important tool for understanding the evolution of HIV viral load (numbers of HIV-1 RNA copies in plasma) and interactions between HIV and its target cells. Most mathematical models developed prior to the mid-1990s were created for computer simulations and for interpreting declines in CD4 cell count after HIV infection. Due to the availability of HAART and methods of providing sensitive measurement of blood plasma HIV-1 RNA concentrations, it is possible to use viral load as a surrogate marker for the health status of HIV-infected individuals. The mathematical modelings of HIV dynamics on the cellular or molecular level are based on a similar principal used in large-scale epidemiological modeling.

Studies of viral dynamics have a common design, in which the viral load, targeted cells, pharmacokinetic and pharmacodynamic factors are repeatedly measured since treatment. The viral load trajectory is complex and has multiple phases of change (Ho et al., 1995; Maldarelli et al., 2007; Perelson et al., 1997; Wei et al., 1995). Data from A5055 (Acosta et al., 2004) (Figure 2.1) shows that: (i) within the first 2 weeks after the initial treatment, the viral load (transformed in natural log scale) dropped linearly and sharply, therefore, the change of viral load can be approximated by an exponential function; (ii) within the first 2–3 months but after the first 2 weeks, the relationship between the viral load and time was still linear but the slope became flatter, which indicates a slower decay rate; (iii) between the third to eighth month, the viral load either decreased more slowly, remained at a constant low level, or started to increase up to the level measured before treatment was initiated. The possible reasons for viral load rebound are development of resistance to the medications, and other clinical issues such as lack of adherence. There is no clear cutoff among the phases, not every subject will have all of these phases and the length of the phases may vary among individuals. Therefore, the associated decay rate in the models for the viral load trajectories is expected to vary over time and can be individually specific.

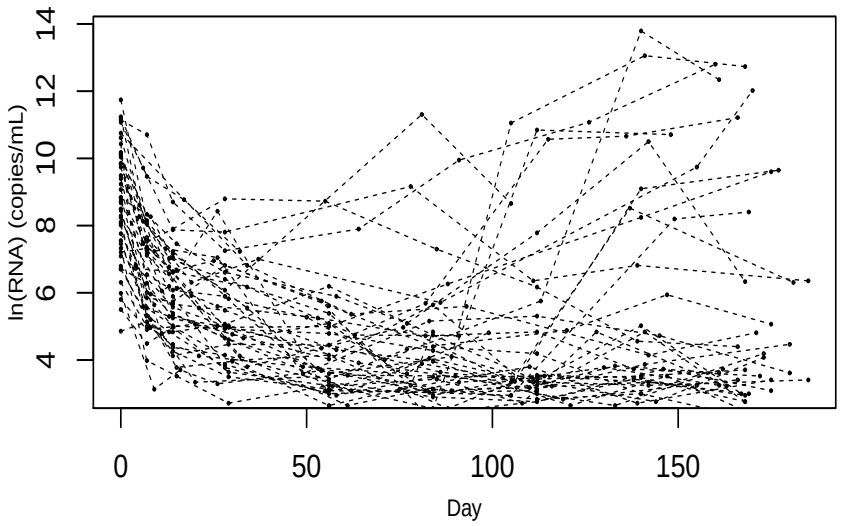

Figure 2.1: Profile of viral load in natural log scale from a clinical trial study-A5055

For the first phase of HIV viral load dynamics (i.e., the first 1 to 2 weeks), we can apply a uniexponential equation (Ho et al., 1995; Wei et al., 1995) as,

$$V(t) = V(0)\exp(-\lambda t) \tag{2.1}$$

where $V(t)$ is total viral load at time t, $V(0)$ is the baseline viral load at $t = 0$ and λ is a constant viral change rate which is the speed of the loss of viral load after initiation of potent antiviral treatment. Although equation (2.1) can precisely describe the phenomenon of a linear decrease of logarithm transformed viral load within approximately one to two weeks since treatment is initiated, we cannot apply it to the whole trajectory because the viral load is only allowed to decrease at a constant rate in this equation. Besides that, there are at least three unsolved issues.

First, in order to use entire HIV follow-up data, extended from equation (2.1), different models have been proposed in the literature, it is unclear which one is more appropriate.

Second, in mixed-effects models for longitudinal data analysis, random errors and/or random-effects are usually assumed to have a normal distribution. Although the normality assumption is satisfied in many situations, it may cause biased or misleading inference if the data include extreme values or show skewness with heavy tails, which are commonly seen in virological responses (Huang and Dagne, 2011; Sahu et al., 2003; Verbeke and Lesaffre, 1996). Figure 2.2 displays the histogram of repeated viral load in natural $\log$ scale for 44 subjects enrolled in the A5055 trial. The

skewness, which is still obvious even after the transformation, is positive and ranges from 0.8 and 2.15 at each of the follow-up measurements. If the ratio between skewness value and standard error of skewness is greater than 2, the data may be regarded as having unignorable skewness (Gardner, 2001). In the A5055 study, the ratio is 4 , which indicates skewness needs to be accounted for.

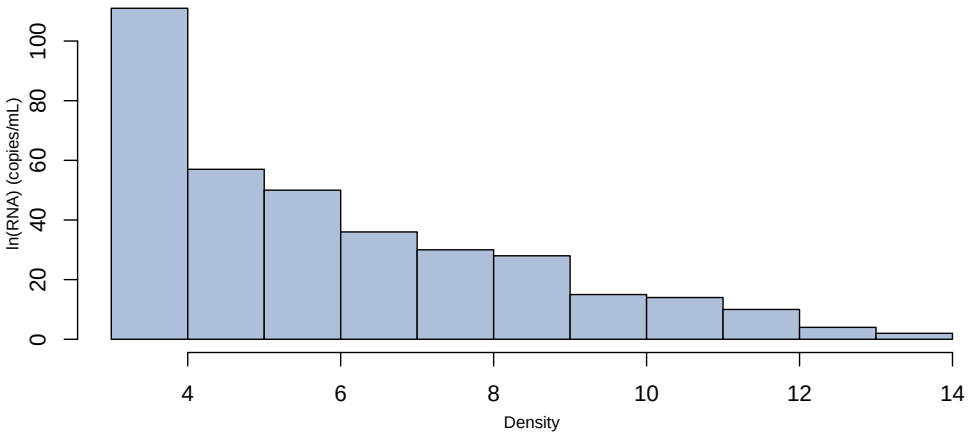

Figure 2.2: The histogram of viral load (in natural log scale) for 44 patients in a clinical trial study-A5055

Third, computational infeasibility can be a challenge. Frequentist and Bayesian are two major approaches used in studies of HIV dynamics. In the frequentist approach, based on the maximum likelihood estimation (MLE), different extensions have been proposed, such as Laplace approximation of the numerical integrals (Beal and Sheiner, 1982; Lindstrom and Bates, 1990; Wu and Zhang, 2002), stochastic approximation EM (SAEM) algorithm (Kuhn and Lavielle, 2005; Lavielle et al., 2011), joint model via Monte Carlo EM algorithm (Liu and Wu, 2007; Wu, 2004) and asymptotic distribution of the maximum h-likelihood estimators (MHLE) (Commenges et al., 2011). The second approach is Bayesian mixed-effects modeling via MCMC (Huang et al., 2006; Huang and Dagne, 2011; Putter et al., 2002). The Bayesian approach is an efficient way to incorporate prior information, both point estimates and uncertainties (variances), into analysis to identify more unknown parameters in complex models.

Via Bayesian approach, the main focus of this chapter is to provide a comprehensive comparison of five commonly used HIV dynamic models with SE distribution in random errors. The rest of the

chapter is organized as follows: Section 2.2 presents the HIV dynamic models that have a time-varying decay rate function so they can be applied to the entire HIV follow-up data. In Section 2.3, we describe a general Bayesian mixed-effects modeling approach. In Section 2.4, we present the motivated AIDS data and results of model comparisons. Section 2.5 includes the conclusion and discussion.

2.2. HIV dynamic models with time-varying decay rate function

As mentioned in Section 2.1, there is a multiphasic change in HIV viral load after the initiation of HAART. One potential interpretation of this phenomenon is that the process involves distinct populations with different homogenous behaviors. For example, the fast decreasing decay rate observed in the first phase is due to the treatment effect on productively infected CD4 cells, while the slower decay rate in the second phase is primarily due to the effect on the latently or long-lived infected cells (Perelson et al., 1997). However, some phenomena can't be explained by this theory. For example, there can be large differences in mean decay rates in response to different treatment regimens: during the first week, the death of infected cells may be substantially slower during days 3–6 than during days 2–3 (Grossman et al., 1999). If different decay rates reflect the rates at which different infected cells died, it is unexpected to see that the decay rate should depend on the type or concentration of the treatment regimen (Mueller et al., 1998). Based on the assumption that reduced production during immune activation events and fewer cycles account for the observed multiphasic HIV decrease, Following uniexponential equation (2.1), Grossman et al.(1999) proposed an equation for viral load as:

$$\begin{aligned} V(t) &= V(0)\exp\{-\tfrac{\log(R)}{\tau}t\} \\ R(v) &= R(1)+\alpha(1-v)\exp(-\rho v) \\ v(t) &= V(t)/V(0) \end{aligned} \tag{2.2}$$

where τ is the average infection cycle time, ρ is adjustable parameter, and R is reproduction ratio. At steady state, $R = 1$, which means each infected cell is replaced, on average, by one newly infected cell; $V(0)$ is the baseline viral load; $v(t)$ is the ratio between the viral load at time *t* and the baseline. Since $V(t)$ and R depend on each other, this equation may not be easy to solve. Based on Zhang and Wu (2011), equation (2.2) is equivalent to the equation $V(t) = V(0)\exp(-\lambda(t)t)$, and

$$\lambda(t) = \log\{R(v)\}/\tau = \log\{R[V(t)/V(0)]\}/\tau$$

$\lambda(t)$ is time varying and may be interpreted as the average relative loss rate of the viral load $V(t)$. Let $\triangle t$ denote a small time period, then the relative loss rate $\alpha(t)$ is:

$$\alpha(t) = \lim_{\triangle t \to \infty} \{\frac{V(t) - V(t+\triangle t)}{V(t)}\} / \triangle t = -V(t)' / V(t)$$

Solving the above differential equation yields:

$$V(t) = V(0) \exp\left\{ - \int_0^t \alpha(\tau) d\tau \right\} = V(0) \exp[-\lambda(t) t] \tag{2.3}$$

where $\lambda(t) = \int_0^t \alpha(\tau) d\tau / t$ is the average relative lost rate of the viral load $V(t)$. $\lambda(t)$ can be positive (if $R < 1$), zero (if $R = 1$), or negative (if $R > 1$). If $\lambda(t) < 0$, the decay rate $\lambda(t)$ at time *t* actually is a growth rate. Therefore, by including a time-varying decay rate function $\lambda(t)$, compared with equation (2.1), equation (2.3) is more flexible and can be applied to include the entire follow-up data without need to arbitrarily truncate the data.

A unified model with a time-varying viral decay rate function can be expressed as:

$$y(t) = \ln[V(t)] + \epsilon = \ln\{V(0) \exp[-\lambda(t) t]\} + \epsilon = \beta_1 - \lambda(t) t + \epsilon \tag{2.4}$$

where $y(t)$ is the natural logarithm transformation of the number of HIV-1 RNA copies per mL of plasma, ϵ is the measurement error, $\ln(V(0)) = \beta_1$ and is the macro-parameter for initial viral load in natural $\log$ scale. In addition to the simplicity of this model's structure, it also indicates that the pattern of HIV decrease may be a physiologically structured, local non-equilibrium dynamic interaction between HIV and immune activated cells after treatment initiation. Therefore, the overall decay rate is the weighted average that is proportional to the local level of infection.

Among the different decay rate functions proposed in the literature, we select five representatives as follow (Dagne and Huang, 2012; Grossman et al., 1999; Wu, 2004; Zhang and Wu, 2011),

$$\begin{aligned}
\text{I:} \quad \lambda(t) &= \beta_2 + \beta_3 t \\
\text{II:} \quad \lambda(t) &= \beta_2 \exp(-\beta_3 t) + \beta_4 \\
\text{III:} \quad \lambda(t) &= \beta_2 \exp(-\beta_3 t) + \beta_4 + \beta_5 t \\
\text{IV:} \quad \lambda(t) &= \beta_2 \exp(-\beta_3 t) + \beta_4 \exp(-\beta_5 t) \\
\text{V:} \quad \lambda(t) &= v[w(t), h_i(t)]
\end{aligned}$$

where the last one, $\lambda(t) = v[w(t), h_i(t)]$, is a nonparametric function.

2.3. Bayesian mixed-effects models with skewed distribution

To account for the skewness observed in the data, the random errors in mixed-effects models can be assumed to follow an SE distribution (see Section 1.4 in detail). The SE distribution is a family of distributions that is not only mathematically tractable but also flexible in its possible shapes. Because in the SE family, skew-normal (SN), normal and Student-*t* distribution are all a special case of skew-*t* (ST) distribution, therefore, in this section, we present a general form of a mixed-effects model with an ST distribution under the Bayesian approach. A general mixed-effects model with an ST distribution can be expressed as:

$$\begin{aligned} \boldsymbol{y}_i &= \boldsymbol{g}_i(\boldsymbol{t}_i, \boldsymbol{\beta}_i) + \boldsymbol{e}_i, \quad \boldsymbol{e}_i \overset{\text{iid}}{\sim} ST_{n_i,\nu}(0, \boldsymbol{\Sigma}, \boldsymbol{\Delta}), \\ \boldsymbol{\beta}_i &= \boldsymbol{d}(\boldsymbol{\beta}, \boldsymbol{b}_i), \qquad \boldsymbol{b}_i \overset{\text{iid}}{\sim} N(\boldsymbol{0}, \boldsymbol{\Sigma}_b), \end{aligned} \tag{2.5}$$

$\boldsymbol{y}_i = (y_{i1}, \ldots, y_{in_i})^T$ with y_{ij} being the response value for the ith individual at the jth time ($i = 1, 2, ..., n,\ \ j = 1, 2, ..., n_i$), $\boldsymbol{g}_i(\boldsymbol{t}_i, \boldsymbol{\beta}_i) = (g(t_{i1}, \boldsymbol{\beta}_{i1}), \ldots, g(t_{in_i}, \boldsymbol{\beta}_{in_i}))^T$, $\boldsymbol{t}_i = (t_{i1}, \ldots, t_{in_i})^T$, $\boldsymbol{\beta}_i = (\boldsymbol{\beta}_{i1}, \ldots, \boldsymbol{\beta}_{in_i})^T$, $\boldsymbol{\beta}_{ij}$ are individual-specific time-dependent parameter vectors and $\boldsymbol{\beta}$ is population parameter vector, $g(\cdot)$ and $d(\cdot)$ are linear or nonlinear known parametric functions, $\boldsymbol{b}_i$ is normal random-effect vector with $\boldsymbol{\Sigma}_b$ being an unstructured covariance matrix. The vector of random errors $\boldsymbol{e}_i = (e_{i1}, \ldots, e_{in_i})^T$ follows a multivariate ST distribution with degrees of freedom ν, within-subject covariance matrix $\boldsymbol{\Sigma}$ and we usually can assume $\boldsymbol{\Sigma} = \sigma^2 \boldsymbol{I}_{n_i}$, and unknown $n_i \times n_i$ skewness diagonal matrix such that $\boldsymbol{\Delta} = diag(\delta_{i1}, \ldots, \delta_{in_i})$, skewness parameter vector $\boldsymbol{\delta}_i = (\delta_{i1}, \ldots, \delta_{in_i})^T$. In particular, if $\delta_{i1} = \cdots = \delta_{in_i} \hat{=} \delta$, then $\boldsymbol{\Delta} = \delta \boldsymbol{I}_{n_i}$ and $\boldsymbol{\delta}_i = \delta \boldsymbol{1}_{n_i}$, where $\boldsymbol{1}_{n_i} = (1, \ldots, 1)^T$, indicating that we are interested in skewness of overall data set.

Following discussion in Section 1.4.1, to implement an MCMC procedure to model (2.5), by introducing one $n_i \times 1$ random vector $\boldsymbol{w}_i$, based on the stochastic representation, the model can be hierarchically formulated as follows.

$$\begin{aligned} \boldsymbol{y}_i | \boldsymbol{b}_i, \boldsymbol{w}_i &\overset{\text{iid}}{\sim} t_{n_i, n_i+\nu}(\boldsymbol{g}(\boldsymbol{t}_i, \boldsymbol{\beta}_i) + \delta \boldsymbol{w}_i, w_i \sigma^2 \boldsymbol{I}_{n_i}), \\ \boldsymbol{w}_i &\overset{\text{iid}}{\sim} t_{n_i, n_i+\nu}(\boldsymbol{0}, \boldsymbol{I}_{n_i}) I(\boldsymbol{w}_i > \boldsymbol{0}), \\ \boldsymbol{b}_i &\overset{\text{iid}}{\sim} N(\boldsymbol{0}, \boldsymbol{\Sigma}_b), \end{aligned} \tag{2.6}$$

where $w_i = (\nu + \boldsymbol{w}_i^T \boldsymbol{w}_i)/(\nu + n_i)$, $t_{n_i,\nu}(\boldsymbol{\mu}, \boldsymbol{A})$ denotes the n_i-variate Student-*t* distribution with parameters $\boldsymbol{\mu}$, $\boldsymbol{A}$ and degrees of freedom ν, $I(\boldsymbol{w} > \boldsymbol{0})$ is an indicator function and $\boldsymbol{w} = |\boldsymbol{X}_0|$ with $\boldsymbol{X}_0 \sim t_{n_i,\nu}(\boldsymbol{0}, \boldsymbol{I}_{n_i})$. Note that the hierarchical model above under Bayesian framework will allow

researchers to easily implement the methods using the freely available WinBUGS software (Lynn, et al., 2000) and the computational effort for the model with an ST distribution is almost equivalent to that of the model with a Student-*t* distribution.

The unknown population parameters in the model (2.5) are $\boldsymbol{\theta} = \{\boldsymbol{\beta}, \sigma^2, \boldsymbol{\Sigma}_b, \delta, \nu\}$, and we assume they are independent of one another. Under Bayesian framework, we also need to specify prior distributions for unknown parameters as follows.

$$\begin{aligned} \boldsymbol{\beta} &\sim N(\beta_0, \boldsymbol{\Lambda}), \quad \sigma^2 \sim IG(\omega_1, \omega_2), \quad \boldsymbol{\Sigma}_b \sim IW(\boldsymbol{\Omega}, v), \\ \delta &\sim N(\mathbf{0}, \gamma), \quad \nu \sim Exp(\nu_0) I(\nu > 2) \end{aligned} \tag{2.7}$$

where the mutually independent Normal (N), Inverse Gamma (IG), Exponential (Exp) and Inverse Wishart (IW) prior distributions are chosen to facilitate computations (Davidian and Giltinan, 1995). The super-parameter matrices $\boldsymbol{\Lambda}$ and $\boldsymbol{\Omega}$ can be assumed to be diagonal for convenient implementation.

Let $\pi(.)$ be a prior density function, so $\pi(\boldsymbol{\theta}) = \pi(\boldsymbol{\beta})\pi(\sigma^2)\pi(\boldsymbol{\Sigma}_b)\pi(\nu)\pi(\delta)$. Denote the observed data by $\mathcal{D} = \{\boldsymbol{y}_i, i = 1, ..., n\}$, and $f(\cdot|\cdot)$ as a conditional density function. Based on Bayesian inference, the posterior density of $\boldsymbol{\theta}$ is proportional to the observed data and prior distribution as:

$$f(\boldsymbol{\theta}|\mathcal{D}) \propto \{\prod_i^n \int f(\boldsymbol{y}_i|\boldsymbol{b}_i, \boldsymbol{w}_i; \boldsymbol{\beta}, \sigma^2, \nu, \delta)\, f(\boldsymbol{w}_i|\boldsymbol{w}_i > \mathbf{0})\, f(\boldsymbol{b}_i|\boldsymbol{\Sigma}_b) d\boldsymbol{b}_i\} \pi(\boldsymbol{\theta}) \tag{2.8}$$

In general, the integral in (2.8) is of high dimension and does not have any closed form. Analytic approximations to the integral may not be sufficiently accurate. Therefore, it is prohibitive to directly calculate the posterior distribution of $\boldsymbol{\theta}$ based on the observed data. As an alternative, MCMC procedures can be used to sample based on (2.8) by the Gibbs sampling along with the Metropolis-Hastings (M-H) algorithm.

2.4. Application: AIDS clinical trial data

2.4.1. AIDS clinical trial data and specific models

We used two AIDS clinical trials to explore the best fit among the models with different time-varying decay rate functions and different model random errors distribution assumption such as normal, SN, Student-*t* and ST distribution. The first trial, A5055, is the focus. Further, we used data from another clinical trial, A398 (Pfister et al., 2003), to validate the conclusions obtained from A5055.

A5055 was a phase I/II, randomized, open-label, 24-week comparative study. It included 44 HIV-1 infected patients who failed their first protease inhibitor treatment. Subjects were randomly assigned into one of the two arms. Subjects were scheduled for follow up visits at study day 0, weeks 1, 2, and 4, and every 4 weeks thereafter through week 24. RNA viral load was measured (copies$/mL$) in blood samples collected at study days 0, 7, 14, 28, 56, 84, 112, 140 and 168. The nucleic acid sequence-based amplification assay (NASBA) was used to measure plasma HIV-1 RNA, with a low limit of quantification of 50 copies$/mL$. HIV-1 RNA measures below this limit are not considered reliable, therefore we imputed such values as 25 copies$/mL$ (Acosta et al., 2004; Davidian and Giltinan, 1995). The mean, minimum and maximum values for the baseline viral load were 6.09 $\times$ 10^3, 199 and 1.07 $\times$ 10^5 $/mL$, respectively. The average age of subjects was 37.8 years (SD=8.1) and approximately 80% of subjects had at least 8 measurements (including the initial measurement). The mean and median number of days of follow-up were 155 and 168, respectively.

A398 was a phase II trial that included 481 HIV-1 positive patients with prior exposure to approved PIs and loss of virological suppression. All patients were assigned to receive routine ART. Besides these medications, depending on the dose and type of PIs to which the patients previously exposed, they were selectively randomly assigned into one of four arms. HIV-1 RNA levels were measured at the time of entry into the study (day 0), at study weeks 2, 4, 8, 16, 24, 32, 40, and 48, every 8 weeks thereafter, and at the time of confirmed virological failure. The mean, minimum and maximum values for the baseline viral load were 2.76 $\times$ 10^4, 260 and 1.32 $\times$ $10^7/mL$, respectively. The low limit of quantification is 100 copies$/mL$ and the HIV-1 RNA measures below this limit are not considered reliable and 50 copies$/mL$ was used instead. The average age of subjects was 40.1 years (SD=19) and approximately 74% subjects had at least 8 measurements (including the initial measurement). The mean and median of follow-up is 168 and 144 days, respectively. We draw two samples from A398 based on the method of simple random sampling without replacement, one sample includes 44 subjects and the other includes 100 subjects. We also used all of the 481 subjects in A398 in the model comparisons.

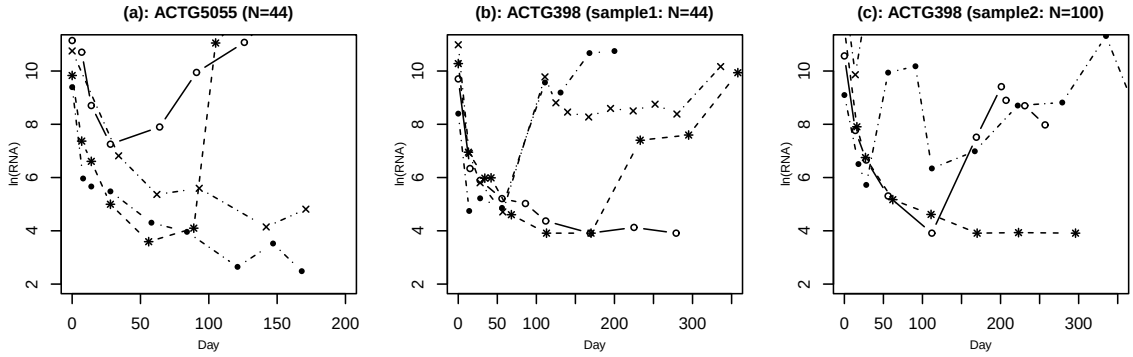

Figure 2.3: Profiles of viral load in natural log scale for four randomly selected patients among A5055 and A398, respectively

Figure 2.3 shows the measurements of viral load natural log scale for four randomly selected patients from A5055 and two sample data sets from A398. We can see that viral load trajectories vary widely and they are substantially different across individuals. To account for this time-varying viral load change, we applied a mixed-effects model with a time-varying decay rate function, as discussed in Section 2.2. In addition, we assumed the model errors followed an ST distribution in order to make the model flexible in considering the skewness observed in the data. The exact day of viral load measurement was used to compute study day in our analysis.

Under the general layout as model (2.4), corresponding to the five time-varying decay functions presented in Section 2.2, the mixed-effects models can be expressed as follow.

Model I: Quadratic linear mixed-effects model:

$$\begin{aligned} y_{ij} &= \beta_{1i} - [\beta_{2i} + \beta_{3i} t_{ij}] t_{ij} + e_{ij} \\ \boldsymbol{e}_i &\overset{\text{iid}}{\sim} ST_{n_i,v}(\mathbf{0}, \sigma^2 \boldsymbol{I}_{n_i}, \delta \boldsymbol{I}_{n_i}) \\ \beta_{1i} &= \beta_1 + b_{1i}, \;\; \beta_{2i} = \beta_2 + b_{2i}, \;\; \beta_{3i} = \beta_3 + b_{3i} \end{aligned} \tag{2.9}$$

where $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3)^T$ and $\boldsymbol{b}_i = (b_{1i}, b_{2i}, b_{3i})^T \overset{\text{iid}}{\sim} N_3(\mathbf{0}, \boldsymbol{\Sigma}_b)$.

Model II: Nonlinear mixed-effects model (uniexponential plus a constant):

$$\begin{aligned} y_{ij} &= \beta_{1i} - [\beta_{2i} \exp(-\beta_{3i} t_{ij}) + \beta_{4i}] t_{ij} + e_{ij} \\ \boldsymbol{e}_i &\overset{\text{iid}}{\sim} ST_{n_i,v}(\mathbf{0}, \sigma^2 \boldsymbol{I}_{n_i}, \delta \boldsymbol{I}_{n_i}) \\ \beta_{1i} &= \beta_1 + b_{1i}, \;\; \beta_{2i} = \beta_2 + b_{2i}, \;\; \beta_{3i} = \beta_3 + b_{3i}, \;\; \beta_{4i} = \beta_4 + b_{4i} \end{aligned} \tag{2.10}$$

where $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3, \beta_4)^T$ and $\boldsymbol{b}_i = (b_{1i}, b_{2i}, b_{3i}, b_{4i})^T \overset{\text{iid}}{\sim} N_4(\mathbf{0}, \boldsymbol{\Sigma}_b)$.

Model III: Nonlinear mixed-effects model (uniexponential plus a linear function):

$$\begin{aligned} y_{ij} &= \beta_{1i} - [\beta_{2i}\exp(-\beta_{3i}t_{ij}) + \beta_{4i} + \beta_{5i}t_{ij}]t_{ij} + e_{ij} \\ \boldsymbol{e}_i &\overset{\text{iid}}{\sim} ST_{n_i,v}(\mathbf{0}, \sigma^2\boldsymbol{I}_{n_i}, \delta\boldsymbol{I}_{n_i}) \\ \beta_{1i} &= \beta_1 + b_{1i},\ \beta_{2i} = \beta_2 + b_{2i},\ \beta_{3i} = \beta_3 + b_{3i},\ \beta_{4i} = \beta_4 + b_{4i},\ \beta_{5i} = \beta_5 + b_{5i} \end{aligned} \tag{2.11}$$

where $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3, \beta_4, \beta_5)^T$ and $\boldsymbol{b}_i = (b_{1i}, b_{2i}, b_{3i}, b_{4i}, b_{5i})^T \overset{\text{iid}}{\sim} N_5(\mathbf{0}, \boldsymbol{\Sigma}_b)$.

Model IV: Nonlinear mixed-effects model (two uniexponential):

$$\begin{aligned} y_{ij} &= \beta_{1i} - [\beta_{2i}\exp(-\beta_{3i}t_{ij}) + \beta_{4i}\exp(-\beta_{5i}t_{ij})]t_{ij} + e_{ij} \\ \boldsymbol{e}_i &\overset{\text{iid}}{\sim} ST_{n_i,v}(\mathbf{0}, \sigma^2\boldsymbol{I}_{n_i}, \delta\boldsymbol{I}_{n_i}) \\ \beta_{1i} &= \beta_1 + b_{1i},\ \beta_{2i} = \beta_2 + b_{2i},\ \beta_{3i} = \beta_3 + b_{3i},\ \beta_{4i} = \beta_4 + b_{4i},\ \beta_{5i} = \beta_5 + b_{5i} \end{aligned} \tag{2.12}$$

where $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3, \beta_4, \beta_5)^T$ and $\boldsymbol{b}_i = (b_{1i}, b_{2i}, b_{3i}, b_{4i}, b_{5i})^T \overset{\text{iid}}{\sim} N_5(\mathbf{0}, \boldsymbol{\Sigma}_b)$.

Model V: Semiparametric mixed-effects model:

$$y_{ij} = \beta_{1i} - v[w(t_{ij}), h_i(t_{ij})]t_{ij} + e_{ij}$$

where $w(t)$ and $h_i(t)$ are unknown nonparametric smooth fixed-effects and random-effects functions, respectively, and $h_i(t)$ are *iid* realizations of a zero-mean stochastic process. Model V is a semiparametric mixed-effects model if $w(t)$ and $h_i(t)$ are modeled non-parametrically such as splines or local polynomials. There are several ways to approximate these nonparametric functions. Following the similar approach as Shi et al.(1996), Rice and Wu (2001), Huang and Dagne (2010), we used natural cubic basis function instead of smoothing splines (Ke and Wang, 2001; Wang 1998; Zhang et al., 1998) or kernel methods (Wu and Zhang, 2002) for two reasons: this method is more straightforward in application and we can select the bases by Akaike information criterion (AIC) or the Bayesian information criterion (BIC) to balance the goodness-of-fit and model complexity. A linear combination of base function can be expressed as:

$$w(t) \approx w_p(t) = \textstyle\sum_{l=0}^{p-1} \mu_l\psi_l(t) = \boldsymbol{\mu}_p\boldsymbol{\Psi}_p(t)^T \quad h_i(t) \approx h_{iq}(t) = \textstyle\sum_{l=0}^{q-1} \xi_{il}\phi_l(t) = \boldsymbol{\xi}_{iq}\boldsymbol{\Phi}_q(t)^T$$

where $\boldsymbol{\mu}_p$ and $\boldsymbol{\xi}_{iq}$ ($q \leq p$) are the unknown vectors of fixed and random coefficients, respectively. We set $\psi_0 = \phi_0 \equiv 1$ and took the same natural cubic splines in the approximations with $p \leq q$, based on the AIC and BIC values, selected the following:

$$w(t_{ij}) + h_i(t_{ij}) \approx \mu_0 + \mu_1\psi_1(t_{ij}) + \mu_2\psi_2(t_{ij}) + \xi_{i0}$$

where $p = 3$ and $q = 1$. Model V, therefore, can be expressed as,

$$\begin{aligned} y_{ij} &= \beta_{1i} - [\mu_0 + \mu_1\psi_1(t_{ij}) + \mu_2\psi_2(t_{ij}) + \xi_{i0}]t_{ij} + e_{ij} \\ \boldsymbol{e}_i &\overset{\text{iid}}{\sim} ST_{n_i,v}(\mathbf{0}, \sigma^2\boldsymbol{I}_{n_i}, \delta\boldsymbol{I}_{n_i}) \\ \beta_{1i} &= \beta_1 + b_{1i}, \quad \beta_{2i} = \mu_0 + \xi_{i0}, \quad \beta_3 = \mu_1, \quad \beta_4 = \mu_2 \end{aligned} \tag{2.13}$$

where $\boldsymbol{\beta} = (\beta_1,\ \mu_0,\ \mu_1,\ \mu_2)^T$ and $\boldsymbol{b}_i = (b_{1i}, \xi_{i0})^T \overset{\text{iid}}{\sim} N_2(\mathbf{0}, \boldsymbol{\Sigma}_b)$.

In each of the five models above, besides the ST distribution assumption, the model random errors can also be assumed to follow other more specific distributions as normal, SN and Student-*t*. We used several criteria to check the model fit by applying the models on the data mentioned above.

We first used deviance information criterion (DIC) (Spiegelhalter, 2002) to compare models. DIC is a generalization of AIC that can be directly obtained from WinBUGS, it consists of two components:

$$\begin{aligned} DIC &= D(\bar{\theta}) + 2p_D \\ p_D &= E_{\theta|y}[-2\log p(y|\theta)] - (-2\log p(y|\hat{\theta}(y))) \\ &= E_{\theta|y}[D(\theta)] - D[E_{\theta|y}(\theta)] = \bar{D} - D(\bar{\theta}) \\ so,\ DIC &= \bar{D} + p_D \end{aligned}$$

where $D(\theta)$ is deviance and defined as $-2\log p(y|\theta)$ and $p(y|\theta)$ is likelihood, $\bar{D}$ is posterior mean deviance which measures "goodness-of-fit", the larger the value of $\bar{D}$, the worse of the fit; p_D is the effective number of parameters that indicates "complexity", the larger the value of p_D, the more complex is the model. Therefore, DIC = "goodness-of-fit" + "complexity". Since $\bar{D}$ will decrease as the number of parameters in a model increases, the p_D term compensates for this effect by favoring models with fewer parameters. Unlike AIC and BIC that require calculating the likelihood at its maximum over θ, which is not readily available from MCMC, DIC is easily calculated from the samples generated by MCMC. Same as AIC and BIC, the smaller the value of DIC, the better of the model fit. DIC is not intended for identification of the 'correct' model, but rather merely as a way to compare a collection of alternative formulations.

Because model comparisons are critical for our study, besides DIC, we also compared the values of expected predictive deviance (EPD) and residual sum of squares (RSS) that obtained from each model. EPD is formulated by $EPD = E\{\boldsymbol{\Sigma}_{i,j}(y_{rep,ij} - y_{obs,ij})^2\}$, where the predictive value $y_{rep,ij}$ is a replicate of the observed $y_{obs,ij}$ and the expectation is taken over the posterior distribution of the model parameters $\boldsymbol{\theta}$ (Gelman et al., 2003). RSS is given by $\boldsymbol{\Sigma}_{i,j}(y_{obs,ij} - y_{fitted,ij})^2$ and it is a measure of the discrepancy between the data and an estimation model. The smaller the value of

DIC, EPD and RSS, the better fit of the model to the data. Besides these statistical criteria, two diagnostic plots, Quantile-Quantile plot (Q-Q plot) and plots of observed values vs. fitted values, were also reported to give a visualized goodness-of-fit in the model comparisons.

We re-scaled the original time t (days) so that the time scale was between 0 and 1. We used the entire follow-up data in all of the models. In the Bayesian inferential approach, we also need to specify values of the hyper-parameters at the population level. Weakly informative prior distributions are taken for all the parameters: (i) for each component of fixed-effect vector of $\boldsymbol{\beta}$, the prior was assumed to follow independent normal distribution as $N(0, 100)$; (ii) for the scale parameter σ^2, we assumed a limiting non-informative inverse gamma prior distribution as $IG(0.01, 0.01)$, therefore, the mean is 1 and variance is 100; (iii) the prior for the variance-covariance matrix for the random-effect $\boldsymbol{\Sigma}_b$ was taken to be inverse Wishart distribution as $IW(\boldsymbol{\Omega}, v)$, the degree of freedom, $v = 5$, and $\boldsymbol{\Omega}$ is diagonal matrix with diagonal elements being 0.01; (iv) for the skewness parameter δ, we chose normal distribution; (v) the degree of freedom ν followed truncated exponential distribution with $\nu_0 = 0.5$

The MCMC sampler was implemented using WinBUGS software. The code for Model IV is available in Appendix A. The posterior means and quantiles were drawn after the collecting the final MCMC samples. We used one long chain. Convergence, which refers the algorithm has reached its equilibrium target distribution, was closely watched by using the standard tools within WinBUGS such as trace plots, the MC error and depicting the evolution of the ergodic means of a quantity over the number of iterations. After an initial 100,000 burn-in iterations, every 50*th* MCMC sample was retained from the next 200,000. Thus, we obtained 4,000 samples of targeted posterior distribution of the unknown parameters for statistical inference.

2.4.2 Results

Step 1: in Section 2.4.2.1, we determine when the model errors are assumed to have an ST distribution, among the five models presented in Section 2.4.1, which one has the best fit. **Step 2**: in Section 2.4.2.2, because normal, SN and Student-*t* distribution are all a special case of an ST distribution, focusing on the model selected from Step 1, we compare the results based on random errors with a normal, SN and Student-*t* and ST distribution. The model comparisons are carried on A5055 data and confirmed by A398 data. **Step 3**: Section 2.4.2.3 presents the results based on the best model

selected. **Step 4**: Section 2.4.2.4 includes simulation study to validate the conclusions made from Step 1 and Step 2.

Table 2.1: DIC, EPD and RSS among the five models, random errors are assumed to follow ST.

Data set		Model I	Model II	Model III	Model IV	Model V
A5055:	DIC	1192	401.1	669.2	21.7	1015.7
All subjects	EPD	0.33	0.12	0.15	0.05	0.20
(n=44)	RSS	59.2	20.7	27.1	8.4	35.3
A398:	DIC	1239.4	1122.6	1162.9	576.9	1200.5
Sample 1	EPD	3.17	3.48	3.63	1.04	3.04
(n=44)	RSS	396	436	454	134	381
A398:	DIC	2264.9	1947.5	2855.3	1757.7	2204.5
Sample 2	EPD	0.44	0.51	3.45	0.33	0.39
(n=100)	RSS	130.6	158.3	1083	99.4	116.7
A398:	DIC	12900.6	10891.2	10763.4	8819.3	14217
All subjects	EPD	1.49	1.42	1.31	0.93	1.84
(n=481)	RSS	2665	2185	2059	1470	2703

2.4.2.1. Comparison of five models under an ST distribution

When we compare the models that have different components, we should not directly compare the estimated parameters' values because they have different meanings. However, because the models were applied to the same HIV viral load data, we could use DIC, EPD and RSS to find out which model had the best fit. The comparison results were shown in Table 2.1. Among all of the data sets: A5055, the two randomly selected samples from A398 and A398 that includes all of the 481 subjects, Model IV constantly has the lowest DIC, EPD and RSS. For example, in A5055, DIC value for Model IV is 21.7, while it is 1192, 401.1, 669.2, and 1015.7 in Models I, II, III and V, respectively, which indicates Model IV is superior to the other models tested.

2.4.2.2. Comparison of four distributions about random errors in the best fitted model - Model IV

For Model IV, we further investigated how different distributions about random errors would affect the model fit and DIC values are shown in Table 2.2 below. Among normal, Student-*t*, SN and ST distribution, the model with either ST (A5055 and the two samples of A398) or SN (whole A398) has the lowest DIC. Because SN has a simpler structure than ST, and the larger the degree of freedom, the closer the Student-*t* distribution is to the normal distribution, it is not surprising to see when the sample size is big (e.g. A398, n=481), SN has a smaller DIC value than ST.

Table 2.2: For Model IV, DIC values under different distribution assumptions.

	Data set:			
Distribution	A5055 (n=44)	A398 (n=44)	A398 (n=100)	A398 (n=481)
Normal	1133.3	961.4	2158.5	10222.0
SN	222.8	727.3	1937.2	**7788.2**
Student-*t*	1004.9	949.8	2144.3	8993.9
ST	**21.7**	**576.9**	**1757.7**	8819.3

We also calculated EPD and RSS, which provide the equivalent conclusions: for example, in A5055 data (n=44), EPD is 0.05, 0.10, 2.19 and 2.29 for ST, SN, Student-*t* and normal, respectively; in A398 data (n=481), EPD is 0.24, 0.93, 2.74 and 3.17 for ST, SN, Student-*t* and normal, respectively;

We applied Model IV on A5055 to further compare the estimation results got from different distribution assumptions. The population posterior mean (PM), the corresponding standard deviation (SD) and 95% credible interval (CI) for fixed-effect parameters are presented in Table 2.3. Table 2.3 shows: (i) except β_5 based on the normal distribution assumption, all of the other estimates were significant since the 95% CIs don't include zero; (ii) for variance σ^2, the estimated value based on the SN (0.05) and ST (0.01) models were much smaller than that based on the model with normal (1.15) or Student-*t* (0.38) distribution; (iii) among all of the parameters estimated, the related SD obtained from ST was the smallest; (iv) the estimates were similar between normal and Student-*t* distribution model, but they were substantially different to those obtained from SN or ST

model. For example, β_2 based on normal or Student-*t* distribution was 34.67 and 38.10, respectively, while it was 24.53 and 27.89 in SN and ST model, respectively; (v) the skewness parameter δ was significantly positive in SN and ST, confirming the positive skewness of the viral load in the natural logarithm transformed as observed in Figure 2.2; (vi) compared to the model with normal or Student-*t* distribution assumption for random errors, the models with an SN or ST distribution fit the data better. For example, in A5055, for DIC value, 11333.3 (normal) vs. 222.8 (SN), 1004.9 (Student-*t*) vs. 21.7 (ST), it indicates that consideration of a departure from normality will improve the model fit.

Table 2.3: A summary of the estimated posterior values (based on A5055 data).

Model IV		β_1	β_2	β_3	β_4	β_5	σ^2	δ	DIC	EPD	RSS
Normal	PM	8.30	34.67	5.95	6.77	0.38	1.15	–	1133.3	2.29	1133.3
	L_{CI}	7.81	26.86	4.01	2.65	-0.40	0.96	–			
	U_{CI}	8.77	43.04	8.54	12.92	1.01	1.37	–			
	SD	0.26	4.15	1.26	2.52	0.35	0.11	–			
SN	PM	6.69	24.53	6.38	13.49	1.55	0.05	2.22	222.8	0.10	17.5
	L_{CI}	6.03	12.66	4.10	7.10	0.89	0.01	1.97			
	U_{CI}	7.40	35.03	11.22	20.19	2.15	0.16	2.51			
	SD	0.35	5.69	1.72	3.85	0.37	0.04	0.14			
Student-*t*	PM	8.33	38.10	9.41	11.97	0.94	0.38	–	1004.9	2.19	420.8
	L_{CI}	7.84	30.31	7.34	8.99	0.66	0.28	–			
	U_{CI}	8.83	46.61	11.64	15.59	1.24	0.50	–			
	SD	0.25	4.11	1.13	1.66	0.15	0.06	–			
ST	PM	7.18	27.89	7.19	13.69	1.73	0.01	1.17	21.7	0.05	8.4
	L_{CI}	6.66	21.53	6.34	11.72	1.55	0.00	0.99			
	U_{CI}	7.67	33.93	8.63	16.22	1.98	0.04	1.34			
	SD	0.25	3.26	0.66	1.28	0.13	0.01	0.08			

Note: L_{CI} and U_{CI} are lower limit and upper limit of 95% equal-tail credible interval, respectively

Several diagnostic plots for goodness-of-fit are also applied. First, we randomly select three subjects from A5055. The individual estimates of viral load trajectories are shown in Figure (2.4). The following findings are observed: (i) the estimated individual trajectories got from SN and ST fit the originally observed data much closer than those got from the model where the random errors are assumed to be normal or Student-*t*; (ii) all of the 95% CI from the model with an SN or ST distribution cover the observed viral load, where 21% and 19% of the 95% CI from normal and Student-*t*, respectively, doesn't include the observed values; (iii) the average SD got from ST is the smallest, which is 0.15, while the mean of SD for the individual estimation got from SN, N and Student-*t* is 0.22, 0.52 and 0.46, respectively. Note that the lack of smoothness in the SN and ST Model estimates of individual trajectories is understandable since a random component $\boldsymbol{w}_i$ was incorporated in the expected function (see equation (2.6) for details) according to the stochastic representation feature of the SN and ST distribution for "chasing the data" to this extent.

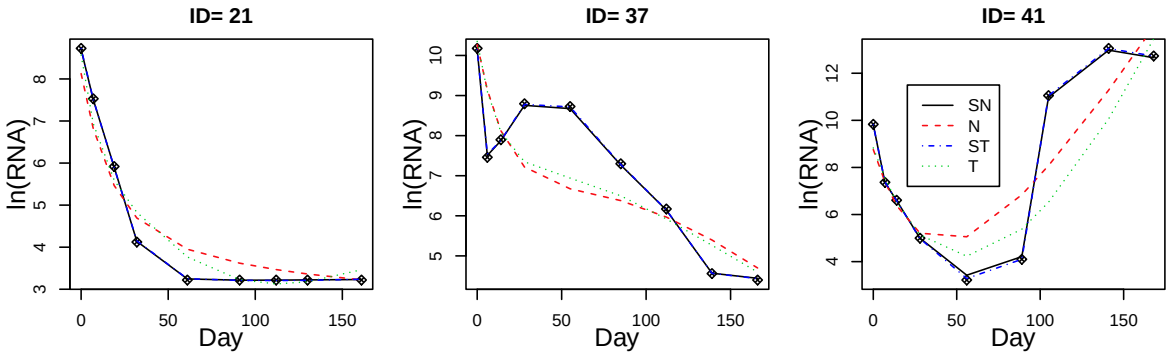

Figure 2.4: Individual estimates of viral load trajectories for three randomly selected patients based on normal, t, SN and ST distribution assumption in Model IV. The observed values are indicated by diamond ♢

Second, we created two diagnostic plots: plots of the observed values versus the fitted values (Figure 2.5) and Q-Q plot (Figure 2.6).

The findings of these two plots agree with that from DIC: the models with SN and ST distribution provided better fit to the observed data than the ones with normal or Student-*t* distribution assumption. Based on the results from DIC, EPD, RSS and the diagnostic plots, we conclude that Model IV with the ST distribution assumption fits the data better than the other combinations be-

tween a different time-varying viral decay rate functions and distribution assumption of random errors.

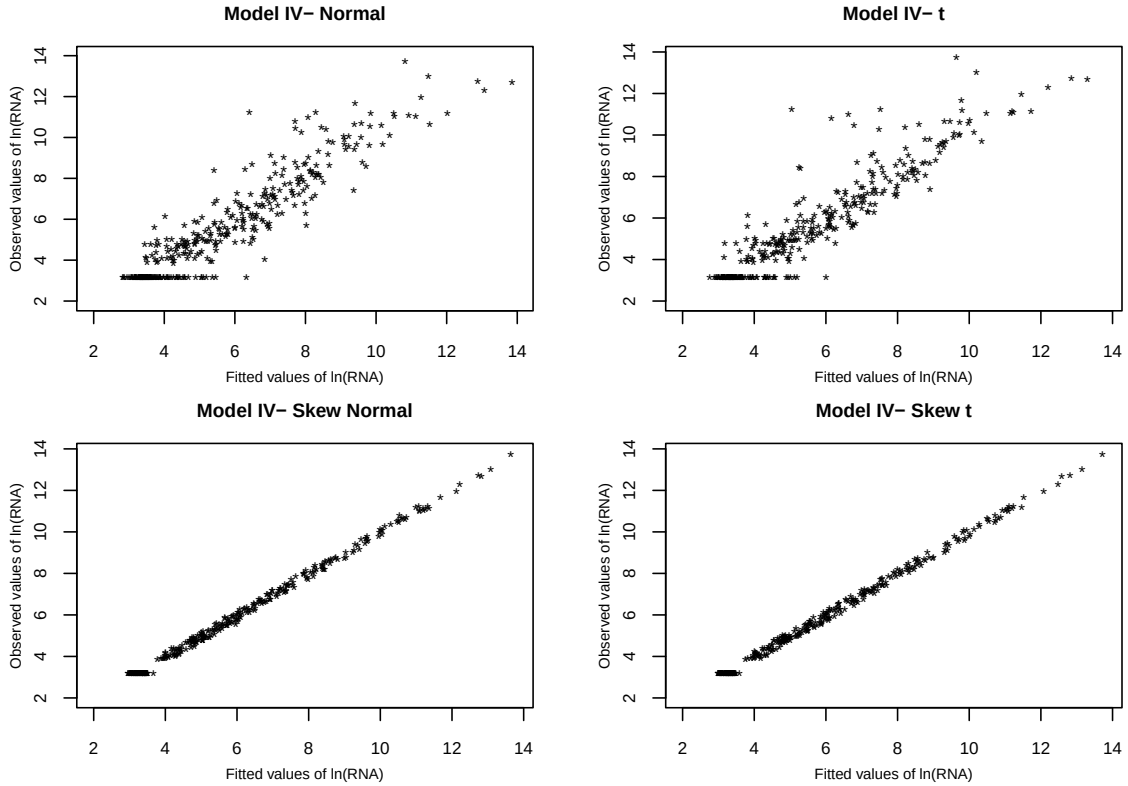

Figure 2.5: The observed values versus fitted values of ln(RNA) based on N, Student-*t*, SN or ST distribution for random errors

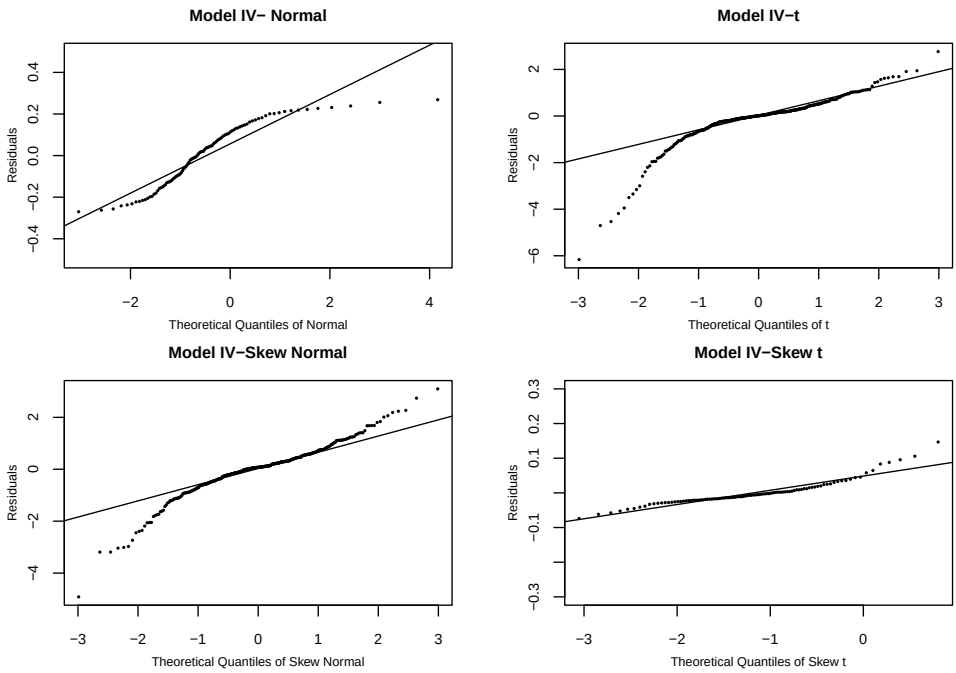

Figure 2.6: Q-Q plot

2.4.2.3. Simulation study

In order to validate the conclusions obtained from Steps 1 and 2 in Sections 2.4.2.1 and 2.4.2.2, resampling method is used via a simple random sampling algorithm. Fifteen additional samples are created from A398 and named as "A398-s1" to "A398-s15", while each sample includes 44 subjects. The validations are carried out in two scenarios. Firstly, we evaluate whether Model IV with the ST distribution is the most appropriate among five models presented in Section 2.4.1. Secondly, we assess whether Model IV with the ST distribution will provide better model fit in comparison of Model IV with the N, Student-*t* and SN distribution assumptions.

Scenario one. Under the ST distribution assumption, Model I $\sim$ Model V are applied to the fifteen samples created from A398 data set. The DIC values are shown in Table 2.4 and the related Boxplots are presented in Figure 2.7. Model IV consistently has the smallest DIC value among the five models. The average DIC in Model IV is 528.68, while it is 978.00, 859.38, 893.65 and 1051.66 in Model I, Model II, Model III and Model V, respectively. The boxplots shown in Figure 2.7 indicate that the median of DIC in Model IV is smaller than the values in the other four models, while the medians of DIC are similar in Model I, II and III but smaller than the value in Model V.

Scenario two. After confirming that Model IV is the most appropriate model among the five

models, we investigate among the fifteen samples, whether Model IV with the ST distribution provides better model fit than that with the N, Student-t and SN distributions. The DIC values are shown in Table 2.5 and Figure 2.8. Among the four distribution assumptions, the median value of DIC in Model IV with an ST distribution assumption is the lowest (641.47), followed by SN (741.10), Student-t (856.60) and Normal (857.07).

In summary, we conclude that, based on the two resampling simulation studies, Model IV with the ST distribution is the best model among the models with five time-varying decay rate functions and/or four distributions. We will further report analysis results based on Model IV with the ST distribution below.

Table 2.4: DIC values among the five models in the 15 samples from A398, random errors are assumed to follow ST distribution.

Data set	Model I	Model II	Model III	Model IV	Model V
A398-s1	1369.62	1335.26	1290.19	981.52	1379.38
A398-s2	931.63	348.82	509.20	171.86	1082.11
A398-s3	792.31	737.33	817.19	446.53	631.61
A398-s4	885.84	288.09	834.81	6.22	850.03
A398-s5	1199.30	1115.39	1132.25	778.70	1215.81
A398-s6	863.41	1105.55	1269.64	725.03	1208.31
A398-s7	844.77	636.58	700.77	557.90	906.25
A398-s8	1150.53	1059.46	1032.95	748.52	1155.85
A398-s9	1276.01	1102.07	1097.55	80.80	1273.38
A398-s10	860.40	954.13	871.74	779.33	1087.35
A398-s11	821.84	477.31	420.93	184.95	810.25
A398-s12	998.68	863.03	837.20	727.37	1002.01
A398-s13	1081.95	1189.50	1085.36	788.59	1317.84
A398-s14	1208.26	1010.11	920.54	788.37	1208.26
A398-s15	385.43	668.15	584.42	164.45	646.43
Mean	978.00	859.38	893.65	528.68	1051.66
SD	246.48	308.37	254.44	311.03	228.36
Median	931.63	908.58	882.7	641.47	1084.73

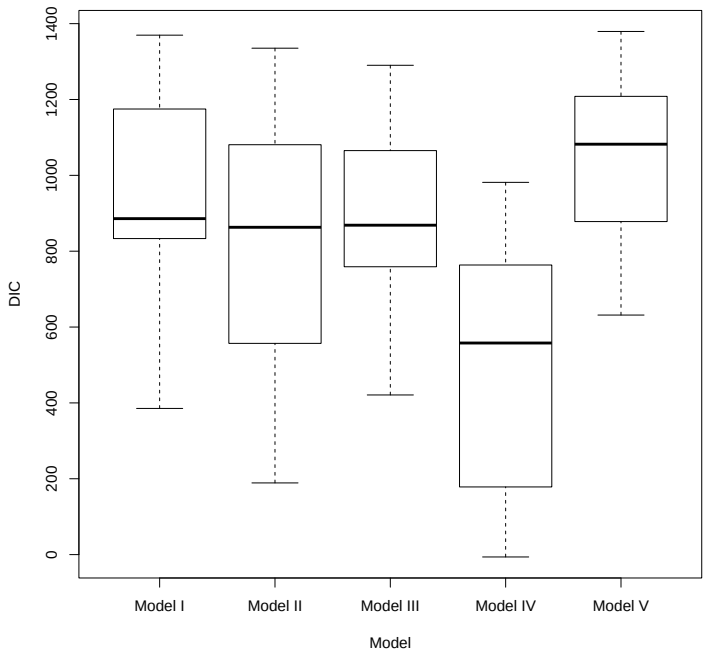

Figure 2.7: Boxplot based on the DIC values from 15 samples from A398

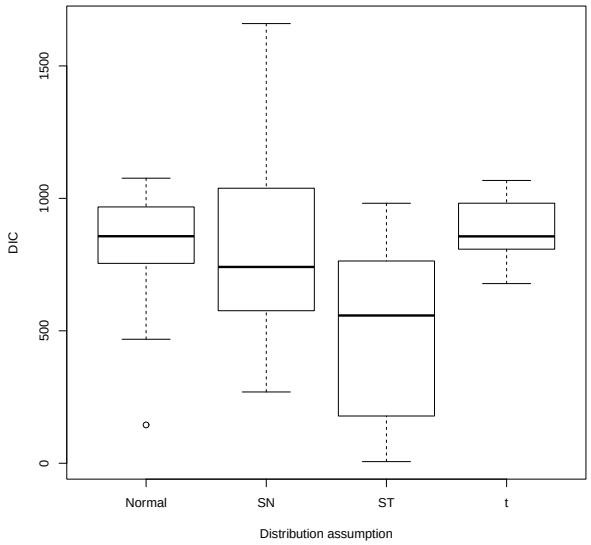

Figure 2.8: Boxplot based on the DIC values from 15 samples in A398, in Model IV with different distribution assumptions

Table 2.5: For Model IV, among the 15 samples in A398, DIC values under different distribution assumptions.

Statistics	Normal	SN	Student-*t*	ST
Mean	805.08	810.06	872.90	509.50
SD	240.54	355.03	115.70	313.82
Median	857.07	741.10	856.60	557.90

2.4.2.4. Results based on Model IV with an ST distribution

Based on Model IV with an ST distribution, the estimated population decay rate function for A5055 data is

$$\hat{\lambda}(t) = 27.89\exp(-7.9t) + 13.69\exp(-1.73t).$$

Because the estimated $\hat{\lambda}(t)$ is always positive, the population viral load would always decrease in this specific HIV/AIDS data set.

The individual time-varying decay rate function is given by,

$$\hat{\lambda}(t_{ij}) = \hat{\beta}_{i2}\exp(-\hat{\beta}_{i3}t_{ij}) + \hat{\beta}_{i4}\exp(-\hat{\beta}_{i5}t_{ij})$$

where the individual estimated decay rate $\hat{\lambda}(t_{ij})$ is considered to be dependent on both subjects and time. We found that the individual decay rate at initial treatment, $\hat{\lambda}(t_{i0})$, was positively correlated with baseline viral load (Spearman correlation coefficient r = 0.769, $p < 0.0001$) and negatively associated with baseline CD4 cell count (r = -0.447, $p = 0.0025$). Overall, the individual decay rate, $\hat{\lambda}(t_{ij})$, was positively associated with viral load ($p < 0.0001$) and negatively associated with CD4 cell count ($p < 0.0001$).

Because 30 $\sim$ 60% (Havlir et al., 2000) of patients eventually will have viral rebound, it is important to have a model that can reasonably predict this type of treatment failure in the long term. Following Wu et al.(2008), we defined rebound as, comparing with the HIV-1 viral load (natural log transformed) from the previous measurement, if there was $\geq$ 1.15 increase at one time point or $\geq$ 0.46 increase at two or more consecutive time points. In ACTG5055, there were 11 (26.2%) subjects had rebound. There was no significant difference in the baseline viral load (natural $\log$ (RNA)) between the rebound and no rebound group (median was 9.18 and 8.78, respectively, p =

0.8610), while the median of baseline CD4 cell count intended to be higher in the no rebound group than that in the rebound group (285 vs. 253/mL, p = 0.1169).

The trend of the changes in decay rates during the treatment was different between the rebound and no rebound group (Figure 2.9). For example, every individual decay rate was positive in the no rebound group, while some individual decay rates in the rebound group became negative, especially after the 3*rd* month of the treatment, which corresponding to the viral load rebound.

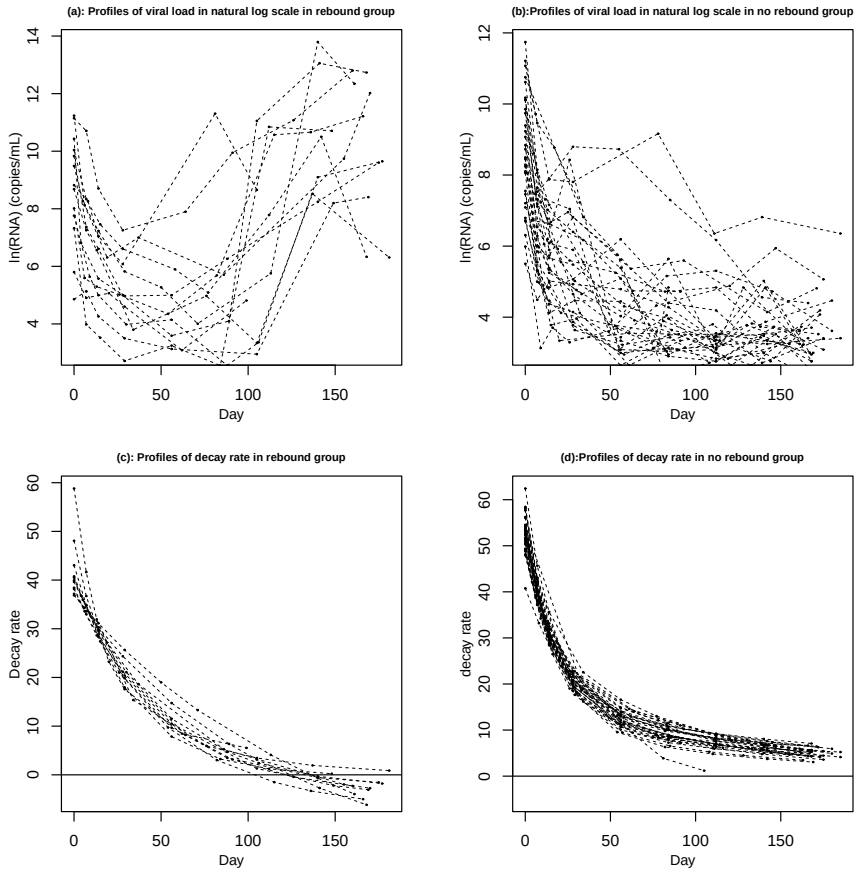

Figure 2.9: Profile of viral load in $\ln$ scale and decay rate in rebound and no rebound group

Based on the results of Model IV under an ST distribution in A5055 data, we also find that: (i) overall, the average value of individual decay rates, $\hat{\lambda}(t_{ij})$, was bigger in the no rebound group (14.97) than that in the rebound group (12.93); (ii) the initial individual decay rates, $\hat{\lambda}(t_{i1})$, were significantly bigger in the no rebound group than that in the rebound group (mean is 53.16 and 40.95, respectively); (iii) $\hat{\lambda}(t_{i1})$ was significantly associated with the rebound status in the long term (OR = 0.703, 95% CI is 0.580 – 0.853, p = 0.0003) and this association was still significant even after controlling the baseline viral load and CD4 cell count (OR = 0.717, 95% CI is 0.588 –

0.875, p = 0.0010). (iv) the average individual decay rate at the last visit ($\hat{\lambda}(t_{in_i})$) among the no rebound subjects was 4.67, while it was -2.28 in the rebound group which indicates the viral load actually increasing instead of decreasing in this group. Among these findings, the most interesting is the associated relationship between initial individual decay rate ($\hat{\lambda}(t_{i0})$) and the rebound status: the results indicated that the odds of rebound decreased about 30% with each one unit increased in the initial decay rate. This may be helpful for clinicians to predict the long term results based on the information at early stage of the disease.

2.5. Conclusion and discussion

With an ST distribution assumption for model random errors to account skewness observed in viral load responses, we compared five commonly used mixed-effects models in HIV dynamics via the Bayesian approach. We also investigated the impact of the four distributions in the skew-elliptical family on the model fit. The results indicate that with the ST distribution, there is potential gain of efficiency and accuracy in estimating certain parameters when the normality assumption does not apply to the data. The skew-elliptical modeling via the Bayesian approach proposed in this study can be easily carried out via the WinBUGS package. Because the proposed model is quite general in theory and accessible to the existing software, it will allow statisticians to apply this method in other fields.

In any discussions of mathematical modeling of complex systems, it is important to point out that, while complex models may be needed to provide accurate descriptions of the underlying dynamics, the models are most useful when they can be compared to clinical and/or experimental data. In developing models for HIV infection and treatment, this requires a balance between complexity and utility. After finding the best fitted model, we estimated the relationship between the viral decay rate and some clinical important variables. Based on the results from the best fitted model with an ST distribution assumption, we found the initial estimated decay rate was positively correlated with the baseline viral load and negatively associated with baseline CD4 cell count. We also found that, overall, the average decay rate was lower in the rebound than that in the no rebound group. A more interesting finding is the significant association between the initial decay rate and the rebound status in the long term, even after controlling for the baseline viral load and CD4 cell count. This finding is clinically important because it may enable physicians to predict the long-term

outcome based on the estimated decay rate at an early stage. Because we didn't find such kinds of associations based on the model with Student-*t* or normal distribution assumption, it is important to consider non-normality into the modeling when the normality assumption cannot be satisfied even after transformation.

Using the model with a time-varying decay rate function has some advantages over the biphasical models. (i) In the biphasical models, the association between the first decay rate and baseline viral load could be positive (Notermans et al., 1998; Wu et al., 2004) or negative (Wu et al., 1999); no significant association was found between the rebound and the first decay rate either (Wu et al., 2008); (ii) although the second decay rate in the biphasical models is supposed to be associated with long-term treatment status such as rebound (Ding and Wu, 1999; Wu et al.,2005), no significant association was found between the second decay rate and the viral replication in long term (Sedaghat et al., 2008; Wu et al., 2003).

This chapter has some limitations. Usually, covariates are included in the mixed-effects models to control within- and between-subject variation, and CD4 cell count is a commonly used covariate in HIV dynamic models. However, in order to use the original proposed models in the comparisons, we did not include any covariates such as CD4 cell count or demographic information. For the viral load, the values below detection limit (BDL) are usually considered as inaccurate. Instead of treating these values as censored, we computed them by half the value of the detectable level. The issue of missing values is not considered in this study either.

This chapter compared commonly used HIV dynamic models and the estimation was through Bayesian statistical inference. The mathematically modeling was extended from a normality assumption and a general skew-elliptical distribution was used in order to account the skewness observed in the data. New technologies were applied to facilitate the computation challenges related to the complex nature of HIV/AIDS data. Furthermore, a more flexible distribution such as skew-normal independent distribution, can be assumed; CD4 cell count, which can either be treated as a covariate or an outcome in the HIV research, needs to be considered too, while the measurement errors, skewness of CD4, and correlation with other factors such as CD8 are all worth to explore.

In conclusion, the skewness parameter in the model with SN or ST distribution is significantly positive, which confirms the positive skewness observed in the viral load data even after natural log transformed. The model fit is the best in the model with skewed (SN or ST) distribution. Because

estimated parameters can be considerably different between the models with skewed distribution and normal or Student-*t* distribution, it is important to account for skewness in the model when data exhibits noticeable skewness. Different models may yield different conclusions about the relationship between the decay rate with viral load, CD4 cell count and rebound status in HIV dynamics, therefore, it is also critical to choose a reasonable model that can balance between complexity and utility.

3 Simultaneous Bayesian inference for linear, nonlinear and semiparametric mixed-effects models with skew-normality and measurement errors in covariates

Disclaimer

This chapter has already been published as: Huang, Yangxin; Chen, Ren; and Dagne, Getachew (2011) "Simultaneous Bayesian Inference for Linear, Nonlinear and Semiparametric Mixed-Effects Models with Skew-Normality and Measurement Errors in Covariates". *The International Journal of Biostatistics*: Vol. 7 : Iss. 1, Article 8. The permission of including it in the dissertation is in the Appendix D. Except the section numbers and question numbers changing in order to make the chapters labeling constant, all of the remaining contents in Chapter 3 are in the original format.

Abstract

In recent years various mixed-effects models have been suggested for estimating viral decay rates in HIV viral dynamic models for complex longitudinal data. Among those models are linear mixed-effects (LME), nonlinear mixed-effects (NLME), and semiparametric nonlinear mixed-effects (SNLME) models. However, a critical question is whether these models produce coherent estimates of viral decay rates, and if not, which model is appropriate and should be used in practice. In addition, one often assumes that model random errors are normally distributed, but the normality assumption may be unrealistic, particularly, if the data exhibit skewness. Moreover, some covariates such as CD4 cell count may be often measured with substantial errors. This paper addresses these issues simultaneously by jointly modeling the response variable with skewness and a covariate process with measurement errors using a Bayesian approach to investigate how estimated parameters are changed or different under these three models. A real data set from an AIDS clinical trial study was used to illustrate the proposed models and methods. It was found that there was a significant incongruity in the estimated decay rates in viral loads based on the three mixed-effects models, suggesting that the decay rates estimated by using Bayesian LME or NLME joint models should be

interpreted differently from those estimated by using Bayesian SNLME joint models. The findings also suggest that the Bayesian SNLME joint model is preferred to other models because an arbitrary data truncation is not necessary; and it is also shown that the models with a skew-normal distribution and/or measurement errors in covariate may achieve reliable results when the data exhibit skewness.

3.1. Introduction

Modeling of HIV dynamics in AIDS research has greatly improved our understanding of the pathogenesis of HIV-1 infection and guided for the treatment of AIDS patients and evaluation of antiretroviral (ARV) therapies (Perelson, 1997; Wu and Ding, 1999; Wu et al., 2005). Such models often incorporate repeated measures over a period of treatment to assess rates of changes in viral load (number of HIV RNA copies in plasma). Recent research indicates that the first-phase decay rate of viral response to treatment is a potentially useful biomarker for ARV potency (Ding and Wu, 1999). Even though various statistical modeling and analysis methods have been applied for estimating the parameters of HIV dynamics, such as, linear and nonlinear regression (Perelson, 1997), linear mixed-effects (LME) and nonlinear mixed-effects (NLME) modeling approach (Wu and Ding, 1999; Wu, et al., 1998, 2004; Wu, 2004), nonparametric NLME modeling approach (Liu and Wu, 2007; Wu and Zhang, 2002; Wu, et al., 2004), joint model approach via Monte Carlo EM algorithm (Liu and Wu, 2007; Wu, 2002; Wu, 2004) and Bayesian NLME modeling approach via Markov chain Monte Carlo (MCMC) procedure (Huang et al., 2006; Huang and Dagne, 2010), it is not clear which model should be used to estimate the first-phase decay rate. More importantly, in all of these models at least one of the following three important issues standout.

Firstly, the common assumption of distributions for (within-subject) random error is normal in those statistical models. This assumption may lack the robustness against departures from normality and/or outliers as discussed by Verbeke and Lesaffre (1996), and may also lead to misleading results (Verbeke and Lesaffre, 1996; Ghosh et al., 2007). Very often in AIDS studies, the virologic responses exhibit skewness. For example, Figure 3.1 displays the histograms of repeated viral load (in natural $\log$ scale) and CD4 cell count measurements for 44 subjects enrolled in the AIDS clinical trial study–A5055 (Acosta et al., 2004). It seems that for these data sets to be analyzed in this paper, the viral load responses are highly skewed even after natural $\log$-transformation. Thus, a normality assumption is not quite realistic and may be too restrictive to provide an accurate rep-

resentation of the structure of the data. One way to incorporate skewness is to use a skew-normal (SN) distribution for (within-subject) random errors. *Secondly*, the mixed-effects models have been used in the previous studies to account for both between-subject and within-subject variations in viral load measurements which are associated with covariates including CD4 cell count. However, the covariates such as CD4 cell count which were considered in those studies are often measured with substantial errors and highly skewed as shown in Figure 3.1(down panel). *Thirdly*, a major challenge with these modeling approaches is the associated intensive computation burden in the inference, and in some cases it can even be computationally infeasible. Particularly, for nonlinear longitudinal models in the presence of measurement errors in covariates, the computational problem becomes much worse. In addition, there may exist the problem of non-convergence of these algorithms under the framework of likelihood estimation. To the best of our knowledge, there is little literature on simultaneously addressing measurement errors in covariates and an SN distribution for random errors to compare performance of the various mixed-effects models under the framework of Bayesian mixed-effects modeling approach. This article provides a unified approach to investigate SN Bayesian mixed-effects models with covariate measurement errors.

In this paper, the main focus is to provide a comprehensive comparison of three mixed-effects models (LME, NLME and semiparametric NLME) with an SN distribution and measurement errors in covariates for estimated viral decay rates in viral dynamic models. We consider a multivariate SN distribution introduced by Sahu et al.(2003) which is suitable for a Bayesian inference since it is built using conditional method and is defined in Section 1.4. The rest of the paper is organized as follows. Section 3.2 presents a general modeling approach for SN Bayesian semiparametric nonlinear mixed-effects (SN-BSNLME) joint models which include three specific models as special cases to be discussed in Section 3.3. We describe data from an AIDS clinical study that motivated this research, discuss the specific models for HIV dynamics and reports the results obtained by using the three different methods or models In Section 3.3. Finally, the paper concludes with some discussions in Section 3.4.

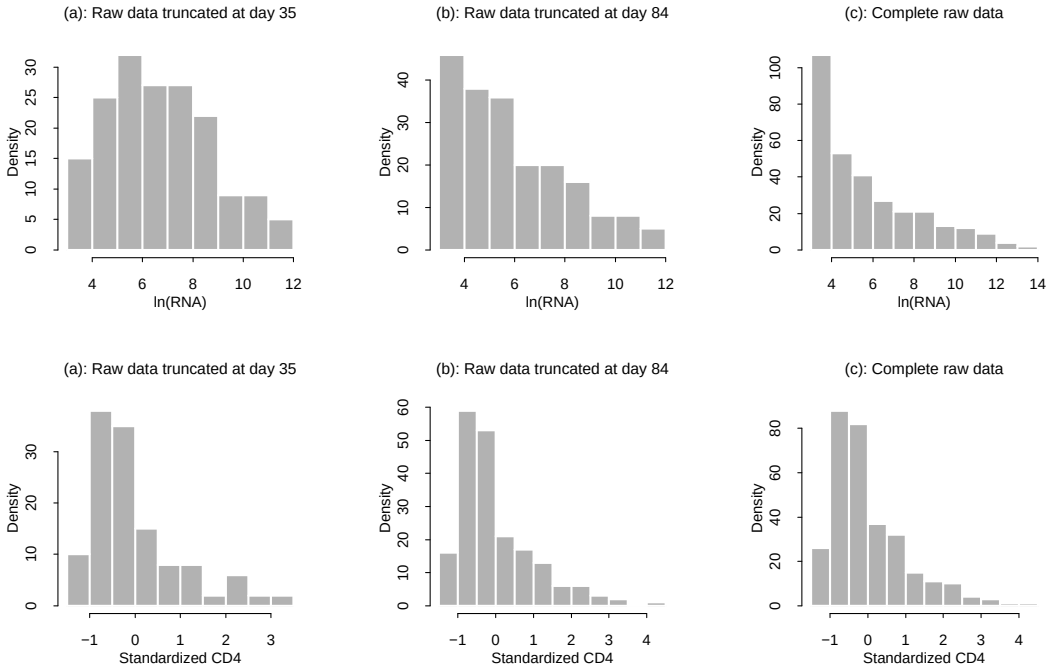

Figure 3.1: The histograms of viral load (in $\ln$ scale) and standardized CD4 cell count measured from day 0 to day 35 [Set (a): data cut at truncation day 35], day 84 [Set (b): data cut at truncation day 84] and the end of study period [Set (c): complete data] for 44 patients in an AIDS clinical trial study.

3.2. Bayesian inference on joint models with skew-normal distributions

3.2.1. Measurement error models with a skew-normal distribution

This section will briefly discuss measurement error joint models with a skew-normal distribution. Various covariate models were investigated in the literature (Carroll et al., 2006; Higgins et al., 1997; Liu and Wu, 2007; Wu, 2002). However, the fundamental assumption of distributions for measurement random errors is normal in these statistical covariate models and this assumption may lack the robustness against departures from normality and/or may violate the agreement with observed data. Thus statistical inference and analysis with normal assumption may lead to misleading results. In this paper, we extend the covariate models to have a skew-normal distribution for measurement errors. For simplicity, we consider a single time-varying covariate. Let z_{ik} be the observed covariate value for individual i at time s_{ik} ($i = 1, \ldots n; k = 1, \ldots, m_i$). Note that for each individual, we allow the covariate measurement times s_{ik} to differ from the response measurement times t_{ij} ($j = 1, \ldots, n_i$). In the presence of measurement errors in covariate, we need to model the covariate process. We consider the following LME model with an SN distribution

$$z_{ik} = \boldsymbol{u}_{ik}^T\boldsymbol{\alpha} + \boldsymbol{v}_{ik}^T\boldsymbol{a}_i + \epsilon_{ik} \ \ (\equiv z_{ik}^* + \epsilon_{ik}), \ \ \boldsymbol{\epsilon}_i \ \ iid \sim SN_{m_i}\left(\mathbf{0}, \tau^2\boldsymbol{I}_{m_i}, \boldsymbol{\Delta}_\epsilon\right), \tag{3.1}$$

where $\boldsymbol{z}_i = (z_{i1}, \ldots, z_{im_i})^T$ with z_{ik} being the covariate value for individual i at time s_{ik}, $\boldsymbol{u}_{ik}$ and $\boldsymbol{v}_{ik}$ are $l \times 1$ design vectors, $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_l)^T$ and $\boldsymbol{a}_i = (a_{i1}, \ldots, a_{il})^T$ are unknown population (fixed-effects) and individual-specific (random-effects) parameter vectors, respectively, and $\boldsymbol{\epsilon}_i = (\epsilon_{i1}, \ldots, \epsilon_{im_i})^T$ is a multivariate skew-normal distribution with ϵ_{ik} being the measurement error for individual i at time s_{ik}, τ^2 is the unknown within-individual variance. The $m_i \times m_i$ skewness diagonal matrix $\boldsymbol{\Delta}_\epsilon = diag(\delta_{\epsilon_{i1}}, \ldots, \delta_{\epsilon_{im_i}})$ and $m_i \times 1$ skewness parameter vector $\boldsymbol{\delta}_{\epsilon_i} = (\delta_{\epsilon_{i1}}, \ldots, \delta_{\epsilon_{im_i}})^T$. In particular, if $\delta_{\epsilon_{i1}} = \cdots = \delta_{\epsilon_{im_i}} \hat{=} \delta_\epsilon$, then $\boldsymbol{\Delta}_\epsilon = \delta_\epsilon\boldsymbol{I}_{m_i}$ and $\boldsymbol{\delta}_{\epsilon_i} = \delta_\epsilon\mathbf{1}_{m_i}$ with $\mathbf{1}_{m_i} = (1, \ldots, 1)^T$; this indicates that we are interested in skewness of overall data set which is the case to be used in real data analysis in the next section. $\boldsymbol{z}_i^* = (z_{i1}^*, \ldots, z_{im_i}^*)^T$ and $z_{ik}^* = \boldsymbol{u}_{ik}^T\boldsymbol{\alpha} + \boldsymbol{v}_{ik}^T\boldsymbol{a}_i$ may be viewed as the true (but unobservable) covariate values at time s_{ik} under normal distribution of model errors in which case skewness parameter $\delta_{\epsilon_{ik}} = 0$. We assume that $\boldsymbol{a}_i \ iid \ \sim N_l(0, \boldsymbol{\Sigma}_a)$, where $\boldsymbol{\Sigma}_a$ is the unrestricted covariance matrix. Note that the model (3.1) may be interpreted as a skew-normal (SN) covariate measurement error model which incorporates the correlation of the repeated measurements on each individual.

3.2.2. Skew-normal Bayesian semiparametric nonlinear mixed-effects joint models

In this section, we present the joint models and methods in general forms, illustrating that our methods may be applicable in other applications. Denote the number of subjects by *n* and the number of measurements on the *i*th subject by n_i. For the response process, we consider a general semiparametric NLME (SNLME) model which is similar to Wu and Zhang (2002) but incorporates possibly mis-measured time-varying covariates and model random errors with a skew-normal distribution.

$$\begin{aligned} y_{ij} &= g(t_{ij}, \boldsymbol{\beta}_{ij}^\dagger, \phi(t_{ij})) + e_{ij}, \quad \boldsymbol{e}_i \ \ iid \sim SN_{n_i}\left(\mathbf{0}, \sigma^2\boldsymbol{I}_{n_i}, \boldsymbol{\Delta}\right), \\ \boldsymbol{\beta}_{ij}^\dagger &= \boldsymbol{d}^\dagger[z_{ij}^*, \boldsymbol{\beta}^\dagger, \boldsymbol{b}_i^\dagger], \quad \boldsymbol{b}_i^\dagger \ \ iid \sim N_{s_3}(\mathbf{0}, \boldsymbol{\Sigma}_b^\dagger), \\ \phi(t_{ij}) &= v[w(t_{ij}), h_i(t_{ij})], \quad i = 1, 2, \ldots, n; \ \ j = 1, 2, \ldots, n_i, \end{aligned} \tag{3.2}$$

where $\boldsymbol{y}_i = (y_{i1}, \ldots, y_{in_i})^T$ with y_{ij} being the response value for individual i at t_{ij}, $g(\cdot)$, $\boldsymbol{d}^\dagger(\cdot)$ and $v(\cdot)$ are known parametric functions, $w(t)$ and $h_i(t)$ are unknown nonparametric smooth fixed-effects and random-effects functions, respectively, $h_i(t)$ are iid realizations of a zero-mean stochastic process, $\boldsymbol{\beta}_{ij}^\dagger$ are $s_1 \times 1$ individual-specific time-dependent parameter vectors, $\boldsymbol{\beta}^\dagger$ are

$s_2 \times 1$ population parameter vectors ($s_2 \geq s_1$), σ^2 is the unknown within-subject variation, $\boldsymbol{e}_i = (e_{i1}, \ldots, e_{in_i})^T$ is the vector of random errors; $\boldsymbol{b}_i^\dagger$ are $s_3 \times 1$ vector of random-effects ($s_3 \leq s_1$) and $\boldsymbol{\Sigma}_b^\dagger$ is the unrestricted covariance matrix. The $n_i \times n_i$ skewness diagonal matrix $\boldsymbol{\Delta} = diag(\delta_{e_{i1}}, \ldots, \delta_{e_{in_i}})$ and the $n_i \times 1$ skewness parameter vector $\boldsymbol{\delta}_{e_i} = (\delta_{e_{i1}}, \ldots, \delta_{e_{in_i}})^T$. In particular, if $\delta_{e_{i1}} = \cdots = \delta_{e_{in_i}} \hat{=} \delta_e$, then $\boldsymbol{\Delta} = \delta_e \boldsymbol{I}_{n_i}$ and $\boldsymbol{\delta}_{e_i} = \delta_e \mathbf{1}_{n_i}$. In the model (3.2), we assume that the individual-specific parameters $\boldsymbol{\beta}_{ij}^\dagger$ depend on the true (but unobservable) covariate z_{ij}^* rather than the observed covariate z_{ij}, which may be measured with errors.

Semiparametric NLME model (3.2) is more flexible than parametric NLME models. It reverts to a parametric NLME model when the nonparametric parts $w(t)$ and $h_i(t)$ are constants. To fit model (3.2), we apply the regression spline method. The working principle is briefly described as follows and more details can be found in Wu and Zhang (2002). The main idea of regression spline is to approximate $w(t)$ and $h_i(t)$ by using a linear combination of spline basis functions. For instance, $w(t)$ and $h_i(t)$ can be approximated by a linear combination of basis functions $\boldsymbol{\Psi}_p(t) = \{\psi_0(t), \psi_1(t), ..., \psi_{p-1}(t)\}^T$ and $\boldsymbol{\Phi}_q(t) = \{\phi_0(t), \phi_1(t), ..., \phi_{q-1}(t)\}^T$, respectively. That is,

$$w(t) \approx w_p(t) = \sum_{l=0}^{p-1} \mu_l \psi_l(t) = \boldsymbol{\mu}_p \boldsymbol{\Psi}_p(t)^T, \quad h_i(t) \approx h_{iq}(t) = \sum_{l=0}^{q-1} \xi_{il} \phi_l(t) = \boldsymbol{\xi}_{iq} \boldsymbol{\Phi}_q(t)^T, \quad (3.3)$$

where $\boldsymbol{\mu}_p$ and $\boldsymbol{\xi}_{iq}$ ($q \leq p$) are the unknown vectors of fixed and random coefficients, respectively. Based on the assumption of $h_i(t)$, we can regard $\boldsymbol{\xi}_{iq}$ as iid realizations of a zero-mean random vector. For our model, we consider natural cubic spline bases with the percentile-based knots. To select an optimal degree of regression spline and numbers of knots, i.e., optimal sizes of p and q, the Akaike information criterion (AIC) or the Bayesian information criterion (BIC) is often applied (Davidian and Giltinan, 1995; Wu and Zhang, 2002). Substituting $w(t)$ and $h_i(t)$ by their approximations $w_p(t)$ and $h_{iq}(t)$, we can approximate model (3.2) in a compact way as follows.

$$y_{ij} = g\left(t_{ij}, \boldsymbol{d}^\dagger[z_{ij}^*, \boldsymbol{\beta}^\dagger, \boldsymbol{b}_i^\dagger], v[\boldsymbol{\Psi}_p(t_{ij})^T \boldsymbol{\mu}_p, \boldsymbol{\Phi}_q(t_{ij})^T \boldsymbol{\xi}_{iq}]\right) + e_{ij} \equiv g\left(t_{ij}, \boldsymbol{d}[z_{ij}^*, \boldsymbol{\beta}, \boldsymbol{b}_i]\right) + e_{ij} \quad (3.4)$$

where $\boldsymbol{\beta} = (\boldsymbol{\beta}^{\dagger T}, \boldsymbol{\mu}_p^T)^T$ and $\boldsymbol{b}_i = (\boldsymbol{b}_i^{\dagger T}, \boldsymbol{\xi}_{iq}^T)^T$ are the vectors of fixed-effects and random-effects, respectively, and $\boldsymbol{d}(\cdot)$ is a known but possible nonlinear function. By doing so, the randomness of the nonparametric mixed-effects is transferred to the randomness of the associated coefficients, whereas the nonparametric feature is represented by the basis functions. Thus, for given $\boldsymbol{\Psi}_p(t)$ and

$\boldsymbol{\Phi}_q(t)$, we approximate the SN semiparametric NLME model (3.2) by the following SN parametric NLME model.

$$\begin{aligned} \boldsymbol{y}_i &= \boldsymbol{g}_i(t_{ij}, \boldsymbol{\beta}_{ij}) + \boldsymbol{e}_i, \qquad \boldsymbol{e}_i \;\; iid \sim SN_{n_i}\left(\mathbf{0}, \sigma^2 \boldsymbol{I}_{n_i}, \boldsymbol{\Delta}\right), \\ \boldsymbol{\beta}_{ij} &= \boldsymbol{d}[z^*_{ij}, \boldsymbol{\beta}, \boldsymbol{b}_i], \qquad \boldsymbol{b}_i \;\; iid \sim N_{s_4}(\mathbf{0}, \boldsymbol{\Sigma}_b), \end{aligned} \tag{3.5}$$

where $s_4 = s_3 + q$, $\boldsymbol{g}_i(t_{ij}, \boldsymbol{\beta}_{ij}) = (g(t_{i1}, \boldsymbol{\beta}_{i1}), \ldots, g(t_{in_i}, \boldsymbol{\beta}_{in_i}))^T$ with $g(\cdot)$ being a known linear or nonlinear function, and $\boldsymbol{\Sigma}_b$ is an unstructured covariance matrix.

Under Bayesian framework, we still need to specify prior distributions for unknown parameters in the models (3.1) and (3.5) as follows.

$$\begin{aligned} &\boldsymbol{\alpha} \sim N_r(\boldsymbol{\alpha}_0, \boldsymbol{\Lambda}_1), \quad \tau^2 \sim IG(\omega_1, \omega_2), \quad \boldsymbol{\Sigma}_a \sim IW(\boldsymbol{\Omega}_1, \nu_1), \quad \boldsymbol{\delta}_{\epsilon_i} \sim N_{m_i}(\mathbf{0}, \boldsymbol{\Gamma}_1), \\ &\boldsymbol{\beta} \sim N_{s_5}(\boldsymbol{\beta}_0, \boldsymbol{\Lambda}_2), \quad \sigma^2 \sim IG(\omega_3, \omega_4), \quad \boldsymbol{\Sigma}_b \sim IW(\boldsymbol{\Omega}_2, \nu_2), \quad \boldsymbol{\delta}_{e_i} \sim N_{n_i}(\mathbf{0}, \boldsymbol{\Gamma}_2), \end{aligned} \tag{3.6}$$

where $s_5 = s_2 + p$, the mutually independent Inverse Gamma (IG), Normal (N) and Inverse Wishart (IW) prior distributions are chosen to facilitate computations (Davidian and Giltinan, 1995). The super-parameter matrices $\boldsymbol{\Lambda}_1$, $\boldsymbol{\Lambda}_2$, $\boldsymbol{\Omega}_1$, $\boldsymbol{\Omega}_2$, $\boldsymbol{\Gamma}_1$ and $\boldsymbol{\Gamma}_2$ can be assumed to be diagonal for convenient implementation.

We assume that $\boldsymbol{e}_i$, $\boldsymbol{\epsilon}_i$, $\boldsymbol{b}_i$ and $\boldsymbol{a}_i$ are independent of each other. Following Sahu et al.(2003) and properties of skew-normal distribution, it can be shown that $\boldsymbol{z}_i$ and $\boldsymbol{y}_i$ in the models (3.1) and (3.5) follow the following distributions

$$\begin{aligned} \boldsymbol{y}_i | \boldsymbol{a}_i, \boldsymbol{b}_i, \boldsymbol{w}_{e_i} &\sim N_{n_i}\left(\boldsymbol{g}_i + \delta_e \boldsymbol{w}_{e_i}, \sigma^2 \boldsymbol{I}_{n_i}\right), \qquad \boldsymbol{w}_{e_i} \sim N_{n_i}(\mathbf{0}, \boldsymbol{I}_{n_i}) I(\boldsymbol{w}_{e_i} > \mathbf{0}), \\ \boldsymbol{z}_i | \boldsymbol{a}_i, \boldsymbol{w}_{\epsilon_i} &\sim N_{m_i}\left(\boldsymbol{z}^*_i + \delta_\epsilon \boldsymbol{w}_{\epsilon i}, \tau^2 \boldsymbol{I}_{m_i}\right), \qquad \boldsymbol{w}_{\epsilon_i} \sim N_{m_i}(\mathbf{0}, \boldsymbol{I}_{m_i}) I(\boldsymbol{w}_{\epsilon_i} > \mathbf{0}), \end{aligned} \tag{3.7}$$

where $I(\boldsymbol{w}_k > 0)$ is an indicator function and $\boldsymbol{w}_k = |\boldsymbol{\xi}|$ with $\boldsymbol{\xi} \sim N_k(0, \boldsymbol{I}_k)$.

Let $\boldsymbol{\theta} = \{\boldsymbol{\alpha}, \boldsymbol{\beta}, \tau^2, \sigma^2, \boldsymbol{\Sigma}_a, \boldsymbol{\Sigma}_b, \boldsymbol{\delta}_{\epsilon_i}, \boldsymbol{\delta}_{e_i}; i = 1, \ldots, n\}$ be the collection of unknown parameters in the models (3.1) and (3.5), and $f(\cdot|\cdot)$ and $\pi(\cdot)$ be a conditional density function and a prior density function, respectively. Denote the observed data by $\mathcal{D}$ = (($\boldsymbol{y}_i$, $\boldsymbol{z}_i$), $i = 1, ..., n$). We assume that $\boldsymbol{\alpha}, \boldsymbol{\beta}, \tau^2, \sigma^2, \boldsymbol{\Sigma}_a, \boldsymbol{\Sigma}_b, \boldsymbol{\delta}_{\epsilon_i}, \boldsymbol{\delta}_{e_i}$ ($i = 1, \ldots, n$) are independent of each other, and thus we have $\pi(\boldsymbol{\theta}) = \pi(\boldsymbol{\alpha})\pi(\boldsymbol{\beta})\pi(\tau^2)\pi(\sigma^2)\pi(\boldsymbol{\Sigma}_a)\pi(\boldsymbol{\Sigma}_b) \prod_i \pi(\boldsymbol{\delta}_{\epsilon_i})\pi(\boldsymbol{\delta}_{e_i})$. After we specify the models for the observed data and the prior distributions for the unknown model parameters, we can make statistical inference for the parameters based on their posterior distributions under Bayesian framework. The joint posterior density of $\boldsymbol{\theta}$ based on the observed data can be given by

$$\begin{aligned} f(\boldsymbol{\theta}|\mathcal{D}) \;\propto\; &\{\prod_i^n \int\int f(\boldsymbol{y}_i | \boldsymbol{a}_i, \boldsymbol{b}_i, \boldsymbol{w}_{e_i}; \boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma^2, \boldsymbol{\delta}_{e_i})\, f(\boldsymbol{w}_{e_i} | \boldsymbol{w}_{e_i} > \mathbf{0}) \times \\ &f(\boldsymbol{z}_i | \boldsymbol{a}_i, \boldsymbol{w}_{\epsilon_i}; \boldsymbol{\alpha}, \tau^2, \boldsymbol{\delta}_{\epsilon_i}) f(\boldsymbol{w}_{\epsilon_i} | \boldsymbol{w}_{\epsilon_i} > \mathbf{0})\, f(\boldsymbol{a}_i | \boldsymbol{\Sigma}_a)\, f(\boldsymbol{b}_i | \boldsymbol{\Sigma}_b) d\boldsymbol{a}_i d\boldsymbol{b}_i\} \pi(\boldsymbol{\theta}). \end{aligned} \tag{3.8}$$

In general, the integrals in (3.8) are of high dimension and do not have closed form. Analytic approximations to the integrals may not be sufficient accurate. Therefore, it is prohibitive to directly calculate the posterior distribution of $\boldsymbol{\theta}$ based on the observed data. As an alternative, MCMC procedures can be used to sample based on (3.8) using the Gibbs sampler along with the Metropolis-Hastings (M-H) algorithm. The above representations based on the models are useful as it allows to implement easily using the WinBUGS codes (Lunn et al., 2000).

3.3. Analysis of AIDS clinical data

3.3.1. Data and models

We illustrate our methods using a real AIDS clinical data. The study consists of 44 HIV-infected patients who were treated with a potent ARV regimen. Viral load, CD4 cell count and other variables were repeatedly measured over a period of 24 weeks. RNA viral load was measured in copies/mL at study days 0, 7, 14, 28, 56, 84, 112, 140 and 168 of follow-up. The nucleic acid sequence-based amplification assay (NASBA) was used to measure plasma HIV-1 RNA, with a lower limit of quantification, 50 copies/mL. the HIV-1 RNA measures below this limit are not considered reliable, therefore we simply imputed such values as 25 copies/mL (Acosta et al., 2004; Davidian and Giltinan, 1995). Covariates such as CD4 cell count including associated baseline data were also measured throughout the study on similar schemes. Figure 3.2 shows the measurements of viral load in natural log scale and CD4 cell count for the three randomly selected patients. Both viral load and CD4 cell count trajectories exhibit distinctive and important patterns throughout the time course. We can see that the rate change in viral load appears to vary substantially across patients, reflecting both biological variation and systematic associations with subject-level covariates.

The baseline value of HIV-1 RNA in plasma (ln scale) ranged from 5.296 to 11.582 with mean 8.715 and standard deviation 1.531. As is evident from Figure 3.3(c), RNA levels varied widely during the early treatment stage. For some patients, the RNA level decreased rapidly with treatment (described as the first-phase decay rate); for others it decreased slowly. Studying the relation between baseline RNA and the first-phase decay rate will provide useful information for clinical decision-making and treatment individualization. The first-phase decay rate indicates the potency of ARV therapies (Ding and Wu, 1999). If we know the potency of a treatment at an earlier stage, we may be able to avoid the less potent regimens for a particular patient. This will help clinicians

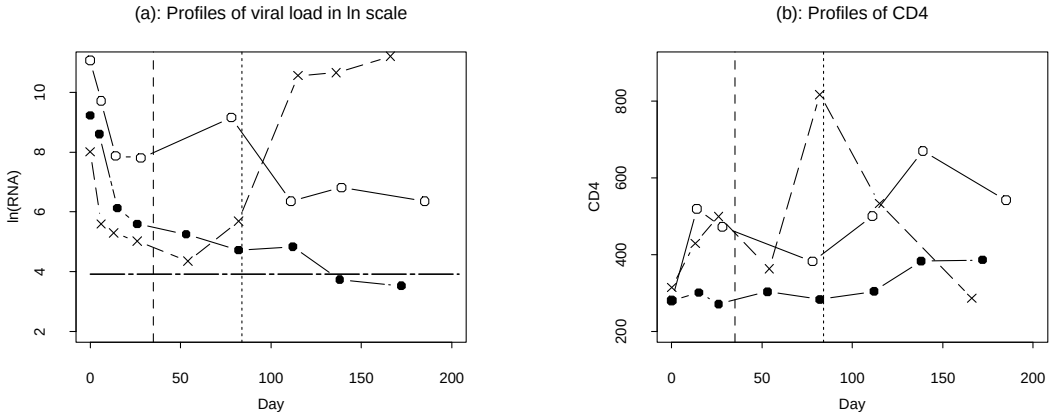

Figure 3.2: Profiles of viral load (response) in natural $\log$ scale and CD4 cell count (covariate) for three randomly selected patients. The horizontal line is below the detectable level of viral load (3.91=$\log(50)$) and the two vertical lines represent truncation days 35 and 84, respectively.

to select a treatment for their patients. Although the patients may receive the same treatment, there may still exist the difference in the potency for different patients with the same regimen. This is because the patients may absorb the drug differently, and patients' immune systems and other factors related to response of the drugs may be different.

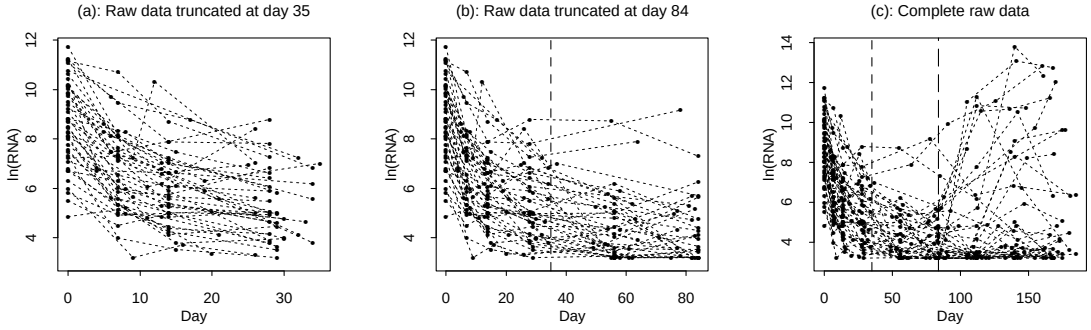

Figure 3.3: Profiles of viral load in $\ln$ scale from an AIDS clinical trial study. Change in HIV-1 load, measured from RNA levels in plasma, with time during treatment from day 0 to (a) day 35, (b) day 84 and (c) the end of study period.

Viral dynamic models can be formulated through a system of ordinary differential equations (ODE) (Huang et al., 2006; Wu et al., 1998; Wu and Ding, 1999). In practice, some investigators have used a LME model simplified from an ODE system to fit viral dynamic data from the very early time period such as displayed in Figure 3.3(a). Their response model can be described by the

following linear model

$$y(t) \quad = \quad \ln\{V(t)\} = \beta_1 - \beta_2 t + e(t) \tag{3.9}$$

where $y(t)$ is the natural logarithm of the number of HIV-1 RNA copies per mL of plasma, $e(t)$ is the measurement error, β_2 is the first-phase viral decay rate which may represent the minimum turnover rate of productively infected cells (Perelson et al., 1997) and β_1 is macro-parameter with $\exp(\beta_1)$ being the baseline viral load at time $t = 0$.

Due to the limitations of current assays, only two infected cell compartments can be identified which are believed to produce a biphasic viral decay (Perelson et al., 1997). Based on biological and clinical arguments, an effective model used to estimate viral dynamic parameters is the biphasic model approximated from a compartmental analysis-based an ODE system (Perelson et al., 1997; Wu and Ding, 1999). This model plays an important role in modeling HIV dynamics and has demonstrated promise in studying HIV response process. The model is described as follows.

$$y(t) \quad = \quad \ln\{V(t)\} = \ln\left\{\exp[\beta_1 - \beta_2 t] + \exp[\beta_3 - \beta_4 t]\right\} + e(t) \tag{3.10}$$

where β_4 is the second-phase viral decay rate which may represent the minimum turnover rate of latently or long-lived infected cells (Perelson et al., 1997) and $\exp(\beta_1) + \exp(\beta_3)$ is the baseline viral load at time $t = 0$. It is generally assumed that $\beta_2 > \beta_4$, which assures that the model is identifiable and is appropriate for empirical studies. It is of particular interest to estimate the viral decay rates β_2 and β_4 because they quantify the antiviral effect and hence can be used to assess the efficacy of the antiviral treatments (Ding and Wu, 1999). In estimating these decay rates, only the early segment of the viral load trajectory data (for example, data shown in Figure 3.3(b)) can be used (Perelson et al., 1997; Wu and Ding, 1999), because the viral load trajectory may have a different shape in later stages (see Figure 3.3(c)).

Although the models (3.9) and (3.10) are widely used in HIV dynamic studies and have shown promise, there are some concerns. For example, when different models give different conclusions, how do we know which is correct and should be used? In our analysis of the clinical data shown above, the models (3.9) and (3.10) produce incongruous estimates of viral decay rates of β_2 and provide conflicting results on their correlations with baseline viral load: one indicates a strongly positive correlation between baseline HIV-1 RNA levels and the first-phase decay rate and the other indicates that these two factors are negatively correlated. In addition, both models are applied to the early segment of the viral load data. That means one has to cut the data at some arbitrary time

point. It is not obvious what time point should be the cut-point or whether we should use different cut-points. To avoid these problems, a semiparametric biexponential model was proposed as follows (Wu and Zhang, 2002).

$$y(t) \quad = \quad \ln\{V(t)\} = \ln\left\{\exp[\beta_1 - \beta_2 t] + \exp[\beta_3 - \beta_4(t)t]\right\} + e(t) \tag{3.11}$$

where the second-phase decay rate $\beta_4(t)$ is a smooth unknown function. Intuitively, model (3.11) is more reasonable because it assumes that the viral decay rate in the second-phase can vary with time as a result of drug resistance, pharmacokinetics, medication adherence and other relevant clinical factors likely to affect changes in the viral load during the late stage of treatment. Therefore, all data obtained during ARV treatment can be used by fitting model (3.11). We also assume that $\beta_2 > \beta_4(t)$ for all time in order to guarantee that there is the first phase of curve decay. This is a semiparametric model because of the mechanistic structure (two-exponential) with constant parameters ($\beta_1, \beta_2, \beta_3$) and a time-varying parameter ($\beta_4(t)$) to capture the time-varying effects of the treatment and over a longer period. This semiparametric model preserves compartmental mechanistic interpretation (Perelson et al., 1997; Wu and Ding, 1999) of the original parametric model under the biexponential form. In particular, the turnover rate of productively infected cells, β_2, can still be estimated. Actually, by including long-term viral load data, the estimate of β_2 may be more accurate and reasonable compared with those obtained in previous studies (Perelson et al., 1997; Wu and Ding, 1999) after excluding long-term viral load data for modeling and analysis by some *ad hoc* rules (that is, the screening and inclusion of viral load data are quite arbitrary). In the mean time, this model enjoys the flexibility of a semiparametric function for the second-phase decay rate $\beta_4(t)$. The estimate of $\beta_4(t)$ provides not only an approximate turnover rate over time of long-lived/latently infected cells at the early stage of treatment as the standard parametric model does, but also more importantly describes how it may change over a long treatment period as driven by, presumably, drug resistance, non-compliance and other clinical determinants. Most importantly, the semiparametric model is capable of modeling long-term viral load data of which the trajectory may vary substantially among different patients (Wu and Zhang, 2002). It is noted that the three different models are applied to different segments of the viral dynamic data. Therefore, the standard goodness-of-fit or model selection methods cannot be used to identify the appropriate model.

To characterize skewness appeared often in viral loads and CD4 cell count, we will develop

SN Bayesian mixed-effects joint models in conjunction with the three dynamic response models and LME covariate model. To model the covariate CD4 process, we consider empirical polynomial LME models and choose the best model (3.1) with quadratic ($l = 3$) based on AIC/BIC model selection criteria as suggested by Liu and Wu (2007) and Wu (2002). We thus adapted the quadratic polynomial as the SN covariate model (3.1) with $\boldsymbol{u}_{ik} = \boldsymbol{v}_{ik} = (1, s_{ik}, s_{ik}^2)^T$ for the CD4 trajectory as follows.

$$z_{ik} = z_{ik}^* + \epsilon_{ik}, \;\; \boldsymbol{\epsilon}_i \sim SN_{m_i}\left(\mathbf{0}, \tau^2 \boldsymbol{I}_{m_i}, \delta_\epsilon \boldsymbol{I}_{m_i}\right), \tag{3.12}$$

where $z_{ik}^* = (\alpha_1 + a_{i1}) + (\alpha_2 + a_{i2})s_{ik} + (\alpha_3 + a_{i3})s_{ik}^2$, $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3)^T$ is population (fixed-effects) parameter vector, and $\boldsymbol{a}_i = (a_{i1}, a_{i2}, a_{i3})^T$ is individual-specific (random-effects) vector with normal distribution $N_3(\mathbf{0}, \boldsymbol{\Sigma}_a)$. Special cases of the model (3.2), which are offered to jointly model HIV dynamics in the presence of CD4 covariate process with measurement errors described in the model (3.12), are discussed below.

Model I: SN Bayesian semiparametric nonlinear mixed-effects (SN-BSNLME) joint model (3.5) based on the semiparametric biexponential model (3.11) in conjunction with the SN covariate model (3.12) along with specified prior distributions (3.6) can be expressed as

$$\begin{aligned} y_{ij} &= \ln\{\exp[\beta_{i1} - \beta_{i2}t_{ij}] + \exp[\beta_{i3} - \beta_{ij4}(t_{ij})t_{ij}]\} + e_{ij}, \;\; \boldsymbol{e}_i \sim SN_{n_i}\left(\mathbf{0}, \sigma^2 \boldsymbol{I}_{n_i}, \delta_e \boldsymbol{I}_{n_i}\right), \\ \beta_{i1} &= \beta_1 + b_{i1}, \;\; \beta_{i2} = \beta_2 + b_{i2}, \;\; \beta_{i3} = \beta_3 + b_{i3}, \;\; \beta_{ij4}(t_{ij}) = \beta_4 + \beta_5 z_{ij}^* + w(t_{ij}) + h_i(t_{ij}), \end{aligned} \tag{3.13}$$

where $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \boldsymbol{\mu}_p^T)^T$, $\boldsymbol{b}_i = (b_{i1}, b_{i2}, b_{i3}, \boldsymbol{\xi}_{iq}^T)^T \sim N_{3+q}(\mathbf{0}, \boldsymbol{\Sigma}_b)$; see equation (3.16) below for detailed specification about $\boldsymbol{\mu}_p$ and $\boldsymbol{\xi}_{iq}$. We can see that the SN-BSNLME joint model above reverts to an SN Bayesian nonlinear mixed-effects (SN-BNLME) models when the nonparametric parts $w(t)$ and $h_i(t)$ become constants. More specifically, the SN-BNLME model reduces to an SN Bayesian linear mixed-effects (SN-BLME) model when the response function is linear. Thus, we next present the following two simplified mixed-effects models.

Model II: SN Bayesian Nonlinear Mixed-Effects (SN-BNLME) joint model based on the biexponential model (3.10) in conjunction with the SN covariate model (3.12) can be expressed as

$$\begin{aligned} y_{ij} &= \ln\{\exp[\beta_{i1} - \beta_{i2}t_{ij}] + \exp[\beta_{i3} - \beta_{ij4}t_{ij}]\} + e_{ij}, \;\; \boldsymbol{e}_i \sim SN_{n_i}\left(\mathbf{0}, \sigma^2 \boldsymbol{I}_{n_i}, \delta_e \boldsymbol{I}_{n_i}\right), \\ \beta_{i1} &= \beta_1 + b_{i1}, \;\; \beta_{i2} = \beta_2 + b_{i2}, \;\; \beta_{i3} = \beta_3 + b_{i3}, \;\; \beta_{ij4} = \beta_4 + \beta_5 z_{ij}^* + b_{i4}, \end{aligned} \tag{3.14}$$

where $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3, \beta_4, \beta_5)^T$ and $\boldsymbol{b}_i = (b_{i1}, b_{i2}, b_{i3}, b_{i4})^T \sim N_4(\mathbf{0}, \boldsymbol{\Sigma}_b)$.

Model III: SN Bayesian Linear Mixed-Effects (SN-BLME) model based on the linear model (3.9) can be written as

$$\begin{aligned} y_{ij} &= \beta_{i1} - \beta_{i2} t_{ij} + e_{ij}, \;\; \boldsymbol{e}_i \sim SN_{n_i}\left(\mathbf{0}, \sigma^2 \boldsymbol{I}_{n_i}, \delta_e \boldsymbol{I}_{n_i}\right), \\ \beta_{i1} &= \beta_1 + b_{i1}, \;\; \beta_{i2} = \beta_2 + b_{i2}, \end{aligned} \tag{3.15}$$

where $\boldsymbol{\beta} = (\beta_1, \beta_2)^T$ and $\boldsymbol{b}_i = (b_{i1}, b_{i2})^T \sim N_2(\mathbf{0}, \boldsymbol{\Sigma}_b)$.

3.3.2. Results of analysis

In this section, we report the results of our analysis of the three data sets (mentioned in Figure 3.3) using SN-BLME, SN-BNLME and SN-BSNLME joint models, respectively. A natural $\log$ transformation for viral load data was used in the analysis in order to stabilize the variation of measurement error and speed up estimation algorithm. To avoid very small (large) estimates which may be unstable, we standardize the time-varying covariate CD4 cell count (each CD4 value is subtracted by mean 375.46 and divided by standard deviation 228.57) and rescale the original time t (in days) so that the time scale is between 0 and 1. To fit the SN-BLME model and the SN-BNLME joint model, we included only the viral load data from day 0 to day 35 (Figure 3.3(a)) and day 0 to day 84 (Figure 3.3(b)), respectively, because the SN-BLME model could only be used to fit linear trajectories of viral load and the SN-BNLME assumptions might be violated after long-term treatment if there are rebounds of viral load (i.e., we excluded the significant rebound data from the analysis). In fitting the SN-BSNLME joint model, we use all viral load data shown in Figure 3.3(c) and employ the linear combinations of natural cubic splines with percentile-based knots to approximate the nonparametric functions $w(t)$ and $h_i(t)$. Following studies in (Liu and Wu, 2007; Wu and Zhang, 2002), we set $\psi_0(t) = \phi_0(t) \equiv 1$ and take the same natural cubic splines in the approximations (3.3) with $q \leq p$. The values of p and q are determined based on the AIC/BIC which suggest the following function for $\beta_{ij4}(t_{ij})$ with $p = 3$ and $q = 1$ in the model (3.13).

$$\beta_{ij4}(t_{ij}) \approx \beta_4 + \beta_5 z^*_{ij} + \mu_0 \psi_0(t_{ij}) + \mu_1 \psi_1(t_{ij}) + \mu_2 \psi_2(t_{ij}) + \xi_{i0}. \tag{3.16}$$

To carry out Bayesian inference, we need to specify the values of the hyper-parameters in the prior distributions. In Bayesian approach, we only need to specify the priors at the population level. We take weakly informative prior distributions for all the parameters. In particular, (i) fixed-effects were taken to be independent normal distribution $N(0, 100)$ for each component of the population parameter vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. (ii) For the scale parameters σ^2 and τ^2 we assume a limiting non-

informative inverse gamma prior distribution, $IG(0.01, 0.01)$ so that the distribution has mean 1 and variance 100. (iii) The priors for the variance-covariance matrices of the random-effects $\mathbf{\Sigma}_a$ and $\mathbf{\Sigma}_b$ are taken to be inverse Wishart distributions $IW(\mathbf{\Omega}_1, \nu_1)$ and $IW(\mathbf{\Omega}_2, \nu_2)$, where degree of freedom $\nu_1 = \nu_2 = 5$, and $\mathbf{\Omega}_1$ and $\mathbf{\Omega}_2$ are diagonal matrices with diagonal elements being 0.01. (iv) For each of the skewness parameters δ_e and δ_ϵ, we choose independent normal distribution $N(0, 100)$, where we specify that $\boldsymbol{\delta}_{e_i} = \delta_e \mathbf{1}_{n_i}$ and $\boldsymbol{\delta}_{\epsilon_i} = \delta_\epsilon \mathbf{1}_{m_i}$ to indicate that we are interested in skewness of both overall viral load data and overall CD4 cell count data.

The MCMC sampler was implemented using WinBUGS software (Lunn et al., 2000), and the program codes are available in Appendix B. In particular, the MCMC scheme for drawing samples from the posterior distributions of all parameters in the both response and covariate models is obtained by iterating between the following two steps: (i) Gibbs sampler is used to update $\boldsymbol{\alpha}, \boldsymbol{\beta}, \tau^2, \sigma^2, \mathbf{\Sigma}_a, \mathbf{\Sigma}_b, \delta_\epsilon, \delta_e$; (ii) we update $\boldsymbol{b}_i$ and $\boldsymbol{a}_i$ ($i = 1, 2, \cdots, n$) using the Metroplis-Hastings (M-H) algorithm. After collecting the final MCMC samples, we are able to draw statistical inference for the unknown parameters. Specifically, we are interested in the posterior means and quantiles. See the articles (Huang et al., 2006; Lunn et al., 2000) for detailed discussions of the Bayesian modeling approach and the implementation of the MCMC procedures, including the choice of the hyper-parameters, the iterative MCMC algorithm, the choice of proposal density related to M-H sampling, sensitivity analysis and convergence diagnostics. When the MCMC implementation is applied to the actual clinical data, convergence of the generated samples is assessed using standard tools within WinBUGS software (such as trace plots). After convergence was achieved, one long chain was run which may be more efficient. We propose that, after an initial number of 50,000 burn-in iterations, every 40th MCMC sample is retained from the next 400,000. Thus we obtain 10,000 samples of targeted posterior distributions of the unknown parameters for statistical inference.

We will investigate the following two scenarios. Firstly, As shown in Figure 3.1, the histograms of viral load and CD4 cell count clearly indicate their asymmetric nature and it seems adequate fitting an SN model to the data set. Since a normal distribution is a special case of an SN distribution when skewness parameter is zero, we will investigate how an SN distribution for random error contributes to modeling results and parameter estimation in comparison with a normal distribution for random error. Secondly, we also estimate the model parameters by using the 'naive' method, which does not separate the measurement errors from the true CD4 values (i.e., completely

ignores measurement errors in CD4 values in the modeling). That is, the 'naive' method only uses the observed CD4 values z_{ij} rather than true (unobservable) CD4 values z_{ij}^* in the response Models I–II in which case the joint models revert to regular models without covariate models involved for inference. We use the 'naive' method as a comparison to the joint modeling method proposed in Section 3.2. This comparison attempts to investigate how the measurement errors in CD4 contribute to parameter estimation.

3.3.2.1. **Comparison of results between models with normal and SN distributions**

As discussed previously, the fundamental assumption of distributions for (within-subject) random error in Models I–III is SN which is different from that of normal distribution in most statistical models in publications (Liu and Wu, 2007; Perelson et al., 1997; Wu et al., 1998; Wu and Ding, 1999; Wu and Zhang, 2002). In this section, we investigate how an SN distribution contributes to modeling results in comparison with a normal distribution for random error.

The population posterior mean (PM), the corresponding standard deviation (SD) and 95% credible interval (CI) for fixed-effect parameters based on the models with an SN or normal random errors are presented in Table 3.1. The following findings are observed based on the estimated results. (i) Firstly, in the response models for the most interested parameters $(\beta_2, \beta_4, \beta_5)$, β_2 based on the three models with a normal random error are larger than the corresponding that based on the three models with an SN random error; all the estimates are statistically significant since the 95% CIs do not contain zero. Secondly, for (β_4, β_5), it can be seen that the estimates based on the SN-BNLME and SN-BSNLME models are substantially different from those based on the N-BNLME and N-BSNLME models. Thirdly, for the variance parameter σ^2, its estimated values (0.04, 0.08, 0.09) based on the three models with a SN random error are much smaller than those (0.96, 0.54, 1.31) based on the three models with a normal random error. Finally, for the skewness parameter δ_e, we found that δ_e associated with the three models with an SN random error is estimated to be significantly positive; this confirms the positive skewness of the viral load data in $\ln$ scale as shown in Figure 3.1, and the estimates of the skewness parameter δ_e based on SN-BLME model (1.57), SN-BNLME model (1.15) and SN-BSNLME model (2.03) are fairly high. (ii) For estimated parameters in the CD4 covariate models, the estimates of intercept α_1 based on the N-BNLME and N-BSNLME models are larger than those based on SN-BNLME and SN-BSNLME models, however the estimates of coefficients α_2 and α_3 are very similar between SN and normal models. For the

variance parameter τ^2, the estimated values (0.10, 0.13) based on the N-BNLME and N-BSNLME models are larger than those (0.01, 0.08) based on SN-BNLME and SN-BSNLME models. The estimates of the skewness parameter δ_ϵ based on SN-BNLME and SN-BSNLME models are significantly positive which is consistent with positive skewness of the CD4 cell count data as shown in Figure 3.1.

Table 3.1: A summary of the estimated posterior mean (PM) of interested population (fixed-effects) and precision parameters

Model		α_1	α_2	α_3	β_1	β_2	β_3	β_4	β_5	τ^2	σ^2	δ_ϵ	δ_e
N-BLME	PM	–	–	–	8.01	3.70	–	–	–	–	0.96	–	–
	L_{CI}	–	–	–	7.52	3.16	–	–	–	–	0.74	–	–
	U_{CI}	–	–	–	8.51	4.22	–	–	–	–	1.25	–	–
	SD	–	–	–	0.25	0.27	–	–	–	–	0.13	–	–
SN-BLME	PM	–	–	–	5.42	3.43	–	–	–	–	0.04	–	1.57
	L_{CI}	–	–	–	4.80	2.85	–	–	–	–	0.01	–	1.34
	U_{CI}	–	–	–	6.03	4.04	–	–	–	–	0.18	–	1.83
	SD	–	–	–	0.31	0.30	–	–	–	–	0.05	–	0.13
N-BNLME	PM	-0.20	0.68	-0.42	8.52	21.0	5.59	1.26	0.04	0.10	0.54	–	–
	L_{CI}	-0.47	0.15	-0.98	7.96	16.6	5.01	0.60	0.01	0.08	0.40	–	–
	U_{CI}	0.07	1.19	0.14	9.09	26.3	6.15	1.91	0.33	0.13	0.73	–	–
	SD	0.14	0.26	0.28	0.28	2.44	0.29	0.33	0.19	0.13	0.08	–	–
SN-BNLME	PM	-1.01	0.70	-0.50	6.70	20.9	3.82	1.35	0.02	0.01	0.08	0.53	1.15
	L_{CI}	-1.30	0.24	-0.99	5.97	16.6	3.03	0.64	0.01	0.01	0.01	0.38	0.79
	U_{CI}	-0.70	1.16	0.04	7.50	26.0	4.60	1.97	0.22	0.05	0.33	0.65	1.42
	SD	0.16	0.24	0.27	0.39	2.39	0.39	0.34	0.15	0.01	0.09	0.07	0.16
N-BSNLME	PM	-0.22	0.66	-0.28	8.29	26.6	3.66	-2.56	0.88	0.13	1.31	–	–
	L_{CI}	-0.48	0.13	-0.85	7.75	17.5	1.12	-10.6	0.31	0.10	1.06	–	–
	U_{CI}	0.03	1.20	0.29	8.83	36.2	5.27	2.90	1.77	0.15	1.62	–	–
	SD	0.13	0.27	0.29	0.28	4.77	1.02	3.33	0.45	0.01	0.14	–	–
SN-BSNLME	PM	-0.62	0.65	-0.27	4.60	17.1	-0.09	-4.44	0.15	0.08	0.09	0.25	2.03
	L_{CI}	-1.13	0.10	-0.85	3.76	9.97	-2.76	-10.9	0.06	0.04	0.01	0.16	1.64
	U_{CI}	0.37	1.19	0.31	5.63	23.0	1.50	-0.60	0.94	0.13	0.38	0.52	2.34
	SD	0.41	0.27	0.29	0.47	3.41	1.14	2.83	0.39	0.03	0.10	0.25	0.17
SN-BSNLME	PM	–	–	–	5.34	28.1	1.57	0.39	0.78	–	0.12	–	1.84
(NV)	L_{CI}	–	–	–	4.55	20.7	0.22	-3.98	0.17	–	0.01	–	1.45
	U_{CI}	–	–	–	6.18	36.7	2.73	3.57	1.37	–	0.43	–	2.18
	SD	–	–	–	0.41	3.99	0.64	1.92	0.32	–	0.12	–	0.18

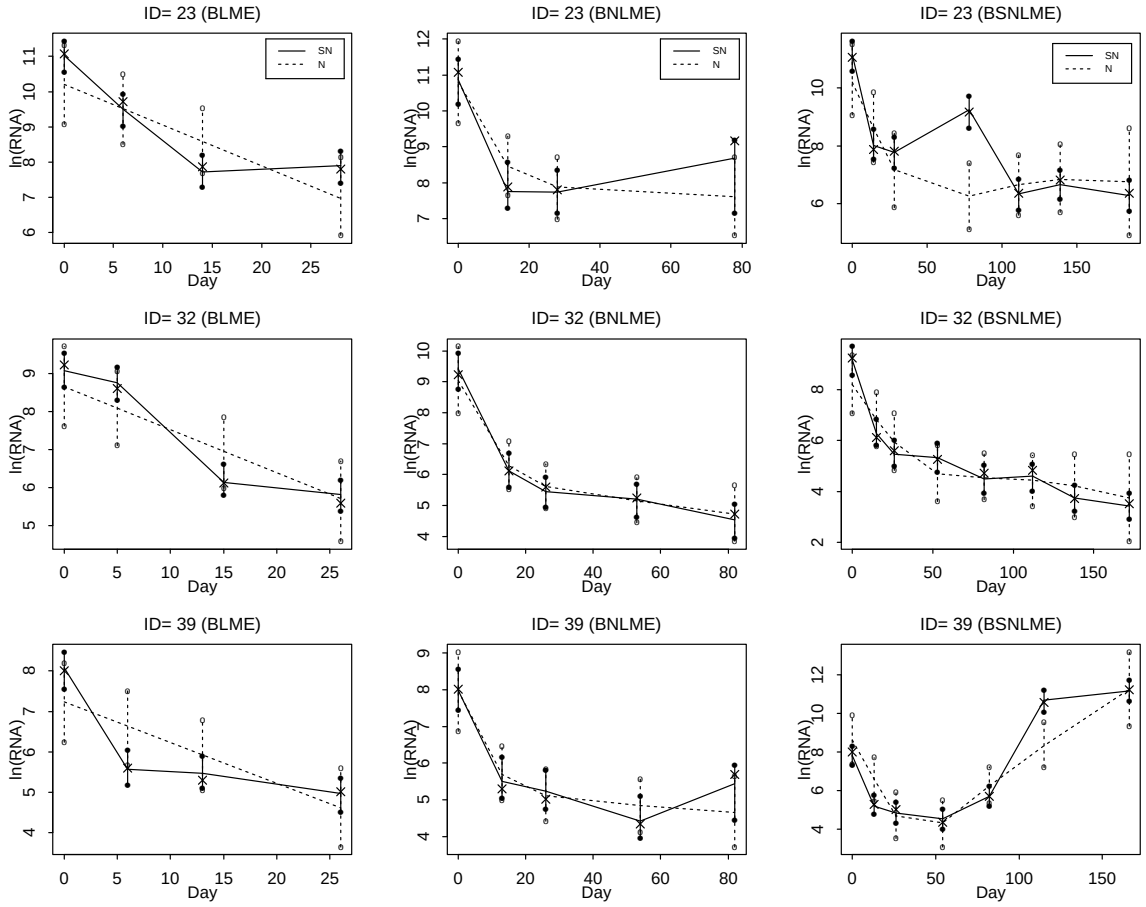

Figure 3.4: The individual estimates of viral load trajectories for three randomly selected patients based on the BLME (left), BNLME (center) and BSNLME (right) models with a normal (dotted line) or SN (solid line) random errors. The respective vertical dotted line (normal) ended with '$\circ$' and solid line (SN) ended with '$\bullet$' on each fitted value are the 95% credible interval (CI) associated with the fitted value. The observed values are indicated by sign crosses ($\times$).

Figure 3.5 displays the three randomly selected individual estimates of viral load trajectories along with the associated 95% CIs on each fitted value obtained based on the BLME (left), BNLME (center) and BSNLME (right) models with a Normal (dotted line) or SN (solid line) random error, respectively. The following findings are observed from joint modeling results. (i) The estimated individual trajectories for the models where the random error is assumed to be SN fit the originally observed values much closer than those for the corresponding models where the random error is assumed to be normal. (ii) Overall, the 95% CI associated with each of fitted values from the normal models is wider than that from the corresponding SN models. (iii) All the 95% CIs from

three SN models cover the true (observed) viralr load values, while some of 95% CIs from three normal models do not. For example, for patient 39 whose observed value at day 115 is 10.57, the 95% CI from the SN-BSNLME model is (9.90,11.02) with the fitted value 10.51, while the corresponding 95% CI from the N-BSNLME model is (7.19, 9.53) with the fitted value 8.33 which does not cover the observed value 10.57.

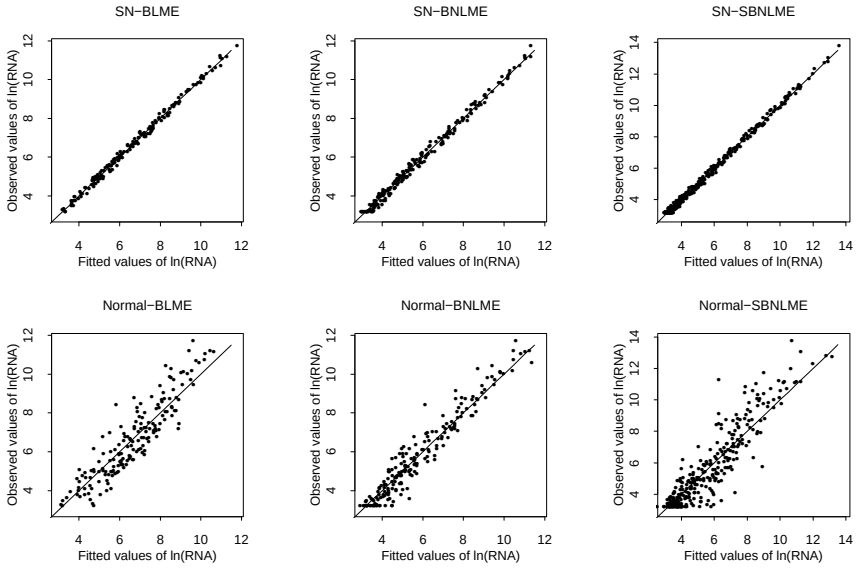

Figure 3.5: The observed values versus fitted values of $\ln$(RNA) based on the BLME (left), BNLME (center) and BSNLME (right) models with a normal or SN random error.

We also investigate the model fitting results for each of the three mixed-effects models with normal and SN random errors, respectively. We have seen that, in general, all the models provided a reasonably good fit to the observed data for most patients in our study, although the fitting for a few patients was not completely satisfactory due to unusual viral load fluctuation patterns for these patients. To assess the goodness-of-fit of each of the three mixed-effects models with normal and SN random errors, the diagnostic plots of the observed values versus the fitted values are presented in Figure 3.5. It was seen from Figure 3.5 that the models where the random error is assumed to be SN provided better fit to observed data, compared with the models where the random error is assumed to be normal. This finding is further confirmed by their residual sum of squares (RSS). That is, for the BLME model the RSSs are 3.62 (SN random error) and 117.37 (normal random error); for the BNLME model the RSSs are 6.64 (SN random error) and 63.78 (normal random

error); for the BSNLME model the RSSs are 7.49 (SN random error) and 312.58 (normal random error).

For selecting the best model that fits the data adequately, a Bayesian selection criterion is used. This criterion, known as deviance information criterion (DIC), was first suggested in a recent publication by Spiegelhalter et al.(2002). As with other model selection criteria, we caution that DIC is not intended for identification of the 'correct' model, but rather merely as a method of comparing a collection of alternative formulations. In each of the three models with the specification of different distributions for the random error, DIC can be used to find out how assumption of an SN distribution contributes to virologic responses and parameter estimation in comparison with that of a normal distribution. We found that the DIC values, 335.14 (SN-BLME), 607.24 (SN-NLME) and 1051.21 (SN-SNLME) for the three models with an SN random error are smaller than the corresponding ones, 524.54 (normal-BLME), 701.17 (normal-NLME) and 1355.45 (normal-SNLME) for the three models with a normal random error, respectively.

As mentioned before, it is hard to tell which model is 'correct' but which one fits data better. Therefore, based on the DIC criterion, the results indicate that each of the three models with an SN random error fits the data better, supporting the contention of a departure from normality. These results are consistent with those from diagnosis of the goodness-of-fit displayed in Figure 3.5. In summary, our results may suggest that it is very important to assume an SN distribution for the response models in order to achieve reliable results, in particular if data exhibit skewness. Along with these observations, we further report our results in details only for the three models with an SN random error.

3.3.2.2. **Results of analysis based on the SN models**

The population (average) estimates of the viral dynamic parameters presented in Table 4 based on the three (SN-BLME, SN-BNLME, SN-BSNLME) models indicate that the estimates of β_1 from the different models agree fairly well. However, the estimates (20.9 and 17.1) of the first decay rate β_2 by SN-BNLME and SN-BSNLME modeling methods are significantly different from that (3.43) obtained by SN-BLME modeling method. Although the estimates of β_2 by SN-BNLME and SN-BSNLME modeling methods are comparable, one problem is that we considered only 84-day data for SN-BNLME model fit. This means that only 68% of the data from the 168-day period were included due to arbitrary truncation of data. Therefore, the SN-BNLME modeling may not

be efficient. In this case, we prefer to use the SN-BSNLME model in which a smooth function of treatment time is incorporated into the second-phase decay rate to better catch rebound viral load data and all data measured during the treatment period can be used.

In the SN-BLME and SN-BNLME model fittings, individual curves (solid lines in Figure 3.5) for each subject follow a similar trend; that is, the trajectories of viral load decay in all 3 subjects decrease rapidly in the first-phase, then flatten in general. When the entire treatment period is considered, the viral loads of subject 39 rebound after the second-phase, whereas the viral loads of subjects 23 and 32 remain low until the end of the treatment period. Obviously, the SN-BLME and SN-BNLME model fittings are reasonable for data cutting at days 35 and 84, but they do not represent data measured over the whole treatment period.

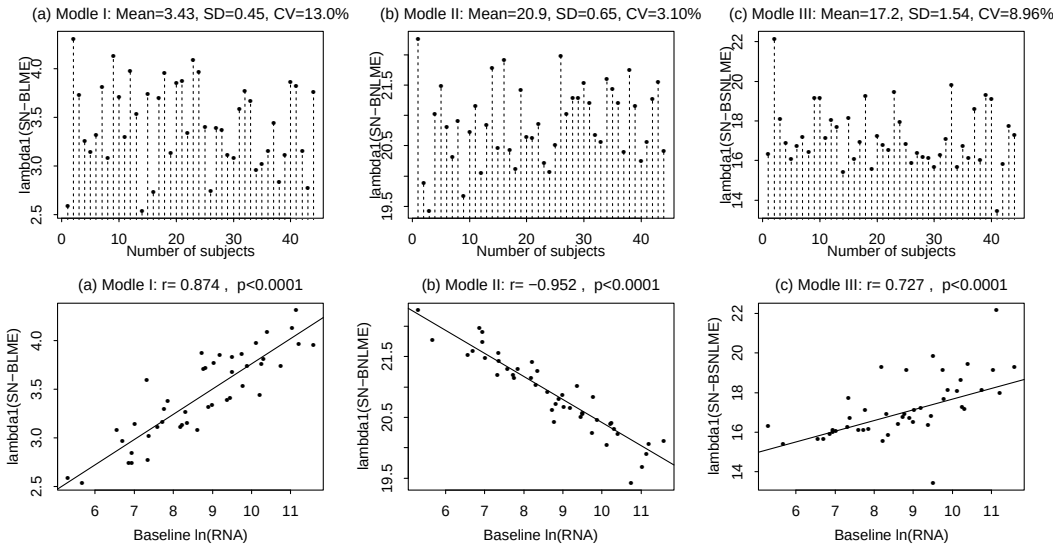

Figure 3.6: Correlations between baseline $\ln$(RNA) levels and the subject-specific first phase viral decay rates, β_{i1} estimated by using each of the three different methods. The solid lines are robust (MM-estimator) linear regression fit. The correlation coefficients (r) and p-values were obtained from Spearman rank correlation test.

It is also worth noting the differences among the estimated values of the first-phase decay rate (β_2) for each subject based on the three models. We can see from Figure 3.6 (top panel) that the individual estimated values of β_2 obtained by the SN-BNLME model fitting are consistently much greater than those obtained by the SN-BLME model fitting, but are slightly different from those

obtained by the SN-BSNLME model fitting. Although for each model the individual estimates of $\ln$(RNA) levels approximately follow the observed values, the differences in β_2 values obtained with each model may suggest a completely different relation between β_2 and baseline viral loads. When we investigate the relation between the estimated individual first-phase decay rates β_2 and baseline $\ln$(RNA) levels, the results are incongruous. The correlations between the subject-specific viral decay rates β_2 estimated by each method and baseline $\ln$(RNA) levels are shown in Figure 3.6 (bottom panel). The subject-specific estimates of β_2 obtained from the SN-BSNLME and SN-BLME modeling methods show significantly positive correlations ($r_I = 0.727$ and $r_{III} = 0.874$ with p-value $p < 0.0001$) with baseline $\ln$(RNA) levels. However, the estimates of β_2 obtained by using the SN-BNLME modeling method are negatively correlated with baseline $\ln$(RNA) levels ($r_{II} = -0.952$ with p-value $p < 0.0001$).

The incongruity in the individual estimates of the first-phase viral decay rates and their correlations with baseline $\ln$(RNA) levels, as determined by the SN-BLME, SN-BNLME and SN-BSNLME modeling methods, is significantly different with the following two observed scenarios: (i) Although the individual estimates of β_2 obtained by the both SN-BLME and SN-BSNLME modeling methods are positively correlated with baseline $\ln$(RNA) levels, the the individual estimates of β_2 from the SN-BSNLME method are, in general, five times larger than those from the SN-BLME method. (ii) The individual estimates of β_2 by SN-BNLME and SN-BSNLME modeling methods are fairly comparable, but the correlations between baseline $\ln$(RNA) levels and the subject-specific viral decay rates β_2 estimated by these two methods are completely opposite. These inconsistences are presumably caused by data truncation. From the above results, it may suggest that the estimates obtained from the both SN-BLME and SN-BNLME models might be not reliable and the estimates based on the SN-BSNLME model may be favorable.

To fit the SN-BLME model, we truncated the data at day 35 in this study. However, it is not clear where one should cut the data between the first- and second-phases of decay. Also, different subjects may have different change points between the two phases. For example, truncation at day 35 may cause data from the second-phase to be included with first-phase data. It is for this reason that SN-BLME models underestimate the first-phase decay rates (β_2). The SN-BSNLME modeling method is preferable to the SN-BLME and SN-BNLME modeling methods, especially for sparse individual data. We believe that estimates obtained from the SN-BSNLME modeling method and

their correlations with baseline $\ln$(RNA) levels are reliable since the complete data are used. Conversely, the both SN-BLME and SN-BNLME model fittings may result in misleading conclusions, as shown above, perhaps because it is impossible to find a truncation point to data for these two model fittings that would be suitable for all patients. The estimated values of β_2 would be affected by the inclusion of second-phase data if truncation occurred too late, and by the loss of data if truncation was too early.

For comparison, as an example, we also employed the 'naive' method-based the SN-BSNLME model to estimate the model parameters presented in Table 3.1 using the observed CD4 values and ignoring the CD4 measurement errors. It can be seen from the estimated results that the estimates of the parameters from the naive method are significantly larger than those from the joint modeling method. It indicates that the naive method may produce overestimated results with substantial biases; in particular, the estimated covariate CD4 effect (β_5) from the naive method is 5 times greater than that from the joint modeling method. The joint modeling method appears to give larger SDs in most cases, probably because it incorporates the variation from fitting the CD4 covariate process. Further, the estimate of the model skewness parameter δ_e for the naive method is slightly smaller than that for the joint modeling method; this result suggests that the naive method may underestimate the skewness parameter due to ignoring measurement errors in CD4 values. Thus the difference of the naive estimates and the joint modeling estimates, due to whether or not ignoring potential CD4 measurement errors in conjunction with the SN-BSNLME model, indicates that it is important to take the measurement errors into account in the analysis when covariates are measured with errors.

3.4. Discussion

For viral dynamic models with skewness characteristics of viral load responses and CD4 measurement errors in covariate, we have investigated and compared the three Bayesian mixed-effects models with an SN distribution that may be preferred over those with a normal distribution in the sense that it produces less biased parameter estimates and provides better fit to observed data. The proposed method may have a significant impact on AIDS research because, in the presence of skewness in the data, appropriate statistical inference for HIV dynamics is important for making robust conclusions and reliable clinical decisions. Our proposed method is quite general and so can be used

to other applications. This kind of SN modeling approach is important in many biostatistical application areas, allowing accurate inference of parameters while adjusting for the data with skewness. The SN distribution is shown to provide an alternative to normal (symmetric) distribution that is often assumed in statistical models. The results indicate that with SN distribution assumption, there is potential to gain efficiency and accuracy in estimating certain parameters when the normality assumption does not hold in the data. The models considered in this paper can be easily fitted using MCMC procedure. Moreover, the proposed modeling approach is fitted using the WinBUGS package that is available publicly. This makes our approach quite powerful and accessible to practicing statisticians in the fields.

To estimate the viral dynamic parameters and study the relation of the baseline level of HIV-1 RNA in plasma with the decay rate of the first-phase of response to treatment, we compared the results of SN-BLME, SN-BNLME and SN-BSNLME model fittings, and found that the both SN-BLME and SN-BNLME model fittings in short-term dynamics may result in misleading conclusions due to data truncation. Of particular interest is that when long-term dynamics are considered, SN-BNLME may also become unreliable because of the complexity of the second-phase of decay. The foregoing results indicate that in a two-phase HIV dynamic model, the first decay rate may remain constant, while the second decay rate may change which depends on time-varying CD4 covariate during the period of study. The analysis results suggest that there may be a significantly positive correlation between the first-phase viral decay and the baseline HIV-1 RNA levels based on the SN-BSNLME modeling method. This finding is consistent with those reported in publications (Notermans et al.,1998; Wu et al., 2004). This positive correlation may be partially explained by the fact that the higher baseline viral load value, which is equivalent to the lower baseline CD4 value due to a negative relation between these two baseline factors, suggests a lower turnover rate of hymphocyte cells, which may lead to a positive correlation between the first-phase viral decay rate (β_2) and the baseline HIV-1 RNA levels. Higher baseline HIV-1 RNA levels reflect more productively infected cells distributed at different sites; thus, greater drug potency or exposure may be required to achieve a similar decay rate to that seen in patients with lower levels of viral replication. This finding is very interesting and clinically important. Since the viral decay rates may reflect the efficacy of ARV treatment, the lower baseline HIV-1 RNA levels may need less potent drug efficacy to suppress virus replication so that a strong treatment strategy is not necessary to avoid side-effect of

drug use. This may help improve understanding of the pathogenesis of HIV infection and evaluation of ARV treatments.

We would point out that the problems we have addressed in this paper cannot be resolved by the standard goodness-of-fit or model selection methods, which are usually used when applying different models to the same data set. From Figure 3.5, we can see that all three models (BLME, BNLME and BSNLME) fitted the corresponding data very well. Goodness-of-fit or model selection methods are unable to identify the right model. We have found that the different models should not be applied to the same, entire data set, but applied simultaneously to appropriate segments of the data set such that the results will be biologically meaningful.

In conclusion, BLME fitting may be misleading and its use should be avoided; BNLME fitting may work well but is subject to data truncation; BSNLME fitting works in a similar way to BNLME fitting but has no data-screening problems associated with its use. Care is necessary in the implementation of BNLME and BSNLME fittings. With the introduction of SN distribution in the models, the estimated results suggest that the skewness parameters in viral load and CD4 cell count are estimated to be significantly positive for each of the three models. This confirms the positive skewness of the viral load and CD4 data presented in Figure 3.1. Thus, we may conclude that accounting for significant skewness is required when one models a data which exhibits skewness.

This paper combines new technologies in mathematical modeling and statistical inference with advances in viral dynamics and ARV treatment to quantify complex HIV disease mechanisms. The complex nature of HIV/AIDS will naturally pose some challenges such as nonignorable missing data and data with detection limit problems. These complicated problems are beyond the purpose of this article, but a further study may be warranted. We are actively investigating these problems now. We hope that we could report these interesting results in the near future.

4 Bivariate linear mixed-effects models with an application to AIDS study using skew-elliptical distributions

4.1. Introduction

Multivariate or bivariate outcomes are used as primary endpoints in many longitudinal studies. For example in AIDS studies, not only the RNA viral load, but also the CD4 and CD8 cell count are measured (Acosta et al., 2004). HIV-1 infection results in a progressive destruction of immune function, which may be indicated by a decrease of CD4 cell count and an increase of CD8 cell count. Studies of changing immunologic CD4 and CD8 responses may be important to identifying indicators for quantifying treatment effect and to improving management of patient care. With this type of multiple outcomes (CD4 and CD8 cell count), the underlying statistical question is to estimate the functions that model their dependence on co-variates and to investigate the relationships between these functions. Similar clinical and epidemiological studies often generate clustered as well as longitudinal follow-up data with bivariate or multivariate outcomes as primary endpoints, which are routinely analyzed using multivariate linear mixed-effects models (Matsuyama and Ohashi, 1997).

In this chapter, we focus on a bivariate LME (BLME) model on the situation where two response variables (CD4 and CD8 cell count) are observed simultaneously for each subject to accommodate individual clustering within subjects as well as the correlation between bivariate measures. BLME can facilitate borrowing of strength across all subjects when assessing the effects of co-variates through treatment time, baseline age, treatment group, viral load at baseline and time-varying treatment efficacy, etc. on AIDS progression. Thus, in particular, a BLME model is used to estimate various parameters including the correlation coefficients between CD4 and CD8 cell count. CD4 cell count is an important indicator of the strength of the immune system. CD4 cells are ‘helper’ cells that lead the attack against infections and are considered as the HIV main target cells. CD8 cells are ‘cytotoxic’ cells and are inappropriately “on” or active in terms of the immune hormones secreted during HIV infection. The hyperdynamic or over-stimulated CD8 immune response, reflected by activation of CD8 subsets as well as elevated total CD8 cell count, may accelerate immune dysfunc-

tion and certain disease processes. The ratio of CD4 and CD8 cell count is important in monitoring the function of the immune system in patients who have viral infections or who have undergone tissue transplantation.

In traditional LME model analysis (Laird and Ware, 1982), the correlation due to clustered/repeated measures on a subject is usually accounted for through the inclusion of random-effects and within-subject measurement errors, which are often assumed to be normally distributed due to the mathematical tractability and computational convenience. While such an assumption makes data analysis amenable to popular software such as SAS and R/Splus, the usual fidelity to normality assumptions has been questionable (Ghidey et al., 2004; Verbeke and Lesaffre, 1996) when data exhibit non-normal behavior. A violation of the assumption could lead to misleading inferences. In fact, observed data in AIDS studies are often far from being "symmetric" and asymmetric patterns of observations usually occur. A common approach adopted for data analysis in these situations is reverting back to usual multivariate normality assumptions after suitable transformation of the response on a continuous scale (e.g. square-root transformation of CD4 and CD8 cell count). Although they may lead to reasonable empirical results, they may be avoided when a suitable alternative theoretical model is available because data transformation hinds underlying data generation mechanisms due to reduced information and often component-wise transformation does not lead to joint normality (Jara et al., 2008). Besides, the transformations are not universal, i.e. transforms used for one particular data may not be adapted for a different data. Moreover, the results may be difficult to interpret based on transformed data. This motivates researchers to consider exploration of a more general mixed effects framework that takes into account the flexibility in distributional assumptions of random-effects and measurement errors to produce robust parameter estimates. The term 'robustness' is quite extensive; here robustness is achieved with respect to parameter estimation.

Considerable research has been done by introducing more flexible parametric families that can accommodate normality departures (skewness and heavy tails) and hence eliminate the need of ad hoc data transformations (Azzalini and Capitanio, 1999). In the context of LME models, the random-effects distribution was relaxed using finite normal mixtures (Verbeke and Lesaffre, 1996), smoothing (Ghidey et al., 2004), a semi-nonparametric density (Zhang and Davidian, 2001) or a thick-tailed normal/independent (NI) densities (Rosa et al., 2003). Much of recent frequentist and

Bayesian advances in regression problems revolve around the attractive and popular skew-elliptical distributions (Azzalini and Capitanio, 1999, 2003; Azzalini and Dalla-Valle, 1996; Sahu et al., 2003). The related literature in this context is very rich (Arellano-Valle et al., 2006; Azzalini, 2005; De la Cruz, 2008; Lin, 2009) and the curious reader might choose to venture an entire monograph (Genton, 2004) dedicated to discuss recent developments. A common feature of these classes of models is that the normal linear mixed model is a special member in each class. In this chapter, we propose a parametric modeling of LME model for robust estimation using SE distributions under a Bayesian paradigm. The multivariate SE distributions used in this chapter are developed primarily from the multivariate SE density proposed by Sahu et al.(2003) for Bayesian regression problems and is different from the SE version proposed by Azzalini and Dalla-Valle (1996). However, the differences are only due to the various parameterizations (Arellano-Valle and Azzalini, 2006) used and an unification of all SE variants is presented by Arellano-Valle and Genton (2005). Recent Bayesian implementation of multivariate SE distributions (Jara et al., 2008) involves SN and ST densities using a conditional stochastic representation. Three distributions, N, SN and ST will be considered in this chapter.

The rest of the chapter proceeds as follows. In Section 4.2, we describe the data set that motivated this research and introduce BLME models. Section 4.3 presents the associated Bayesian inference method and related model comparison techniques. In Section 4.4, we apply the proposed method to the real data set described in Section 4.2 and report the analysis results. We conclude the chapter with discussion in Section 4.5.

4.2. Data and models with the skew-elliptical distributions

4.2.1. Motivating data set

The data set that motivated this research is from A5055 and some detail information about this data set can be found in Section 2.4.1. and 3.3.1. Besides the HIV-1 viral load in plasma, CD4 and CD8 cell count in peripheral blood were designed to be measured in cells/mL at the same schedule as HIV, which was days 0, 7, 14, 28, 56, 84, 112, 140 and 168 of follow-up.

CD4 and CD8 cell count: CD4 and CD8 cell count were measured in cells/mL at designed study days. The median value at the baseline (t = 0) is 262/mL for CD4 and 880/mL for CD8

cell. The exact day of CD4 and CD8 cell count measurements (not predefined study day) was used to compute study day in our analysis. It is noted that observed data in the AIDS studies are often far from being "symmetric" and skewed heavy tailed patterns in CD4 and CD8 cell count usually occur (Figure 4.1). Thus, an asymmetric distribution (such as SE) should be more appropriate than a symmetric distribution, and statistical analysis must take these features of the data into account. Figure 4.2 shows the trajectories of observed CD4 and CD8 cell count (in standardized scale) after the initiation of an ARV treatment for 44 patients.

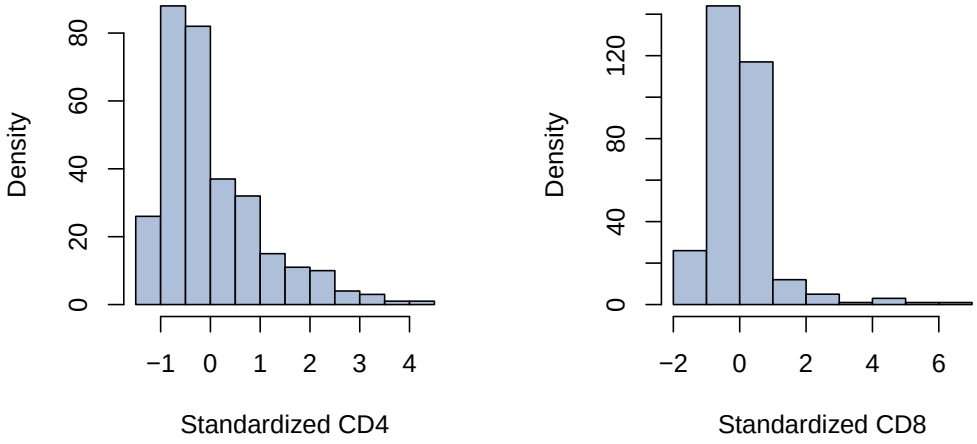

Figure 4.1: The histogram of CD4 and CD8 cell count (standardized scale) measured in peripheral blood for 44 patients.

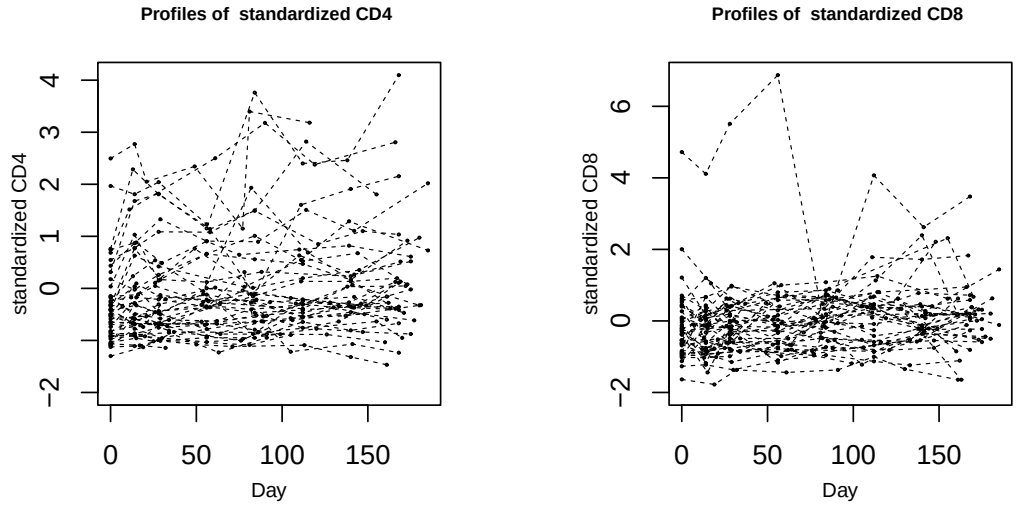

Figure 4.2: The trajectory profiles of CD4 and CD8 cell count (standardized scale) measured in peripheral blood for 44 patients.

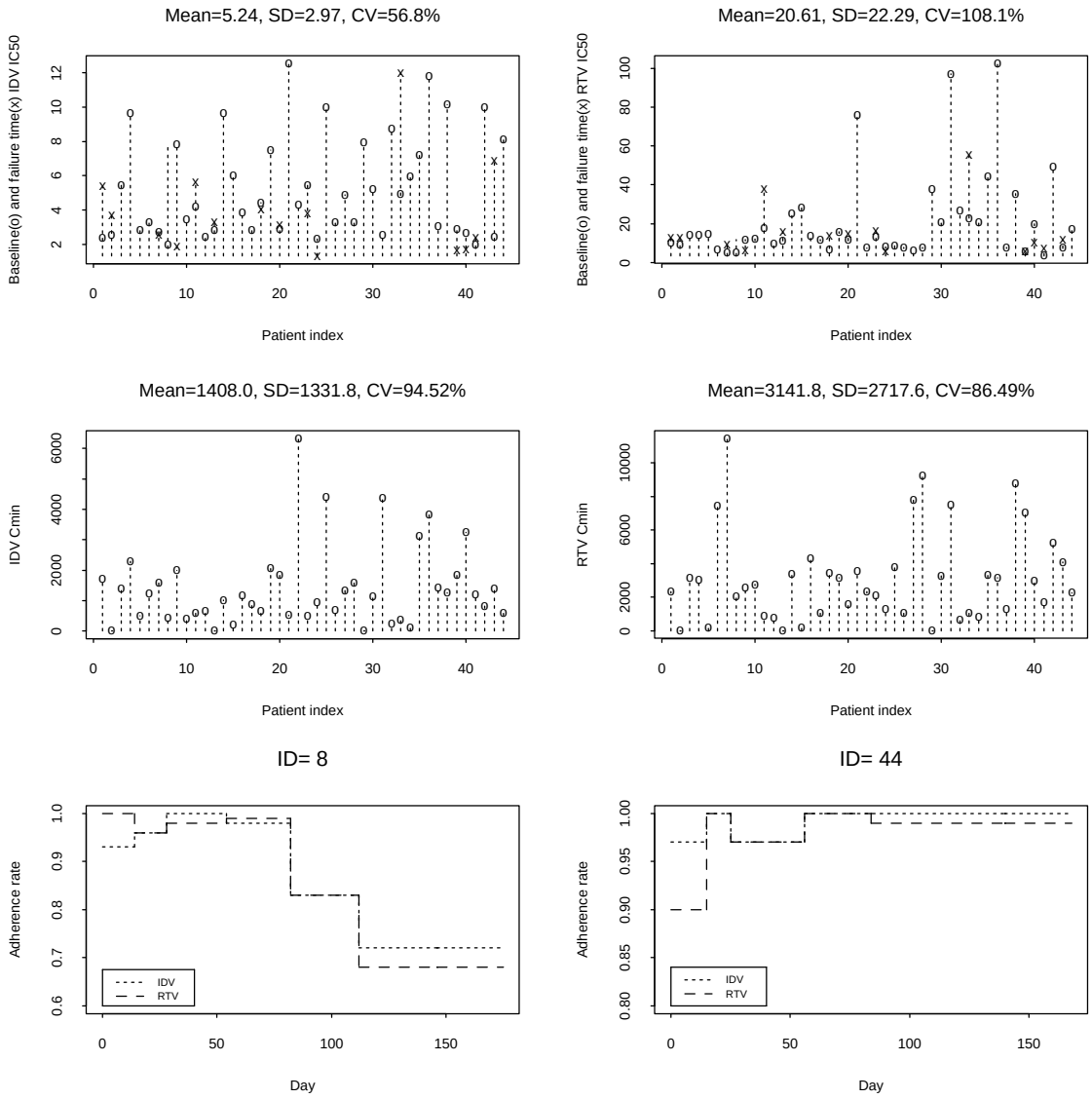

Figure 4.3: The baseline (○) and failure time(×) IC_{50} for IDV/RTV drugs (top panel), the minimum drug concentration (C_{min}) for two drugs (middle panel) for the 44 individual patients and adherence rates of IDV/RTV drugs (bottom panel) over time for the two representative patients. SD and CV denote the standard deviation and coefficient of variation, respectively.

Phenotypic drug susceptibility: Phenotypic drug susceptibilities were retrospectively determined from baseline samples. Phenotypic determination of ARV drug resistance was performed at baseline and/or at the time of virological failure (viral load rebounds). Some patients had virologic failure and phenotypic susceptibility testing done on samples at the time of failure. For analysis, we used the phenotype marker, the median inhibitory concentration (IC_{50}) (Molla et al., 1996) to quantify agent-specific drug susceptibility. We refer to this marker as the median inhibitory concentration. The baseline ($\circ$) and failure time($\times$) IC_{50} from 44 individuals for the two agents used in the A5055 trial, ritonavir (RTV) and indinavir (IDV), are displayed in Figure 4.3 (upper panel) which were used to construct $IC_{50}(t)$. Note that for patients without virological failure, $IC_{50}(t)$ was held by a constant with the baseline IC_{50} over time.

Pharmacokinetics variation: Plasma for intensive Pharmacokinetics (PK) analysis was obtained at pre-dose, and 0.5, 1, 2, 3, 4, 5, 6, 8, 10, and 12 hours following an observed IDV/RTV dose. PK parameters of IDV and RTV were determined using non-compartmental methods. Calculated PK parameters included maximum (C_{max}), minimum (C_{min}) drug concentration, and area under the curve (AUC). Wu et al.(2006) compared these PK parameters as predictors of virological responses and no significant differences were found. Thus, C_{min} displayed in Figure 4.3 (middle panel) was used in our analysis because it is easily obtained in clinical studies.

Medication adherence: Medication adherence was measured by the use of questionnaires. It was completed by the study participant and/or by a face-to-face interview with study personnel. As an example, the adherence rates over time based on questionnaire data for IDV (dotted stair-step line) and RTV (dashed stair-step line) drugs from the two representative patients are presented in Figure 4.3 (bottom panel).

Time-varying drug efficacy: We briefly discuss the drug efficacy function with two or more agents. In clinical practice, genotypic or phenotypic tests can be performed to determine the sensitivity of HIV-1 to ARV agents before a treatment regimen is selected. Here we use the phenotypic marker, IC_{50}, half maximal inhibitory concentration, to quantify agent-specific drug susceptibility. Because experimental data tracking development of resistance suggest that the resistant fraction of the viral population that grows exponentially, we propose a $\ln$-linear function to model the within-

host changes over time based on available IC_{50} observations as follows.

$$IC_{50}(t) = \begin{cases} \ln(S_0 + \frac{S_r - S_0}{t_r} t) & \text{for } 0 < t < t_r \\ \ln(S_r) & \text{for } t \geq t_r \end{cases} \tag{4.1}$$

where S_0 and S_r are respective exponential values of IC_{50} at baseline and time point t_r at which the resistant mutations dominate. In our study, t_r is the observed time of virologic failure from clinical studies. Given that IC_{50} is measured only at baseline and at the time of treatment failure (Molla et al., 1996), $IC_{50}(t)$ remains practical although more complex models for $IC_{50}(t)$ can be considered. For patients without a failure time IC_{50}, baseline IC_{50} was held constant over time. In other words, if $S_r = S_0$, no new drug resistant mutation is developed during treatment.

Poor adherence to a treatment regimen is one of the major causes of treatment failure (Ickovics and Melisler, 1997). Patients may occasionally miss doses, misunderstand prescription instructions or miss multiple consecutive doses for various reasons. These deviations from prescribed dosing affect drug exposure in predictable ways. We use the following model to represent medication adherence,

$$A(t) = \begin{cases} 1 & \text{for } T_k < t \leq T_{k+1}, \text{ if all doses are taken in } [T_k, T_{k+1}] \\ R & \text{for } T_k < t \leq T_{k+1}, \text{ if } 100R\% \text{ doses are taken in } [T_k, T_{k+1}] \end{cases} \tag{4.2}$$

where $0 \leq R < 1$, with R indicating the adherence rate for a drug (in our study, we focus on the two PI drugs discussed previously). T_k denotes the adherence evaluation time at the kth clinical visit. HAART contains two or more reverse transcriptase inhibitors (RTIs) and protease inhibitors (PIs) has proven to be effective at reducing the amount of virus in the blood and tissues of HIV-infected patients. In most viral dynamic studies (Ding and Wu, 2000; Perelson et al., 1996), investigators assumed that the drug efficacy was constant over treatment time. Drug efficacy may actually vary, however, because the concentrations of ARV drugs and other factors (i.e., emergence of drug-resistant mutations) vary during treatment. Also, patients' viral load may rebound because of drug resistance, non-adherence, and other factors. A simple pharmacodynamic (PD) sigmoidal Emax model for the dose effect relation follows (Gabrielsson and Weiner, 2000): $E = E_{max}C/(EC_{50} + C)$, where E_{max} is the maximal effect that can be achieved, C is the drug concentration, and EC_{50} is the drug concentration that induced an effect equivalent to 50% of the maximal effect. Many different variations of the E_{max} model have been developed by pharmacologists to model PD effects. More detailed discussions on E_{max} models can be found in the book

by Gabrielsson and Weiner (2000) and in the article by Huang et al.(2003). To model the relationship of multiple treatment factors with ARV drug efficacy, we employ the following modified E_{max} equation to represent the time-varying drug efficacy for two ARV agents within a class,

$$\gamma(t) = \frac{1}{2}\left\{\frac{C^1_{min}A_1(t)}{IC^1_{50}(t) + C^1_{min}A_1(t)} + \frac{C^2_{min}A_2(t)}{IC^2_{50}(t) + C^2_{min}A_2(t)}\right\} \tag{4.3}$$

where $\gamma(t)$ ranges from 0 to 1; $A_d(t)$, C^d_{min} and $IC^d_{50}(t)$ $(d = 1, 2)$ are the adherence profile, the minimum drug concentration in plasma and the time-course of median inhibitory concentrations for the two agents, respectively. Note that C_{min} could be replaced by other PK parameters such as AUC and C_{max}.

4.2.2. Bivariate linear mixed-effects models with ST distribution

Now we summarize the LME model for the AIDS data with bivariate correlated responses (CD4 and CD8 cell count). Let $\boldsymbol{y}_i^{(c)} = (y_{i1}^{(c)}, y_{i2}^{(c)}, \ldots, y_{in_i}^{(c)})^T$, $(c = 4, 8)$ be the repeated measurements (in cells/mm^3) of the CD4 and CD8 cell count, respectively, for the i^{th} subject at time t_{ij} $(i = 1, 2, ..., n,\ j = 1, 2, ..., n_i)$. Associated with each vector of measurements is a vector of times $\boldsymbol{t}_i = (t_{i1}, \ldots, t_{in_i})^T$ at which subject's measurements were taken. Let $\boldsymbol{x}_i^{(4)}$ and $\boldsymbol{x}_i^{(8)}$ be the $n_i \times p$ design matrices associated with the fixed-effects $\boldsymbol{\beta}^{(4)}$ and $\boldsymbol{\beta}^{(8)}$ of the two markers, respectively, and $\boldsymbol{z}_i^{(4)}$ and $\boldsymbol{z}_i^{(8)}$ be the corresponding $n_i \times q$ design matrices associated with the random-effects $\boldsymbol{b}_i^{(4)}$ and $\boldsymbol{b}_i^{(8)}$, respectively. To make notation more compact, let $\boldsymbol{Y}_i = \left(\boldsymbol{y}_i^{(4)T}, \boldsymbol{y}_i^{(8)T}\right)^T$, $\boldsymbol{X}_i$ =Diag$\left(\boldsymbol{x}_i^{(4)}, \boldsymbol{x}_i^{(8)}\right)$, $\boldsymbol{Z}_i$ =Diag$\left(\boldsymbol{z}_i^{(4)}, \boldsymbol{z}_i^{(8)}\right)$, $\boldsymbol{\beta} = \left(\boldsymbol{\beta}^{(4)T}, \boldsymbol{\beta}^{(8)T}\right)^T$, $\boldsymbol{b}_i = \left(\boldsymbol{b}_i^{(4)T}, \boldsymbol{b}_i^{(8)T}\right)^T$, $\boldsymbol{e}_i = \left(\boldsymbol{e}_i^{(4)T}, \boldsymbol{e}_i^{(8)T}\right)^T$, where $\boldsymbol{e}_i^{(4)}$ and $\boldsymbol{e}_i^{(8)}$ are the within-subject residuals for CD4 and CD8 cell count, respectively. As suggested by Lachos et al. (2009, 2010), we consider a skew-t bivariate LME (ST-BLME) model in which the within-subject errors are assumed to follow a normal distribution and the random-effects (latent random variables) are assumed to have an ST distribution which may be more reasonable. Thus, we have a skew-*t* bivariate LME (ST-BLME) model as follows,

$$\begin{aligned} \boldsymbol{Y}_i &= \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i + \boldsymbol{e}_i \\ \boldsymbol{b}_i &\sim ST_{2q,\nu}(-J(\nu)\boldsymbol{\delta}, \boldsymbol{\Sigma}_b, \boldsymbol{\Delta}) \\ \boldsymbol{e}_i &\sim N_{2n_i}(\boldsymbol{0}_{2n_i}, \boldsymbol{\Sigma}) \end{aligned} \tag{4.4}$$

where $\boldsymbol{\Sigma}_b = (\lambda_{kl})_{4\times4}$ is the dispersion matrix corresponding to between-subject variability for random-effects, $\boldsymbol{\Sigma} = (\sigma^2_{kl})_{2\times2}$ is variance-covariance matrix for model errors, $\boldsymbol{\Delta}$ = Diag($\boldsymbol{\delta}$) and

$\boldsymbol{\delta} = (\delta_1, \ldots, \delta_{2q})^T$, and $J(\nu) = (\nu/\pi)^{1/2}[\Gamma((\nu-1)/2)/\Gamma(\nu/2)]$. Note that, in the model (4.4), the correlation between the bivariate responses is accommodated and incorporated by both random-effects $\boldsymbol{b}_i$ and model errors $\boldsymbol{e}_i$.

4.3. Bayesian Inference

In this section, we implement the Bayesian methodology using MCMC techniques for the ST-BLME model. A key feature of this model, which allows writing easily WinBUGS codes, is that the model can be formulated in a flexible hierarchical representation. By introducing one random variable vector, $\boldsymbol{w}_i = (w_{i1}, ..., w_{i2q})^T$ and one random variable, $\xi_i, (i = 1, ..., n)$, based on the stochastic representation for the ST distribution (see Section 1.4 in detail), $\boldsymbol{Y}_i$ in the ST-BLME model (4.4) can be written hierarchically as

$$\begin{aligned} \boldsymbol{Y}_i|\boldsymbol{w}_i, \xi_i &\sim N_{2n_i}(\boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i, \xi_i^{-1}\boldsymbol{\Sigma}) \\ \boldsymbol{b}_i &\sim N_{2q}(\boldsymbol{\Delta}\boldsymbol{w}_i - J(\nu)\boldsymbol{\delta}, \xi_i^{-1}\boldsymbol{\Sigma}_b) \\ \boldsymbol{w}_i &\sim N_{2q}(\boldsymbol{0}, \boldsymbol{I}_{2q})I(\boldsymbol{w}_i > \boldsymbol{0}) \\ \xi_i &\sim \Gamma(\nu/2, \nu/2) \end{aligned} \tag{4.5}$$

where $(i = 1, \ldots, n)$. An important advantage of the above representations based on the hierarchical model is that they allow us to easily implement the ST-BLME model via the freely available WinBUGS software (Lunn et al., 2000), and the computational effort is similar to the one necessary to fit the models with the standard normal distribution. Our methodology can be widely applied to real problems for longitudinal studies as long as they meet the specifications proposed in this chapter.

Let $\boldsymbol{\theta} = \{\boldsymbol{\beta}^{(4)}, \boldsymbol{\beta}^{(8)}, \boldsymbol{\Sigma}, \boldsymbol{\delta}, \boldsymbol{\Sigma}_b, \nu\}$ be the collection of unknown parameters in model (4.4), $\boldsymbol{y} = (\boldsymbol{y}_1^T, ..., \boldsymbol{y}_n^T)^T$, $\boldsymbol{b} = (\boldsymbol{b}_1^T, ..., \boldsymbol{b}_n^T)^T$, $\boldsymbol{w} = (\boldsymbol{w}_1^T, ..., \boldsymbol{w}_n^T)^T$ and $\boldsymbol{\xi} = (\xi_1, ..., \xi_n)^T$, then the full likelihood function is given by

$$\begin{aligned} L(\boldsymbol{\theta}|\boldsymbol{y}, \boldsymbol{b}, \boldsymbol{w}, \boldsymbol{\xi}) \propto\ & \Pi_{i=1}^n\{\phi_{2n_i}(\boldsymbol{y}_i; \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i\boldsymbol{b}_i, \xi_i^{-1}\boldsymbol{\Sigma})\phi_{2q}(\boldsymbol{b}_i; \boldsymbol{\Delta}\boldsymbol{w}_i - J(\nu)\boldsymbol{\delta}, \xi_i^{-1}\boldsymbol{\Sigma}_b) \\ & \phi_{2q}(\boldsymbol{w}_i; \boldsymbol{0}, \boldsymbol{I}_{2q})\mathrm{I}(\boldsymbol{w}_i > \boldsymbol{0})\Gamma(\nu/2, \nu/2)\} \end{aligned} \tag{4.6}$$

To complete the Bayesian formulation, we specified the values of the hyper-parameters in the prior distributions and took weakly informative prior distribution for the parameters as follows

$$\begin{aligned} &\boldsymbol{\beta} \sim N_{2p}(\boldsymbol{\beta}_0, \boldsymbol{G}_1), \qquad \boldsymbol{\Sigma}_b \sim IW(\boldsymbol{G}_2, \eta_1), \qquad \boldsymbol{\Sigma} \sim IW(\boldsymbol{G}_3, \eta_2), \\ &\boldsymbol{\delta} \sim N_{2q}(\boldsymbol{0}, \boldsymbol{G}_4), \qquad \nu \sim Exp(\tau)I(\nu > 2) \end{aligned} \tag{4.7}$$

where the mutually independent Inverse Gamma (IG), Normal (N), Exponential (Exp) and Inverse Wishart (IW) prior distributions are chosen to facilitate computations (Gelman et al., 2003). The super-parameter matrices $\boldsymbol{G}_1$, $\boldsymbol{G}_2$, $\boldsymbol{G}_3$ and $\boldsymbol{G}_4$ can be assumed to be diagonal for convenient implementation.

One usually assumes that elements of the parameter vector $\boldsymbol{\theta}$ are independent of each other, i.e., $\pi(\boldsymbol{\theta}) = \pi(\boldsymbol{\beta})\pi(\boldsymbol{\Sigma})\pi(\boldsymbol{\delta})\pi(\boldsymbol{\Sigma}_b)\pi(\nu)$. After we specify the models for the observed data and the prior distributions for the unknown model parameters, we can make statistical inference for the parameters based on their posterior distributions under a Bayesian framework. Combining the likelihood function (4.6) and the prior distributions, the joint posterior density of $\boldsymbol{\theta}$ based on the observed data $\mathcal{D}$ can be given by

$$f(\boldsymbol{\theta}|\mathcal{D}) \quad \propto \quad L(\boldsymbol{\theta}|\boldsymbol{y}, \boldsymbol{b}, \boldsymbol{w}, \boldsymbol{\xi})\pi(\boldsymbol{\theta}) \tag{4.8}$$

Distribution (4.8) is analytically intractable, and it is prohibitive to directly calculate the posterior distribution of $\boldsymbol{\theta}$ based on the observed data. As an alternative, MCMC procedures can be used to sample based on (4.8) using the Gibbs sampler, from which features of marginal posterior distribution of interest can be inferred.

4.4. Data analysis

4.4.1. Specific model and implementation

We illustrate our method by applying it to the AIDS clinical data described in Section 4.2.1. We consider the following BLME model for CD4 and CD8 cell count.

$$\begin{aligned} y_{ij}^{(c)} &= \beta_{0i}^{(c)} + \beta_{1i}^{(c)} t_{ij} + \beta_4^{(c)} Age_i + \beta_5^{(c)} \log_{10}(RNA)_i + \beta_6^{(c)} \gamma(t_{ij}) + e_{ij}^{(c)}, \\ \beta_{0i}^{(c)} &= \beta_0^{(c)} + \beta_1^{(c)} g_i + b_{0i}^{(c)}, \\ \beta_{1i}^{(c)} &= \beta_2^{(c)} + \beta_3^{(c)} g_i + b_{1i}^{(c)}, \end{aligned} \tag{4.9}$$

The BLME hierarchal models (4.9) can be formulated as follows.

$$\begin{aligned} y_{ij}^{(c)} &= \beta_0^{(c)} + \beta_1^{(c)} g_i + \beta_2^{(c)} t_{ij} + \beta_3^{(c)} (g_i \times t_{ij}) + \beta_4^{(c)} Age_i \\ &\quad + \beta_5^{(c)} \log_{10}(RNA)_i + \beta_6^{(c)} \gamma(t_{ij}) + b_{0i}^{(c)} + b_{1i}^{(c)} t_{ij} + e_{ij}^{(c)}, \end{aligned} \tag{4.10}$$

where $c = 4$ and 8 correspond to CD4 and CD8 cell count, respectively, $y_{ij}^{(4)}$ and $y_{ij}^{(8)}$ are the respective standardized CD4 and CD8 cell count for the i^{th} subject at time t_{ij}, Age_i and $\log_{10}(RNA)_i$ are

age and viral load (in $\log_{10}$ scale) covariates for the i^{th} subject at baseline, $\gamma(t_{ij})$ is drug efficacy for the i^{th} subject at time t_{ij}, $g_i = 1$ if the i^{th} subject was treated in group one and 0 in group two. The corresponding regression coefficient $\beta_3^{(c)}$ can be interpreted as the treatment effect between groups one and two. $b_{0i}^{(c)}$ and $b_{1i}^{(c)}$ are random-effects representing a random intercept and a random slope, respectively. This model assumes that the mean baseline measurement (intercept) and mean rate of change (slope) are different between two treatment groups.

Several statistical models with different distribution from the SE class for the latent random-effects and random errors are compared. These models are as follows.

- **N Model:** Normal distribution for the random-effects and random errors.
- **SN Model:** Skew-normal distribution for the random-effects and normal distribution for the random errors.
- **ST Model:** Skew-t distribution for the random-effects and normal distribution for the random errors.

Note that random-errors can also be assumed to follow skew-normal or skew-*t* distribution, however, once the random-effects are assumed to be SN or ST distribution, the results are similar between normal and skewed distribution for random errors assumption. In order to make the comparison more straightforward, the random-errors are kept under the same assumption of normal distribution among the three models. In the absence of historical data/experiment, we specify practical weakly informative priors for all model parameters to obtain well-defined (proper) posteriors following the recommendations in (Hobert and Casella, 1996; Zhao et al., 2006). In particular, (i) fixed-effects are taken to be independent normal distribution $N(0, 100)$ for each component of the population parameter vector $\boldsymbol{\beta}$. (ii) The prior for the variance-covariance matrix of the random-effects $\boldsymbol{\Sigma}_b$ is taken to be inverse Wishart distributions $IW(\boldsymbol{G}_2, \eta_1)$ with covariance matrix $\boldsymbol{G}_2 = Diag(0.01, 0.01, 0.01, 0.01)$ and degree of freedom $\eta_1 = 5$. (iii) The prior for the variance-covariance matrix of the model errors $\boldsymbol{\Sigma}$ is taken to be inverse Wishart distributions $IW(\boldsymbol{G}_3, \eta_2)$ with covariance matrix $\boldsymbol{G}_3 = Diag(0.01, 0.01)$ and degree of freedom $\eta_2 = 3$. (iv) For each of the skewness parameters δ_1 and δ_2, independent normal distribution $N(0, 100)$ is used to accommodate either positive or negative skewness, and it allows the data to determine which one is more

appropriate. (v) Prior choice for ν is chosen as $\nu \sim Exp(0.5)\mathrm{I}(\nu > 2)$ (i.e., exponential density truncated at 2) to reflect a prior on ν with a well-defined and finite variance.

The MCMC sampler was implemented using WinBUGS software (Lunn et al., 2000), and the program codes are available in Appendix C. In particular, the MCMC scheme for drawing samples from the posterior distributions of all parameters is obtained by the Gibbs sampler. After collecting the final MCMC samples, we are able to draw statistical inference for the unknown parameters. Specifically, we are interested in the posterior means and quantiles. See the articles (Huang et al., 2006; Lunn et al., 2000) for detailed discussions of the Bayesian modeling approach and the implementation of the MCMC procedures, including the choice of the hyper-parameters, the iterative MCMC algorithm, sensitivity analysis and convergence diagnostics. We propose that, after an initial number of 50,000 burn-in iterations, every 20th MCMC sample is retained from the next 200,000. Thus, we obtain 10,000 samples of targeted posterior distributions of the unknown parameters for statistical inference.

4.4.2. Model comparison results

Bayesian modeling approach based on the specific model (4.10) with different model distribution specifications from the SE class was used to fit the data. For selecting the best model that fits the data adequately, a Bayesian selection criterion, DIC, is used. As with other model selection criteria, we caution that DIC is not intended for identification of the "correct" model, but rather merely as a method of comparing a collection of alternative formulations. As an alternative, we also evaluate EPD and RSS, while the detail information related to DIC, EPD and RSS can be found in Section 2.4.1.

Table 4.1 presents the DIC, EPD and RSS values among the three competing models. It is seen that the SN and ST Model produce better fit than the N Model in terms of DIC, EPD and RSS. In particular, the ST Model (with the smallest DIC) offers the best fit among the N, SN and ST Modal; these findings are consistent to those obtained by EPD and RSS criteria. Thus, we select the ST Model as our best fit model.

Figure 4.4 shows the box-plots for the skewness parameter, δ_1 and δ_2, among the SN and ST Model. The 95% CI of skewness parameters in the ST Model for both CD4 and CD8 cell count do not include zero, which confirm significantly positive skewness of the bivariate responses. In

Table 4.1: Model comparison using DIC, EPD and RSS criteria.

	Model		
Criterion	N	SN	ST
$\bar{D}$	979.14	945.09	600.97
p_D	140.58	141.91	186.02
DIC	1119.72	1086.86	**788.99**
EPD	0.840	0.825	**0.821**
RSS	690.4	128.6	**127.1**

the SN Model, the 95% CI of skewness parameter for CD8 cell count doesn't include zero, while it includes zero for CD4 cell count.

4.4.3. Estimation results based on the ST Model

Based on DIC, EPD and RSS, the best fit model is the ST Model, in which random-effects are assumed to have an ST distribution. Figure 4.5 plots the marginal posterior densities of the parameter ν for the ST Model. It shows some degree of right skewness confirming non-normal nature.

Compared with the N and SN Models, the SD for the fixed-effects parameters for both CD4 and CD8 cell count ($\beta_0 \sim \beta_6$) from the ST Model were smaller (Table 4.2). In the ST Model, the estimated skewness parameters for the CD4 and CD8 cell count are 0.441 and 0.518, respectively (Table 4.2). Because the 95% posterior credible interval for both skewness parameters do not include zero (95% CI is 0.098 $\sim$ 0.678 for the CD4 cell count, and is 0.219 $\sim$ 0.833 for the CD8 cell count), this confirms the positive skewness observed from the data. We also found the estimated skewness parameter for the CD8 cell count was significantly higher than that for the CD4 cell count ($p < 0.0001$). As expected, the results shown in Table 4.2 also indicate that there is a negative association between the CD4 cell count and baseline $\log_{10} RNA$ ($\beta_5^{(4)} = -0.113$, 95% CI is -0.208 $\sim$ -0.016). But we did not find any significant association between the CD8 cell count and the baseline $\log_{10} RNA$ ($\beta_5^{(8)} = 0.004$, 95% CI is -0.116 $\sim$ 0.124). Both CD4 and CD8 cell count significantly increase with the treatment time (95% CI is 0.282 to 0.850 for the CD4 cell count, and is 0.210 to

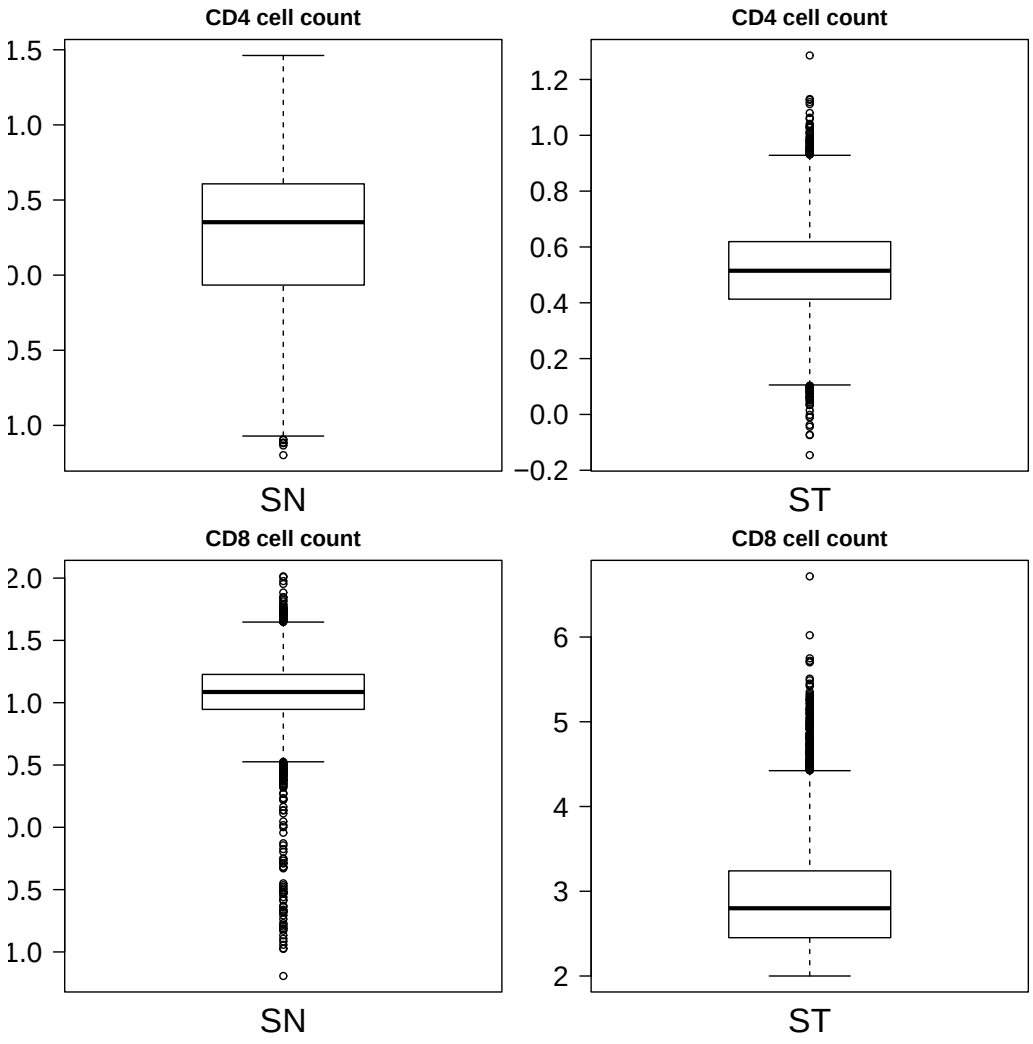

Figure 4.4: Box plot of skewness parameter for the SN and ST Models. The upper panel is for CD4 cell count and the lower panel for CD8 cell count

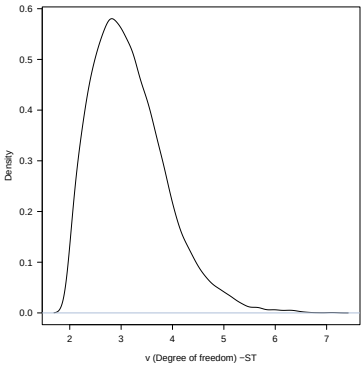

Figure 4.5: Marginal posterior densities estimates of parameter ν for the ST Model.

0.950 for the CD8 cell count). However, no significant group difference was found.

Table 4.2: A summary of the estimated posterior mean (PM) of population (fixed-effects) parameters, as well as the corresponding standard deviation (SD) and lower limit (L_{CI}) and upper limit (U_{CI}) of 95% equal-tail credible interval (CI).

CD4		$\beta_0^{(4)}$	$\beta_1^{(4)}$	$\beta_2^{(4)}$	$\beta_3^{(4)}$	$\beta_4^{(4)}$	$\beta_5^{(4)}$	$\beta_6^{(4)}$	δ_1	σ_{11}^2	ν
N	PM	-0.056	0.368	0.363	0.112	0.007	-0.233	1.546	–	0.131	–
	L_{CI}	-6.914	-0.100	0.053	-0.336	-0.024	-0.381	-5.618	–	0.109	–
	U_{CI}	7.066	0.837	0.678	0.545	0.038	-0.056	8.864	–	0.156	–
	SD	3.506	0.238	0.158	0.224	0.016	0.081	3.516	–	0.012	–
SN	PM	0.433	0.301	0.410	0.070	0.001	-0.212	1.083	0.559	0.128	–
	L_{CI}	-5.893	-0.154	0.059	-0.407	-0.029	-0.337	-4.000	-0.469	0.107	–
	U_{CI}	5.639	0.750	0.762	0.544	0.029	-0.054	7.500	1.135	0.153	–
	SD	3.118	0.229	0.179	0.242	0.014	0.071	3.137	0.441	0.012	–
ST	PM	-1.334	0.148	0.559	-0.103	0.016	-0.113	1.437	0.407	0.051	2.914
	L_{CI}	-3.837	-0.199	0.282	-0.466	-0.007	-0.208	-0.809	0.098	0.035	2.072
	U_{CI}	0.841	0.513	0.850	0.259	0.035	-0.016	3.967	0.678	0.071	4.306
	SD	1.261	0.179	0.144	0.184	0.011	0.049	1.298	0.148	0.009	0.602

CD8		$\beta_0^{(8)}$	$\beta_1^{(8)}$	$\beta_2^{(8)}$	$\beta_3^{(8)}$	$\beta_4^{(8)}$	$\beta_5^{(8)}$	$\beta_6^{(8)}$	δ_2	σ_{22}^2	σ_{12}^2
N	PM	-2.108	-0.370	0.173	0.469	0.001	0.051	1.764	–	0.290	0.062
	L_{CI}	-7.435	-0.949	-0.244	-0.130	-0.027	-0.115	-4.275	–	0.239	0.036
	U_{CI}	4.045	0.246	0.598	1.061	0.033	0.231	7.257	–	0.350	0.090
	SD	3.097	0.304	0.213	0.303	0.015	0.087	3.254	–	0.028	0.014
SN	PM	-0.457	-0.286	0.227	0.404	-0.008	0.028	0.561	1.017	0.283	0.061
	L_{CI}	-4.523	-0.808	-0.223	-0.257	-0.041	-0.123	-3.351	0.605	0.236	0.037
	U_{CI}	3.595	0.236	0.685	1.043	0.023	0.174	4.871	1.416	0.340	0.088
	SD	2.159	0.266	0.233	0.332	0.016	0.078	2.230	0.228	0.027	0.013
ST	PM	-2.078	0.020	0.565	-0.027	0.002	0.004	1.860	0.518	0.120	0.037
	L_{CI}	-6.275	-0.383	0.210	-0.525	-0.025	-0.120	-1.569	0.219	0.083	0.023
	U_{CI}	0.998	0.435	0.950	0.468	0.032	0.137	6.480	0.833	0.166	0.054
	SD	2.072	0.211	0.188	0.252	0.014	0.062	2.177	0.157	0.021	0.008

Although the increase of CD4 cell count is expected during an effective HAART, elevated CD8 cell count is associated with HIV virologic treatment failure (Krantz et al., 2011). Figure 4.6 shows the estimated individual coefficient of time for CD4 and CD8 cell count in the rebound and no rebound group. The detailed definition of rebound can be found in Section 2.4.2.3. Compared with

the rebound group, in the no rebound group, the average individual coefficient of time for CD4 cell count is significantly higher (median value is 0.476 and 0.133, for the no rebound and rebound, respectively. $p = 0.010$). While no significant difference was found in the CD8 between the two groups (median value is 0.379 vs. 0.260, for the no rebound and rebound, respectively. $p = 0.5420$). Because the maximum follow-up of A5055 study was only 6 months and even an interval of 12 months is considered as too premature to evaluate immune response to HAART (Dronda et al., 2002), a longer follow-up might be needed in order to draw a conclusion.

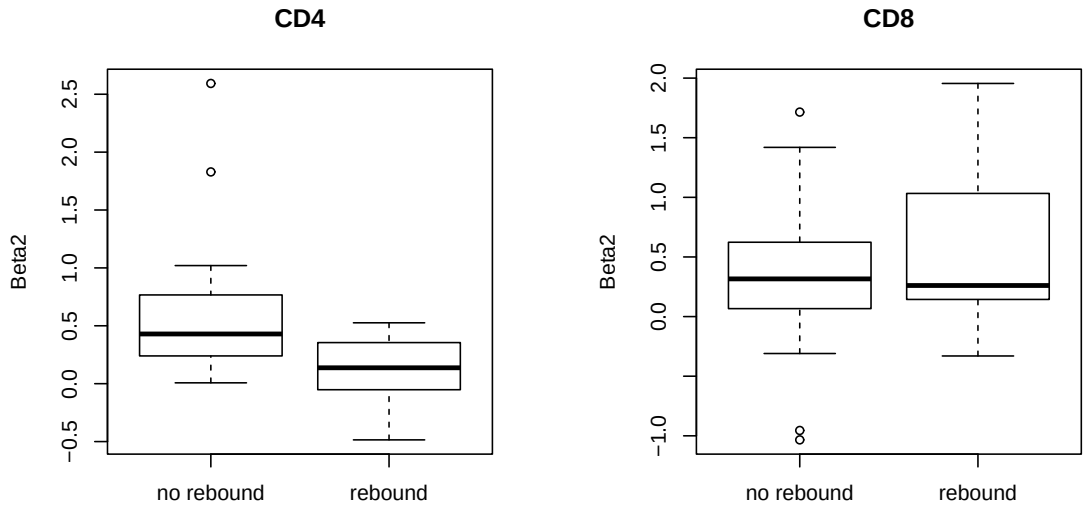

Figure 4.6: The coefficient of time for CD4 and CD8 cell count in rebound and no rebound group .

The estimated of SDs for the dispersion matrix parameter in the ST Model are smaller compared to the Normal or SN Model (Table 4.3). This is expected because high variability, heaviness of the tails and the skewness are interrelated to a certain extent.

Table 4.3: A summary of the estimated posterior mean (PM) of dispersion matrix parameter, as well as the corresponding standard deviation (SD) and lower limit (L_{CI}) and upper limit (U_{CI}) of 95% equal-tail credible interval (CI).

Model		λ_{11}	λ_{22}	λ_{33}	λ_{44}	λ_{12}	λ_{13}	λ_{14}	λ_{23}	λ_{24}	λ_{34}
N	PM	0.556	0.323	0.904	0.528	0.245	0.088	-0.035	0.060	0.062	-0.397
	L_{CI}	0.342	0.162	0.554	0.198	0.088	-0.149	-0.303	-0.155	-0.137	-0.810
	U_{CI}	0.885	0.567	1.437	1.048	0.449	0.360	0.205	0.285	0.291	-0.109
	sd	0.140	0.106	0.231	0.219	0.092	0.127	0.127	0.111	0.105	0.179
SN	PM	0.340	0.235	0.320	0.314	0.116	0.071	0.048	0.022	0.091	-0.146
	L_{CI}	0.091	0.082	0.097	0.103	-0.053	-0.088	-0.166	-0.141	-0.070	-0.492
	U_{CI}	0.719	0.506	0.809	0.740	0.355	0.275	0.266	0.193	0.301	0.051
	sd	0.174	0.112	0.183	0.170	0.111	0.090	0.107	0.082	0.091	0.143
ST	PM	0.181	0.148	0.241	0.257	0.028	0.061	-0.023	0.002	0.056	-0.113
	L_{CI}	0.078	0.071	0.103	0.097	-0.050	-0.024	-0.152	-0.099	-0.041	-0.300
	U_{CI}	0.354	0.279	0.459	0.542	0.118	0.175	0.083	0.097	0.175	0.009
	sd	0.071	0.054	0.092	0.118	0.042	0.051	0.058	0.049	0.054	0.079

4.4.4. Comparison between bivariate (CD4 and CD8) model and univariate (CD4 or CD8) model

It is also important to compare estimations from the bivariate linear mixed-effects models and the ones from univariate model that CD4 and CD8 are modeled separately. Table 4.4 shows the estimated population parameters based on the bivariate model and univariate model under the ST distribution.

For CD4 cell count, except the SD for the intercept, the SD of the population parameters from the bivariate model is bigger than that from the corresponding univariate model. For example, the SD of the skewness parameter from the bivariate model is 0.148, while it is 0.119 in the univariate model. The estimated parameters are also different. For example, the intercept from the bivariate is -1.334, while it is 0.175 in the univariate model; the coefficient for the time-varying drug efficacy is only 24% in the univariate as the value in the bivariate model (0.353 vs. 1.437).

For CD8 cell count, some estimated parameters show differently between the univariate and bivariate models. For example, the coefficient for the group is 0.020 in the bivariate, while it is

Table 4.4: Bivariate and univariate mixed-effect models: a summary of the estimated PM of population (fixed-effects) parameters, as well as the corresponding SD and 95% equal-tail CI.

CD4		$\beta_0^{(4)}$	$\beta_1^{(4)}$	$\beta_2^{(4)}$	$\beta_3^{(4)}$	$\beta_4^{(4)}$	$\beta_5^{(4)}$	$\beta_6^{(4)}$	δ_1
Bivariate-ST	PM	-1.334	0.148	0.559	-0.103	0.016	-0.113	1.437	0.407
	L_{CI}	-3.837	-0.199	0.282	-0.466	-0.007	-0.208	-0.809	0.098
	U_{CI}	0.841	0.513	0.850	0.259	0.035	-0.016	3.967	0.678
	SD	1.261	0.179	0.144	0.184	0.011	0.049	1.298	0.148
Univariate-ST	PM	0.175	0.071	0.529	-0.113	0.006	-0.127	0.353	0.409
	L_{CI}	-3.226	-0.214	0.269	-0.441	-0.014	-0.219	-3.091	0.176
	U_{CI}	3.557	0.366	0.818	0.218	0.027	-0.039	3.878	0.650
	SD	1.732	0.149	0.139	0.167	0.010	0.046	1.786	0.119

CD8		$\beta_0^{(8)}$	$\beta_1^{(8)}$	$\beta_2^{(8)}$	$\beta_3^{(8)}$	$\beta_4^{(8)}$	$\beta_5^{(8)}$	$\beta_6^{(8)}$	δ_2
Bivariate-ST	PM	-2.078	0.020	0.565	-0.027	0.002	0.004	1.860	0.518
	L_{CI}	-6.275	-0.383	0.210	-0.525	-0.025	-0.120	-1.569	0.219
	U_{CI}	0.998	0.435	0.950	0.468	0.032	0.137	6.480	0.833
	SD	2.072	0.211	0.188	0.252	0.014	0.062	2.177	0.157
Univariate-ST	PM	-1.700	-0.016	0.493	0.081	-0.003	-0.025	1.897	0.502
	L_{CI}	-6.133	-0.420	0.139	-0.387	-0.029	-0.152	-2.719	0.126
	U_{CI}	2.719	0.377	0.861	0.569	0.022	0.102	6.610	0.843
	SD	2.234	0.202	0.184	0.241	0.013	0.064	2.367	0.183

-0.016 in the univariate model. The estimated coefficient for time for CD4 and CD8 got from the univariate models are slightly smaller than that for the bivariate model.

4.5. Conclusion and discussion

In this chapter, we consider a Bayesian bivariate SE approach to jointly model the CD4 and CD8 cell count in peripheral blood. The hierarchical representations given in equation 4.4 and 4.5 provide easy model implementation by using the conventional Bayesian software WinBUGS. Using suitable model choice criteria (DIC, EPD and RSS), among the three models, the N, SN and ST Model, we found the ST Model had the best model fit and the related SDs for the fixed-effects parameters were also the smallest.

The results from the ST Model confirm the positive skewness and heavy tails observed from the raw data. Because the estimated skewness parameter for CD8 cell count was bigger than that for

CD4 cell count, it also confirms that the distribution of the CD8 cell count in peripheral blood are more skewed than that of the CD4 cell count. The baseline viral load ($\log_{10}$ RNA) was negatively associated with the CD4 cell count, while we didn't find such a significant association between the baseline viral load and the CD8 cell count. Both CD4 and CD8 cell count significantly increased with the time of the treatment. Compared with the rebound group, the average increase rate of the CD4 cell count over time was higher in no rebound group, while no such significant difference between the rebound and no rebound group was found in the CD8 cell count.

For hierarchical LME via Bayesian approach, one concern is that using weakly informative priors may lead to inconsistent inferential results (Natarajan and Kass, 2000; Zhao et al., 2006). In order to check whether the results were sensitive to the prior choice, we conducted sensitive analysis and recomputed the posterior quantities of interest. Although slight changes in the values of population parameters were noticed, the results were quite robust overall and the conclusions among the three models were kept the same.

This chapter has some limitations. The AIDS clinical trial data we used only included the total number of CD8 cells in peripheral blood. Because HIV-1 virus will not directly 'kill' CD8 cells, the qualitative factors within the CD8 cells response and sub-groups of CD8 cells such as HIV specific CD8 , naive CD8, or activated CD8 cells, may be the principal determinants of HIV/AIDS disease progression (Migueles et al., 2002). Also, a clinical trial study with longer follow-up and 'naive' patients who never got HAART treatment may be a better data set for testing the proposed models and methods.

The skewed-elliptical models applied in this study are quite flexible and can be easily extended to a more general distribution family such as skew-normal independent (SNI) by changing the distribution of ξ_i in equation (4.5), for example, if we assume $\xi_i \sim Beta(\nu, 1)$ and $\nu > 0$, then it will be multivariate skew-slash distribution. This kind of skewed modeling approach is important in many biostatistical applications areas where either skewness should be considered or transformation can be avoided. The method is useful for the exploring a "clustered" data, regardless of whether it is cross-sectional or longitudinal one. We can easily apply the proposed method in WinBUGS and building in model checking will facilitate model comparisons.

5 Overall discussion and conclusions

This study has relaxed the normal distribution assumption in longitudinal data by using a multivariate SE distribution family via Bayesian nonlinear or linear mixed-effects modeling approach. This chapter summarizes the new development arising from this study, the contributions of this study in terms of methodology and application, the study's limitations, and further research goals.

In linear and nonlinear mixed-effects models, random-effects and within-subject random errors are commonly assumed to follow a normal distribution. Although this assumption will bring convenience in the computation, it may be an unrealistic and obscure important features among and within subject variation. During recent years, modifying well-known distributions by conditioning and transforming in order to include skewness has received much attention. Among different versions of modified skewed distributions, we selected Sahu's version because it can be easily applied via Bayesian approach in WinBUGS. The following summarizes the main contributions of this dissertation.

Multiphasic HIV viral load changes since the treatment indicates that the viral decay rate is a time-varying process. Mixed-effects models with different time-varying decay rate functions have been proposed in the literature, however, there are two critical issues: (i) it has not been determined which model is more appropriate for practical application; (ii) the model random errors are commonly assumed to follow a normal distribution, which may be unrealistic and can cause biased inference. Because skewness of HIV viral load is still noticeable even after transformation, it is important to use a more general distribution family that can allow us to relax the normal assumption. In Chapter 2, we developed the skew-elliptical Bayesian mixed-effects models by considering that model random errors have an SE distribution. We compared the performance of different SE models with time-varying decay rate function. We also compared the performance among the models with normal, Student-*t*, SN and ST distribution. Two AIDS clinical trial data sets were used to test the proposed models and methods. The results indicate the ST distribution model with a time-varying function that includes two exponential components is superior to the other models. The model fit is

better with the assumption of ST and SN than with an assumption of normal or Student-*t* distribution. This finding suggests that it is important to assume a model with a skewed distribution in order to achieve reasonable results when the data exhibit non-normality characteristic. Based on the best fitting model under the ST distribution, we found the time-varying viral decay rate was significantly associated with the CD4 cell count and HIV-1 viral rebound status, which may provide important clinical information such as prediction of long-term outcome based on the early stage response.

Since we didn't consider any covariates in Chapter 2, we extended the SE mixed-effects models in Chapter 3 such that the CD4 cell count was included as a covariate in order to account for between- and within-subject variation. Among the models that can be used for different lengths of HIV follow-up, we compared LME, NLME, and SNLME models. The critical question that needed to be answered was whether these models produce coherent estimates of viral decay rates, and if not, which model is appropriate and should be used in practice. Besides the skewed distribution observed in the HIV viral load, CD4 cell count also shows skewness that should not be ignored and they may be often measured with substantial errors or at different measurement schedules, in Chapter 3 these issues are addressed simultaneously by jointly modeling the response variable with skewness and a covariate process with measurement errors using a Bayesian approach to investigate how estimated parameters are changed or different under these three models. We found that there was a significant incongruity in the estimated decay rates in viral loads based on the three mixed-effects models, suggesting that the decay rates estimated by using Bayesian LME or NLME joint models should be interpreted differently from those estimated by using Bayesian SNLME joint models. The findings also suggested that the Bayesian SNLME joint model is preferred to other models because an arbitrary data truncation is not necessary; and it was further shown that the models with a skew-normal distribution and/or measurement errors in covariates may achieve reliable results when the data exhibit skewness.

In Chapter 4, multiple correlated outcomes should be estimated in a model where their dependence on the independent variables can be considered. To accomplish this, the bivariate outcomes of CD4 and CD8 cell count were jointly analyzed under BLME, while the baseline viral load, age, time-varying drug efficacy and treatment groups were included as covariates. In HIV immunologic responses, such as CD4 and CD8 cell count, the corresponding values often show noticeable non-normal characteristics such as skewness with a heavy right tail. There are several limitations to

using transformation, including reduced information, no guarantee of joint normality, no universal transformation, or difficultly in interpreting of the results. Therefore, we applied a more flexible class of parametric multivariate SE distributions to model CD4 and CD8 cell count. This approach provides an appealing robust alternative to the symmetric normal process in an LME model framework. The estimated results confirmed the positive skewness and heavy tails observed from the data, while the DIC value indicated the model fitting was improved by considering these two issues. One of the advantages of the proposed methods is the easy extension to a more general family such as skew-normal independent distribution.

The proposed methods may have a significant impact on AIDS research from a methodology and application standpoint of view. Given the fact that HIV-1 viral load, CD4 and CD8 cell count all show noticeable skewness, relaxing normality assumption with skew-elliptical distribution will allow more accurate inference of parameters with adjusting for the data with skewness. The estimation is made via Bayesian MCMC approach that can be carried out by using publicly available WinBUGS software. The theoretical and technical solutions proposed in this research are quite general, so they can be used in other biological fields where skewness should be considered. From an application point of view and to our best knowledge, few published articles have: (i) compared different HIV dynamic models that can be used for the entire follow-up data with the normality assumption being relaxed, (ii) compared HIV dynamic models used in different lengths of HIV follow-up through a joint model that can simultaneously consider the measurement errors in the covariates such as CD4 and skewness observed in the outcome and covariates, (iii) used CD4 and CD8 as bivariate outcomes and accommodated normality departures (skewness and heavy tails) and hence eliminated the need for ad hoc data transformations.

Until the late 1980s, Bayesian statistics were only considered as an interesting alternative to the "classical" frequentist approach for several reasons. The main objection from "classical" statisticians was the subjective view point via the prior distribution in Bayesian approach, which can be easily arguable since either non-informative prior distributions can be used when no strong previous information exists or reasonable informative prior distributions should be selected so the knowledge is a developing process, therefore, no need of "building from scratch" every time. The real obstacle that prevented Bayesian theory from being as one of the main stream approaches was the intractability involved in the calculation of the posterior distribution. Implementation of the MCMC

methods equipped with powerful personal computers made Bayesian statistics applicable. In HIV dynamic models, the estimation of the parameters is complex because many times, the likelihood has no closed form, even for simple models. Asymptotic approximation methods had provided some solutions but may have issues such as lack of generalization or inconsistent estimates. The Bayesian approach provides an attractive alternative solution in this field, and it allows us to build more realistic models with more complicated structure that may be prohibitive or at least very cumbersome in frequentist approach.

Although the Bayesian approach has proved to be a very useful tool in the complex model inference, a main concern is the uncertainty regardling the prior distributions and the initial values selection. The basic tool for investigating model uncertainty is the sensitivity analysis. Based on different initial values, re-computing the posterior quantities of interest can indicate whether the results have changed in a way that will significantly affect the results interpretations and conclusions. If the results are robust against the different initial values, we can report the results with confidence and be assured that the related conclusions are solid. The results should be explained with caution if they are sensitive to the initial values. The sensitivity analyses in this dissertation show that the estimated parameters were not sensitive to both prior parameters and the initial values, so the results were reasonable and robust. One thing needs to be pointed out: after we found the results were robust to the priors and initial values, we used the same priors and initial values in the model comparisons.

The complicated HIV pathogenesis motivated us to combine a new general distribution family and Bayesian inference. The proposed methods enhance the modeling flexibility and allow researchers to analyze longitudinal and multiple treatment factors in a wide variety of considerations. In addition, the proposed hierarchical modeling approach can be easily implemented using the WinBUGS package that is publicly available. These factors make our approach quite powerful and easily accessible to statisticians. In many biological and medical fields, non-normality is a commonly seen phenomenon while the underlying mechanism of the outcome is always complex, which requires advanced mathematical modeling. We believe the proposed models and methods are quite general and helpful and can be used in other biostatistical applications.

There are several limitations in this dissertation. We did not consider missing data issue, and due to the complex nature of HIV/AIDS and the toxicity of medications, the assumption of missing

at random may not be a reasonable assumption. This issue may be addressed using two-step or multiple computation methods. Below detection limit, which is left censoring, was not considered in this study either. Another limitation involves the related biomarkers used: we examined a pooled host cell population and did not consider the individual compartments of short-lived productively infected cells, long-lived or latently infected cells; the virus compartment was not further decomposed into infectious versus noninfectious virions.

Although the SE distributions used in this dissertation belong to a general family that includes many commonly used distributions, ideally, it would be helpful to have an even more general family that can be used to develop a Bayesian analysis of censored mixed-effects models. Skew-normal/independent (SNI) distribution is an attractive class of asymmetric heavy-tailed distributions that includes the skew-normal, the skew *t*, skew slash and the skew-contaminated normal distributions as special cases. As viral decay rates reflect the potency of antiviral regimens, it is important to evaluate the regimen's effect on long term responses. Besides the binary endpoint used in this research (rebound vs. no rebound), the long term responses can be survival endpoints such as time to HIV-1 RNA falling below the detection limit, time to virological failure or time to progression to AIDS. The estimated viral decay rates can be treated as covariates and because frequent evaluations may not be common during long-term follow-up, the event will be known only to have occurred within some interval of time (interval censoring). Rebound is generally caused by emergence of a drug-resistant virus strain, and it is important to develop a flexible, yet parsimonious mechanistic model to predict the rebound. By Extending basic HIV dynamic model, the latent HIV dynamic model can be use via Bayesian MCMC-SAEM algorithm. There is marked variation in how the body responds to the virus , and in the time-course of progression to AIDS. It is known that host genetic differences contribute to this variation, but our knowledge of the relevant host genetic factors is currently limited. Systematic searches of the genome to identify common genetic variants (genotypes) that influence the human response to HIV-1 is promising approach.

List of References

Acosta EP, Wu H, Hammer SM, Yu S, Kuritzkes DR, Walawander A,Eron JJ, Fichtenbaum CJ, Pettinelli C, Neath D, Ferguson E, Saah AJ, Gerber JG, the adult AIDS clinical trial group 5055 protocol team (2004). Comparison of two indinavir/Ritonavir Regimens in the treatment of HIV-infected individuals. *Journal of Acquired Immune Deficiency Syndromes*. **37(3)**:1358–1366.

Adams BM, Banks HT, Davidian M, Kwon HD, Tran HT, Wynne SN. (2005). HIV dynamics: modeling, data analysis, and optimal treatment protocols. *Journal of Computational and Applied Mathematics*. **184**:10–49.

Alimonti JB, Ball TB, Fowke KR. (2003) Mechanisms of CD4+ T lymphocyte cell death in human immunodeficiency virus infection and AIDS.*Journal of General Virology*. **84**:1649–1661.

Arellano-Valle RB, Bolfarine H, Lachos VH. (2005). Skew-normal linear mixed models. *Journal of Data Science*. **3**, 415–438.

Arellano-Valle RB, Genton MG (2005). On fundamental skew distribution. *Journal of Multivariate Analysis*. **96**:93–116.

Arellano-Valle RB, Azzalini A (2006). On the unification of families of skew-normal distributions. *Scandinavian Journal of Statistics*. **33**(3): 561–574.

Arellano-Valle RB, Branco MD, and Genton MG (2006). A unified view on skewed distributions arising from selections. *Canadian Journal of Statistics*. **75(24)**: 11983–11991.

Arellano-Valle RB, Bolfarine H, and Lachos VH (2007). Bayesian inference for skew-normal linear mixed models. *Journal of Applied Statistics*. **34**, 663–682.

Azzalini A, Dalla-Valle A. (1996). The multivariates skew-normal distribution. *Biometrika*. **8**:715–726.

Azzalini A, Capitanio A (1999). Statistical applications of the multivariate skew normal distribution. *Journal of the Royal Statistical Society Series B*. **61**:579–602.

Azzalini A, Capitanio A (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t distributions. *Journal of Royal Statistical Society, Series B*. **65**, 367–389.

Azzalini A. (2005). The skew-normal distribution and related multivariate families (with discussion by Marc G. Genton and a rejoinder by the author). *Scandinavian Journal of Statistics*. **32**, 159–200.

Beal SL, Sheiner LB. (1982). Estimating population kinetics. *Critical Review in Biomedical Engineering*. **8**:192–222.

Betts MR, Ambrozak DR, Douek DC, Bonhoeffer S, Brenchley JM, Casazza JP, Koup RA, Picker LJ. (2001) Analysis of total Human Immunodeficiency Virus (HIV)-specific CD4+ and CD8+ T-cell responses: relationship to viral load in untreated HIV infection. *Journal of Vriology*. **6(11)**:1349–1352.

Burnham KP, Anderson DR. (2002). Model selection and multimodel inference: a practical informationtheoretic approach (2nd edn). *Springer:* New York, 2002.

Carroll RJ, Ruppert D, Stefanski LA., Crainiceanu CM (2006). *Measurement Error in Nonlinear Models: A Modern Perspective* (2nd edition). London: Chapman and Hall.

Chevret S, Roguin H, Ganne P, Lefrere JJ. (1992) Prognostic value of an elevated CD8 lymphocyte count in HIV infection. Results of a prospective study of 152 asymptomatic HIV-positive individuals. *AIDS*. **6(11)**:1349–1352.

Commenges D, Jolly D, Drylewicz J, Putter H, Thiebaut R. (2011) Inference in HIV dynamics models via hierarchical likelihood. *Biometrics*. **55(1)**:446–456.

Dagne G, Huang Y. (2012) Mixed-effects Tobit joint models for longitudinal data with skewness, detection limits and measurement errors. *Journal of Probability and Statistics*. Volume 2012 (2012): Article ID 614102.

Davidian M. Gilitinan DM. (1995) Nonlinear models for repeated measurement data. *London: Chapman & Hall*.

De la Cruz R (2008). Bayesian non-linear regression models with skew-elliptical errors: applications to the classification of longitudinal profiles. *Computational Statistics and Data Analysis*. **53**(2), 436–449.

Donnet S. Samson A. (2007). Estimation of parameters in incomplete data models defined y dynamical systems. *Journal of Statistical Planning and Inference*. **137)**:2815–2831.

Ding AA, and Wu H. (1999). Relationships between antiviral treatment effects and biphasic viral decay rates in modeling HIV dynamics. *Mathematical Biosciences*. **160**, 63–82.

Ding AA, and Wu H (2000). A comparison study of models and fitting procedures for biphasic viral decay rates in viral dynamic models. *Biometrics*. **56**, 16–23.

Dragic T, Litwin V, Allaway GP, Martin SR, Huang Y, Nagashima KA, Cayanan C, Maddon PJ, Koup R, Moor JP, Paxton WA. (1996). HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5. *Nature*. **381**:667–673.

Dronda F, Moreno S, Moreno A, Casado JL, Perez-Elias MJ, Antela A(2002). A. Long-term outcomes among antiretroviral-naive human immunodeficiency virusinfected patients with small increases in CD4+ cell count after successful virologic suppression. *Clin Infect Dis* . **35**: 1005-9.

Embretson J, Zupancic M, Ribas J, Burke A, Raca P, Tenner-Racz K, Haase AT. (1993). Massive covert infection of helper T lymphocytes and macrophages by HIV during the incubationperiod of AIDS. *Nature*. **362**:359–362.

Evan G, Littlewood, T. (1998). A matter of life and cell death. *Science*. **281**: 1317–1322.

Fauci AS. (1988). The human immunodeficiency virus; infectivity and mechanisms of pathogenesis. *Science*. **239**:617–622.

Ferguson TS (1973). A Bayesian analysis of some nonparametric problems. *Annals of Statistics*. 1, 209 – 230.

Fernandez C , Steel MFJ (1998). On Bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association*. **93**:359–371.

Fitzgerald A, DeGruttola V, Vaida F. (2002). Modeling HIV viral rebound using non-linear mixed effects models. *Statistics in Medicine*. **21**: 2093–2108.

Gabrielsson J, and Weiner D (2000). *Pharmacokinetic and Pharmacodynamic Data Analysis: Concepts and Applications* (4th ed.). Taylor & Francis Group.

Gardner RC. (2001) Pyschological statistics using SPSS for Windows. *McGraw-Hill*.

Gelman A, Carline JB, Stern HS, Rubin DB. (2003). *Bayesian data analysis*. Chapman & Hall/CRC. London.

Geman S, Geman D (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on pattern Recognition and machine Intelligence*. **6**, 721–741.

Genton, M.G. (2004). *Skew-elliptical Distributions and their Applications: A Journey Beyond Normality*. Chapman & Hall: FL, USA

Ghidey W, Lesaffre E, and Eilers P (2004). Smooth random effects distribution in a linear mixed model. *Biometrics*. **60**: 945–953.

Ghosh P, Branco M, Chakraborty H. (2007). Bivariate random effect model using skew-normal distribution with application to HIV-RNA. *Statistics in Medicine*. **26**: 1255–1267.

Guidelines (2012) $http://aidsinfo.nih.gov/guidelines/$ [Accessed Aug 1 2012].

Grossman Z, Polis M, Feinberg M, Grossman Z, Levi I, Jankelevich S. Yarchoan R. Boon J, Wolf FD, Lang JMA, Goudsmit J, Dimitrov DS, Paul WE. (1999). Ongoing HIV dissemination during HAART. *Nature Medicine*. **5(10)**: 1099–1104.

Guedj J, Thiébaut R, Commenges D. (2007) Maximum likelihood estimation in dynamical models of HIV. *Biometrics*. **69**: 2493–2513.

Havlir D, Hellmann NS, Petropoulos CJ, Whitcomb JM, Collier AC, Hirsch MS, Tebas P, Sommadossi JP, Richman DD. Drug susceptibility in HIV infection after viral rebound in patients receiving Indinavir-containing regimens. *Journal of the American Medical Association*. **283**: 229–234.

Helsel DR (1990). Less than obvious-statistical treatment of data below the detection limit. *Environmental Science and Technology*. **24(12)**, 1766–1774.

Hobert JP, and Casella G (1996). The effect of improper priors on Gibbs sampling in hierarchical linear mixed models. *Journal of the American Statistical Association*. **91**: 1461–1473.

Harrington M.; Carpenter CC. (2000). "Hit HIV-1 hard, but only when necessary". *The Lancet*. **355 (9221)**: 2147.

Havlir DV, Hellmann NS, Petropoulos CJ, Whitcomb JM, Collier AC, Hirsch MS, Richman DD (2000). Drug Susceptibility in HIV Infection After Viral Rebound in Patients Receiving

Indinavir-Containing Regimens. *JAMA*. **283**: 229–234.

Higgins M, Davidian M, Gilitinan DM (1997). A two-step approach to measurement error in time-dependent covariates in nonlinear mixed-effects models, with application to IGF-I pharmacokinetics. *Journal of the American Statistical association*. **92**, 436 – 448.

Ho DD., Neumann AU, Perelson AS, Chen W, Leonard JM, Markowitz M. (1995). Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature*. **373**:123–126.

Ho DD. (1995). "Time to hit HIV, early and hard". *The New England Journal of Medicine*. **333(7)**: 450–451.

Hobert JP, Casella G. (1996). The effect of improper priors on Gibbs sampling in hierarchical linear mixed models. *Journal of the American Statistical Association*. **91**: 1461–1473.

Huang Y, Rosenkranz S, Wu, H. (2003). Modeling HIV dynamics and antiviral responses with consideration of time-varying drug exposures, sensitivities and adherence. *Mathematical Biosciences*. **184**:165–186.

Huang Y, Liu D, Wu H. (2006). Hierarchical Bayesian Methods for Estimation of Parameters in a Longitudinal HIV Dynamic System. *Biometrics*. **62(2)**:413–423.

Huang Y, Dagne G. (2010). Skew-normal Bayesian nonlinear mixed-effects models with application to AIDS studies. *Statistics in Medicine*. **29**:2384–2398.

Huang Y, Dagne G. (2011). A Bayesian approach to joint mixed-effects models with a skew-normal distribution and measurement errors in covariates. *Biometrics*. **67(1)**: 260-269.

Huang Y, Dagne G. (2012a). Simultaneous Bayesian inference for skew-normal semiparametric nonlinear mixed-effects models with covariate measurement errors. *Bayesian Analysis*. **7(1)**: 189-210.

Huang Y, Dagne G. (2012b). Comparison of mixed-effects models for skew-normal responses with an application to AIDS data: A Bayesian approach. *Communications in Statistics-Simulation and Computation*. **7(1)**: 189210.

Hughes J. (1999). Mixed effects models with censored data with applications to HIV RNA levels. *Biometrics*. **55**: 625–629.

Ickovics JR, and Melisler AW (1997). Adherence in AIDS clinical trial: a framework for clinical

research and clinical care. *Journal of Clinical Epidemiology*. **50**: 385–391.

Ir M, Mayer V, Sobota K. (1990). CD4/CD8 counts and immunoglobulin levels in anti-HIV-1 positive people with and without lymphadenopathy. *AIDS*. **4(2)**: 168.

Iweala OI. (2004). HIV diagnostic tests: an overview. *Contraception*. **70**:141–147.

Jara A, Quintana F, and Martin ES (2008). Linear mixed models with skew-elliptical distributions: a Bayesian approach. *Computational Statistics and Data Analysis*. **52**:5033–5045.

Jones MC, Faddy MJ (2003) A skew extension of the *t*-distribution, with applications. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*. **65(1)**:159–174.

Kaplan D, Sieg S. (1993). Role of the FAS/FAS ligand apoptotic pahtway in HIV-1 disease. *Journal of Molecular Biology*. **72**:6279–6282.

Ke C. Wang Y. (2001). Semiparametric nonlinear mixed-effects models and their application with discussion. *Journal of the American Statistical Association*. **96**:1272–1298.

Krantz EM, Hullsiek KH, Okullicz JF, Weintrob AC, Agan BK, Crum-Cianflone NE, Ferguson TM, Hale BR. (2011) Elevated CD8 count during HAART are associated with HIV virologic treatment failure. *Journal of Acquired Immune Deficiency Syndromes*. **57(5)**:369–403.

Kuhn E. Lavielle M. (2005). Maximum likelihood estimation in nonlinear mixed effects models. *Computational Statistics and Data Analysis*. **49**:1020–1038.

Kuritzkes DR, Ribaudo HJ, Squires KE, Koletar SL, Santana J, Riddler SA, Reichman R, Shikuma C, Meyer WA 3rd, KL, Gulick RM; ACTG A5166s Protocol Team. (2007). Plasma HIV-1 RNA dynamics in antiretroviral-naive subjects receiving either triple-nucleoside or efavirenz-containing regimens: ACTG A5166s. *The Journal of Infectious Diseases*. **195(8)**:1167-1-176.

Laird NM and Ware JH. (1982). Random-effects models for longitudinal data. *Biometrics*. **38**: 963-974.

Lachos VH, Dey DK, and Cancho VG (2009). Robust linear mixed models with skew-normal independent distributions from a Bayesian perspective. *Journal of Statistical Planning and Inference*. **139**: 4098–4110.

Lachos VH, Ghosh P, Arellano-Valle RB. (2010) Likelihood based inference for skew-normal independent linear mixed models. *Statistica Sinica*. **20**: 303–322.

Laird, N.M., and Ware, J.H. (1982). Random effects models for longitudinal data. *Biometrics*. **38**, 963–974.

Lavielle M. Samson A, Karine-Fermin A, Mentre E. (2011). Maximum likelihood estimation of long-term HIV dynamic models and antiviral response. *Biometrics*. **67(1)**:250–259.

Levy JA. (1993). Pathogenesis of human immunodeficiency virus infection. *Microbiology Review*. **57**:183–289.

Lin TI (2010). Robust mixture modeling using multivariate skew-t distributions. *Statistics and Computing*. **20**: 343–356.

Liu W, Wu, L. (2007). Simultaneous inference for semiparametric nonlinear mixed-effects models with covariate measurement errors and missing responses. *Biometric*. **63**: 342–350.

Louie M, Hogan C, Di Mascio M, Hurley A, Simon V, Rooney J, Ruiz N, Brun S, Sun E, Perelson AS, Ho DD, Markowitz M. (2003). Determining the relative efficacy of highly active antiretroviral therapy. *The Journal of Infectious Diseases*. **187(6)**:896–900.

Lindstrom M, Bates D. (1990). Nonlinear mixed effects models for repeated measures data. *Biometrics*. **46**:673–687.

Lunn D, Thomas A, Best N, Spiegelhalter D. (2000). WinBUGS-A Bayesian modeling framework: concepts, structures, and extensibility. *Statistics and Computing*. **10**:325–337.

Maldarelli F, Palmer S, King MS, Wiegand A, Polis MA, Mican J, Kovacs JA, Davey RT, Rock-Kress D, Dewar R, Liu S, Metcalf JA, Rehm C, Brun SC, Hanna GH, Kempf DJ, Coffin JM, Mellors JM. (2007). ART suppresses plasma HIV-1 RNA to a stable set point predicted by pretherapy viremia. *PloS Clinical Trials*. **3(4)**:e46–e87.

Matsuyama Y, Ohashi Y (1997). Mixed models for bivariate response repeated measures data using Gibbs sampling. *Statistics in Medicine*. **16**: 1587–1601.

Merrill S. (1987). AIDS: background and the dynamics of the decline of immunocompetence. Part 2. *Addison-Wesley, Redwood City, Calif*: 59–75.

Merrill S. (1989). Modeling the interaction of HIV with the cells of the immune system. *Mathematical and Statistical Approaches to AIDS Epidemiology*. **83**: 371–385.

Migueles SA, Laborico AC, Shupert WL, Sabbaghian MS, Rabin R, Hallahan CW, Van Baarle D, ..., Connors M. (2002). HIV-specific CD8+ T cell proliferation is coupled to perforin expression and is maintained in nonprogressors. *Nature Immunology*. **3(11)**: 1061–1068.

Müller P, Rosner GL (1997). A Bayesian population model with hierarchical mixture priors applied to blood count data. *Journal of the American Statistical Association*. **92**, 1279 – 1292.

Mueller BU, Zeichner SL, Kuznetsov VA, Heath-Chiozzi M, Pizzo PA, Dimitrov D. (1998). Individual prognoses of long-term responses to antiretroviral treatment based on virological, immunological and pharmacological parameters measured during the first week under therapy. *AIDS*. **12(15)**:F191–F196.

Natarajan R, Kass RE. (2000) Reference Bayesian methods for generalized linear mixed models. *Journal of the American Statistical Association*. **95**: 227–237.

Nie Z, Phenix BN, Lum JJ, Alam A, Lynch DH, Beckett B, Krammer PH, et al, Badley AD. (2002) HIV-1 protease processes procaspase 8 to cause mitochondrial release of cytochrome c, caspase cleavage and nuclear fragmentation. *Cell Death Differ*. **9**:1172–1184.

Notermans DW, Goudsmit J, Danner SA, De Wolf F, Perelson AS, Mittler J. (1998). Rate of HIV-1 decline following antiretroviral therapy is related to viral load at baseline and drug regimen. *AIDS*. **12**:1483–1490.

Nowak K, May R. (1991) Mathematical biology of HIV infections: antigenic variation and diversity threshold. *Mathematical Biosciences*. **106**:1–21.

Merrill S. (1989). Modeling the Interaction of HIV with the Cells of the Immune System, in Mathematical and Statistical Approaches to AIDS Epidemiology. *Lecture Notes in Biomath*. **83**. New York: Springer-Verlag.

Palmer S, Wiegand AP, Maldarelli F, Bazmi H, Mican JM, Polis M, Dewar RL, Planta A, Liu S, Metcalf JA, Mellors JW, Coffin JM. (2003). New real-time reverse transcriptase-initiated PCR assay with single-copy sensitivity for human immunodeficiency virus type I RNA in plasma. *Journal of Clinical Microbiology*. **41(10)**:4531–4536.

Parekh BS, Kennedy MS, Dobb T, Pau CP, Byers R, Green T, et al, McDougal S. (2002). Quantitative detection of increasing HIV type I antibodies after seroconversion: a simple assay for detecting recent HIV infection and estimating incidence. *AIDS Research and Human Retroviruses*. **18**:295–307.

Perelson AS, Kischner D, deBoer R. (1993). Dynamics of HIV infection of CD4 T cells. *Mathematical Biosciences*. **114**:81–125.

Perelson AS, Neumann AU, Markowitz M, Leonard JM, Ho DD. (1996) HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. *Science*. **271(5255)**: 1582–1586.

Perelson AS, Essunger P, Cao Y, Vesane M, Hurley A, Saksela K, Markowitz M, Ho DD. (1997). Decay characteristics of HIV-1 infected compartments during combination therapy. *Nature*. **387**: 188–191.

Perelson AS, Nelson PW. (1999). Mathematical analysis of HIV-1 dynamics in vivo. *Society for Industrial and Applied Mathematics Review*. **41(1)**: 3–44.

Pfister M, Labbe L, Hammer SM, Mellors H, Bennett KK, Rosenkranz S, Sheiner LB and the AIDS clinical trial group protocol 398 investigators. Population pharmacokinetics and pharmacodynamics of Efavirenz, Nelfinavir and Indinavir: adult AIDS clinical trial group study 398. *Antimicrobial Agents and Chemotherapy*. **47(1)**:130–137.

Putter H. Heisterkamp SH. Lange JMA, Wolf F. (2002) A Bayesian approach to parameter estimation in HIV dynamic models. *Statistics in Medicine*. **21**: 2199–2214.

Rice JA, Wu CO. (2001). Nonparametric mixed-effects models for unequally sampled noisy curves. *Biometrics*. **57**:253–259.

Rodriguez B, Sethi A, Cheruvu V. (2006) Predicative value of plasma HIV RNA level on rate of CD4 T-cell decline in untreated HIV infection. *Journal of the American Medical Association*. **296**: 1523–1525.

Rosa GJM, Padovani CR, and Gianola D (2003). Robust linear mixed models with Normal/Independent distributions and Bayesian MCMC implementation. *Biometrical Journal*. **45**, 573–590.

Sahu S, Dey DK, Branco MD (2003). A new class of multivariate skew distribution with application to Bayesian regression models. *The Canadian Journal of Statistics*. **31(2)**: 129–150.

Samson A. Lavielle M. Mentré. (2006). Extension of the SAEM algorithm to left-censored data in non-linear mixed-effects models: application to HIV dynamics models. *Computational Statistical and Data Analysis*. **51**: 1562–1574.

Sedaghat AR, Dinoso JB, Shen L, Wilke CO, Siliciano RF. (2008) Decay dynamics of HIV-1 depend on the inhibited stages of the viral life cycle. *Proceedings of the National Academy of Sciences (PNAS)*. **150(12)**, 4832–4837.

Shah A, Laird D, Schoenfeld A. (1997). A random-effects models for multiple characteristics with possibly missing data. *Journal of the American Statistical Association*. **92**: 775–779.

Shi, M, Weiss, RE, Taylor JMG (1996). An analysis of pediatric CD4+ counts for acquired immune deficiency syndrome using flexible random curves. *Applied Statistics*. **45**, 151–163.

Spiegelhalter D. (2002) Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. **64(4)**:583-639.

Sy JP, Taylor JM, Cumberland WG. (1997). A stochastic model for the analysis of bivariate longitudinal AIDS data. *Biometrics*. **53**: 542–555.

Sy AA, Freed BA, Chau FK, Marcus M. (2007). National estimates of the characteristics of individuals infected with HIV who are likely to report and receive treatment for painful bleeding gums. *Special Care in Dentistry*. **31(5)**:162–169.

Thiébaut R, Jacqmin-Gadda H, Babiker A, Commenges D.(2005). Joint modeling of bivariate longitudinal data with informative dropout and left-censoring with application to the evolution of CD4+ cell count an HIV RNA viral load in response to treatment of HIV infection. *Statistics in Medicine*. **24**: 65–82.

UNAIDS 2011 Report on the Global AIDS Epidemic.
$http://www.unaids.org/en/media/unaids/contentassets/documents/unaidspublication/2011/JC2216\_WorldAIDSday\_report\_2011\_en.pdf$ [Accessed August 17 2012].

U.S. Department of Health and Human Services Panel on Antiretroviral Guidelines for Adults and Adolescents – A Working Group of the Office of AIDS Research Advisory Council (OARAC). March 27, 2012.

$http://aidsinfo.nih.gov/guidelines/html/1/adult-and-adolescent-treatment-guidelines/0/$. [Accessed August 21 2012]

Varbanov M, Espert L, Biard-Piechaczyk B. (2006). Mechanisms of CD4 T-cell depletion trigged by HIV-1 viral infections. *AIDS Reviews*. **8**: 221-236.

Verbeke G, Lesaffre E. (1996). A liner mixed-effects models with heterogeneity in the random-effects population. *Journal of the American Statistical Association*. **433**:217–221.

Vonesh EF, Chinchilli JM. (1996). *Linear and nonlinear models for the analysis of repeated measurements*. New York: Marcel Dekker.

Wang Y. (1998). Mixed-effects smoothing spline ANOVA. *Journal of the Royal Statistical Society, Series B*. **60**:159–174.

Wei X, Ghosh SK, Taylor ME, Johnson, VA, Emini EA, Deutsch PD, et al, Shaw GM. (1995). Viral dynamics in human immunodeficiency virus type 1 infection. *Nature*. **373(12)**: 117–122.

Wu H, Ding A, DeGruttola V. (1998). Estimation of HIV dynamic parameters. *Statistics in Medicine*. **17**: 2463–2485.

Wu H, Kuritzkes DR, McClernon DR, Kessler H, Connick E, Landay A, Spear G, Heath-Chiozzi M, Rousseau F, Fox L, Spritzler J, Leonard JM, Lederman MM. (1999). Characterization of viral dynamics in human immunodeficiency virus type-1 infected patients treated with combination antiretroviral therapy. *Journal of Infectious Diseases*. **179**:799-807.

Wu H, Ding A. (1999). Population HIV-1 dynamics in vivo: applicable models and inferential tools for virological data from AIDS clinical trials. *Biometric*. **55**: 410–418.

Wu H. (2002) A joint model for nonlinear mixed-effects models with censoring and covariates measured with error, with application to AIDS studies. *Journal of the American Statistical Association*. **97**: 955–964.

Wu H, Zhang JT (2002). The study of long-term HIV dynamics using semi-parametric non-linear mixed-effects models. *Statistics in Medicine*. **21**: 3655–3675.

Wu H, Mellors J, Ruan P, McMahon D, Kelleher D, Lederman MM; CNAA2004 Study Investigators. (2003). Viral dynamics and their relations to baseline factors and longer term virologic responses in treatment-naive HIV-1-infected patients receiving abacavir in combination with HIV-1 protease inhibitors. *Journal of AIDS*. **33(5)**:557–63.

Wu H, Zhao C, Liang H. (2004). Comparison of linear, nonlinear and semiparametric mixed-effects models for estimating HIV dynamics parameters. *Biometrical Journal*. **6**: 233–245.

Wu L. (2004). Simultaneous inference for longitudinal data with detection limits and covariates measured with errors, with application to AIDS studies. *Statistics in Medicine*. **23**: 1715–1731.

Wu H, Huang Y, Acosta EP, Rosenkranz SL, Kuritzkes DR, Eron JJ, Perelson AS, Gerber JG. (2005). Modeling long-term HIV dynamics and antiretroviral response. *Journal of AIDS*. **39**: 272–283.

Wu H, Zhang JT (2006). Nonparametric models for longitudinal data with application to CD4 cell numbers in HIV seroconverters. *Biometrics*. **50**: 689–699.

Wu L, Hu XJ, Wu H. (2008). Joint inference for nonlinear mixed-effects models and time to event at the presence of missing data. *Biostatistics*. **9(2)**: 308–320.

Zhang D, and Davidian M (2001). Linear mixed models with flexible distributions of random effects for longitudinal data. *Biometrics*. **57**, 795–802.

Zhang D. Din X, Raz J. Sowers M. (1998). Semiparametric stochastic mixed models for longitudinal data. *Journal of the American Statistical Association*. **93**:710–719.

Zhang J. WU H. (2011). Modeling HIV dynamics using unified mixed-effects models. *American Journal of Mathematical and Management Sciences*. **30(2)**: 83–111.

Zhao Y, Staudenmayer H, Coull BA, Wand MP. (2006) General design Bayesian generalized linear mixed models. *Statistical Science*. **21(1)**: 35–51.

Appendices

Appendix A: WinBUGS Code for ST-Model IV-Equation (2.12) in Chapter 2

```
##  Begin of model
model
{
for (i in 1:N)
{
  ## parametric random effects of normal-Response model
   b2[i,1] <- 0
   b2[i,2] <- 0
   b2[i,3] <- 0
   b2[i,4] <- 0
   b2[i,5] <- 0
   b[i,1:5] ~dmnorm(b2[i,1:5],Omega2[,])

   nbeta1[i]<- beta[1]+b[i,1]
   nbeta2[i]<- beta[2]+b[i,2]
   nbeta3[i]<- beta[3]+b[i,3]
   nbeta4[i]<- beta[4]+b[i,4]
   nbeta5[i]<- beta[5]+b[i,5]
 ## Individual  parameter estimates
 }

  for (j in 1 : M)
 {
 ## Main components of response based on ST
beta1[j] <- beta[1]+b[y[j,4],1]
beta2[j] <- beta[2]+b[y[j,4],2]
beta3[j] <- beta[3]+b[y[j,4],3]
beta4[j] <- beta[4]+b[y[j,4],4]
# y[j,4]= id
beta5[j] <- beta[5]+b[y[j,4],5]

    #decay rate for Model IV
decay[j] <- (beta2[j]*exp(-beta3[j]*y[j,6]) +
             beta4[j]*exp(-beta5[j]*y[j,6]) )

dm1[j] <-beta1[j] -decay[j]*y[j,6]
# y[,6]= time(day)

 ## skew-T distribution to t sidtribution
w[j]  ~  dt(0, 1,nu) I(0,)
mu[j]  <- dm1[j] + delta*w[j]
# ST distribution
aau[j] <- (nu+w[j]*w[j])/n1*eta
y[j,12]  ~ dt(mu[j],aau[j],n1)
# y[,12]=logeRNA
Y.pred[j]  ~ dt(mu[j],aau[j],n1)
```

```
# predicted values

## Fitted values and Residuals
fit[j]  <- mu[j]
resid[j]   <-  y[j,12]-fit[j]
sresid[j] <- sqrt(eta)*(y[j,12]-fit[j])
ssr.r[j]  <- pow(resid[j],2)
 # squares of residuals
sssr.r[j]  <- pow(sresid[j],2)
 # squares of SR
ssr.Y.obs[j]<-pow((Y.pred[j]-y[j,12]),2)
 }
SSR <-sum(ssr.r[])
 # sum of squares of residuals
SSSR <- sum(sssr.r[])
# sum of standardized squares of residuals
SSR.pred <- mean(ssr.Y.obs[])
# EPD

 ## Prior distributions of the hyperparameters
  # (0) Degree of freedom
      nu0<-0.1
      nu ~ dexp(nu0) I(2,)
      n1<-nu+1

   # (1) Coefficients
     for (l in 1:5){ beta[l]~dnorm(0,1.0E-2)}

  # (3) Skew-t random effects
 delta ~ dnorm(0.0, 0.01)

  # (4) Variance-covariance matrice
  Omega2[1:5,1:5] ~dwish(R2[,],5)
 v2[1:5,1:5] <- inverse(Omega2[,])

   # (5) Precision parameters
      eta ~dgamma(0.01,0.01)
  sigma <- 1/eta
}
## End of model

## Data inputed
list(N=44, M=274,
       R2 = structure(
.Data = c(1,0, 0, 0, 0,
          0,1, 0, 0, 0,
          0,0, 1, 0, 0,
          0,0, 0, 1, 0,
```

```
          0,0, 0, 0, 1),
.Dim = c(5, 5)))

## Initial values
list(beta=c(4.54, -5.55, 5.00, 0.5, 1),nu=10,delta=1.5, eta=1,
       Omega2= structure(
.Data = c( 1.94 ,   1.00,   -0.56,  -0.58, -0.45,
                        1.00,     7.03, -1.08, -3.93, -1.08,
                         -0.56,    -1.08, 1.30, 2.74, 0.69,
                         -0.58,    -3.93, 2.74, 8.75, 2.07,
                         -0.45,    -1.08, 0.69, 2.07, 0.66),
                         .Dim = c(5, 5)))

 ##  End of program
```

Appendix B: WinBUGS Code for Model I in Chapter 3

```
# Model I: Skew-Normal Bayesian Semiparametric
# Nonlinear Mixed-Effects
# (SN-BSNLME) Model in conjunction with the
# semiparametric biexponential
# model (3.3)}.

model{
       for (i in 1:44) {
      a2[i,1] <- 0
      a2[i,2] <- 0
      a2[i,3] <- 0
  a2[i,4] <- 0

     a3[i,1] <- 0
     a3[i,2] <- 0
     a3[i,3] <- 0

   b[i,1:4]~dmnorm(a2[i,1:4], Sigma2[,])
   a[i,1:3]~dmnorm(a3[i,1:3], Sigma3[,])
                        }

 for (j in 1:310) {
 ## modeling true CD4 via measurement error model
     z.star[j] <-(alpha[1]+a[y[j,4],1]) +
          (alpha[2]+a[y[j,4],2])*y[j,13] +
            (alpha[3]+a[y[j,4],3])*y[j,13]*y[j,13]+
            delta2*w2[j]
     w2[j] ~dnorm(0,1)I(0,)

  y[j,11]~dnorm(z.star[j],tau2)

  ##  Viral load response model associated with true CD4 covariate
  betai1[j]  <-beta[1] +b[y[j,4],1]
  betai2[j]  <-beta[2] +b[y[j,4],2]
  betai3[j]  <-beta[3] +b[y[j,4],3]
  betaij4[j] <-beta[4] +beta[5]*z.star[j]
  +mu.not[1]*Z[j,2]+mu.not[2]*Z[j,3]+b[y[j,4],4]

  dm1[j] <-betai1[j]-step(betai2[j]-betaij4[j])*betai2[j]*y[j,13]
  dm2[j] <-betai3[j]-step(betai2[j]-betaij4[j])*betaij4[j]*y[j,13]
  dm3[j] <-exp(dm1[j])
  dm4[j] <-exp(dm2[j])
  dm5[j] <-dm3[j] +dm4[j]

  mu[j]  <-log(dm5[j]) +delta*w[j]
```

```
  w[j]~dnorm(mu[j],tau)
  y[j,10]~dnorm(mu[j],tau)    # SN

 ## Fitted values and residues
   fit[j] <-mu[j]
   res[j] <-y[j,10]-fit[j]
 }

 ## prior distributions of the hyperparameters:

# 1. coefficients
for(l in 1:5) {beta[l]~dnorm(0, 1.0E-2)}
for(l in 1:2) {mu.not[l]~dnorm(0, 1.0E-2)}
for(k in 1:3) {alpha[k]~dnorm(0, 1.0E-2)}

# 2. Precision parameters
tau~dgamma(0.01, 0.01)
sigma.tau <- 1/tau
tau2~dgamma(0.01, 0.01)
sigma.tau2 <- 1/tau2

# 3. Variance-cvoariance matrices
Sigma2[1:4,1:4]~dwish(R2[,],5)
v2[1:4,1:4] <-inverse(Sigma2[,])
Sigma3[1:3,1:3]~dwish(R3[,],5)
v3[1:3,1:3] <-inverse(Sigma3[,])

# 4. skewness parameters
    delta~dnorm(0.0,0.01)
    delta2~dnorm(0.0,0.01)
  }

## Data
list(R2=structure(.Data=c(1,0,0,0,0,1,0,0,0,0,1,0,0,0,0,1),
.Dim=c(4,4)),R3=structure(.Data=c(1,0,0,0,1,0,0,0,1), .Dim=c(3,3)))

#initial values
list(beta=c(5, 4, 3.4, 4.0, 1.5, 4.2),
 alpha=c(0,0,0),mu.not=c(0,0),delta=0.5,
 delta2=0.5,tau=1, tau2=1,
 Sigma2=structure(.Data=c(1.229,  0.043, -0.750,  0.710,
                          0.043,  0.090, 0.002,  -0.013,
                          -0.750, -0.002,1.059, -0.214,
                          0.170, -0.013,-0.214, 0.120),
                          .Dim=c(4,4)),
Sigma3=structure(.data=c(.1,0, 0, 0,.1, 0, 0, 0,.1), .Dim=c(3,3)))
```

Appendix C: WinBUGS Code for ST Bivariate Model-Equation (4.5) in Chapter 4

```
model
{
for (i in 1:N)
   {  u.b[i] ~ dgamma(db,db)

# random effects for bivariate model
# correlation between Cd4 and CD8 is
 incorporated in random effects and model errors
for (k in 1:2)
 { b1[i,k] <- 0
  w.b1[i,k]~dnorm(b1[i,k], R.a[k,k]) I(0,)
  w.b[i,k]<-w.b1[i,k]/sqrt(u.b[i])
  b1.b[i,k]<-delta1*(w.b[i,k]-mub)
}
 for (k in 3:4)
 { b1[i,k] <- 0
   w.b1[i,k]~dnorm(b1[i,k], R.a[k,k]) I(0,)
   w.b[i,k]<-w.b1[i,k]/sqrt(u.b[i])
   b1.b[i,k]<-delta2*(w.b[i,k]-mub)
 }
for (k in 1:4)
  {
  for (l  in 1:4)
   {
   Omegab[i,k,l]<-Omega[k,l]*u.b[i]
   }
  }
 b[i,1:4] ~dmnorm(b1.b[i,1:4],Omegab[i,,])
   }

 for (j in 1 : M)
   {  # ST-Bivariate CD4 and CD8 LME model
beta1[j] <-beta[1] +beta[2]*y[j,2]+b[y[j,4],1]
 #y[j,2]=arm
beta2[j] <-beta[3] +beta[4]*y[j,2]+b[y[j,4],2]
#y[j,4]=id
alpha1[j] <-alpha[1] +alpha[2]*y[j,2]+b[y[j,4],3]
#y[j,2]=arm
alpha2[j] <-alpha[3] +alpha[4]*y[j,2]+b[y[j,4],4]
#y[j,4]=id

cd[j,1] <-beta1[j] + beta2[j]*y[j,6] +beta[5]*y[j,17]
         +beta[6]*y[j,20] +beta[7]*y[j,18]
cd[j,2] <-alpha1[j] + alpha2[j]*y[j,6] +alpha[5]*y[j,17]
         +alpha[6]*y[j,20] +alpha[7]*y[j,18]
```

```
# y[j,6]=scaled time (0,1); y[j,17] =age;
# y[j,20]=baseline ln(RNA); y[j,18]=eff
for (k in 1:2){
 for (l  in 1:2)
  {
 O1[j,k,l]<-Omega1[k,l]*u.b[y[j,4]]
  }
  }

y[j,15:16]~dmnorm(cd[j,1:2], O1[j,,])
# y[j,15]=CD4, y[j,16]=CD8

Y.pred[j,1:2]~dmnorm(cd[j,1:2], O1[j,,])


 }


 # Prior distributions of the hyperparameters
  # (0) Degree of freedom
     nub~ dexp(0.1) I(2,)
     nb<-nub+1
     mub<- exp(loggam(0.5*(nub-1.))-
          loggam(0.5*nub))*sqrt(nub/3.14159)
     db<-0.5*nub

  # (1) Coefficients
for (l in 1:7)
    {  beta[l]~dnorm(0,0.01)
       alpha[l]~dnorm(0,0.01)}

  # (2). Variance-covariance matrice for model errors
  Omega1[1:2,1:2] ~dwish(R1[,],3)
  v1[1:2,1:2] <- inverse(Omega1[,])

  # (3) Variance-covariance matrice for random effects
   Omega[1:4,1:4] ~dwish(R[,],5)
   v[1:4,1:4] <- inverse(Omega[,])

 # (4). Skewness parameters
  delta1~dnorm(0,0.01)
  delta2~dnorm(0,0.01)

}
## End of model
```

Appendix D: WinBUGS Code for ST Univariate Model for CD4 in Chapter 4

```
##  Begin of model
model
{
for (i in 1:N)
   {
 u.b[i] ~ dgamma(db,db)  # for random effects of bivariate model

# Random effects for bivariate model
            for (k in 1:2)
             { b1[i,k] <- 0
               w.b1[i,k]~dnorm(b1[i,k], R.a[k,k]) I(0,)
              w.b[i,k]<-w.b1[i,k]/sqrt(u.b[i])
               b1.b[i,k]<-delta1*(w.b[i,k]-mub)
              }
            for (k in 1:2){
               for (l  in 1:2)
                    {
                      Omegab[i,k,l]<-Omega[k,l]*u.b[i]
                    }
              }
             b[i,1:2] ~dmnorm(b1.b[i,1:2],Omegab[i,,])
   }
 for (j in 1 : M)
   {    # Univariate LME model for CD4
        beta1[j] <-beta[1] +beta[2]*y[j,2]+b[y[j,4],1]
            #y[j,2]=arm
        beta2[j] <-beta[3] +beta[4]*y[j,2]+b[y[j,4],2]
         #y[j,4]=id
        cd4[j] <-beta1[j] + beta2[j]*y[j,6] +beta[5]*y[j,17]
        +beta[6]*y[j,20] +beta[7]*y[j,18]
        # y[j,6]=scaled time (0,1); y[j,17] =age;
        # y[j,20]=baseline ln(RNA); y[j,18]=eff

        au4[j]<-u.b[y[j,4]]*tau1
        y[j,15] ~dnorm(cd4[j], au4[j])
        # y[,j,15] =standarized CD4
        Y.predcd4[j]~dnorm(cd4[j], au4[j])

  # Fitted values and Residuals
      fitcd4[j] <- cd4[j]
      residcd4[j] <-  y[j,15]-fitcd4[j]
      sresidcd4[j] <- sqrt(tau1)*(y[j,15]-fitcd4[j])*sqrt(1- 2/nub)
      ssr.cd4[j] <- pow(residcd4[j],2)
      # SR of CD4
```

```
      sssr.cd4[j] <- pow(sresidcd4[j],2)
      # SSR of CD4
      ssr.Pcd4[j]<-pow((Y.predcd4[j]-y[j,15]),2)
      # squares of predicted value
    }

SSR.CD4  <- sum(ssr.cd4[])
# sum of SR of CD4
SSSR.CD4 <- sum(sssr.cd4[])
# sum of SSR of CD4
SSRP.CD4<- mean(ssr.Pcd4[])
# EPD of CD4

 ## Prior distributions of the hyperparameters
  # (0) Degree of freedom
nub~ dexp(0.1) I(2,)
nb<-nub+1
mub<- exp(loggam(0.5*(nub-1.))-loggam(0.5*nub))*sqrt(nub/3.14159)
     db<-0.5*nub

  # (1) Coefficients
       for (l in 1:7)
        {beta[l]~dnorm(0,0.01)}

  # (2). Precision parameters
          tau1~dgamma(0.01,0.01)
          sigma1 <- 1/tau1

  # (3) Variance-covariance matrice
      Omega[1:2,1:2] ~dwish(R[,],5)
      v[1:2,1:2] <- inverse(Omega[,])

 # (4). Skewness parameters
          delta1~dnorm(0,0.01)
}
## End of model

# Data
list(N=44, M=310,
       R = structure(
.Data = c(1, 0.5,
         0.5, 1),
.Dim = c(2,2)),
   R.a= structure(
            .Data=c(1,0,
        0,1),
 .Dim = c(2,2)))
```

Appendix E: Permission of reprint for Chapter 3

Chen, Ren

Subject: FW: questions about using one published paper as part of my dissertation

From: Spencer McGrath [mailto:spencer.mcgrath@degruyter.com]
Sent: Thursday, September 06, 2012 2:56 PM
To: Chen, Ren
Subject: RE: questions about using one published paper as part of my dissertation

HI Ren,

Permission granted. Re-use in dissertations is permitted.

All the best,
Spencer

Spencer McGrath
Journals Editor

DE GRUYTER
121 High Street, Third Floor
Boston, MA 02110
USA
T (857) 284-7073 x111
C (914) 434-6976
F (857) 284-7358
Spencer.McGrath@degruyter.com

From: Chen, Ren [mailto:rchen@health.usf.edu]
Sent: Tuesday, September 04, 2012 4:04 PM
To: Spencer McGrath
Subject: questions about using one published paper as part of my dissertation

Dear Mr. Spencer,

We published the following article in your 2011 **International Journal of Biostatistics**:

Huang, Yangxin; **Chen, Ren**; and Dagne, Getachew (2011) "Simultaneous Bayesian Inference for Linear, Nonlinear and Semiparametric Mixed-Effects Models with Skew-Normality and Measurement Errors in Covariates," *The International Journal of Biostatistics*: Vol. 7 : Iss. 1, Article 8.
I am the 2nd author and hope to use this paper as one part of my Ph.D. dissertation. I already got permission from the two authors, Dr. Huang and Dr. Dagne. I hope to get the permission from the Journal. Could you please tell me what is the process that I should follow to apply the permission?

Since I will include three manuscripts in my dissertation, I intend to include this paper as its published form with the full title of the journal and periodical.

Thanks.

Ren Chen

About the Author

Ren Chen originally came from Xi'an, China. She worked as an internal physician for three years after graduating with M.D. from the College of Medicine, Xi'an Jiaotong University, Xi'an, China. Ms. Chen also got an M.P.H in Biostatistics and M.H.A, both from the College of Public Health, University of South Florida (USF).

Ms. Chen has been working at USF for over twelve years as statistical data analyst and Biostatistician, at Policy and Services Research Data Center , Department of Mental Health Law & Policy (2000 $\sim$ 2007), and Clinical & Translational Science Institute, Morsani College of Medicine (2007 $\sim$ current). She, as a biostatistician, has participated in many clinical trials and administrative studies, and provided critical statistical support.