©Copyright 2015
Gordon Chau

Bayesian estimation of participants' adherence in an HIV prevention trial using multiple data sources and pharmacokinetic models

Gordon Chau

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Master of Science

University of Washington

2015

Reading Committee:

Jim Hughes, Chair

Jon Wakefield

Program Authorized to Offer Degree:
UW Biostatistics

University of Washington

Abstract

Bayesian estimation of participants' adherence in an
HIV prevention trial using multiple data
sources and pharmacokinetic models

Gordon Chau

Chair of the Supervisory Committee:
Ph.D Jim Hughes
UW Biostatistics

Pre-Exposure Prophylaxis (PrEP) is an HIV prevention practice which uses antiretrovial drugs to keep HIV-negative individuals from infection with HIV. One of the antiretroviral drugs that is usually used for HIV treatment but has been widely adopted in PrEP is Truvada by Gilead. In the recent studies of using Truvada as PrEP, the results have been variable regarding to the efficacy of Truvada for HIV prevention. However, the adherence to pill-taking varies on a large scale across those studies, which can potentially underestimate the drug efficacy.

The HPTN 067 trial is an open-label, phase II HIV prevention trial, that studies the coverage and adherence associated with different PrEP pill-taking strategies. Motivated by the HPTN 067 trial, a statistical model has been developed to estimate the individual-level adherence, misreporting and pharmacokinetic (PK) parameters using self-reported adherence and drug concentrations in plasma. The goal of this thesis is to investigate the estimation method used in this model and evaluate/improve its performance in terms of computational efficiency and data with detection of limits. In addition, a new statistical model (two-biological-measure model) is developed to estimate the individual-level adherence, misreporting and PK parameters from two biological samples (i.e. plasma and PBMC).

Pharmacokinetic models are used to estimate the individual-level PK parameters. A one

compartment IV bolus model is used to estimate the PK parameters for plasma. The two-biological-measure model incorporates a second biological measure (i.e. PBMC) with different half-life into the previously developed statistical model, using a one compartment first-order absorption model. The statistical models are analyzed using MCMC methods. All the data are simulated in this study. The observed data include self-reported adherence, drug concentrations measured in plasma at multiple time points, drug concentration measured in PBMC at multiple time points and lower limits of detection of drug concentrations in plasma and PBMC.

In this study, a shorter look-back period is used to improve the computational efficiency and correct the sensitive estimation. Estimation accuracy is also improved using detection limits. The two-biological-measure model improves the accuracy in estimating adherence but has a non-identifiability issue which can be fixed using informative priors for the population PBMC absorption rate parameter. However, the subject-specific absorption rate parameters in PBMC still cannot be accurately estimated.

A two-biological-measure model adopting a multi-compartment system is also evaluated. Although the model still has non-identifiability issue, it gives good estimation of individual adherence probabilities and some PK parameters.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to wholeheartedly thank my thesis advisor, Dr. Jim Hughes for his keen interest and persistent guidance on me at every stage of my research. The value of his suggestions and assistance to my thesis preparation is unmeasurable.

I would like to thank my reading committee member, Dr. Jon Wakefield, who has offered me many helps with pharmacokinetics and computer programming. He has devoted lots of time to read my thesis which I cannot be more appreciated.

I would also like to thank Chloe Krakauer, a graduate student from University of Washington who studied this project prior to me. I thank her for introducing me to this project and answering my questions with great patience and careful instruction.

In addition, I would like to thank my beloved mother, Bei and my father Alex who have never stopped giving me guidance and encouragement.

In the end, I would like to acknowledge the faculty, staff and fellow students of the Biostatistics and Statistics departments at University of Washington. It is my honor to spend time with you through my graduate study. Thank you all.

DEDICATION

to my parents, Bei and Alex.

Chapter 1
INTRODUCTION

1.1 The Battle of HIV Prevention: Adherence vs. Efficacy

Pre-Exposure Prophylaxis (PrEP) is an HIV prevention approach in which people who do not have HIV use antiretroviral (ARV) drugs to reduce the risk of HIV infection. This concept has gained attention since using Truvada as PrEP was approved by the FDA for the prevention of HIV infection in July 2012[1]. Truvada, a type of ARV medicine that contains two drugs (Tenofovir and Emtricitabine), is believed to be a revolutionary and extremely beneficial drug for HIV-uninfected individuals who are at high risk of HIV infection. While several clinical trials have showed that Truvada can reduce the risk of HIV infection[2], other studies have shown that Truvada did not significantly lower the risk of infection[3]. One of the major issues in these trials is believed to be the inconsistent adherence of the treatment drugs which could potentially underestimate the true efficacy of Truvada[4].

1.2 HTPN 067 Trial

HTPN 067 trial (The ADAPT Study) is a phase II, open-label randomized clinical trial of dosing strategy for oral Truvada. The study enrolled approximately 540 participates from Africa, Thailand and the USA. One of the objectives of this trial is to estimate adherence associated with different PrEP pill-taking schedules[5].

The dosing period of this trial consists at two phases: a directly observed therapy period (DOT) and a self-administrated period (SA). The DOT phase includes time 0 (enrollment) and weeks 1 - 5 of the study, during which participants take one directly observed dose of Truvada[1] weekly at enrollment and weeks 1 - 4, followed by one week without dosing. Bio-

[1]One dose of Truvada contains 300 mg of Tenofovir Disoproxil Fumarate and 200 mg of Emtricitabine

logical samples are collected during weeks 4 and 5 in this phase. Since the rate of adherence is known to be 100% in the DOT phase, researchers are able to estimate the individual-level pharmacokinetic (PK) parameters after collecting drug concentrations in the biological samples which includes plasma, peripheral blood mononuclear cell (PBMC), dried blood spot (DBS) and hair samples. Participants are then randomized into three treatment groups: daily dosing, time-driven dosing or event-driven dosing at week 6. Participants in the daily dosing group are instructed to take one dose of Truvada daily. In the time-driving dosing group, participants are instructed to take two doses per week plus a post-exposure boost within two hours after potential exposure to HIV infection. The event-driven group are instructed to take one dose within 24 - 48 hours before and within two hours after a potential exposure to HIV infection. Randomization is followed by entry into the 24 weeks self-administrated (SA) phase. During the SA phase, participants do not consume Truvada under direct observation but instead are interviewed weekly and asked to self-report their pill-taking and sexual behavior. The schedules of the biological sample collection in the DOT and SA phase are shown in table 1.1. In addition, participants are asked to take their pills only from an EDM device/case, which is an electronic device that transmits the medicine case opening times and dates to the investigators.

The EDM might underestimate or overestimate the adherence[6] since a case opening is assumed the same as taking one dose of Truvada[5]. For instance, participants could take out more than one pill or not take any pill when opening the EDM case. Although self-reported adherence through weekly interview may help adjusting the EDM accuracy, there is still a possibility of recall bias and prevarication (lying about adherence). Thus we propose developing a model that will allow us to estimate the individual-level probability of adherence and misreporting using both self-report and biological measures.

Table 1.1: HPTN 067 schedule of biological sample collection

Week since time 0	0	1	2	3	4	5	6	10	14	18	22	26	30
DOT phase	X	X	X	X	X	X[1]							
SA phase							X[2]	X	X	X	X	X	X
Plasma sample					X	X	X	X	X	X	X	X	X
PBMC sample					X	X	X	X		X			X
DBS sample					X	X	X	X		X			X
Hair sample					X	X	X	X		X			X

[1] No drug is administrated in week 5

[2] Randomization

1.3 Thesis Overview

In 2014, Chloe Krakauer, a Biostatistics M.S. student from University of Washington developed a statistical model that estimates the individual-level PK parameters and probabilities of adherence and misreporting using simulated drug concentration measurements in a single biological measure and self-reported adherence data[7].

We incorporate the research of Krakauer into our study and improve the model speed and capability of dealing data with a lower limit of detection. In addition, to investigate the hypothesis that estimation of adherence and misreporting would be more accurate using data from multiple biological measures, we expand Krakauer's model to include drug concentrations in plasma and PBMC (two-biological-measure model).

The general approach in this thesis is based on the assumption that each participant has a true unobserved adherence which can be estimated using Bayes' theorem. Let A be the unobserved adherence and R be the collected self-reported adherence. Let $C.pl$ and $C.pb$ represent the concentration measurements in plasma and PBMC. Also, consider $p(A)$ as the prior distribution of adherence. Let $p(R|A)$, $p(C.pl|A)$, $p(C.pb|A)$ represent the likelihood of

self-reported adherence and drug concentrations in plasma and PBMC which are specified by the statistical model. The distribution of A given R, $C.pl$ and $C.pb$ (posterior) can be updated using Bayes' theorem,

$$p(A|R, C.pl, C.pb) = \frac{p(R, C.pl, C.pb|A)p(A)}{p(R, C.pl, C.pb)} \tag{1.1}$$

$$\propto p(R|A)p(C.pl|A)p(C.pb|A)p(A) \tag{1.2}$$

assuming R, $C.pl$ and $C.pb$ are conditionally independent, given A. More detailed methods of the statistical model are described in Chapter 2.

Chapter 2

METHODS

The methods chapter is divided into four sections: the methods of data simulation, an introduction to the statistical model, a description of the model improvements/extensions and the methods of the two-biological-measure model with a multi-compartment system. All the data used in this thesis are simulated. In the data simulation section, we describe how these data are generated. We then discuss the model and statistical estimation using Bayesian Inference MCMC. In the model improvement section, we will focus on three topics: improving model speed, censoring and a two-biological-measure model. In the last section, we describe how to simulate data using a multi-compartment PK model and how to build a two-biological-measure model incorporating the multi-compartment system.

2.1 Methods of Data Simulation

2.1.1 Design of Pill-taking and Biological Samples Collection

The simulated data include drug concentrations and participants' self-reported adherence. Additional information on the lower limit of detection and prior information about PK parameters are also incorporated into the model. The design of the simulated data is based on the pill-taking and sample collection schedule of the HPTN 067 trial which includes two phases of study: DOT phase and SA phase. In the simulated trial, we assume there are 100 participants and the trial is designed to be 73 days long in total with the length of the DOT and the SA phase being 43 days and 30 days respectively[1]. During the DOT phase, all participants are scheduled to take one pill weekly for 5 weeks starting on the first day (see

[1]The SA phase in the simulations is shorter than the 24 week SA phase in HPTN 067 trial to facilitate the simulation studies.

table 2.1). The rate of the adherence is 100% and the rate of falsely report adherence given non-adherence is 0% in this phase.

Table 2.1: Simulated pill-taking and sample collection schedule

Day in the trial	1	8	15	22	29	36	43	48	53	58	63	68	73
Directly observed pill-taking	X	X	X	X	X								
Plasma sample collection					X	X	X	X	X	X	X	X	X
PBMC sample collection					X	X	X	X	X	X	X	X	X

Let A_{ij} be the true adherence and R_{ij} be the self-reported adherence for subject i at the j^{th} prescribed pill-taking event. Let t_{ij} be the day of the j^{th} prescribed pill-taking event for subject i. Then, for the DOT period

$$A_{ij} = 1 \text{ for } t_{ij} \leqslant 43 \tag{2.1}$$

$$R_{ij} = 1 \text{ for } t_{ij} \leqslant 43 \tag{2.2}$$

for i in (1,...,100).

Biological samples are measured three times during the DOT phase (day 29[2], 36 and 43). The SA phase starts immediately after the DOT phase. The data in the SA phase are simulated in line with the study design of daily dosing treatment group in the HPTN 067 trial so that participants are scheduled to take medicine on a daily base starting on day 43[3]. The biological samples are collected every 5 days starting on day 48 till the end of the study (see table 2.1). In total, 9 measurements are designed to be collected for each participant. The drug concentrations in biological samples are modeled using pharmacokinetic models (See section 2.1.2).

[2]The biological samples are collected before the prescribed pill-taking time on day 29.

[3]The biological samples are collected before the prescribed pill-taking time on day 43.

2.1.2 Drug Concentration

In order to simulate the drug concentrations in plasma and PBMC, we use pill-taking times, biological sample collection times, individual-level pharmacokinetic parameters and true adherence probabilities. Assuming two biological samples (i.e. plasma and PBMC) are collected and assuming a one compartment first-order IV bolus model for plasma and a one compartment first-order absorption model for PBMC, there are five pharmacokinetic parameters for each subject (see section 2.1.3): elimination rate constant for plasma ($k^e.pl$), creatinine clearance in plasma ($Cl.pl$), elimination rate constant for PBMC ($k^e.pb$), creatinine clearance in PBMC ($Cl.pb$) and absorption rate constant for PBMC ($k^a.pb$). Although a multi-compartment system describes the plasma pharmacokinetics and PBMC pharmacokinetics in a more realistic way[8], it is also more complicated. We concern that using a multi-comparment system to describe drug concentrations at the stage of investigating the performance of the statistical model (see section 2.2) can be time-consuming since multiple experiments are required to be implemented. Therefore, simpler pharmacokinetic models are used to facilitate the computational efficiency of the statistical model.

As noted above, the drug concentrations in plasma are simulated using a one compartment first-order IV bolus model and drug concentrations in PBMC are simulated using a one compartment first-order absorption model. The pharmacokinetic models are used to compute the subject-specific expected drug concentrations in plasma and PBMC for taking one dose of Truvada t time units ago. Theoretically, after drug concentration has reached the peak in the body system, it decays as a function of the time since the drug is taken. For example, assuming the drug concentration in plasma reaches the peak immediately after a subject has consumed 300 mg of Tenofovir disoproxil fumarate[4] (i.e. IV bolus model), the expected drug concentration in plasma may be 90 ng/ml at the peak, 47 ng/ml a day later and 24 ng/ml in two days[5]. Since each participant takes pills multiple times in the HPTN 067 trial,

[4]Tenofovir disoproxil fumarate (TDF) is an inactive oral formulation of Tenofovir that is contained in Truvada[5].

[5]The values vary depending on PK parameters.

we assume drug concentrations have an additive effect. Let j be the index for counts of prescribed pill-taking events and let k be the index for counts of biological sample collection. Then, the expected drug concentration (C_{ik}) at time t_{ik} for subject i who has taken pills at times t_{ij} are simulated using formula:

$$C_{ik} = \sum_{ij|t_{ij}<t_{ik}} A_{ij} h(\boldsymbol{\theta_i}, ij, ik) \tag{2.3}$$

A_{ij} is an indicator for true adherence for subject i at t_{ij}, where t_{ij} is the time of the j^{th} prescribed pill-taking event. t_{ik} is the time of the k^{th} biological sample collected for subject i. Only the pills taken prior to the k^{th} measurement contribute to the computation of C_{ik}. $\boldsymbol{\theta_i}$ is a vector of all the individual-level pharmacokinetic parameters for subject i and $h(\boldsymbol{\theta_i}, t_{ij}, t_{ik})$ represents the pharmacokinetic model of drug concentration in a biological sample t_{ik} - t_{ij} since one dose of Truvada is taken by subject i.

For a one-compartment IV bolus model:

$$h(\boldsymbol{\theta_i}, t_{ij}, t_{ik}) = \frac{Dk^e.pl_i}{Cl.pl_i} e^{-k^e.pl_i(t_{ik}-t_{ij})} \tag{2.4}$$

For a one-compartment first-order absorption model:

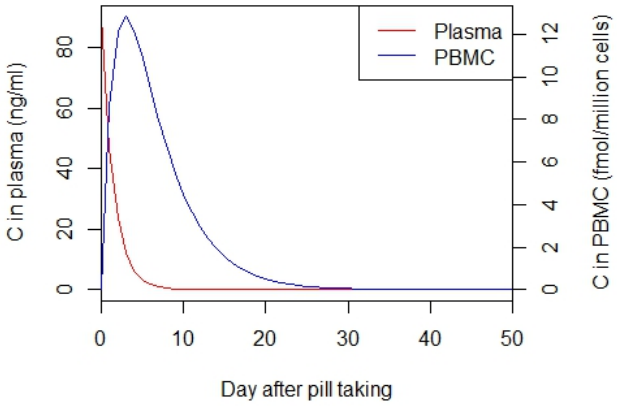
$$h(\boldsymbol{\theta_i}, t_{ij}, t_{ik}) = \frac{Dk^e.pb_i k^a.pb_i}{Cl.pb_i(k^a.pb_i - k^e.pb_i)} (e^{-k^e.pb_i(t_{ik}-t_{ij})} - e^{-k^a.pb_i(t_{ik}-t_{ij})}) \tag{2.5}$$

assuming $k^a.pb_i > k^e.pb_i$.

Figure 2.1 shows the distribution of mean drug concentration in plasma and PBMC as a function of time since pill-taking (time 0). According to figure 2.1, the drug concentration in plasma, which is modelled with the one compartment IV bolus model, hits the peak immediately after the pill is taken and drop rapidly from then on. This is reasonable for plasma with a daily scale since the half-lives of Tenofovir and Emtricitabine in plasma are both less than a day[6]. The drug concentration in PBMC increases gradually and reaches the peak in approximately 3 - 4 days. Moreover, the drug is almost completely removed from

[6]Tenofovir and Emtricitabine are the components of Truvada. The plasma half-life of Tenofovir is about 17 hours[9] and the plasma half-life of Emtricitabine is about 10 hours[10].

Figure 2.1: Distribution of drug concentration (C) versus time since pill-taking



both plasma and PBMC in 28 days.

Once the expected drug concentration C_{ik} is obtained, the expected log drug concentration for subject i at the k^{th} measurement is computed.

$$E(Y_{ik}) = \log(C_{ik}) \tag{2.6}$$

The measurement error is simulated using

$$\epsilon_{ik} \sim \mathcal{N}(0, 0.1^2) \tag{2.7}$$

The observed log drug concentration is

$$Y_{ik} = E(Y_{ik}) + \epsilon_{ik} = \log(C_{ik}) + \epsilon_{ik} \tag{2.8}$$

2.1.3 Pharmacokinetic Parameters

The drug dosage (D) is set equal to 300 mg[7] for all participants. When simulating only one biological measure (i.e. plasma), the PK parameters $\boldsymbol{\theta_i}$ for subject i are simulated from a

[7]D is chosen to be in line with the amount of TDF in Truvada[5].

multivariate normal distribution.

$$\boldsymbol{\theta_i} = \begin{pmatrix} \log(k^e.pl_i) \\ \log(Cl.pl_i) \end{pmatrix} \sim \mathcal{N}\left[\begin{pmatrix} -.41 \\ 0.79 \end{pmatrix}, \begin{pmatrix} 0.02 & -0.03 \\ -0.03 & 0.07 \end{pmatrix} \right] \tag{2.9}$$

for i in $(1, \ldots, 100)$. The values of the mean and the variance of the multivariate normal distribution are chosen based on the plasma samples collected over 15 days from an experiment with 6 individuals taking one single dose of Truvada[7].

In the simulation when two biological measures (i.e. plasma and PBMC) are collected, the pharmacokinetic parameters are simulated using

$$\boldsymbol{\theta_i} = \begin{pmatrix} \log(k^e.pl_i) \\ \log(Cl.pl_i) \\ \log(k^e.pb_i) \\ \log(Cl.pb_i) \\ \log(k^a.pb_i) \end{pmatrix} \sim \mathcal{N}\left[\begin{pmatrix} -.41 \\ 0.79 \\ -1.5 \\ 1 \\ -0.693 \end{pmatrix}, \begin{pmatrix} 0.02 & -0.03 & 0 & 0 & 0 \\ -0.03 & 0.07 & 0 & 0 & 0 \\ 0 & 0 & 0.04 & -0.05 & 0 \\ 0 & 0 & -0.05 & 0.09 & 0 \\ 0 & 0 & 0 & 0 & 0.05 \end{pmatrix} \right] \tag{2.10}$$

The mean and the covariance matrix for $\log(k^e.pl)$ and $\log(Cl.pl)$ are inherited from equation (2.9). The values of mean of $\log(k^e.pb)$, $\log(Cl.pb)$ and $\log(k^a.pb)$ are chosen so that the drug concentration in PBMC reaches the peak $\approx$ 12 fmol/million cells in 3-4 days, which is reasonable with a dose of 300 mg oral TDF[11]. The covariance matrix of $\log(k^e.pb)$, $\log(Cl.pb)$ and $\log(k^a.pb)$ is determined to be similar to the covariance matrix in equation (2.9). In addition, we assume the PK parameters for plasma are independent from the ones for PBMC and the absorption rate parameter for PBMC $\log(k^a.pb)$ is independent with $\log(k^e.pb)$ and $\log(Cl.pb)$.

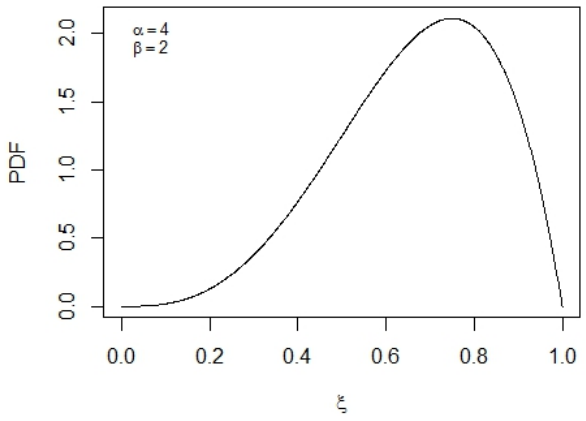
2.1.4 Adherence

Individual-level adherence probability for the subject i, ξ_i, is simulated from a beta distribution with $\alpha = 4$ and $\beta = 2$. This implies the population mean of individual-level adherence probability is 67% and the IQR of ξ_i is 26%. ξ_i is assumed to be fixed for subject i over time. The probability density function of ξ_i is shown in figure 2.2. The daily adherence for

subject i at j^{th} prescribed pill-taking event A_{ij} is determined from a Bernoulli distribution with success probability $= \xi_i$.

$$A_{ij}|\xi_i \sim Bernoulli(\xi_i) \tag{2.11}$$

Figure 2.2: Probability density function of ξ

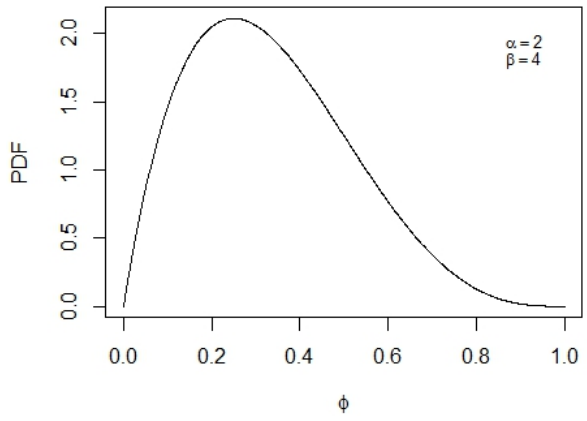


2.1.5 Self-reported Adherence

Individual-level misreporting probability given non-adherence, ϕ_i, for subject i is simulated from a beta distribution with $\alpha = 2$ and $\beta = 4$, implying the mean individual-lvel misreporting probability is 33% given non-compliance (IQR = 26%). ϕ_i is assumed to be fixed for subject i over time. The probability density function of ϕ_i is shown in figure 2.3. The daily self-reported adherence for subject i at j^{th} prescribed pill-taking event R_{ij} is simulated according to

$$R_{ij}|\phi_i, A_{ij} \sim \begin{cases} Bernoulli(1),\ if\ A_{ij} = 1 \\ Bernoulli(\phi_i),\ if\ A_{ij} = 0 \end{cases} \tag{2.12}$$

Figure 2.3: Probability density function of ϕ



2.1.6 Censored Data

Let $\nu.pl$ and $\nu.pb$ be the lower limits of detection for the drug concentrations in plasma and PBMC respectively. Let ***Y.pl*** be a matrix of the log drug concentrations in plasma with each entry $Y.pl_{ik}$ as the log drug concentration in plasma for subject i at k^{th} measurement. Let ***Y.pb*** be a matrix of the log drug concentrations in PBMC with each entry $Y.pb_{ik}$ as the log drug concentration in PBMC for subject i at k^{th} measurement. Let ***W.pl*** and ***W.pb*** be the matrices of the censoring indicator for plasma and PBMC. Thus, the entries of ***W.pl*** and ***W.pb*** are defined as ,

$$W.pl_{ik} = \begin{cases} 0, \ if \ Y.pl_{ik} \geq log(\nu.pl) \\ 1, \ if \ Y.pl_{ik} < log(\nu.pl) \end{cases} \tag{2.13}$$

$$W.pb_{ik} = \begin{cases} 0, \ if \ Y.pb_{ik} \geq log(\nu.pb) \\ 1, \ if \ Y.pb_{ik} < log(\nu.pb) \end{cases} \tag{2.14}$$

Then $Y.pl_{ik}$ and $Y.pb_{ik}$ are re-defined,

$$Y.pl_{ik} = \begin{cases} Y.pl_{ik}, \ if \ W.pl_{ik} = 0 \\ NA, \ if \ W.pl_{ik} = 1 \end{cases} \tag{2.15}$$

$$Y.pb_{ik} = \begin{cases} Y.pb_{ik},\ if\ W.pb_{ik} = 0 \\ NA,\ if\ W.pb_{ik} = 1 \end{cases} \tag{2.16}$$

In order to compare model performance in terms of different censoring levels, we choose multiple values for $\nu.pl$ and $\nu.pb$. Those values are shown in table 2.2.

Table 2.2: Choices of censoring levels

		$\nu.pl$	$\nu.pb$
One-biological-measure model	choice 1	0.7455	-
	choice 2	1.203	-
Two-biological-measure model	choice 1	0.7924	9.009
	choice 2	0.7924	1

2.2 Introduction to Statistical Modeling

The statistical model is designed to estimate the subject-specific adherence probability (ξ_i), misreporting probability given lack of adherence (ϕ_i) and the PK parameters ($\boldsymbol{\theta_i}$) using the observed data. The observed data include the log drug concentrations in plasma and PBMC (***Y.pl*** and ***Y.pb***) and self-reported adherence (***R***) where ***R*** is a matrix of self-report adherence indicator with each entry R_{ij} equal the self-reported adherence for subject i at the j^{th} prescribed pill-taking event. In addition, we also have information about the lower limits of detection ($\nu.pl$ and $\nu.pb$; ***W.pl*** and ***W.pb***) and the prior distribution for the parameters. As described in section 1.3, Bayes' theorem is used to derive the posterior probability distributions and MCMC algorithm is applied to obtain the posterior mean estimates. In this section, we introduce three models that build up the statistical model: the adherence model, the reported adherence model and the concentration model. We also describe how they work

together for estimation in subsection 2.2.4[8].

2.2.1 Adherence model

The distribution of the individual-level adherence A_{ij} for subject i at j^{th} prescribed pill-taking event is specified,

$$A_{ij}|\xi_i \sim Bernoulli(\xi_i) \tag{2.17}$$

The prior distributions of ξ_i is

$$\xi_i \sim Beta(\alpha_1, \beta_1) \tag{2.18}$$

Let η_1 and ζ_1 be the hyper-priors:

$$\eta_1 = \frac{\alpha_1}{\alpha_1 + \beta_1} \sim Beta(1, 1) \tag{2.19}$$

$$\zeta_1 = \alpha_1 + \beta_1 \sim Gamma(3, 1) \tag{2.20}$$

η_1 is the prior mean of ξ_i where $0 \leqslant \xi_i \leqslant 1$. Thus, η_1 is given a Beta distribution with E(η_1) = 0.5 to ensure ξ_i is less informative. ζ_1 is the prior sample size and $\zeta_1 > 0$. η_1 would be less informative as ζ_1 decreases[12]. Since E(ζ_1) is 3, the prior belief about η_1 is relatively weak.

2.2.2 Self-reported adherence model

The distribution of individual-level self-reported adherence for subject i at j^{th} prescribed pill-taking event is specified,

$$R_{ij}|\phi_i, A_{ij} \sim \begin{cases} 1 & if \quad A_{ij} = 1 \\ Bernoulli(\phi_i) & if \quad A_{ij} = 0 \end{cases} \tag{2.21}$$

The prior distributions of ϕ_i is

$$\phi_i \sim Beta(\alpha_2, \beta_2) \tag{2.22}$$

[8]The introduction to statistical modeling is based on the one-biological-measure model. The statistical model for two biological measures is discussed in section 2.3.3

where the hyper-priors η_2 and ζ_2 are defined

$$\eta_2 = \frac{\alpha_2}{\alpha_2 + \beta_2} \sim Beta(1, 1) \tag{2.23}$$

$$\zeta_2 = \alpha_2 + \beta_2 \sim Gamma(3, 1) \tag{2.24}$$

η_2 is the prior mean of ϕ_i and ζ_2 is the sample size that determine our prior belief about η_2[12]. Since the self-reported adherence $\boldsymbol{R}$ is observed, the likelihood of $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_{100})$ and $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_{100})$ is

$$L(\boldsymbol{\xi}, \boldsymbol{\phi}; \boldsymbol{R}) = \sum_{\boldsymbol{A}} p(\boldsymbol{R}|\boldsymbol{A}, \boldsymbol{\phi}) p(\boldsymbol{A}|\boldsymbol{\xi}) = p(\boldsymbol{R}|\boldsymbol{\xi}, \boldsymbol{\phi}) \tag{2.25}$$

$\boldsymbol{A}$ is a matrix with each entry A_{ij} being the adherence for subject i at j^{th} prescribed pill-taking event.

2.2.3 Concentration Model

Let $\boldsymbol{\Theta}$ be a matrix of all the PK parameters for all participants with each row $\boldsymbol{\theta_i}$ being the PK parameters for subject i. The distribution of individual-level log drug concentration in plasma $Y.pl_{ik}$ for subject i measured at k^{th} time is specified,

$$Y.pl_{ik}|A_{ij}, \boldsymbol{\theta_i} \sim \mathcal{N}(\log(C.pl_{ik}), \tau^{-2}) \tag{2.26}$$

where $\log(\tau^2)$, the log of the precision variable, is assigned an improper uniform prior so that

$$p(\log(\tau^2)) \propto 1 \tag{2.27}$$

$\log(\tau^2)$ is exponentiated to ensure the precision (and the variance) of $Y.pl_{ik}|A_{ij}$, $\boldsymbol{\theta_i}$ is positive. An alternative way is to assign τ^2 a gamma distribution so that $\tau > 0$.
$C.pl_{ik}$ is computed using equations 2.3 and 2.4. e.g.,

$$C.pl_{ik} = \sum_{ij|t_{ij}<t_{ik}} A_{ij} \frac{Dk^e.pl_i}{Cl.pl_i} e^{-k^e.pl_i(t_{ik}-t_{ij})} \tag{2.28}$$

D is known to be 300 mg/ml. $\boldsymbol{\theta_i} = \{\log(k^e.pl_i),\ \log(Cl.pl_i)\}$ is assigned a normal prior distribution with mean = $\boldsymbol{\mu}$ and variance = $\boldsymbol{\Sigma^{-1}}$,

$$\boldsymbol{\theta_i} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma^{-1}}) \tag{2.29}$$

where $\boldsymbol{\mu} = \{\mu_1,\ \mu_2\}$ has a hyper-prior distribution

$$\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \sim \mathcal{N}\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1000 & 0 \\ 0 & 1000 \end{pmatrix} \right]$$

and $\boldsymbol{\Sigma}$ is assigned a Wishart(r=2, $\boldsymbol{S}$) distribution as its hyper-prior, where

$$\boldsymbol{S}^{-1} = \begin{pmatrix} 0.18 & 0 \\ 0 & 0.18 \end{pmatrix} \tag{2.30}$$

The values of r and $\boldsymbol{S}$ are chosen to yield a prior expectation of the coefficient of variation = 30% for $\boldsymbol{\theta_i}$ [13]. In general, all the priors and hyper-priors for PK parameters are uninformative. Since we have the log drug concentration in plasma for each subject at each measurement, the likelihood of $\boldsymbol{\xi}$ and $\boldsymbol{\Theta}$ is

$$L(\boldsymbol{\xi}, \boldsymbol{\Theta}; \boldsymbol{Y.pl}) = \sum_A p(\boldsymbol{Y.pl}|\boldsymbol{A}, \boldsymbol{\xi}, \boldsymbol{\Theta}) p(\boldsymbol{A}|\boldsymbol{\xi}) = p(\boldsymbol{Y.pl}|\boldsymbol{\xi}, \boldsymbol{\Theta}) \tag{2.31}$$

2.2.4 Statistical Model

Assuming the data collected include the log drug concentrations in plasma ($\boldsymbol{Y.pl}$) and self-reported adherence ($\boldsymbol{R}$). Consider the individual-level parameters $\boldsymbol{\xi}$, $\boldsymbol{\phi}$ and $\boldsymbol{\Theta}$ that we would like to estimate, then based on Bayes theorem, the posterior joint distribution is

$$\begin{aligned} p(\boldsymbol{\xi}, \boldsymbol{\phi}, \boldsymbol{\Theta}|\boldsymbol{Y.pl}, \boldsymbol{R}) &= \frac{p(\boldsymbol{Y.pl}, \boldsymbol{R}|\boldsymbol{\xi}, \boldsymbol{\phi}, \boldsymbol{\Theta}) p(\boldsymbol{\xi}, \boldsymbol{\phi}, \boldsymbol{\Theta})}{p(\boldsymbol{Y.pl}, \boldsymbol{R})} \\ &= \frac{p(\boldsymbol{Y.pl}|\boldsymbol{\xi}, \boldsymbol{\phi}, \boldsymbol{\Theta}) p(\boldsymbol{R}|\boldsymbol{\xi}, \boldsymbol{\phi}, \boldsymbol{\Theta}) p(\boldsymbol{\xi}) p(\boldsymbol{\phi}) p(\boldsymbol{\Theta})}{p(\boldsymbol{Y.pl}, \boldsymbol{R})} \\ &\propto p(\boldsymbol{Y.pl}|\boldsymbol{\xi}, \boldsymbol{\Theta}) p(\boldsymbol{R}|\boldsymbol{\xi}, \boldsymbol{\phi}) p(\boldsymbol{\xi}) p(\boldsymbol{\phi}) p(\boldsymbol{\Theta}) \end{aligned} \tag{2.32}$$

We assume all observations in the data are independent and the parameters $\boldsymbol{\xi}$, $\boldsymbol{\phi}$ and $\boldsymbol{\Theta}$ are mutually independent. Therefore, we can estimate the joint posterior distribution using the prior distributions of $\boldsymbol{\xi}$, $\boldsymbol{\phi}$ and $\boldsymbol{\Theta}$ and likelihood computed by self-reported adherence model and concentration model.

In order to compute the posterior mean estimates of the parameters of interest, an MCMC algorithm is applied to draw samples from the posterior distribution in n iterations successively. Let m be the length of the burn-in period and k be the every k^{th} iteration to keep in the chain (thinning). Let x_i be a sample drawn from the posterior distribution for an arbitrary parameter X at i^{th} iteration. Then, the posterior mean estimator $\hat{X}$ is

$$\hat{X} = \sum_{j=1}^{(n-m)/k} x_{j\times k+m} \tag{2.33}$$

for j in $(1, \ldots, \frac{n-m}{k})$. For all the experiments in this thesis, n, m are determined by the user, after visually inspecting the chain convergence using the chain history plot[14]. Thinning strategy is used to reduce the sample autocorrelation and provide a reasonable amount of samples to compute the posterior mean estimates[15]. n, m and k are set to be 7000, 4000 and 10, respectively in the one-biological-measure model.

2.3 Model Improvement and Expansion

This section will be split into three sub-sections: model speed improvement, model with censored data and two-biological-measure model.

2.3.1 Model Speed Improvement

We first focused on improving model speed, since fitting the current model (see section 2.2) is time consuming. Slow speed is a huge disadvantage since the model needs to be run multiples times for different simulation experiments. Thus, we investigated approaches to speed up the estimation process.

As noted in section 2.2.3, equations (2.3) and (2.4) are used to estimate the mean of the

distribution of $\boldsymbol{Y}|\boldsymbol{\xi}, \boldsymbol{\theta}$ where $\boldsymbol{\xi}$ and $\boldsymbol{\theta}$ are latent variables. Since the sum in equation (2.3) is over all prior pill-taking events, the program has to compute equation (2.4) at least 462 times in each step of the Markov chain in order to estimate C_{ik} for all nine measurements (see section 2.1.1) of subject i[9]. If chain is designed to run 7000 steps, equation (2.4) would be computed 323,400,000 times in total. This will be even worse when we extend the model to two biological measures since we would have to compute equation (2.4) along with equation (2.5). However, if we only look at the pill-taking events within 4 weeks before each measurement rather than all prior events, almost half of the computation will be eliminated in terms of modeling drug concentrations and this should improve the model speed. This approach is reasonable if we can assume the pills taken beyond the 4 week look-back period have no or very little effect on drug concentrations. According to figure 2.1, the drug concentration in plasma is about 0 ng/ml after 3 weeks. This gives us good support to shorten the look-back period to improve the model performance.

Using plasma only, we examined the performance of estimation using 2 week, 3 week, 4 week and infinite look-back periods respectively. The drug concentrations in plasma without a lower limit of detection are used to isolate the impact of finite look-back on model speed. For each experiment, two chains with different initial values for unknown parameters are used.[10] The individual level estimates of $\log(k^e.pl_i)$, $\log(Cl.pl_i)$, $\boldsymbol{\xi_i}$ and $\boldsymbol{\phi_i}$ are compared to the true parameters using plots with estimated values on Y-axis and true values on X-axis.

---

[9] $\sum_{k=1}^{9} t_{ik} - 1 = 462$, since equation (2.3) goes through the days before each measurement for all subjects

[10] The initial values for each chain include the individual-level and population-level PK estimators; the precision matrix for the population-level PK estimators; the individual-level estimators for probability of adherence and misreporting; the precision of estimated log drug concentrations.

2.3.2 Model with censored data

As mentioned in section 2.1.6, if the observed log drug concentration of participant i at the k^{th} measurement Y_{ik}, is lower than a lower limit of detection, we define an indicator W_{ik}.

$$W_{ik} = \begin{cases} 0 & if \quad Y_{ik} \geq log(\nu) \\ 1 & if \quad Y_{ik} < log(\nu) \end{cases} \tag{2.34}$$

Before we test the model behavior with regard to censored data, we developed BUGS code for censoring in the model (See the code in appendix). We then examine the performance of the one-biological-measure model using plasma data with different lower limits of detection. Essentially, more measurements would be censored as the lower limit of detection is increased. We test the model with two data sets that have different lower limit of detection for drug concentrations: 0.7455 ng/ml and 1.203 ng/ml. For each data set, we run multiple experiments using two week, three week and four week look-back period models to investigate the reliability of shorter look-back periods when data are censored.

2.3.3 Two-biological-measure Model

In order to test if the estimation is improved when a second biological measure with a different half-life is added, a two biological measure statistical model is developed. The drug concentrations in plasma and PBMC are used in the model. Since the data are simulated, we know the drug concentrations in plasma follows a one-compartment IV model and the drug concentration in PBMC follows a one-compartment first-order absorption model and we adopt these models in the estimation procedure.

The concentration model is a little different since we have two biological measures. The distribution of the individual-level log drug concentration in plasma $Y.pl_{ik}$ and individual-level log drug concentration in PBMC $Y.pb_{ik}$ for subject i at the k^{th} measurement are

$$Y.pl_{ik}|\xi_i, \theta_i \sim \mathcal{N}(\log(C.pl_{ik}), \tau.pl^{-2}) \tag{2.35}$$

$$Y.pb_{ik}|\xi_i, \theta_i \sim \mathcal{N}(\log(C.pb_{ik}), \tau.pb^{-2}) \tag{2.36}$$

where $\log(\tau.pl^2)$ and $\log(\tau.pb^2)$ are assigned an improper uniform prior respectively. $C.pl_{ik}$ is computed using equation (2.3) and (2.4) and $C.pb_{ik}$ is computed using equation (2.3) and 2.5. D is known to be 300 mg/ml. $\boldsymbol{\theta_i} = \{\log(k^e.pl_i), \log(Cl.pl_i), \log(k^e.pb_i), \log(Cl.pb_i), \log(k^a.pb_i)\}$ is assigned a normal prior distribution with mean = $\boldsymbol{\mu}$ and variance = $\boldsymbol{\Sigma^{-1}}$,

$$\boldsymbol{\theta_i} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma^{-1}}) \tag{2.37}$$

where $\boldsymbol{\mu} = \{\mu_1, \mu_2, \mu_3, \mu_4, \mu_5\}$ has a hyper-prior distribution

$$\begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \\ \mu_5 \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1000 & 0 & 0 & 0 & 0 \\ 0 & 1000 & 0 & 0 & 0 \\ 0 & 0 & 1000 & 0 & 0 \\ 0 & 0 & 0 & 1000 & 0 \\ 0 & 0 & 0 & 0 & 1000 \end{pmatrix} \right] \tag{2.38}$$

and $\boldsymbol{\Sigma}$ is assigned a Wishart(r=5, $\boldsymbol{S}$) distribution as its hyper-prior, where

$$\boldsymbol{S^{-1}} = \begin{pmatrix} 0.45 & 0 & 0 & 0 & 0 \\ 0 & 0.45 & 0 & 0 & 0 \\ 0 & 0 & 0.45 & 0 & 0 \\ 0 & 0 & 0 & 0.45 & 0 \\ 0 & 0 & 0 & 0 & 0.45 \end{pmatrix} \tag{2.39}$$

One estimation experiment with uncensored data is done using uninformative PK priors. In this experiment, we found that the model recognizes $\log(k^e.pb)$ as $\log(k^a.pb)$ and vice versa if the initial guesses of them are comparatively poor so another experiment with uncensored data is carried out using an informative prior for $log(k^a.pb)$. It is expected that once the mean of $\log(k^a.pb)$ is identifiable, the model should also be able to identify the mean of the $\log(k^e.pb)$. The concentration model is slightly modified for the latter experiment by

changing equation (2.38) into

$$\begin{pmatrix}\mu_1\\ \mu_2\\ \mu_3\\ \mu_4\\ \mu_5\end{pmatrix} \sim N\left[\begin{pmatrix}0\\ 0\\ 0\\ 0\\ -0.693\end{pmatrix}, \begin{pmatrix}1000 & 0 & 0 & 0 & 0\\ 0 & 1000 & 0 & 0 & 0\\ 0 & 0 & 1000 & 0 & 0\\ 0 & 0 & 0 & 1000 & 0\\ 0 & 0 & 0 & 0 & 0.001\end{pmatrix}\right] \tag{2.40}$$

Two experiments are done with different levels of lower limit of detection and informative priors for $\log(k^a.pb)$. The lower limits of detection chosen for the drug concentrations in plasma and PBMC are $\nu.pl = 0.7924$ and $\nu.pb = 9.009$; $\nu.pl = 0.7924$ and $\nu.pb = 1$ (see table 2.2). For each experiment, a 4 week look-back period and two chains with different initial values are used. We computed five posterior pharmacokinetic estimates: $\log(k^e.pl)$, $\log(k^e.pb)$, $\log(Cl.pl)$, $\log(Cl.pb)$ and $\log(k^a.pb)$. The estimates of individual level probability of adherence $\boldsymbol{\xi}$ and probability of misreporting $\boldsymbol{\phi}$ are also computed. In addition, the number of iterations in the Markov chain is increased to 10,000 and the burn-in period is increased to 7000 after visually inspecting the chain convergence (also see section 2.2.4). It is reasonable to expect the chain converges slower in the two-biological-measure model than the one-biological-measure model since the former one has a more complicated model structure.

To evaluate if the estimation of adherence is more accurate using two biological measures than one biological measure, the two-biological-measure model and the one-biological-measure model with 4 weeks look-back-period are both implemented using the same data. Only the plasma data is used for the one-biological-measure model. The lower limits of detection are $\nu.pl = 0.7924$ and $\nu.pb = 1$. Since the two chains obtained in each experiment are similar, the first chain in both experiments are chosen and the posterior estimates are computed. The estimated subject-specific adherence on day t for subject i using one biological measure and two biological measures, $\hat{a}_{m1it}$ and $\hat{a}_{m2it}$ are compared to the true adherence a_{it} respectively, where $i \in (1,\ldots,100)$ and $t \in (1,\ldots,72$[11]$)$. The sum of squared residuals SSR_{m1} and SSR_{m2}

[11]There isn't an estimated adherence on day 73 since measurement isn't collected

are computed.

$$SSR_{m1} = \sum_{i}^{100} \sum_{t}^{72} (\hat{a}_{m1it} - a_{it})^2 \tag{2.41}$$

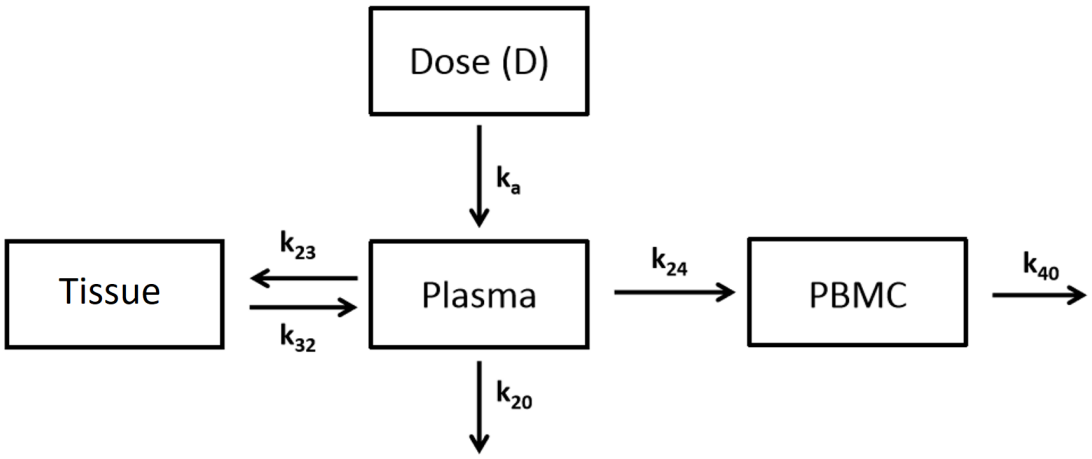
$$SSR_{m2} = \sum_{i}^{100} \sum_{t}^{72} (\hat{a}_{m2it} - a_{it})^2 \tag{2.42}$$

2.4 Two-biological-measure Model Using Multi-compartment System

As noted in section 2.1.2, modeling expected plasma and PBMC drug concentrations with a one-compartment IV bolus model and a one-compartment first-order absorption model might be unrealistic [8]. Therefore, a multi-compartment model that describes the plasma/PBMC pharmacokinetics is adopted from Burns et al.(2015)[16] to model the expected drug concentrations. The expected drug concentration in plasma is modeled using a two-compartment first-order absorption/elimination model and linked to the drug concentration in PBMC by an additional compartment with first-order uptake and first-order elimination rates. The PK model and parameters are shown in figure 2.4. k_a is the absorption rate parameter. Parameters k_{23} and k_{32} are the rate for influx and outflux to/from other tissues. k_{24} represents the uptake rate paramter to the PBMC compartment. k_{20} and k_{40} are the elimination rates for plasma and PBMC.

The population PK parameters are adopted from Burns et al.(2015)[16]. We assume the subject-specific PK parameters follow a multivariate lognormal distribution. Let $\boldsymbol{\theta_i} = \{\log(k_{ai}), \log(k_{23i}), \log(k_{32i}), \log(k_{20i}), \log(k_{24i}), \log(k_{40i})\}$ be the subject-specific log PK pa-

Figure 2.4: Multi-compartment pharmacokinetic model in plasma and PBMC



rameters for subject i. Then,

$$\boldsymbol{\theta_i} = \begin{pmatrix} \log(k_{ai}) \\ \log(k_{23i}) \\ \log(k_{32i}) \\ \log(k_{20i}) \\ \log(k_{24i}) \\ \log(k_{40i}) \end{pmatrix} \sim \mathcal{N}\left[\begin{pmatrix} 2.28 \\ -0.46 \\ -0.93 \\ -2.04 \\ -4.07 \\ -4.34 \end{pmatrix}, \begin{pmatrix} 0.154 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.027 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.057 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.151 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.284 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.281 \end{pmatrix} \right] \tag{2.43}$$

where the coefficients of variation are chosen to be similar to the ones from other data used in this thesis since no information on subject-specific variation is provided in Burns et al.(2015)[16]. In addition, all the log PK parameters are assumed to be independent which is a simplification made for our initial simulations with this model. The correlations of PK parameters and the model performance using correlated data will be emphasized in future study.

Let t be the time since a pill-taking event. The single-dose drug concentrations in plasma and PBMC for subject i can be expressed as a function of t and pharmacokinetic parameters:

$$\begin{aligned} C_{2i}(t) =& Dk_{ai}\left[\frac{k_{32i}-k_{ai}}{(\alpha_i-k_{ai})(\beta_i-k_{ai})}e^{-k_{ai}t}\right. \\ &+\frac{k_{32i}-\alpha_i}{(k_{ai}-\alpha_i)(\beta_i-\alpha_i)}e^{-\alpha_i t} \\ &\left.+\frac{k_{32i}-\beta_i}{(k_{ai}-\beta_i)(\alpha_i-\beta_i)}e^{-\beta_i t}\right] \end{aligned} \tag{2.44}$$

$$\begin{aligned} C_{4i}(t) =& Dk_{ai}k_{24i}\left[\frac{k_{32i}-k_{ai}}{(k_{40i}-k_{ai})(\alpha_i-k_{ai})(\beta_i-k_{ai})}e^{-k_{ai}t}\right. \\ &+\frac{k_{32i}-k_{40i}}{(k_{ai}-k_{40i})(\alpha_i-k_{40i})(\beta_i-k_{40i})}e^{-k_{40i}t} \\ &+\frac{k_{32i}-\alpha_i}{(k_{ai}-\alpha_i)(k_{40i}-\alpha_i)(\beta_i-\alpha_i)}e^{-\alpha_i t} \\ &\left.+\frac{k_{32i}-\beta_i}{(k_a-\beta_i)(k_{40i}-\beta_i)(\alpha_i-\beta_i)}e^{-\beta_i t}\right] \end{aligned} \tag{2.45}$$

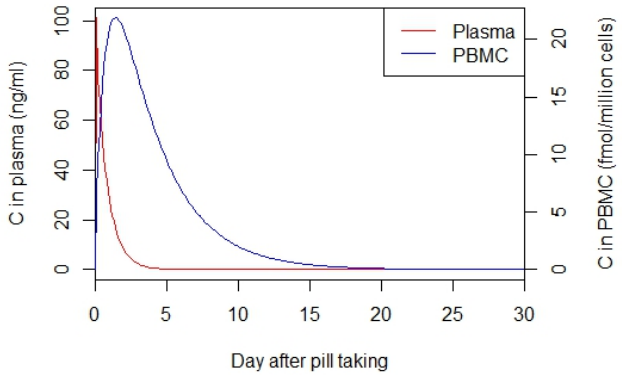
$$\begin{aligned} \alpha_i &= 0.5\times(W_i+\sqrt{W_i^2-4k_{32i}(k_{20i}+k_{24i})}) \\ \beta_i &= 0.5\times(W_i-\sqrt{W_i^2-4k_{32i}(k_{20i}+k_{24i})}) \\ W_i &= k_{23i}+k_{32i}+k_{20i}+k_{24i} \end{aligned}$$

where D is the drug dosage and $C_{2i}(t)$ and $C_{4i}(t)$ are the plasma drug level and PBMC drug level at time t for subject i, respectively[12]. Figure 2.5 shows the distribution of mean drug concentration in plasma and PBMC as a function of time since pill-taking, using PK parameters from equation (2.43) and multi-compartment system. Comparing to figure 2.1, the drug concentrations for both plasma and PBMC in figure 2.5 have higher peaks. However, the drug is eliminated from the plasma and PBMC in a shorter period (about 21 days) when the new PK parameters and multi-compartment system are used.

A new dataset is simulated using the same methods described in section 2.1 and equa-

[12]Note that the unit of time is day in equations (2.4) and (2.5) but is hour in equations (2.44) and (2.45)

Figure 2.5: Drug concentration (C) vs time (multi-compartment system)



tions (2.44) and (2.45) are used to replace the equations (2.4) and (2.5). The multi-compartment model is then incorporated into the BUGS code for a two-biological-measure model to estimate the expected log drug concentrations. The performance of the two-biological-measure model with multi-compartment system is evaluated. We expect the model will be more time-consuming since the multi-compartment system is more complicated than the one-compartment PK model so a two week look-back period is used. The lower limits of detection chosen for the drug concentrations in plasma and PBMC are $\nu.pl = 0.01$ and $\nu.pb = 1$. Uninformative priors for PK parameters and two chains with different initial values are used. Posterior estimates are computed for the six PK parameters shown in figure 2.4. The estimates of individual level probability of adherence $\boldsymbol{\xi}$ and probability of misreporting $\boldsymbol{\phi}$ are also computed. The number of iterations in the Markov chain is set to be 10,000 and the burn-in period is set to be 7000.

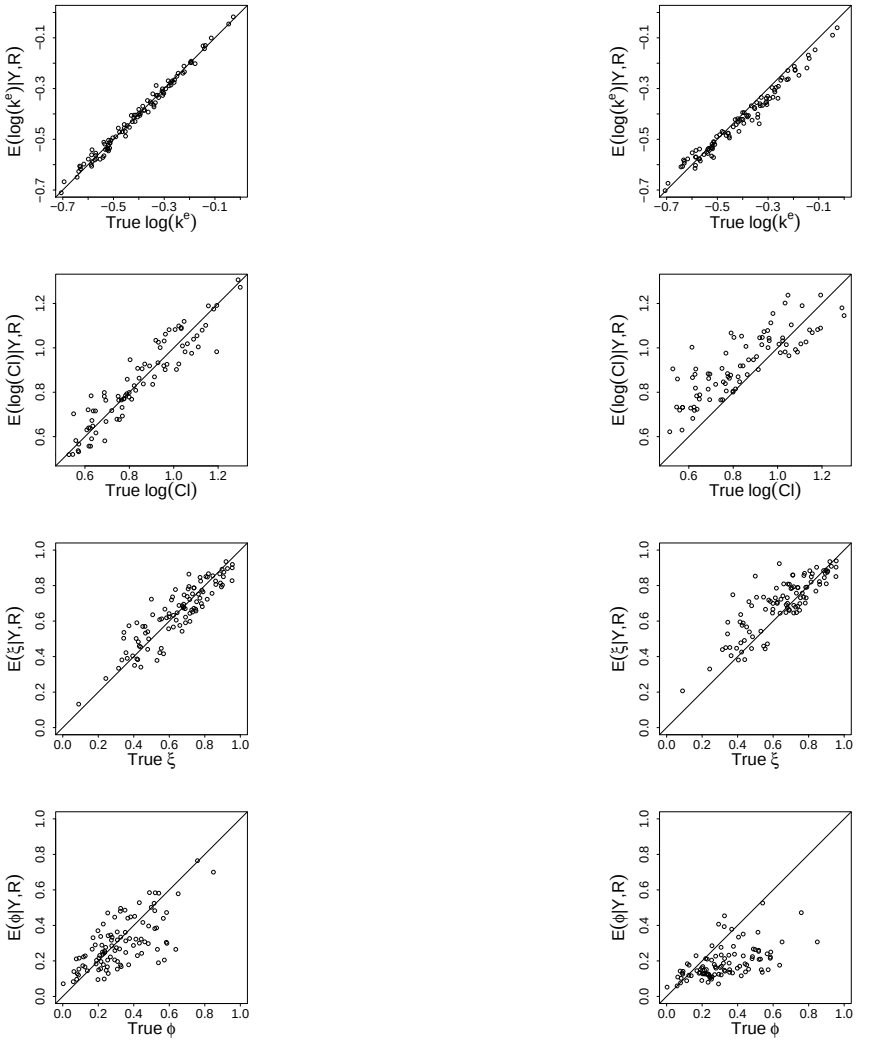
Chapter 3

RESULTS

In this chapter, we first talk the improvement in computational efficiency and estimation results using different look-back periods. We then discuss the results using different censoring levels. We also show the estimation results of the two-biological-measure model with respect to uninformative PK priors and an informative prior for absorption rate using uncensored and censored data. In addition, the estimation accuracy of adherence using two-biological-measure model is compared to the one-biological-measure model and the performance results of the two-biological-measure model is further presented when a multi-compartment system is incorporated.

3.1 Model with Shorter Look-back Period

A shorter look-back period model (see section 2.3.1) allows the statistical model to run faster and gain computational efficiency. The experiments with 4 week, 3 week and 2 week look-back periods are approximately 40%, 50%, and 60% faster comparing to the infinite look back period, respectively, for our 73 day design. Greater gains are expected as the duration of the followup period increases. Figure 3.1 shows the results of estimation using one-biological-measure model with infinite look-back period. Plasma data without censoring are used and the results are generated using 2 chains with different initial values.

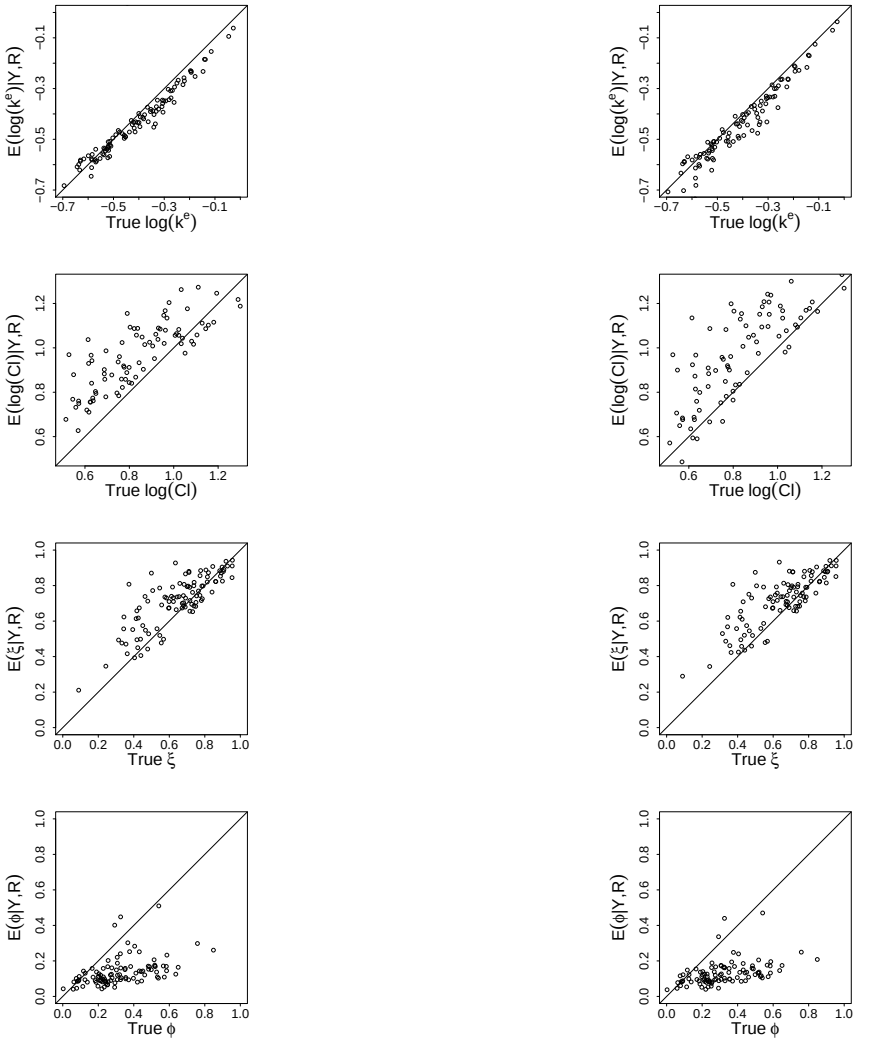


Chain 1 Chain 2

Figure 3.1: Posterior mean estimates of pharmacokinetic parameters ($\log(k^e)$,$\log(Cl)$), adherence probability (ξ) and misreporting probability (ϕ) for 100 individuals, generated using infinite look-back period.

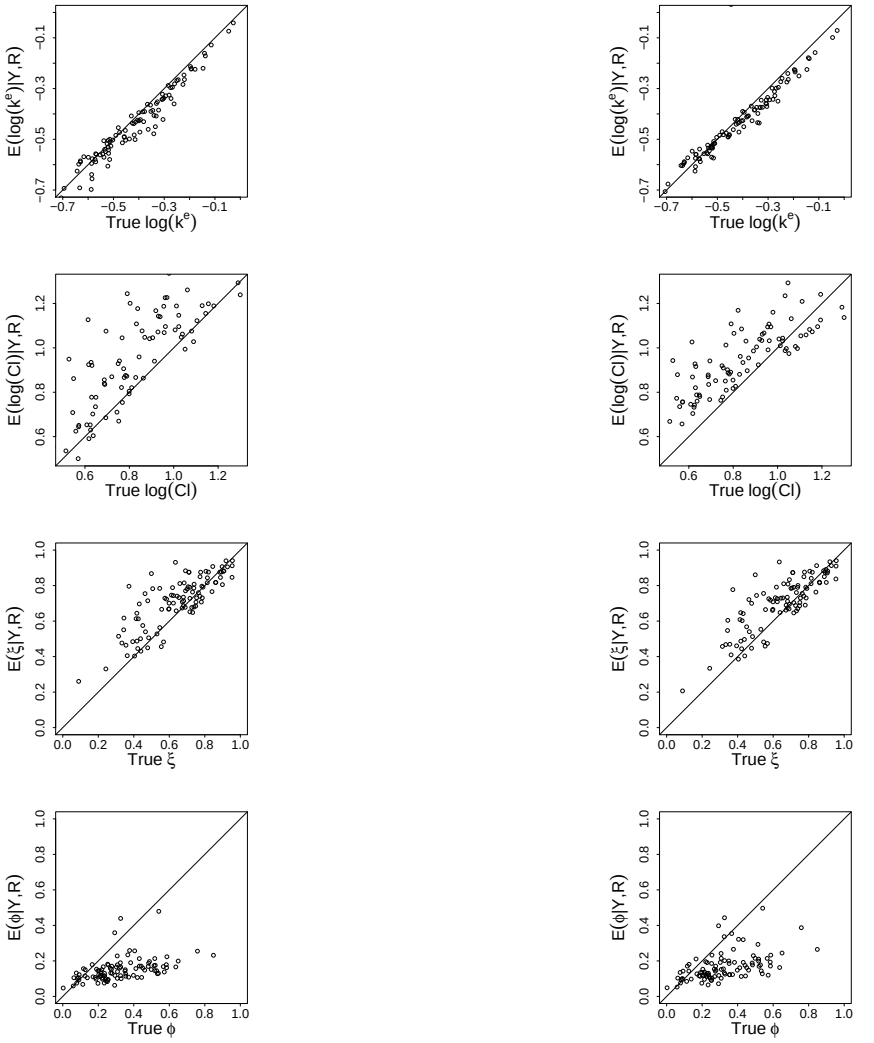
According to the figure 3.1, the estimation are more accurate in chain 1 compared to chain 2. The pharmacokinetic estimates E($\log(k^e)$ and E($\log(Cl)$) are unbiased in chain 1 whereas E($\log(k^e)$ is biased down and E($\log(Cl)$) is biased up in chain 2. This implies the posterior mean estimates are sensitive to initial values.

Figures 3.2, 3.3 and 3.4 show the results of estimation using 2 week, 3 week and 4 week look-back periods.



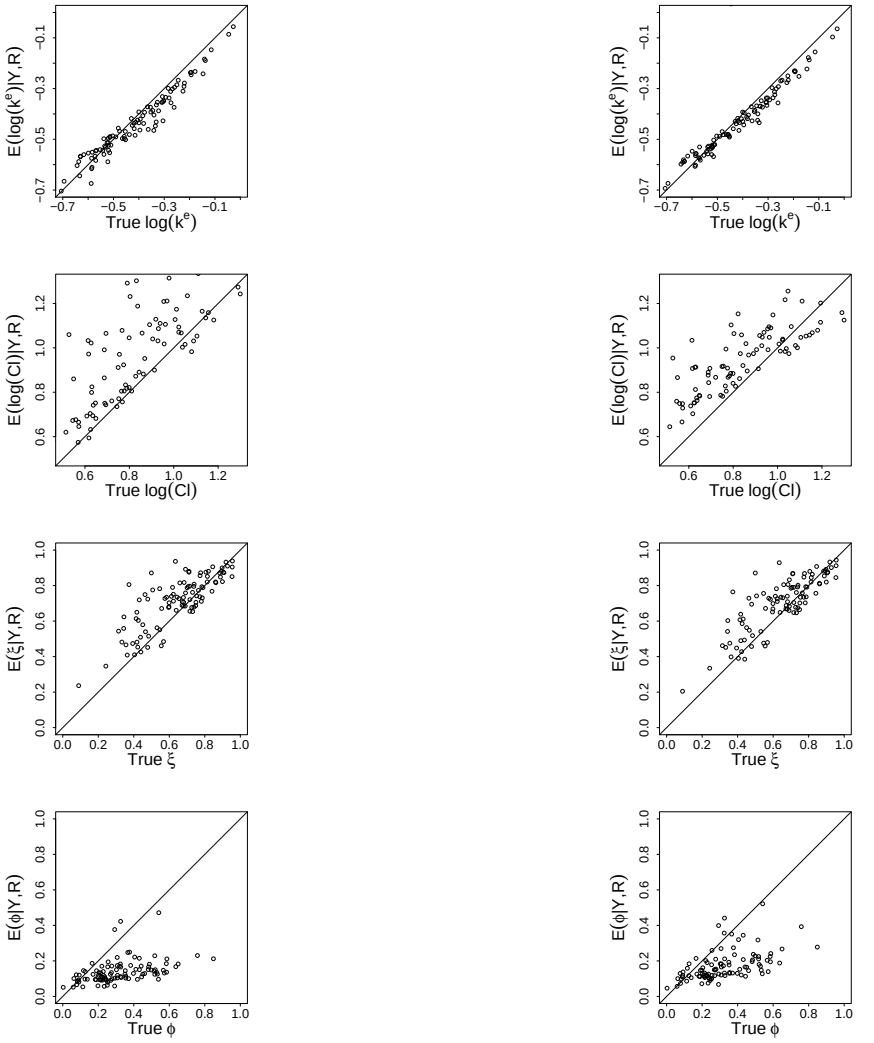
Chain 1 Chain 2

Figure 3.2: Posterior mean estimates of pharmacokinetic parameters ($\log(k^e)$,$\log(Cl)$), adherence probability (ξ) and misreporting probability (ϕ) for 100 individuals, generated using two week look-back period.



Chain 1 Chain 2

Figure 3.3: Posterior mean estimates of pharmacokinetic parameters ($\log(k^e)$,$\log(Cl)$), adherence probability (ξ) and misreporting probability (ϕ) for 100 individuals, generated using three week look-back period.



Chain 1 Chain 2

Figure 3.4: Posterior mean estimates of pharmacokinetic parameters ($\log(k^e)$,$\log(Cl)$), adherence probability (ξ) and misreporting probability (ϕ) for 100 individuals, generated using four week look-back period.

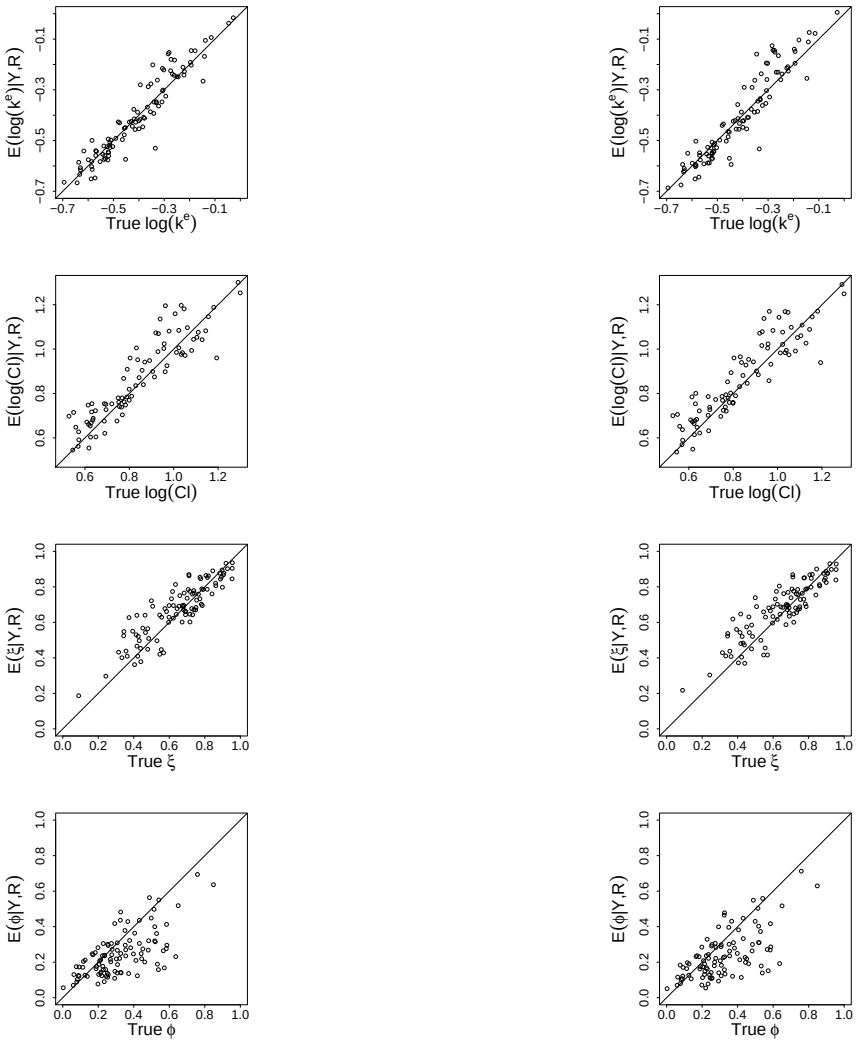
In the experiments with shorter look-back period, the results in the two chains with different initial values are similar, although there isn't significant improvement in terms of the accuracy of estimating subject-specific parameters compared to the results from the experiment with infinite look-back period shown in figure 3.1.

3.2 Estimation with Censored Data

Figures 3.5, 3.6, 3.7, 3.8 and 3.9 shows the estimation results using the one-biological-measure model when censoring is applied to the plasma data. Two lower limits of detection are chosen (see table 2.2) and 2 week, 3 week and 4 week look-back periods models are tested for each lower limit of detection. For $\nu.pl = 0.7455$, 57% of the measurements in DOT phase and <2% of the data in SA phase are censored. For $\nu.pl = 1.203$, 86% of the measurements in DOT phase and 2% of the data in SA phase are censored.

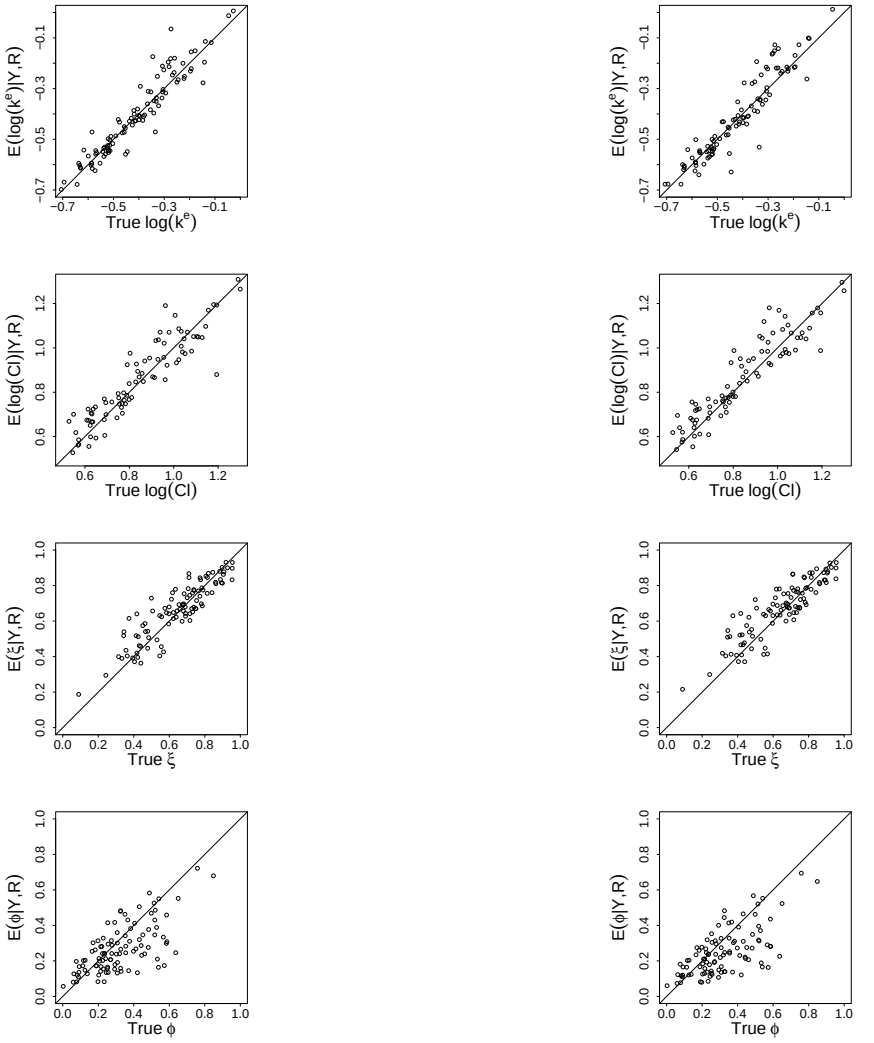
The result of 2 weeks look-back period model with $\nu.pl = 1.203$ is not provided since Openbugs encountered errors when this model was run. Since a large proportion of the drug concentrations in the DOT period is censored, we speculate that the error is due to insufficient data which causes the estimates of parameters of interest unable to converge. Among all the other experiments, the estimations are less biased comparing to the experiments with uncensored data, although the estimate tend to be more scattered (less precise) when the censoring level is higher. Since the log drug concentrations especially in plasma are linearly associated with time since pill-taking (see figure 3.10), we speculate that the estimation bias seen in figures $3.2 - 3.4$ is caused by the high leverage of the low log drug concentration measurements in the DOT phase, in which biological samples are collected at least one week after pill-taking. Since the values of the samples collected in the DOT period are very low (< 1), they are negative values on a log scale. Theoretically, log drug concentration would approach to -∞ as drug concentration approaches to 0. Therefore, the model might overestimate the expected log drug concentrations when the observation is extremely low. In conclusion, accounting for a lower limit of detection improves the statistical accuracy but high levels of censoring especially in the DOT period can cause the estimation to be less

precise. In addition, if the look-back period is short and the censoring level is high, the estimation might fail due to insufficient data.



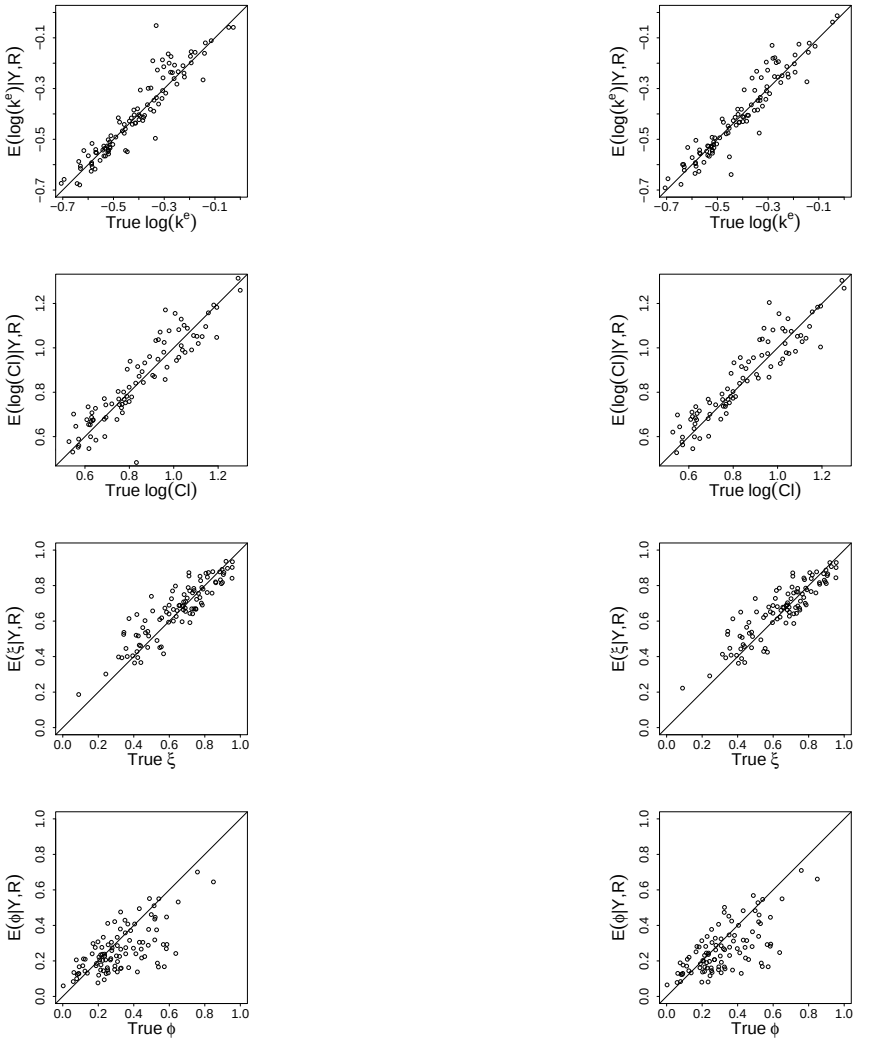
Chain 1 Chain 2

Figure 3.5: Posterior mean estimates of pharmacokinetic parameters ($\log(k^e)$,$\log(Cl)$), adherence probability (ξ) and self-reported adherence probability (ϕ) for 100 individuals with $\nu = 0.7455$, generated using 2 week look-back period.



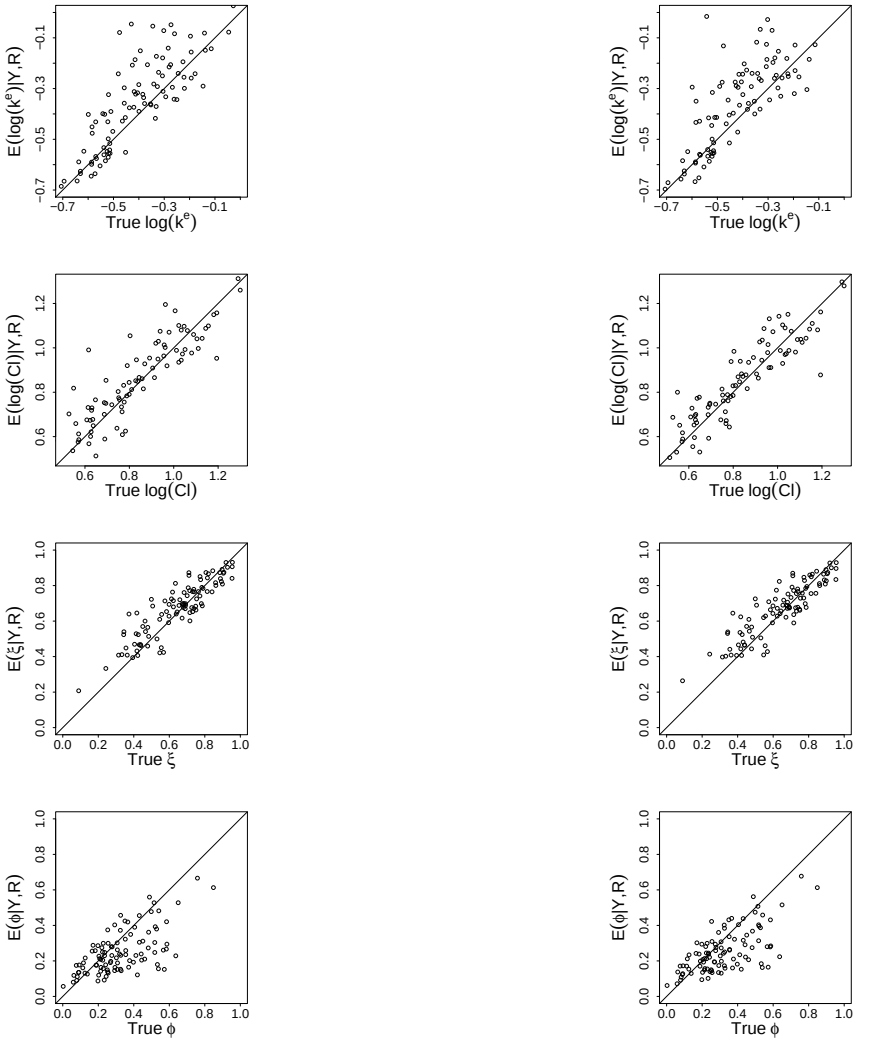
Chain 1 Chain 2

Figure 3.6: Posterior mean estimates of pharmacokinetic parameters ($\log(k^e)$,$\log(Cl)$), adherence probability (ξ) and self-reported adherence probability (ϕ) for 100 individuals with $\nu = 0.7455$, generated using 3 week look-back period.



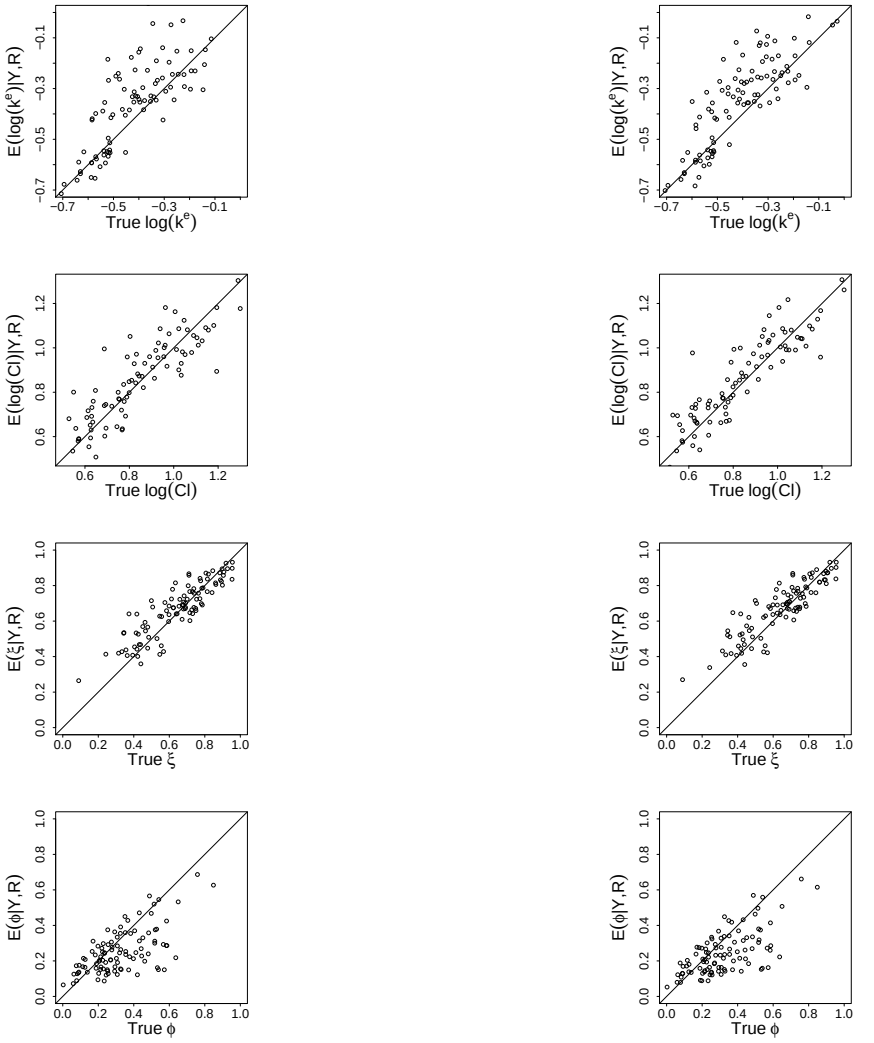
Chain 1 Chain 2

Figure 3.7: Posterior mean estimates of pharmacokinetic parameters ($\log(k^e)$,$\log(Cl)$), adherence probability (ξ) and self-reported adherence probability (ϕ) for 100 individuals with $\nu = 0.7455$, generated using 4 week look-back period.



Chain 1 Chain 2

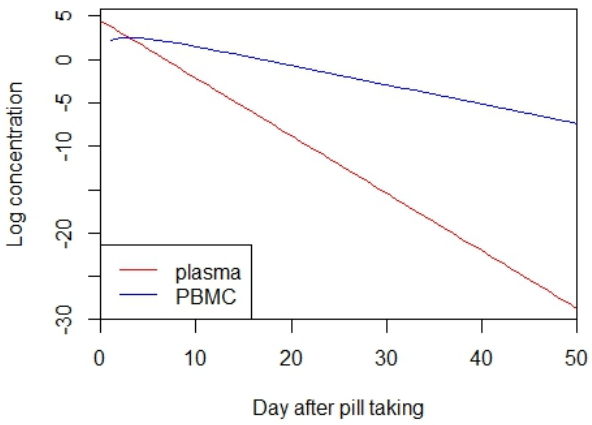
Figure 3.8: Posterior mean estimates of pharmacokinetic parameters ($\log(k^e)$,$\log(Cl)$), adherence probability (ξ) and self-reported adherence probability (ϕ) for 100 individuals with $\nu = 1.203$, generated using 3 week look-back period.



Chain 1 Chain 2

Figure 3.9: Posterior mean estimates of pharmacokinetic parameters ($\log(k^e)$,$\log(Cl)$), adherence probability (ξ) and self-reported adherence probability (ϕ) for 100 individuals with $\nu = 1.203$, generated using four week look-back period.

Figure 3.10: Distribution of log drug concentration vs. time



3.3 Estimation with Two-biological-measure Model

We first test the two-biological-measure model with uncensored data using uninformative pharmacokinetic parameters. Figure 3.11 shows the estimation results for 7 parameters. We see that in chain 2 the estimations of $k^e.pb$ and $k^a.pb$ are way off compared to the true values. This is due to the non-identifiability between $k^e.pb$ and $k^a.pb$ in the one-compartment first-order absorption model which was used for the PBMC outcome. Note that in equation (2.5), parameters k^e and k^a are exchangeable. Therefore, when the initial guesses are far away from the true values, the statistical model might swap the sampling distributions of k^e and k^a. To remedy this we define a better informative prior for $k^a.pb$ to distinguish $k^e.pb$ and $k^a.pb$ in the PBMC model (i.e. the one compartment first-order absorption model). As noted in section 2.3.3, the uninformative PK priors for the individual-level PK parameters $\boldsymbol{\theta_i} = \{\log(k^e.pl_i), \log(Cl.pl_i), \log(k^e.pb_i), \log(Cl.pb_i), \log(k^a.pb_i)\}$ are defined to be a normal distribution with mean = $\boldsymbol{\mu}$ and variance = $\boldsymbol{\Sigma^{-1}}$,

$$\boldsymbol{\theta_i} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma^{-1}}) \tag{3.1}$$

where $\boldsymbol{\mu} = \{\mu_1, \mu_2, \mu_3, \mu_4, \mu_5\}$ has a hyper-prior distribution

$$\begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \\ \mu_5 \end{pmatrix} \sim \mathcal{N}\left[\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1000 & 0 & 0 & 0 & 0 \\ 0 & 1000 & 0 & 0 & 0 \\ 0 & 0 & 1000 & 0 & 0 \\ 0 & 0 & 0 & 1000 & 0 \\ 0 & 0 & 0 & 0 & 1000 \end{pmatrix} \right] \tag{3.2}$$

and $\boldsymbol{\Sigma}$ is assigned a Wishart(r=5, $\boldsymbol{S}$) distribution as its hyper-prior, where

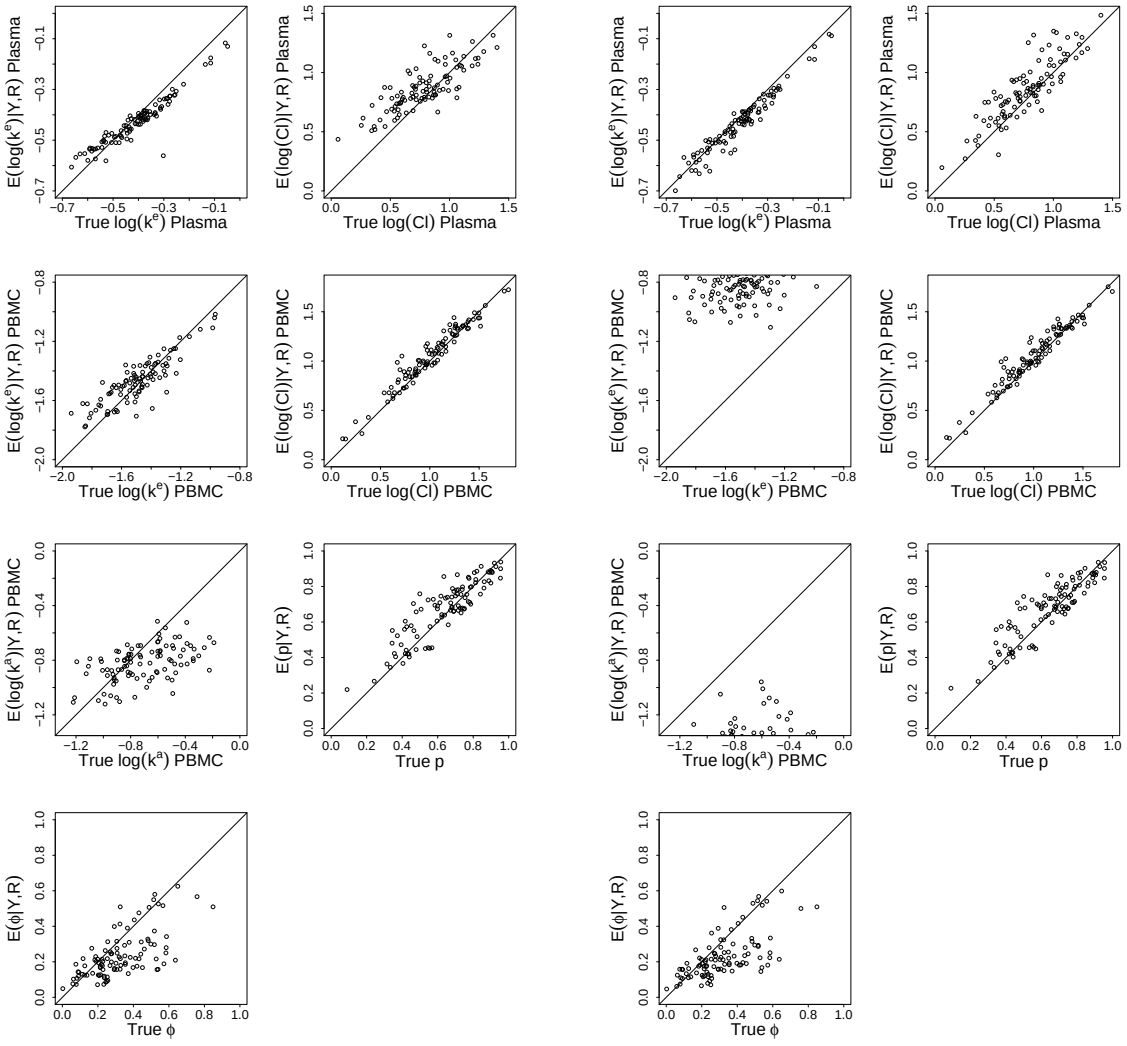
$$\boldsymbol{S^{-1}} = \begin{pmatrix} 0.45 & 0 & 0 & 0 & 0 \\ 0 & 0.45 & 0 & 0 & 0 \\ 0 & 0 & 0.45 & 0 & 0 \\ 0 & 0 & 0 & 0.45 & 0 \\ 0 & 0 & 0 & 0 & 0.45 \end{pmatrix} \tag{3.3}$$

In order to give $\log(k^a.pb)$ a informative prior, the distribution of $\boldsymbol{\mu}$ is adjusted,

$$\begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \\ \mu_5 \end{pmatrix} \sim \mathcal{N}\left[\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -0.693 \end{pmatrix}, \begin{pmatrix} 1000 & 0 & 0 & 0 & 0 \\ 0 & 1000 & 0 & 0 & 0 \\ 0 & 0 & 1000 & 0 & 0 \\ 0 & 0 & 0 & 1000 & 0 \\ 0 & 0 & 0 & 0 & 0.001 \end{pmatrix} \right] \tag{3.4}$$

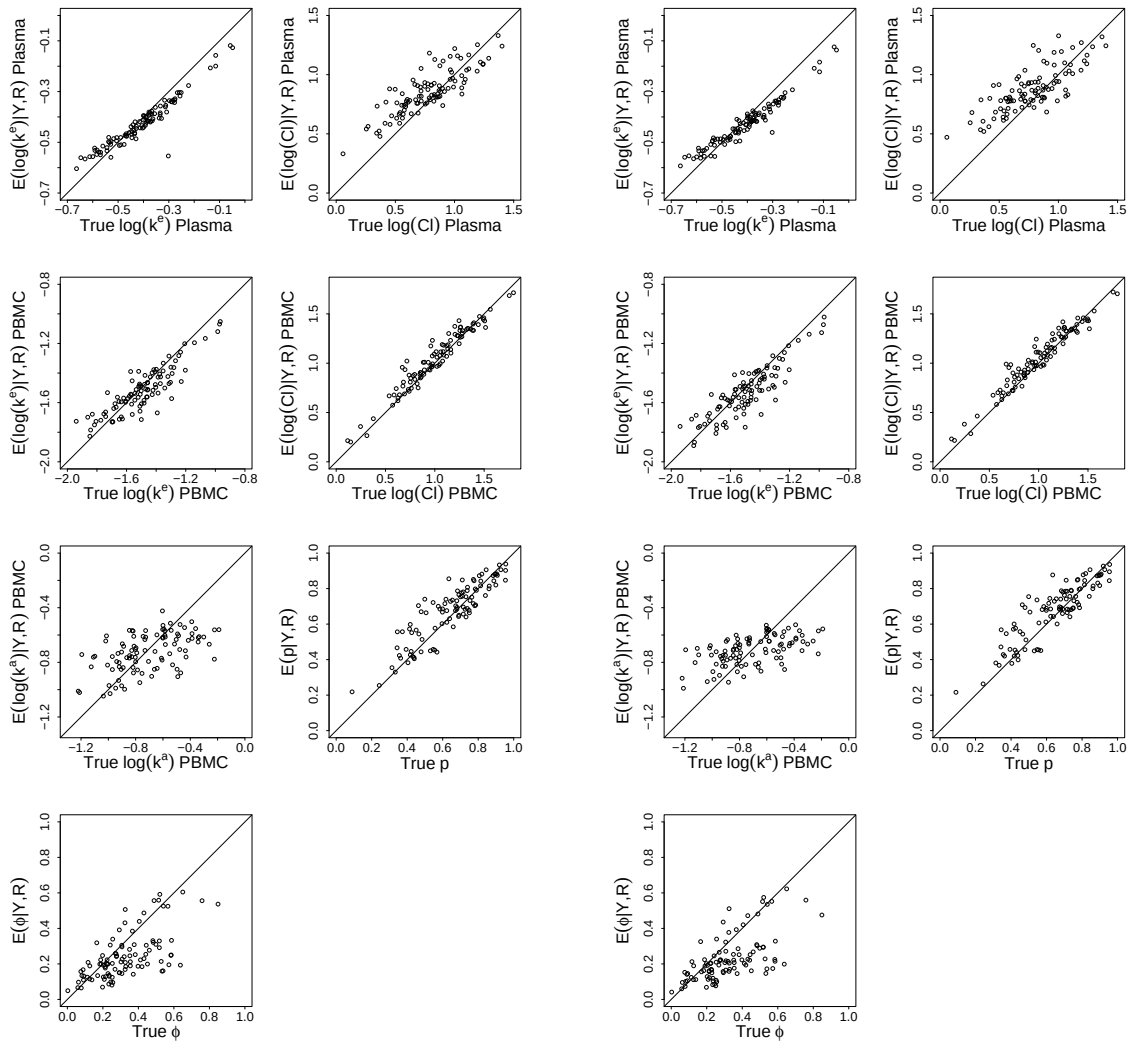
where -0.693 is the true value of population $\log(k^a.pb)$.

An alternative way of enforcing the identifiability is to constrain $k^a.pb_i > k^e.pb_i$ for subject i in the statistical model (see discussion section).



Chain 1 Chain 2

Figure 3.11: Posterior mean estimates of pharmacokinetic parameters, adherence probability (ξ) and self-reported adherence probability (ϕ) for 100 individuals, generated using four week look-back period and data without censoring.

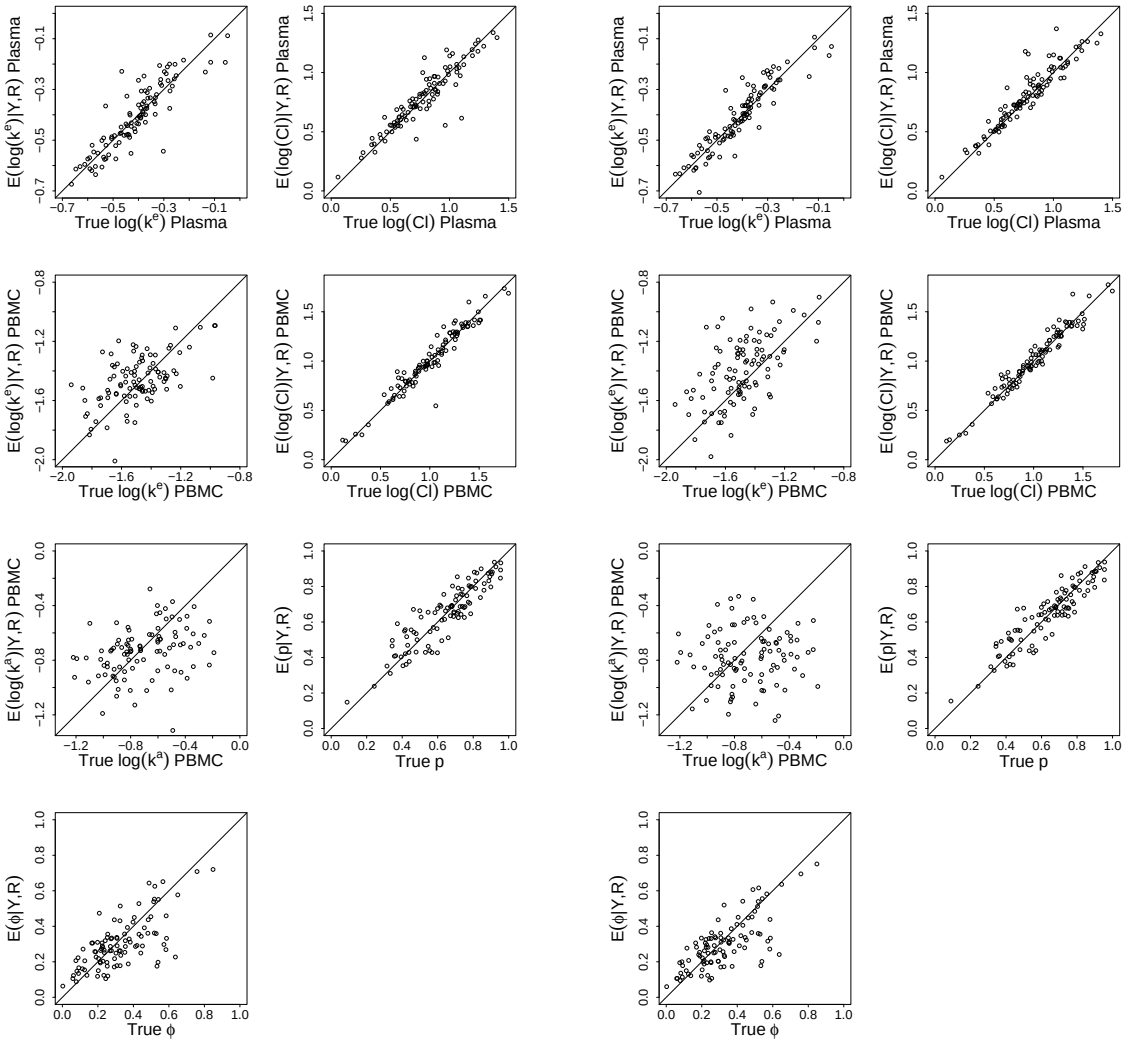


Chain 1 Chain 2

Figure 3.12: Posterior mean estimates of pharmacokinetic parameters, adherence probability (ξ) and self-reported adherence probability (ϕ) for 100 individuals, generated using four week look-back period, informative prior for $\log(k^a.pb)$ and data without censoring.

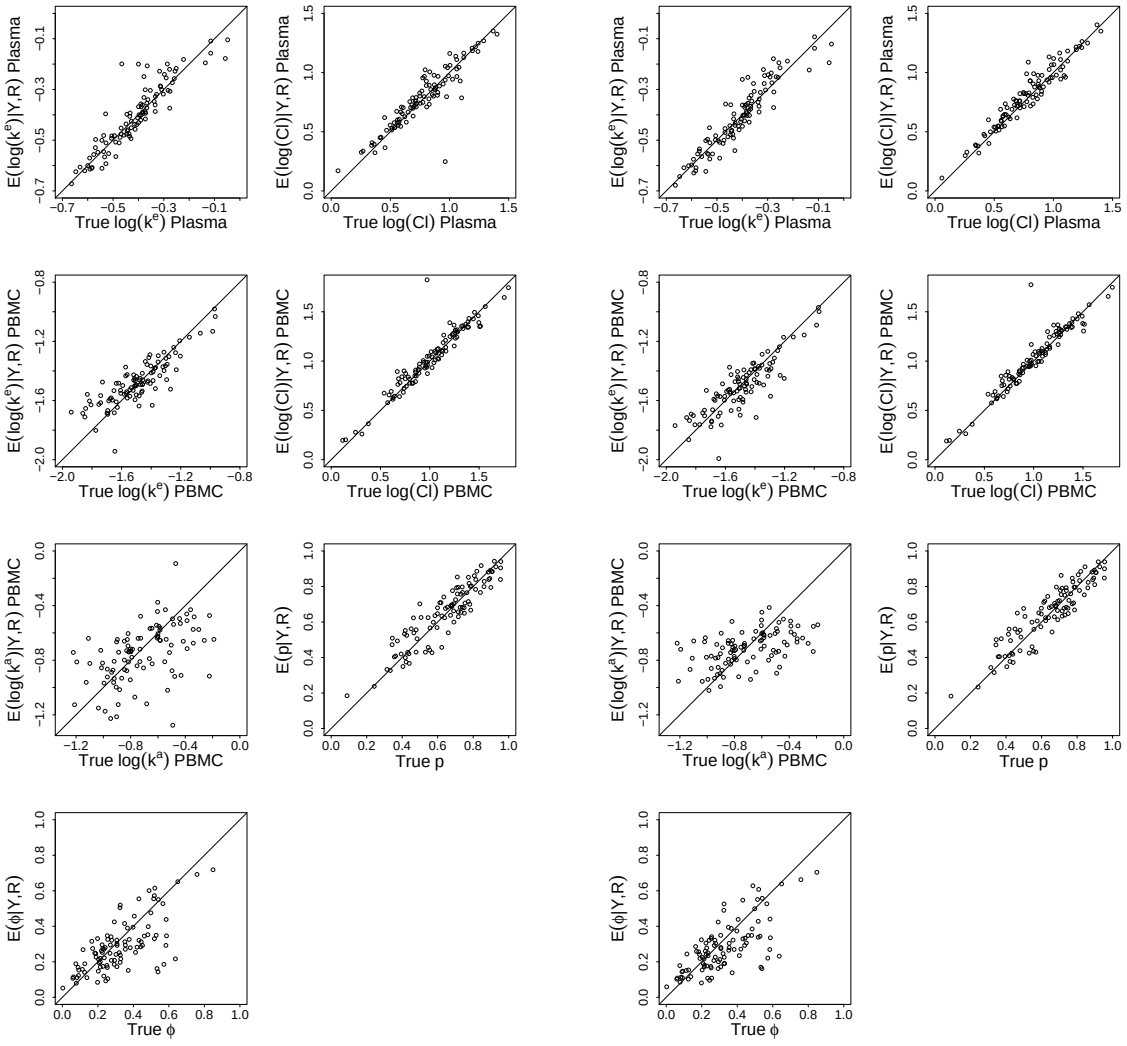
Figure 3.12 shows that the estimation of $\log(k^e.pb)$ is a lot better with an informative prior of $\log(k^a.pb)$. The non-identifiably issue is gone since the model are capable to identify the mean of $\log(k^a.pb)$. However, similar to results showed in figure 3.11, the estimations of $\log(k^a.pl)$ and $\log(Cl.pl)$ are still biased. We expect this problem can be fixed using censored data based on the results from section 3.2. Also the estimation of $\log(k^a.pb)$ is heavily regressed to the mean, which implies that the estimation of subject-specific $\log(k^a.pb)$ is weak even though an informative prior for the population absorption rate parameter is given.

We next test the model with censored data. Figure 3.13 shows the results from the model which uses censoring levels $\nu.pl = 0.7924$ and $\nu.pb = 9.009$. 57% of the drug concentrations in plasma and 56% of the drug concentration in PBMC are censored in the DOT phase. Less than 2% of the drug concentrations in plasma and PBMC are censored in the SA phase. As before, sensitivity to initial values is greatly reduced since 4 week look-back period is applied. We find that the estimations of PBMC parameters are not too bad but getting a little worse than the estimation using uncensored data. However, the bias in the estimations of plasma parameters is fixed when censored data is used.



Chain 1 Chain 2

Figure 3.13: Posterior mean estimates of pharmacokinetic parameters, adherence probability (ξ) and self-reported adherence probability (ϕ) for 100 individuals with $\nu.pl = 0.7924$ and $\nu.pb = 9.009$, generated using four week look-back period and informative prior for $k^a.pb$.



Chain 1 Chain 2

Figure 3.14: Posterior mean estimates of pharmacokinetic parameters, adherence probability (ξ) and self-reported adherence probability (ϕ) for 100 individuals with $\nu.pl = 0.7924$ and $\nu.pb = 1$, generated using four week look-back period and informative prior for $k^a.pb$.

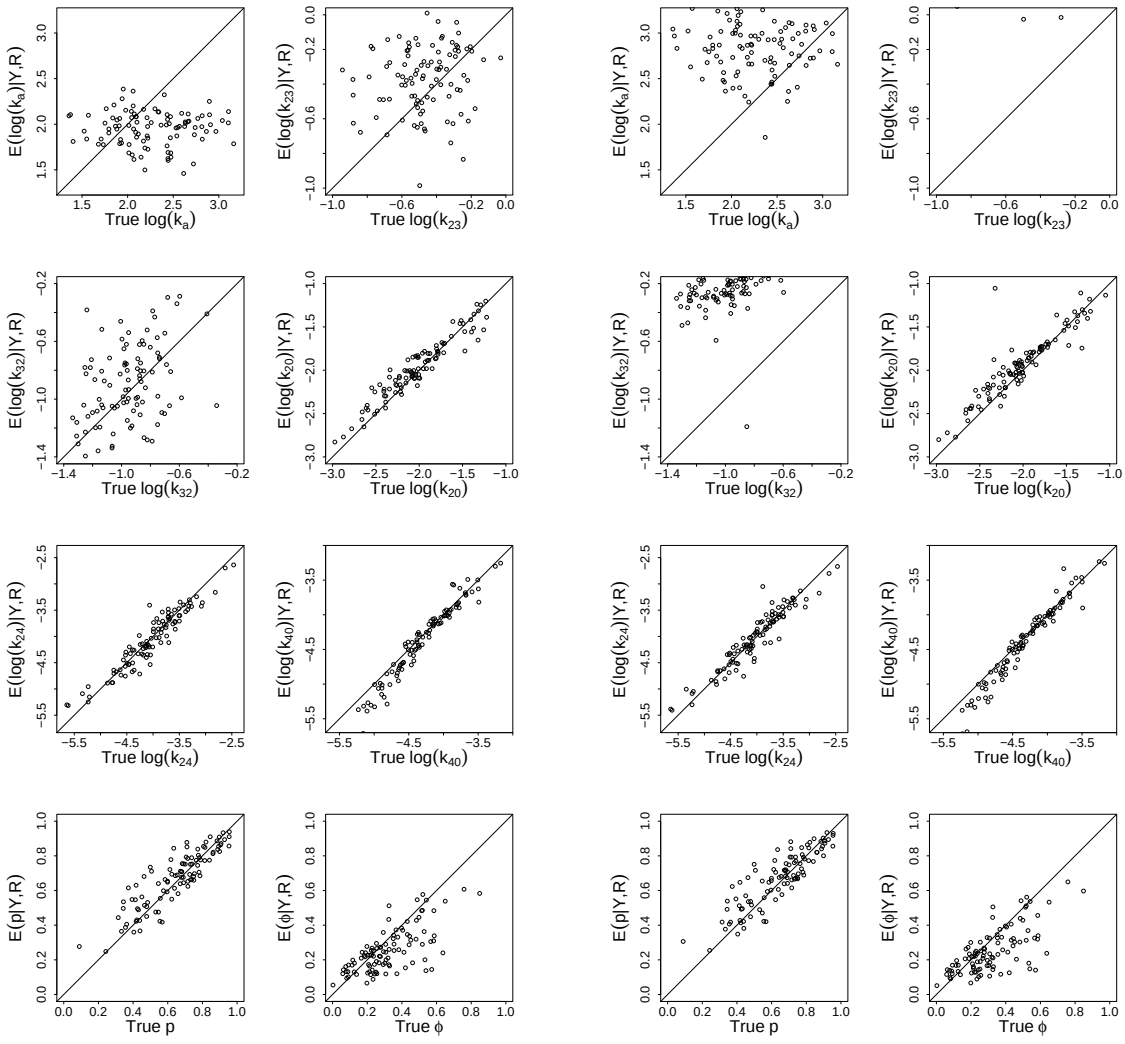
Another experiment with the lower limit of detection for PBMC set to $\nu.pb = 1$ was performed. Figure 3.14 shows the result of the model with adjusted lower limit of detection for PBMC ($\nu.pb = 1$). With this data, 57% of the drug concentrations in plasma and <1% of the drug concentrations in PBMC are censored in the DOT phase. Less than 2% of the drug concentrations in plasma and <1% of the drug concentrations in PBMC are censored in the SA phase. Except for the fact that the estimation of $k^a.pb$ is not ideal, the estimations of the other parameters are less biased and appear accurate. This makes sense due to the design of the DOT phase. $k^a.pb$ represents the rate of drug absorption in the body which normally happen in the first couple days after the pill-taking event. Since the measurements are taken at least a week after the pill-taking in DOT phase, it would be challenging to estimate the absorption rate.

The accuracy of the estimation of adherence is evaluated by comparing the sum of squared residuals of adherence using one biological measure and two biological measures (see section 2.3.3). We notice that the sum of squared residuals obtained from one-biological-measure model SSR_{m1} is 183.8 and the sum of squared residuals obtained from two-biological-measure model SSR_{m2} is 144.5 which is 21% less than SSR_{m1}. The reduction in sum of square residuals implies that the estimation accuracy in adherence is improved when two-biological-measure model is used.

3.4 Performance of Two-biological-measure Model with Multi-compartment System

As noted in section 2.4, an additional experiment of the two-biological-measure model is implemented to evaluate the model performance using a more realistic model of pharmacokinetics of plasma and PBMC (a multi-compartment system). A two week look-back period is used to improve the computational efficiency. The lower limits of detection are $\nu.pl = 0.01$ and $\nu.pb = 1$. 61% of the drug concentrations in plasma and 26% of the drug concentration in PBMC are censored in the DOT phase. 4% of the drug concentrations in plasma and 1% of the drug concentrations in PBMC are censored in the SA phase. The estimation results

from two chains are shown in figure 3.15.



Chain 1 Chain 2

Figure 3.15: Posterior mean estimates of pharmacokinetic parameters, adherence probability (ξ) and self-reported adherence probability (ϕ) for 100 individuals with $\nu.pl = 0.01$ and $\nu.pb = 1$, generated using two week look-back period.

As shown in figure 3.15, the estimation of subject-specific $\log(k_a)$ is not optimal but the model is able to approximately identify the population-level $\log(k_a)$ (the estimated population-level $\log(k_a)$ is 1.94 in the first chain and 2.84 in the second chain compared to 2.28, the population-level $\log(k_a)$ used for simulation). The estimation of $\log(k_{20})$, $\log(k_{24})$ and $\log(k_{40})$ is good whereas the estimation of $\log(k_{23})$ and $\log(k_{32})$ is poor and has non-identifiability issues as seen in the second chain. We speculate that there exists constraint on these two parameters so that $k_{23} > k_{32}$ is the warranted. The constraint is not enforced in the model so the orders of the rate parameters aren't guaranteed. Importantly, even though the model is not able to describe all the subject-specific PK parameters, the estimation of the adherence probability is still good in both chains. The measure of misreporting of adherence, ϕ, tends to be underestimated. In spite of $\log(k_a)$, $\log(k_{23})$ and $\log(k_{32})$, the other estimates are insensitive to initials. In general, we think the two-biological-measure model is practical to estimate the subject-specific probability of adherence and misreporting along with the subject-specific PK parameters in terms of more sophisticated PK model and more realistic data. We believe that the non-identifiability issues can be solved using informative priors or model constraint and the estimation accuracy can be improved using a four week look-back period. Since complete elimination of the drug from PBMC takes about three weeks, using the four week look-back period is a better way to estimate the expected drug concentration in PBMC. However, it will be computationally challenging to use the four week look-back period.

Chapter 4

DISCUSSION

The motivation of our research comes from the fact that we don't have a good method to estimate/identify the true pill-taking adherence of participants in clinical trials. Low adherence can potentially bias the drug efficacy[4]. The HPTN 067 trial (The ADAPT Study) provides an opportunity to look for methods that can be used to estimate individual-level adherence and misreporting probability since the trial aims at estimating participants' adherence with different pill-taking schedules. As mentioned in section 1.3, Chloe Krakauer from University of Washington developed a statistical model, that is designed to estimate subject-specific PK parameters, adherence and misreporting probability in HTPN 067 trial, based on simulated data[7]. Due to the design of the model, we realize this model is very computational inefficient. Therefore, we started our study by investigating this model and looking for methods to improve the computational efficiency. We also investigate the estimation performance with data censored at different levels assuming the real data collected in HTPN 067 trial have lower limits of detection. In addition, the original model developed by Krakauer[7] can only use one biological measure which could be insufficient since more than one biological samples might be collected. Therefore, we develop a model to estimate the probability of adherence and misreporting from two biological measures with different half-lives.

The results of improving the computational efficiency are discussed in section 3.1. A shorter look-back period is developed to increase the model running speed. Experiments with 2 week, 3 week and 4 week look-back periods are performed and provide a 60%, 50% and 40% improvement in model running time respectively. It is also noted that the estimations are not sensitive to the initial values when the shorter look-back period is applied.

The major finding during the investigation of model performance with censored data (see the results in section 3.2) is that the estimation bias of the PK parameters is fixed when detection limits (censoring) is applied to the data. We believe the bias is due to the high leverage of low log drug concentrations (see figure 3.10) in the DOT phase with sparse sampling.

The results of the two-biological-measure model are described in section 3.3. Just like the one-biological-measure model, we find that using detection limits can fix the bias of the plasma PK estimates but can reduce the estimation precision if censoring level is too high. The model has non-identifiablity problems when uninformative priors are used. We are able to fix this by giving the population absorption rate parameter $\log(k^a.pb)$ an informative distribution so that the mean of $\log(k^a.pb)$ is identifiable. However, we find that we are unable to estimate the individual-level absorption rate due to the design of the DOT phase. In addition, the sum of squared residual for adherence using two biological measures is 21% less than using one biological measure, indicating the estimation of adherence is more accurate using a two-biological-measure model.

It is concerning that the pharmacokinetic models incorporated into the two-biological-measure model is unrealistic since the plasma compartment is assumed to be independent to the PBMC compartment. Therefore, a two-biological-measure model with multi-compartment system is tested to see if estimation is practical using more realistic model of plasma/PBMC pharmacokinetics and drug concentrations. We see good estimation of adherence when this more complicated PK model is incorporated. However, some of the subject-specific PK parameters are poorly estimated and may be unidentifiable.

There are a few more things we would like to study in the future to improve the two-biological-measure model with multi-compartment system. First, a four week look-back period should be used to improve estimation of the expected drug concentration in PBMC. Secondly, it is critical to fix the non-identifiability issues due to the pharmacokinetic formulation. According to other experiments in this thesis, it is expected that the issues can be solved using informative PK priors. We would also like to study alternative approach that can be used

to fix the non-identifiability such as setting a constraint on pharmacokinetic formulas. In addition, we would like to study the correlations among the subject-specific PK parameters. Since the estimation obtained using this model is based on the assumption that all the PK parameters are independent, a sensitivity analysis is expected to be performed using correlated PK parameters. Lastly, pill-taking events are assumed to have a purely additive effect both in the simulation and the statistical model. With the real data, it may be necessary to define a maximum steady state drug concentration[17].

The design of the simulated data could also be improved to be in line with the real data. For example, more detailed sample collection time can be specified so that it allows sample to be collected after a pill-taking event on the same day. As we mentioned in section 2.1.1, the simulated 30 day SA phase is a lot shorter than the actual 24 weeks (168 days) SA phase in HPTN 067 trial to accommodate the simulation study. It may be worthwhile to investigate the model performance with a longer SA phase which will provide more information on subject-specific adherence and misreporting given non-adherence.

The design of HPTN 067 trial especially the DOT phase add some limitations on the development of the model. Due to the sparse sampling in DOT phase, The subject-specific absorption rate parameters can hardly be estimated. It is recommended to evaluate the model performance with respect to an alternative DOT sampling plan, which can be used to support future study.

We are hoping that our findings are supportive for the future research of this model and some of the strategies mentioned above will be applied and used to fix problems. With some model improvement, we are looking forward to estimating the adherence in HTPN 067 trial in the near future.

BIBLIOGRAPHY

[1] U.S. Food and Drug Administration. (July 16, 2012), "FDA approves first drug for reducing the risk of sexually acquired HIV infection." Retrieved on 7/21/2015 from `http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm312210.htm`.

[2] R. M. Grant, J. R. Lama, P. L. Anderson, V. Mcmahan, A. Y. Liu, L. Vargas, P. Goicochea, M. Casapa, J. V. Guanira Carranza, M. E. Ramirez Cardich, O. Montoya Herrera, T. Fernndez, V. G. Veloso, S. P. Buchbinder, S. Chariyalertsak, M. Schechter, L. G. Bekker, K. H. Mayer, E. G. Kalls, K. R. Amico, K. Mulligan, L. R. Bushman, R. J. Hance, C. Ganoza, P. Defechereux, B. Postle, F. Wang, J. J. Mcconnell, J. H. Zheng, J. Lee, J. F. Rooney, H. S. Jaffe, A. I. Martinez, D. N. Burns, and D. V. Glidden, "Preexposure Chemoprophylaxis for HIV Prevention in Men Who Have Sex with Men," *The New England Journal of Medicine*, vol. 363, no. 27, pp. 2587–2599, 2010.

[3] L. Van Damme, A. Corneli, K. Ahmed, K. Agot, J. Lombaard, S. Kapiga, M. Malahleha, F. Owino, R. Manongi, J. Onyango, L. Temu, M. C. Monedi, P. Mak'oketch, M. Makanda, I. Reblin, S. E. Makatu, L. Saylor, H. Kiernan, S. Kirkendale, C. Wong, R. Grant, A. Kashuba, K. Nanda, J. Mandala, K. Fransen, J. Deese, T. Crucitti, T. D. Mastro, and D. Taylor, "Preexposure Prophylaxis for HIV Infection among African Women," *The New England Journal of Medicine*, vol. 367, no. 5, pp. 411–422, 2012.

[4] I. Holme, M. Szarek, N. B. Cater, O. Faergeman, J. J. Kastelein, A. G. Olsson, M. J. Tikkanen, M. L. Larsen, C. Lindahl, and T. R. Pedersen, "Adherence-adjusted efficacy with intensive versus standard statin therapy in patients with acute myocardial infarction in the ideal study," *European Journal of Cardiovascular Prevention & Rehabilitation*, vol. 16, no. 3, pp. 315–320, 2009.

[5] HIV Prevention Trials Network, "HPTN 067 The ADAPT study: A Phase II, Randomized, Open-Label, Pharmacokinetic and Behavioral Study of the Use of Intermittent Oral Emtricitabine/Tenofovir Disoproxil Fumarate Pre-Exposure Prophylaxis (PrEP) ." Retrieved on 7/21/2015 from `http://www.hptn.org/web%20documents/HPTN067/Protocol/067ProtocolV5_10Jun2013.pdf`.

[6] C. Bova, K. Fennie, G. Knafl, K. Dieckhaus, E. Watrous, and A. Williams, "Use of Electronic Monitoring Devices to Measure Antiretroviral Adherence: Practical Considerations," *AIDS and Behavior*, vol. 9, no. 1, pp. 103–110, 2005.

[7] C. Krakauer, "The Use of Pharmacokinetics to Estimate Adherence During a Clinical Trial Implementing a Bayesian Framework," Master's thesis, University of Washington, 2014.

[8] S. Duwal, C. Schtte, M. von Kleist, and J. Tang, "Pharmacokinetics and Pharmacodynamics of the Reverse Transcriptase Inhibitor Tenofovir and Prophylactic Efficacy against HIV-1 Infection," *PLoS ONE*, vol. 7, no. 7, p. E40382, 2012.

[9] T. N. Gengiah, C. Baxter, L. E. Mansoor, A. B. Kharsany, and S. S. Abdool Karim, "A drug evaluation of 1% tenofovir gel and tenofovir disoproxil fumarate tablets for the prevention of HIV infection," *Expert opinion on investigational drugs*, vol. 21, no. 5, pp. 695–715, 2012.

[10] J. E. Frampton and C. M. Perry, "Emtricitabine: a review of its use in the management of HIV infection.(Adis Drug Evaluation)," *Drugs*, vol. 65, no. 10, p. 1427(22), 2005.

[11] K. Madrasi, R. N. Burns, C. W. Hendrix, M. J. Fossler, and A. Chaturvedula, "Linking the Population Pharmacokinetics of Tenofovir and Its Metabolites With Its Cellular Uptake and Metabolism," *CPT: Pharmacometrics & Systems Pharmacology*, vol. 3, no. 11, pp. 1–6, 2014.

[12] J. K. Kruschke, *Doing Bayesian data analysis : a tutorial with R and BUGS*. MA: Academic Press, 2011.

[13] J. Wakefield, *Bayesian and frequentist regression methods (Springer series in statistics)*. Springer New York, 2013.

[14] P. Shevchenko and P. V. Shevchenko, *Modelling Operational Risk Using Bayesian Inference*. Dordrecht: Springer, 2011.

[15] H. Brown and R. Prescott, *Applied Mixed Models in Medicine*. Third ed., Statistics in practice, 2014.

[16] R. N. Burns, C. W. Hendrix, and A. Chaturvedula, "Population pharmacokinetics of tenofovir and tenofovirdiphosphate in healthy women," *Journal of Clinical Pharmacology*, vol. 55, no. 6, pp. 629–638, 2015.

[17] M.-Y. Liu, S.-N. Meng, H.-Z. Wu, S. Wang, and M.-J. Wei, "Pharmacokinetics of single-dose and multiple-dose memantine in healthy chinese volunteers using an analytic method of liquid chromatography-tandem mass spectrometry," *Clinical Therapeutics*, vol. 30, no. 4, pp. 641–653, 2008.

Appendix A

OPENBUGS CODE

A.1 One-Biological-Measure Model using look-back period

```
model
{
# Data are as follows:
#
# npart = number of participants.
# nplasma = number of plasma measurements (same for now across individuals).
# npill = number of days each subject participated in the trial
#(same for all individuals for now).
# pilltime[i,j] = time ith individual took jth pill.
# plasmatime[i,k] = time of kth plasma measurement for individual i.
# plasma[i,k] = measured value of drug in plasma of kth
#blood sample from ith participant.
# dot[i,j] = 1 if jth pill was taken during the DOT period for
#individual i (=1 for entire DOT period).
# sched[i,j] = 1 for any day during the DOT period where a pill
#was scheduled to be taken.
# r[i,j] = 1 if they self-report taking the pill that day
#(includes DOT period).
# npill2[i,k] = the last day in the look-back period at
#kth measurement for individual i.
# npill3[i,k] = the first day in the look-back period at
```

```
#kth measurement for individual i.
#
#
#
# Self-reporting
for (i in 1:npart)
{
        p[i] ~ dbeta(alpha1,beta1) # Adherence
        phi[i] ~ dbeta(alpha2,beta2)
        # Self-reporting pill-taking given lack of adherence
        for (j in 1:npill)
        {
                prob.adh[i,j] <- (dot[i,j]*sched[i,j]) +
                ((1-dot[i,j])*p[i])
                a[i,j] ~ dbern(prob.adh[i,j])
                prob.self[i,j] <-
                ((a[i,j]+((1-a[i,j])*phi[i]))*(1-dot[i,j]))+
                (dot[i,j]*a[i,j])
                r[i,j] ~ dbern(prob.self[i,j])
        }
}


# Pill-taking.

for (i in 1:npart)
{
```

```
        # Pharmacokinetic parameters.
        theta[i,1:2] ~ dmnorm(mu[],tau[,])
        log.ke[i] <- theta[i,1]
        log.Cl[i] <- theta[i,2]
        Cl[i] <- exp(log.Cl[i])
        ke[i] <- exp(log.ke[i])

        # DOT and self-administered phase.
        for (k in 1:nplasma)
        {
                for (j in npill3[i,k]:npill2[i,k])
                {
                        summand[i,k,j] <-
                        a[i,j]*((dose*ke[i])/Cl[i])*
                        exp(-ke[i]*(plasmatime[i,k]-pilltime[i,j]))
                }
                mplasma[i,k] <-
                log(sum(summand[i,k,npill3[i,k]:npill2[i,k]]))
                pcens[i,k] <-phi((level-mplasma[i,k])/sigma3)
                cens[i,k] ~ dbern(pcens[i,k])
                lim[i,k] <-
                ((1-cens[i,k])*100000000) + (cens[i,k]*level)
                plasma[i,k] ~ dnorm(mplasma[i,k],eps.tau)C(,lim[i,k])
        }
}
```

```
zeta1 ~ dgamma(3,1)
zeta2 ~ dgamma(3,1)
eta1 ~ dbeta(1,1)
eta2 ~ dbeta(1,1)

alpha1 <- eta1*zeta1
alpha2 <- eta2*zeta2
beta1 <- zeta1*(1-eta1)
beta2 <- zeta2*(1-eta2)

# Priors for pharmacokinetic parameters
mu[1] ~ dnorm(0,0.001)
mu[2] ~ dnorm(0,0.001)
tau[1:2,1:2] ~ dwish(R[,],2)
R[1,1] <- 0.18
R[1,2] <- 0
R[2,1] <- 0
R[2,2] <- 0.18
sigma[1:2,1:2] <- inverse(tau[,])
rho <- sigma[1,2]/sqrt(sigma[1,1]*sigma[2,2])

eps.tau <- exp(logtau)
logtau ~ dflat()
sigma3 <- 1/sqrt(eps.tau)
}
```

A.2 Two-biological-measure Model with uninformative PK parameters

```
model
{
# Data are as follows:
#
# npart = number of participants.
# nmeasure = number of measurements (same for now across individuals).
# npill = number of days each subject participated in the trial
#(same for all individuals for now).
# pilltime[i,j] = time ith individual took jth pill.
# measuretime[i,k] = time of kth measurement for individual i.
# plasma[i,k] = measured value of drug in plasma of kth
#blood sample from ith participant.
# pbmc[i,k] = measured value of drug in PBMC of kth
#blood sample from ith participant.
# dot[i,j] = 1 if jth pill was taken during the DOT period for
#individual i (=1 for entire DOT period).
# sched[i,j] = 1 for any day during the DOT period where a pill
#was scheduled to be taken.
# r[i,j] = 1 if they self-report taking the pill that day
#(includes DOT period).
# npill2[i,k] = the last day in the look-back period at
#kth measurement for individual i.
# npill3[i,k] = the first day in the look-back period at
#kth measurement for individual i.
#
# Self-reporting
```

```
for (i in 1:npart)
{
        p[i] ~ dbeta(alpha1,beta1) # Adherence
        phi[i] ~ dbeta(alpha2,beta2)
        # Self-reporting pill-taking given lack of adherence
        for (j in 1:npill)
        {
                prob.adh[i,j] <-
                (dot[i,j]*sched[i,j]) + ((1-dot[i,j])*p[i])
                a[i,j] ~ dbern(prob.adh[i,j])
                prob.self[i,j] <-
                ((a[i,j]+((1-a[i,j])*phi[i]))*(1-dot[i,j]))
                + (dot[i,j]*a[i,j])
                r[i,j] ~ dbern(prob.self[i,j])
        }
}


# Pill-taking.

for (i in 1:npart) {

# Pharmacokinetic parameters.
theta[i,1:5] ~ dmnorm(mu[],tau[,])
log.ke.pl[i] <- theta[i,1]
log.Cl.pl[i] <- theta[i,2]
log.ke.pb[i] <- theta[i,3]
```

```
log.Cl.pb[i] <- theta[i,4]
log.ka.pb[i] <- theta[i,5]
Cl.pl[i] <- exp(log.Cl.pl[i])
ke.pl[i] <- exp(log.ke.pl[i])
Cl.pb[i] <- exp(log.Cl.pb[i])
ke.pb[i] <- exp(log.ke.pb[i])
ka.pb[i] <- exp(log.ka.pb[i])

# DOT and self-administered phase.
for (k in 1:nmeasure)
{
    for (j in npill3[i,k]:npill2[i,k])
    {
            summand.pb[i,k,j] <-
            a[i,j]*((dose*ke.pb[i])/Cl.pb[i])
            *(ka.pb[i]/(ka.pb[i]-ke.pb[i]))
            *(exp(-ke.pb[i]*(measuretime[i,k]-pilltime[i,j]))
            -exp(-ka.pb[i]*(measuretime[i,k]-pilltime[i,j])))
                summand.pl[i,k,j] <-
        a[i,j]*((dose*ke.pl[i])/Cl.pl[i])
        *exp(-ke.pl[i]
        *(measuretime[i,k]-pilltime[i,j]))
        }
        mpbmc[i,k] <-
        log(sum(summand.pb[i,k,npill3[i,k]:npill2[i,k]]))
        mplasma[i,k] <-
        log(sum(summand.pl[i,k,npill3[i,k]:npill2[i,k]]))
```

```
        pcens.pb[i,k] <-phi((level.pb-mpbmc[i,k])/sigma3.pb)
        pcens.pl[i,k] <-phi((level.pl-mplasma[i,k])/sigma3.pl)
        cens.pb[i,k] ~ dbern(pcens.pb[i,k])
        cens.pl[i,k] ~ dbern(pcens.pl[i,k])
        lim.pl[i,k] <- ((1-cens.pl[i,k])*100000000)
        + (cens.pl[i,k]*level.pl)
        plasma[i,k] ~ dnorm(mplasma[i,k],eps.tau.pl)C(,lim.pl[i,k])
        lim.pb[i,k] <- ((1-cens.pb[i,k])*100000000)
        + (cens.pb[i,k]*level.pb)
        pbmc[i,k] ~ dnorm(mpbmc[i,k],eps.tau.pb)C(,lim.pb[i,k])
        }
}


zeta1 ~ dgamma(3,1)
zeta2 ~ dgamma(3,1)
eta1 ~ dbeta(1,1)
eta2 ~ dbeta(1,1)

alpha1 <- eta1*zeta1
alpha2 <- eta2*zeta2
beta1 <- zeta1*(1-eta1)
beta2 <- zeta2*(1-eta2)

# Priors for pharmacokinetic parameters
mu[1] ~ dnorm(0,0.001)
mu[2] ~ dnorm(0,0.001)
```

```
mu[3] ~ dnorm(0,0.001)
mu[4] ~ dnorm(0,0.001)
mu[5] ~ dnorm(0,0.001)

tau[1:5,1:5] ~ dwish(R[,],5)
R[1,1] <- 0.45
R[1,2] <- 0
R[1,3] <- 0
R[1,4] <- 0
R[1,5] <- 0
R[2,1] <- 0
R[2,2] <- 0.45
R[2,3] <- 0
R[2,4] <- 0
R[2,5] <- 0
R[3,1] <- 0
R[3,2] <- 0
R[3,3] <- 0.45
R[3,4] <- 0
R[3,5] <- 0
R[4,1] <- 0
R[4,2] <- 0
R[4,3] <- 0
R[4,4] <- 0.45
R[4,5] <- 0
R[5,1] <- 0
R[5,2] <- 0
```

```
R[5,3] <- 0
R[5,4] <- 0
R[5,5] <- 0.45

sigma[1:5,1:5] <- inverse(tau[,])
rho.pl <- sigma[1,2]/sqrt(sigma[1,1]*sigma[2,2])
#correlation between logke and logcl for pbmc
rho.pb1 <- sigma[3,4]/sqrt(sigma[3,3]*sigma[4,4])
#correlation between logcl and logka for pbmc
rho.pb2 <- sigma[4,5]/sqrt(sigma[4,4]*sigma[5,5])
#correlation between logke and logka for pbmc
rho.pb3 <- sigma[3,5]/sqrt(sigma[3,3]*sigma[5,5])

eps.tau.pb <- exp(logtaupb)
eps.tau.pl <- exp(logtaupl)
logtaupb ~ dflat()
logtaupl ~ dflat()
sigma3.pb <- 1/sqrt(eps.tau.pb)
sigma3.pl <- 1/sqrt(eps.tau.pl)
}
```

A.3 Two-biological-measure Model with informative absorption rate parameter

```
model
{
# Data are as follows:
#
# npart = number of participants.
# nmeasure = number of measurements (same for now across individuals).
# npill = number of days each subject participated in the trial
#(same for all individuals for now).
# pilltime[i,j] = time ith individual took jth pill.
# measuretime[i,k] = time of kth measurement for individual i.
# plasma[i,k] = measured value of drug in plasma of kth
#blood sample from ith participant.
# pbmc[i,k] = measured value of drug in PBMC of kth
#blood sample from ith participant.
# dot[i,j] = 1 if jth pill was taken during the DOT period for
#individual i (=1 for entire DOT period).
# sched[i,j] = 1 for any day during the DOT period where a pill
#was scheduled to be taken.
# r[i,j] = 1 if they self-report taking the pill that day
#(includes DOT period).
# npill2[i,k] = the last day in the look-back period at
#kth measurement for individual i.
# npill3[i,k] = the first day in the look-back period at
#kth measurement for individual i.
#
```

```
# Self-reporting
for (i in 1:npart)
{
        p[i] ~ dbeta(alpha1,beta1) # Adherence
        phi[i] ~ dbeta(alpha2,beta2)
        # Self-reporting pill-taking given lack of adherence
        for (j in 1:npill)
        {
                prob.adh[i,j] <-
                (dot[i,j]*sched[i,j]) + ((1-dot[i,j])*p[i])
                a[i,j] ~ dbern(prob.adh[i,j])
                prob.self[i,j] <-
                ((a[i,j]+((1-a[i,j])*phi[i]))*(1-dot[i,j]))
                + (dot[i,j]*a[i,j])
                r[i,j] ~ dbern(prob.self[i,j])
        }
}


# Pill-taking.

for (i in 1:npart) {

# Pharmacokinetic parameters.
theta[i,1:5] ~ dmnorm(mu[],tau[,])
log.ke.pl[i] <- theta[i,1]
log.Cl.pl[i] <- theta[i,2]
```

```
log.ke.pb[i] <- theta[i,3]
log.Cl.pb[i] <- theta[i,4]
log.ka.pb[i] <- theta[i,5]
Cl.pl[i] <- exp(log.Cl.pl[i])
ke.pl[i] <- exp(log.ke.pl[i])
Cl.pb[i] <- exp(log.Cl.pb[i])
ke.pb[i] <- exp(log.ke.pb[i])
ka.pb[i] <- exp(log.ka.pb[i])

# DOT and self-administered phase.
for (k in 1:nmeasure)
{
    for (j in npill3[i,k]:npill2[i,k])
    {
            summand.pb[i,k,j] <-
            a[i,j]*((dose*ke.pb[i])/Cl.pb[i])
            *(ka.pb[i]/(ka.pb[i]-ke.pb[i]))
            *(exp(-ke.pb[i]*(measuretime[i,k]-pilltime[i,j]))
            -exp(-ka.pb[i]*(measuretime[i,k]-pilltime[i,j])))
                summand.pl[i,k,j] <-
        a[i,j]*((dose*ke.pl[i])/Cl.pl[i])
        *exp(-ke.pl[i]
        *(measuretime[i,k]-pilltime[i,j]))
        }
        mpbmc[i,k] <-
        log(sum(summand.pb[i,k,npill3[i,k]:npill2[i,k]]))
        mplasma[i,k] <-
```

```
        log(sum(summand.pl[i,k,npill3[i,k]:npill2[i,k]]))
        pcens.pb[i,k] <-phi((level.pb-mpbmc[i,k])/sigma3.pb)
        pcens.pl[i,k] <-phi((level.pl-mplasma[i,k])/sigma3.pl)
        cens.pb[i,k] ~ dbern(pcens.pb[i,k])
        cens.pl[i,k] ~ dbern(pcens.pl[i,k])
        lim.pl[i,k] <- ((1-cens.pl[i,k])*100000000)
        + (cens.pl[i,k]*level.pl)
        plasma[i,k] ~ dnorm(mplasma[i,k],eps.tau.pl)C(,lim.pl[i,k])
        lim.pb[i,k] <- ((1-cens.pb[i,k])*100000000)
        + (cens.pb[i,k]*level.pb)
        pbmc[i,k] ~ dnorm(mpbmc[i,k],eps.tau.pb)C(,lim.pb[i,k])
        }
}


zeta1 ~ dgamma(3,1)
zeta2 ~ dgamma(3,1)
eta1 ~ dbeta(1,1)
eta2 ~ dbeta(1,1)

alpha1 <- eta1*zeta1
alpha2 <- eta2*zeta2
beta1 <- zeta1*(1-eta1)
beta2 <- zeta2*(1-eta2)

# Priors for pharmacokinetic parameters
mu[1] ~ dnorm(0,0.001)
```

```
mu[2] ~ dnorm(0,0.001)
mu[3] ~ dnorm(0,0.001)
mu[4] ~ dnorm(0,0.001)
mu[5] ~ dnorm(-0.693,1000)

tau[1:5,1:5] ~ dwish(R[,],5)
R[1,1] <- 0.45
R[1,2] <- 0
R[1,3] <- 0
R[1,4] <- 0
R[1,5] <- 0
R[2,1] <- 0
R[2,2] <- 0.45
R[2,3] <- 0
R[2,4] <- 0
R[2,5] <- 0
R[3,1] <- 0
R[3,2] <- 0
R[3,3] <- 0.45
R[3,4] <- 0
R[3,5] <- 0
R[4,1] <- 0
R[4,2] <- 0
R[4,3] <- 0
R[4,4] <- 0.45
R[4,5] <- 0
R[5,1] <- 0
```

```
R[5,2] <- 0
R[5,3] <- 0
R[5,4] <- 0
R[5,5] <- 0.45

sigma[1:5,1:5] <- inverse(tau[,])
rho.pl <- sigma[1,2]/sqrt(sigma[1,1]*sigma[2,2])
#correlation between logke and logcl for pbmc
rho.pb1 <- sigma[3,4]/sqrt(sigma[3,3]*sigma[4,4])
#correlation between logcl and logka for pbmc
rho.pb2 <- sigma[4,5]/sqrt(sigma[4,4]*sigma[5,5])
#correlation between logke and logka for pbmc
rho.pb3 <- sigma[3,5]/sqrt(sigma[3,3]*sigma[5,5])

eps.tau.pb <- exp(logtaupb)
eps.tau.pl <- exp(logtaupl)
logtaupb ~ dflat()
logtaupl ~ dflat()
sigma3.pb <- 1/sqrt(eps.tau.pb)
sigma3.pl <- 1/sqrt(eps.tau.pl)
}
```

A.4 Two-biological-measure Model with multi-compartment system

```
model
{
# Data are as follows:
#
# npart = number of participants.
# nmeasure = number of measurements (same for now across individuals).
# npill = number of days each subject participated in the trial
#(same for all individuals for now).
# pilltime[i,j] = time ith individual took jth pill.
# measuretime[i,k] = time of kth measurement for individual i.
# plasma[i,k] = measured value of drug in plasma of kth
#blood sample from ith participant.
# pbmc[i,k] = measured value of drug in PBMC of kth
#blood sample from ith participant.
# dot[i,j] = 1 if jth pill was taken during the DOT period for
#individual i (=1 for entire DOT period).
# sched[i,j] = 1 for any day during the DOT period where a pill
#was scheduled to be taken.
# r[i,j] = 1 if they self-report taking the pill that day
#(includes DOT period).
# npill2[i,k] = the last day in the look-back period at
#kth measurement for individual i.
# npill3[i,k] = the first day in the look-back period at
#kth measurement for individual i.
#
# Self-reporting
```

```
for (i in 1:npart)
{
        p[i] ~ dbeta(alpha1,beta1) # Adherence
        phi[i] ~ dbeta(alpha2,beta2)
        # Self-reporting pill-taking given lack of adherence
        for (j in 1:npill)
        {
                prob.adh[i,j] <-
                (dot[i,j]*sched[i,j]) + ((1-dot[i,j])*p[i])
                a[i,j] ~ dbern(prob.adh[i,j])
                prob.self[i,j] <-
                ((a[i,j]+((1-a[i,j])*phi[i]))*(1-dot[i,j]))
                + (dot[i,j]*a[i,j])
                r[i,j] ~ dbern(prob.self[i,j])
        }
}
# Pill-taking.

for (i in 1:npart) {

    # Pharmacokinetic parameters.
    theta[i,1:6] ~ dmnorm(mu[],tau[,])
        log.ka[i] <- theta[i,1]
        log.k23[i] <- theta[i,2]
        log.k32[i] <- theta[i,3]
        log.k20[i] <- theta[i,4]
        log.k24[i] <- theta[i,5]
```

```
log.k40[i] <- theta[i,6]

ka[i] <- exp(log.ka[i])
k23[i] <- exp(log.k23[i])
k32[i] <- exp(log.k32[i])
k20[i] <- exp(log.k20[i])
k24[i] <- exp(log.k24[i])
k40[i] <- exp(log.k40[i])

alp[i] <-
0.5*((k23[i]+k32[i]+k20[i]+k24[i])
+sqrt(pow((k23[i]+k32[i]+k20[i]+k24[i]),2)
-4*k32[i]*(k20[i]+k24[i])))
be[i] <-
0.5*((k23[i]+k32[i]+k20[i]+k24[i])
-sqrt(pow((k23[i]+k32[i]+k20[i]+k24[i]),2)
-4*k32[i]*(k20[i]+k24[i])))

pl.part1[i] <-
dose*ka[i]/((-ka[i]+alp[i])*(-ka[i]+be[i])*(alp[i]-be[i]))

pb.part1[i] <-
((dose*ka[i]*k24[i])/((-ka[i]+k40[i])
*(-ka[i]+alp[i])*(-ka[i]+be[i])
*(alp[i]-be[i])*(k40[i]-alp[i])
*(k40[i]-be[i])))
```

```
# DOT and self-administered phase.
for (k in 1:nmeasure){
for (j in npill3[i,k]:npill2[i,k]){

pl.part2[i,k,j] <-
(k32[i]-ka[i])*(alp[i]-be[i])*exp(-ka[i]*24*(measuretime[i,k]
-pilltime[i,j]))+(k32[i]-alp[i])*(-ka[i]+be[i])
*exp(-alp[i]*24*(measuretime[i,k]-pilltime[i,j]))
-(k32[i]-be[i])*(-ka[i]+alp[i])
*exp(-be[i]*24*(measuretime[i,k]-pilltime[i,j]))

summand.pl[i,k,j] <- a[i,j]*pl.part1[i]*pl.part2[i,k,j]



pb.part2[i,k,j] <-
(k32[i]-ka[i])*(alp[i]-be[i])
*(k40[i]-alp[i])*(k40[i]-be[i])
*exp(-ka[i]*24*(measuretime[i,k]-pilltime[i,j]))
pb.part3[i,k,j] <-
(k32[i]-k40[i])*(alp[i]-be[i])
*(-ka[i]+alp[i])*(-ka[i]+be[i])
*exp(-k40[i]*24*(measuretime[i,k]-pilltime[i,j]))
pb.part4[i,k,j] <-
(k32[i]-alp[i])*(-ka[i]+k40[i])
*(-ka[i]+be[i])*(k40[i]-be[i])
*exp(-alp[i]*24*(measuretime[i,k]-pilltime[i,j]))
```

```
pb.part5[i,k,j] <-
(k32[i]-be[i])*(-ka[i]+k40[i])
*(-ka[i]+alp[i])*(k40[i]-alp[i])
*exp(-be[i]*24*(measuretime[i,k]-pilltime[i,j]))

summand.pb[i,k,j] <-
a[i,j]*pb.part1[i]
*(pb.part2[i,k,j]-pb.part3[i,k,j]
+pb.part4[i,k,j]-pb.part5[i,k,j])
}

mpbmc[i,k] <-
log(sum(summand.pb[i,k,npill3[i,k]:npill2[i,k]]))
mplasma[i,k] <-
log(sum(summand.pl[i,k,npill3[i,k]:npill2[i,k]]))
pcens.pb[i,k] <-phi((level.pb-mpbmc[i,k])/sigma3.pb)
pcens.pl[i,k] <-phi((level.pl-mplasma[i,k])/sigma3.pl)
cens.pb[i,k] ~ dbern(pcens.pb[i,k])
cens.pl[i,k] ~ dbern(pcens.pl[i,k])
lim.pl[i,k] <-
((1-cens.pl[i,k])*100000000) + (cens.pl[i,k]*level.pl)
plasma[i,k] ~
dnorm(mplasma[i,k],eps.tau.pl)C(,lim.pl[i,k])
lim.pb[i,k] <-
((1-cens.pb[i,k])*100000000) + (cens.pb[i,k]*level.pb)
pbmc[i,k] ~ dnorm(mpbmc[i,k],eps.tau.pb)C(,lim.pb[i,k])
}
```

```
    }


zeta1 ~ dgamma(3,1)
zeta2 ~ dgamma(3,1)
eta1 ~ dbeta(1,1)
eta2 ~ dbeta(1,1)

alpha1 <- eta1*zeta1
alpha2 <- eta2*zeta2
beta1 <- zeta1*(1-eta1)
beta2 <- zeta2*(1-eta2)

# Priors for pharmacokinetic parameters
mu[1] ~ dnorm(0,0.001)
mu[2] ~ dnorm(0,0.001)
mu[3] ~ dnorm(0,0.001)
mu[4] ~ dnorm(0,0.001)
mu[5] ~ dnorm(0,0.001)
mu[6] ~ dnorm(0,0.001)

tau[1:6,1:6] ~ dwish(R[,],6)
R[1,1] <- 0.54
R[1,2] <- 0
R[1,3] <- 0
R[1,4] <- 0
R[1,5] <- 0
```

```
R[1,6] <- 0
R[2,1] <- 0
R[2,2] <- 0.54
R[2,3] <- 0
R[2,4] <- 0
R[2,5] <- 0
R[2,6] <- 0
R[3,1] <- 0
R[3,2] <- 0
R[3,3] <- 0.54
R[3,4] <- 0
R[3,5] <- 0
R[3,6] <- 0
R[4,1] <- 0
R[4,2] <- 0
R[4,3] <- 0
R[4,4] <- 0.54
R[4,5] <- 0
R[4,6] <- 0
R[5,1] <- 0
R[5,2] <- 0
R[5,3] <- 0
R[5,4] <- 0
R[5,5] <- 0.54
R[5,6] <- 0
R[6,1] <- 0
R[6,2] <- 0
```

```
R[6,3] <- 0
R[6,4] <- 0
R[6,5] <- 0
R[6,6] <- 0.54

sigma[1:6,1:6] <- inverse(tau[,])

eps.tau.pb <- exp(logtaupb)
eps.tau.pl <- exp(logtaupl)
logtaupb ~ dflat()
logtaupl ~ dflat()
sigma3.pb <- 1/sqrt(eps.tau.pb)
sigma3.pl <- 1/sqrt(eps.tau.pl)
}
```