



Universiteit
Leiden
The Netherlands

A Bayesian population pharmacokinetic meta-analysis of individual participant data

Osch, P. van

Citation

Osch, P. van. (2013). *A Bayesian population pharmacokinetic meta-analysis of individual participant data*.

Version: Not Applicable (or Unknown)

License: [License to inclusion and publication of a Bachelor or Master thesis in the Leiden University Student Repository](#)

Downloaded from: <https://hdl.handle.net/1887/3597311>

Note: To cite this publication please use the final published version (if applicable).

Puck van Osch

A Bayesian population pharmacokinetic
meta-analysis of individual participant
data

Master Thesis, Defend on March 26, 2013

Thesis advisor:
Prof. dr. Emmanuel Lesaffre

Specialization: Statistical Science



Abstract

The effect of a drug is directly related to drug concentration in blood. Therefore, predicting and describing the time course of drug concentration in the body, i.e. the study of pharmacokinetics, is crucial in establishing and optimizing drug therapy.

The immunosuppressant mycophenolate mofetil (with active ingredient mycophenolic acid), is extensively used in renal transplantation in order to prevent acute allograft rejection. Large inter-subject and inter-occasion variability in the exposure of mycophenolic acid in combination with one standard dosing regimen for all renal transplant recipients makes it uncertain if the therapeutic window is reached in every patient. With the pharmacokinetic data of six historical studies, it is tried to find the underlying causes responsible for the inter-subject variability in order to optimize the mycophenolic acid exposure in every patient using Bayesian statistics. As the statistical methods are essential in answering such a research question and these data is already analysed using traditional, maximum likelihood, methods by Van Hest *et al.*, emphasis is also put on the comparison of our and their results.

The data were described with nonlinear mixed effects models using WinBUGS and its interface PKBUGS. Analysing the data of the individual studies resulted in identifiability problems, i.e. the data of most studies were insufficient in estimating all model parameters with non-informative priors. However, advantages of Bayesian methods emerged, as prior information was increased to obtain accurate posterior summary measures.

Combining the data of all studies, i.e. the individual participant data meta-analyses, improved the quality of the data, the analyses and the reliability of the results. Subsequently, the identifiability problems diminished and using close to non-informative prior information resulted in reliable, precise posterior summary measures on almost all model parameters and some covariates were found to explain the inter-subject variability.

Comparing our results with the results obtained from the analysis with non-Bayesian methods turned out to be hard. More precise and very different estimates of the model parameters were obtained with Bayesian methods. Probably caused by differences between the used models, e.g. in contrast to the model used for non-Bayesian methods, we did not consider covariates with missing values and time-dependent covariates but did correct for inter-study variability. Nevertheless, the contribution of prior information is a major advantage in the analysis of pharmacokinetic data, especially because pharmacokinetic data is hard to collect and highly reliable historical pharmacokinetic studies are available for each drug on the market.

Contents

1	Introduction	1
2	The Data Sets and the Research Questions	5
2.1	Introduction	5
2.2	The Studies	6
2.3	The Aims of this Thesis	7
3	Pharmacokinetics	11
3.1	Rate Processes in Pharmacokinetics	11
3.2	ADME	12
3.2.1	Absorption	12
3.2.2	Distribution	13
3.2.3	Elimination = Metabolism and Excretion	13
3.3	Compartment Models	14
3.3.1	Building the Model	16
4	Model Estimation Techniques	19
4.1	Bayesian Methods	20
4.2	PKBUGS	24
4.3	Bayesian Population Pharmacokinetic Models	25
4.3.1	Modeling the Inter-Occasion Variability	26
4.3.2	Modeling the Inter-Study Variability	27
5	Analysis of the Individual Studies	31
5.1	Introduction	31
5.1.1	The Model	31
5.2	Study 1	32
5.2.1	Model with Non-Informative Prior Information	32
5.2.2	Model with Fully-Informative Prior Information	33
5.2.3	Model with Minimally-Informative Prior Information	33
5.3	Study 2	34
5.4	Study 3	34
5.5	Study 4	37
5.6	Study 5	37
5.7	Study 6	40
5.8	Conclusion	43
6	The Individual Participant Data Meta-Analysis	45
6.1	Estimating the Pharmacokinetic Parameters	45
6.1.1	Prior Information	46
6.1.2	The Results	46
6.2	The Regression Analyses	49
6.2.1	The Covariates	49

6.2.2	Prior Information	49
6.2.3	Model Selection	50
6.2.4	The Results	50
6.3	Conclusion	51
7	Conclusion	55
A	PKBUGS Analyses	59
A.1	The Data	59
A.2	Model Specifications	60
A.3	The Analysis	60
A.4	The results	61
B	Priors per Study	63
B.1	Study 1	64
B.2	Study 2	64
B.3	Study 3	64
B.4	Study 4	64
B.5	Study 5	65
B.6	Study 6	65
C	Trace Plots of the Analysis per Study	67
D	Trace Plots of the Meta-Analysis	77
E	Regression Plots of the Covariates	81
E.1	Gender	81
E.2	Race	82
E.3	Weight	83
E.4	Height	84
E.5	Age	85
E.6	MMF dose	86
E.7	Diabetes Mellitus	87
E.8	Creatinine Clearance using Cockcroft & Gault	88
E.9	Hemoglobin	89
E.10	The use of Antacids	90
E.11	The use of Proton Pump Inhibitors	91
E.12	The use of Anti-Viral Agents	92
	Bibliography	93

1 Introduction

Pharmacokinetics (PK) finds its roots in the multi-disciplinary process of drug research. Drug research involves the development of novel therapeutic agents in areas of medical need and can be divided into two stages: the discovery and design stage *vs.* the development stage. The discovery and design stage includes, among other things, the screening of new lead molecules and the research of new targets for a drug to interact with (most often receptors or enzymes). The development stage starts with a new drug molecule and focuses on its safety and efficacy. This is a very long and expensive stage. On average, it takes 12-15 years for a new drug to reach the patient and costs about US\$ 500-2000 million [1]. The study of pharmacokinetics belongs, to a large extent, to the development stage of drug research and goes hand in hand with the study of *pharmacodynamics (PD)*. A very intuitive and famous definition of PK/PD is given by Benet [2]:

”Pharmacokinetics may be simply defined as what the body does to the drug, as opposed to pharmacodynamics which may be defined as what the drug does to the body.”

Pharmacokinetics is the study of predicting and describing the time course of drug concentration in the body, while pharmacodynamics is the study of the time course and the intensity of the drug effect on the organs/physiology. Both disciplines are often combined resulting in PK/PD modeling. This thesis focuses only on pharmacokinetic modeling. [3]

The pharmacokinetics of a drug is characterized by four important processes: the *absorption, distribution, metabolism* and *excretion (ADME)* of a drug. These processes can be illustrated based on the plasma drug concentrations over time (Figure 1): After oral administration of a drug, the plasma drug concentration increases due to the absorption of the drug by the body. At the same time, the already absorbed amount of the drug gets distributed through the body and some of it gets eliminated from the body, whereby the elimination process can be divided in the metabolism and excretion of a drug. When the elimination process exceeds the absorption process, the plasma drug concentration decreases until all drug is eliminated. The drug concentration reaches its peak when the rate of drug absorption is equal to the rate of elimination and when distribution equilibrium has established. These processes result in a characteristic plasma drug concentration time profile. All of the above mentioned processes have their own parameters to describe the characteristics of the drug.

A basic assumption in the study of pharmacokinetics involves the correlation between the drug effect (PD) and the drug concentration at the site of action. The site of action often involves receptors which are located within or at the surface of cells and are widely distributed throughout the body, which makes them inaccessible to observations. To avoid this problem, drug

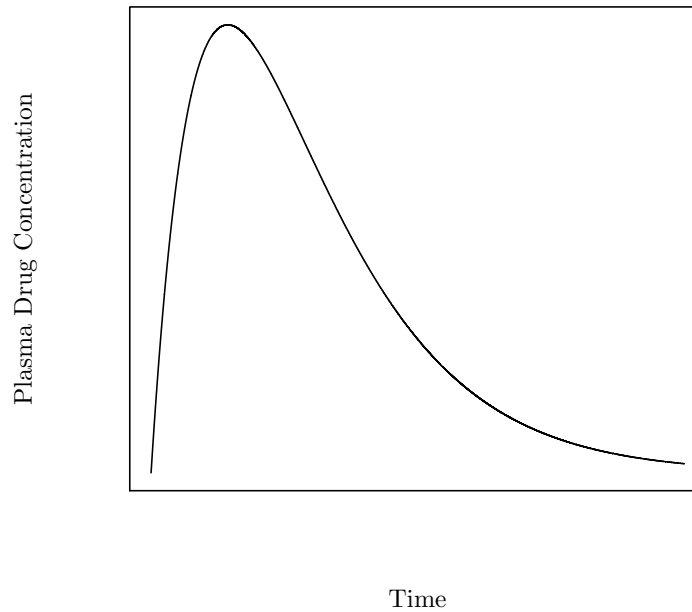


Figure 1: General drug concentration time profile of an oral administered drug.

concentration is most frequently measured in the blood, plasma or urine (for convenient reasons, we will refer to the plasma as the site of measurement throughout this thesis, however, all concepts can be applied to all sites of measurements mentioned before). The drug concentration in these, easy to sample, fluids is proportional to the drug concentration in other tissues, including the site of action. [4,5]

The ultimate goal of drug therapy is the achievement of efficacy without toxicity. To accomplish this, the plasma drug concentration should be high enough to cause the wanted effect but low enough to avoid toxicity. This range is called the therapeutic window, and it depends on the pharmacokinetic properties of a drug. The plasma drug concentration should remain within the therapeutic window of the drug until the desired therapeutic effect is reached, which can be after a single dose of the drug (to relieve a headache) or it can take a life time (in the treatment of diabetes). The characteristics of the pharmacokinetic processes of a drug are important in the correct use of that drug in therapy, and help answering the most important questions in drug therapy: How much? How often? and How long? These questions are highly correlated to each other, and reach the surface in, for instance, the choice of the route of drug administration. The preferred

route is oral administration (because it is easy and cheap), however, some drugs are poorly absorbed by the body resulting in a low plasma drug concentration. The absorption process can be avoided when using intravenous (IV) administration, in which the drug is immediately in the blood of the patient. A good example of such a drug is the analgesic morphine, which is widely used to relieve severe pain; when administered orally, only 40-50% of the dose reaches the site of action. It depends on the needs of the patient which route is more effective. If a patient needs to be treated chronically with high doses of morphine (in cases of severe chronic disease), the therapy will probably involve IV infusion at constant rate. While another patient suffers from severe pain attacks, it is probably advised to use a oral dose of morphine when necessary. Another example of the correct use of a drug in therapy which is based on the pharmacokinetic characteristics of that drug is the choice of the best dosing regimen. Some drugs are completely eliminated from the body in two hours, while other drugs need twelve hours to be completely eliminated from the body. A drug with such a long elimination time should be taken less often by a patient than a drug with a very short elimination time. In general and for convenient reasons, the frequency of drug administration should be kept to a minimum. Two extremes include the benzodiazepines (psychoactive drugs which have sedative, hypnotic, anti-anxiety and muscle-relaxant properties) diazepam, better known as Valium, and triazolam. Fifty percent of diazepam is eliminated from the body in 20-100 hours while the same amount of triazolam takes only 1.5-5.5 hours to be eliminated. Suppose both drugs have the exact same therapeutic effect, it will again depend on the patient's needs (for instance chronic anxiety disorders or acute panick attacks) which drug would be chosen for therapy and would depend on the elimination characteristics of the drug. [5]

The main goal of this thesis is to describe the analysis of complex pharmacokinetic data sets using Bayesian methods. In order to reach this goal, the subsequent sections aim to introduce the reader in all elements needed to achieve this objective. We start with Section 2, which introduces the used data sets and describes the clinical and statistical research questions of this thesis. The subsequent section elaborates on the study of pharmacokinetics. In Section 4, the popular estimation techniques in analysing pharmacokinetic data are discussed with a focus on the techniques and software used in this thesis. The actual analysis are described in Sections 5 and 6. We start with the analysis per study and then describe the analysis of all studies together, i.e. the meta-analysis. Finally, all results, pitfalls and achievements are summarized in the conclusion, Section 7.

2 The Data Sets and the Research Questions

2.1 Introduction

Pharmacokinetic data consist of repeated observations on the same subject over a short period of time. More specifically, the patient receives a fixed amount of a drug, and the concentration of that drug is measured in the patient's blood, plasma or urine over time. The drug of interest in this thesis is mycophenolate mofetil (MMF). MMF is an immunosuppressive drug that is widely used in solid organ transplantations to prevent acute allograft rejection [6]. Six different studies provide the data in this thesis and are discussed in Section 2.2. In these studies, all patients underwent a renal transplantation and received different doses of MMF. MMF is a prodrug, which means that MMF itself is biologically inactive, but after metabolism it produces its active substance mycophenolic acid (MPA). The data consist a total of 1894 MPA concentration time profiles obtained from 468 renal transplant recipients who participated in one of the six studies. The administered MMF dose varies within and between the studies and the plasma MPA concentrations were measured at different time points, varying also within and between the studies. Apart from measuring the MPA concentration, lots of other biological relevant characteristics of the patients were measured. Table 1 shows per study the number of participants, the number of occasions, the number of measurements per occasion, the dosing regimens and summary statistics of the covariates.

When multiple patients receive the same amount of a drug, there is often a high variability between their drug concentration time profiles, the *inter-subject variability (ISV)*. Additionally, when drug concentration time profiles of the same patient are obtained on different occasions, there is often a high variability between the profiles on different occasions, the *inter-occasion variability (IOV; also referred to as intra-subject variability)*. MMF experiences this high ISV and IOV in patients, see Figure 2, and because the recommended dosing regimen of MMF in clinical practice is 1 g twice daily it is uncertain if the therapeutic window is reached in every patient [7–10]. In order to optimize the MPA exposure in every patient and to reduce the number of acute allograft rejections, individualization of the MMF dose may be necessary. This can be achieved if the underlying causes responsible for the fluctuation in drug exposure between and within patients, are found. When their impact is revealed and quantified, physicians may be capable of predicting the MMF exposure in individual patients, and adjust the dosing regimen.

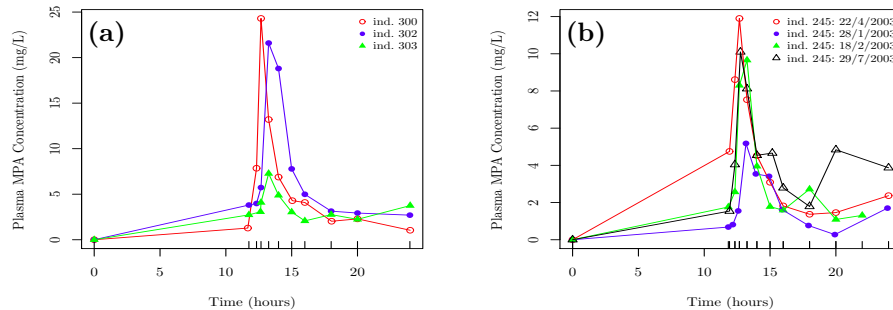


Figure 2: Plasma MPA concentration time profiles. (a) Profiles of three different individuals from study 6 showing inter-subject variability. (b) Profiles of four different occasions of the same individual from study 4 showing inter-occasion variability.

2.2 The Studies

A total of 18 renal transplant recipients participated in the first, unpublished study. The study started one day after the renal transplantation and ended twenty days later. The patients' plasma MPA concentrations were measured on seven time points at baseline and at the end date, see Figures 3(a) and (b) for the ISV and IOV in study 1. Besides MMF, the patients received prednisone (anti-inflammatory) and cyclosporine (immunosuppressant) as co-medication.

The second study is a randomized double-blind multicenter study, with as main goal the evaluation of safety and efficacy of MMF during the first six months after renal transplantation. Three treatment groups were compared based on the incidence of acute rejection, two groups received different dosing regimens of MMF and one group received a recommended dose of azathioprine (immunosuppressant). A total of 499 patients participated in this study and their plasma MPA concentrations were measured on seven time points at the first and the fifth day after transplantation and on the day of hospital discharge (ranging from 6 to 21 days), see Figures 3(c) and (d). The patients also received prednisone and cyclosporine as co-medication. [11,12]

The third study is a randomized double-blind multicenter study. The aim was to investigate the relationship between MPA exposure and acute rejection within the first six months post-transplantation. A total of 154 renal transplantation recipients were randomly allocated to receive one of the three MMF dosing regimens (Table 1). MPA concentration time profiles were obtained at nine fixed occasions (ranging from 3 days to 20 weeks post-transplantation) on eight (the first three occasions) or five time points per day, see Figures 3(e) and (f). Again, all patients were co-medicated with prednisone and cyclosporine. [13,14]

A total of 536 renal transplant recipients enrolled in the fourth study, which was an open-label, multicenter, prospective study with as main clinical goal the improvement of the long term renal function. All patients received the same dose of MMF, but the co-medication varied across the three treatment groups. In two groups, the patients received daclizumab (immunosuppressant) with corticosteroids (anti-inflammatory and immunosuppressant) and cyclosporine, with varying doses of cyclosporine (standard recommendation or a low dose). Patients in the third group received corticosteroids with a standard recommendation of cyclosporine. The MPA exposure was measured at five fixed occasions (ranging from 4 days to 6 months post-transplantation) on ten time points, see Figures 3(g) and 2(b). [15]

In the fifth, unpublished study, 118 renal transplant recipients participated. The MPA exposure was measured on ten time points at four fixed occasions (ranging from 7 days to 12 months), see Figures 3(h) and (i). As co-medication, the patients received either prednisone and cyclosporine or sirolimus (immunosuppressant) and daclizumab.

The sixth study has an open-label, multicenter, parallel group design and aimed to find a correlation between MPA exposure and ethnicity, especially between African-Americans and Caucasians. In total 84 renal transplant recipients participated and the MPA exposure was measured at one occasion (ranging from 6 months to 10 years) on ten time points, see Figure 2(a). The patients received prednisone and cyclosporine as co-medication. [16]

2.3 The Aims of this Thesis

The main clinical goal of this thesis is to explore whether any of the covariates summarized in Table 1 explain the inter-subject variability of MPA exposure. This clinical goal is already stated, investigated and published by Van Hest *et al.* using the same data as the data in this thesis [17]. However, the statistical methods used in this thesis differ from previous research. Here we look at Bayesian methods, while Van Hest *et al.* used maximum likelihood techniques to estimate the parameters. Combining the clinical research question and the Bayesian methods gives rise to the following statistical research question: What is the probability that any of the covariates summarized in Table 1 explain the inter-subject and the inter-occasion variability of the estimated PK parameters given the observed data? And consequently: Do Bayesian methods and inferences have advantages or disadvantages over the classical statistical methods used by Van Hest *et al.* in analysing this data?

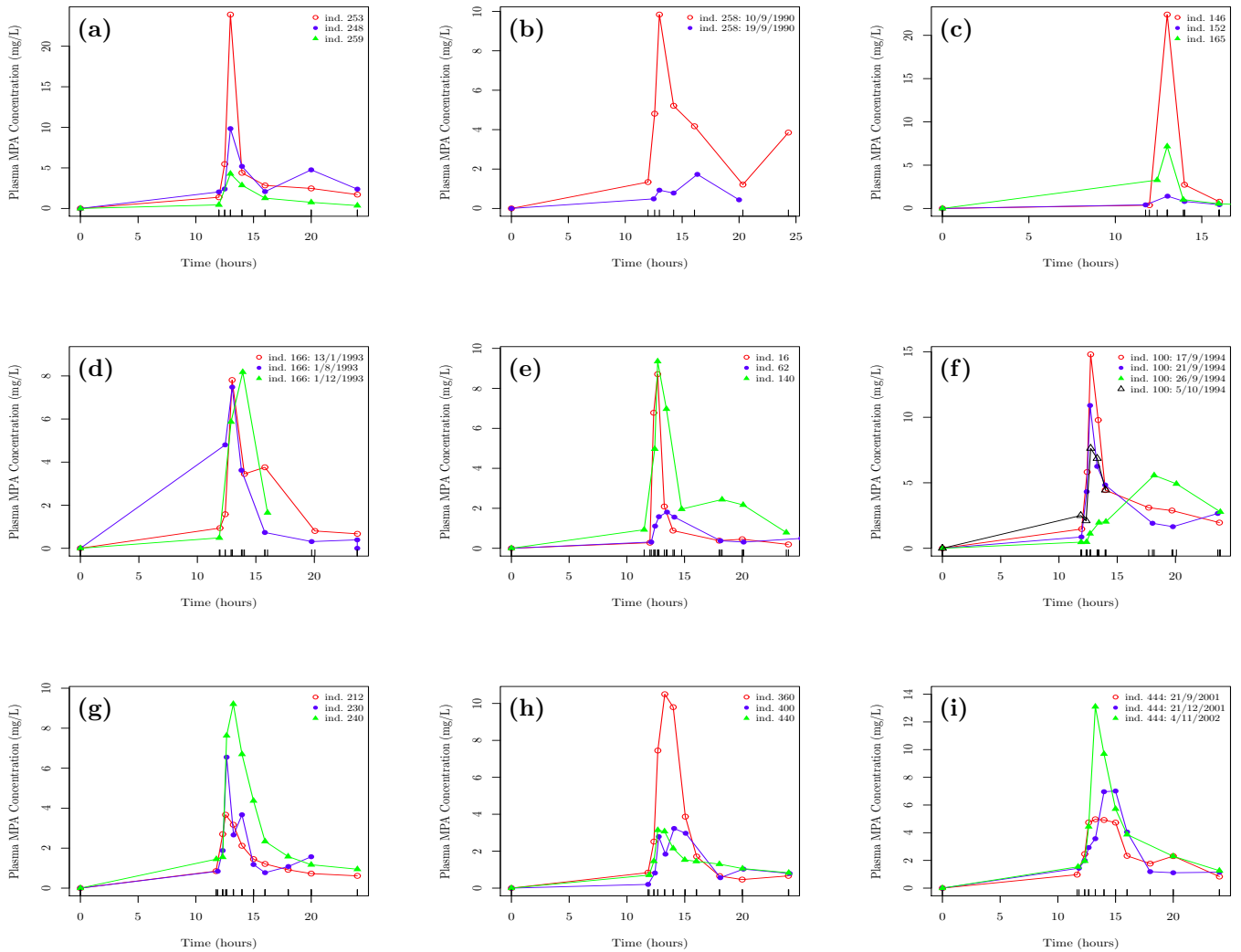


Figure 3: Plasma MPA concentration time profiles showing (a) ISV in study 1, (b) IOV in study 1, (c) ISV in study 2, (d) IOV in study 2, (e) ISV in study 3, (f) IOV in study 3, (g) ISV in study 4, (h) ISV in study 5 and (i) IOV in study 5.

Table 1: Characteristics of the study with summary statistics of the variables per study, median (IQR), and n are based on the first (baseline) measurement of each individual.

Characteristics	All studies	Study 1	Study 2	Study 3	Study 4	Study 5	Study 6
Individuals	467	18	62	141	44	118	84
Number of occasions	3 (5)	2 (0)	2 (1)	9 (2)	5 (2)	4 (2)	1 (0)
Number of measurements per occasion	8.8 (4)	6.25 (1)	7 (0.5)	7 (0.25)	9 (1.1)	11 (0.5)	11 (0)
MMF dose (mg, twice daily)	-	1000, 1500 or 1750	1000 or 1500	450, 950 or 1700	1000	750 or 1000	1000, 1250 or 1500
Gender	-	-	-	-	-	-	-
Female	177	6	21	53	16	45	36
Male	290	12	41	88	28	73	48
Race	-	-	-	-	-	-	-
White	388	11	51	132	37	112	45
Black	57	6	9	3	-	-	39
Other	22	1	2	6	7	6	-
Age(yr)	50 (17)	43.5 (15.8)	49.5 (16.8)	50 (16)	48 (15.8)	53 (16)	51.5 (18.5)
Bodyweight (kg)	73.2 (21.3)	72.4 (25.6)	76 (27)	68 (15.6)	79.8 (30.4)	72.1 (20.2)	79.15 (18.4)
Height(cm)	170 (14)	170 (6.5)	173 (14.8)	170 (13)	172.4 (15.4)	168 (13)	172.7 (12.7)
Diabetes	103	8	25	7	12	24	27
Delayed Graft Function	53	-	8	23	7	15	0
Use of Antacids	63	11	35	5	2	8	2
Use of Proton Pump Inhibitors	15	0	1	7	0	1	6
Use of H2 Antagonist	87	6	1	58	1	0	21
Use of Antiviral Agents	60	9	47	2	0	0	2
Use of Sirolimus	30	0	0	0	30	-	-
Serum Creatinine ($\mu\text{mol/L}$)	246.8 (511.8)	534.8 (643)	505 (444)	212 (353)	526.0 (658.6)	577.6 (519.4)	132.6 (35.4)
Creatinine Clearance (ml/min)	31.0 (43.6)	16.6 (35.4)	16.0 (24.6)	31.5 (32.2)	19.4 (44.2)	13.0 (32.7)	63.6 (20.4)
Plasma Albumin (g/dl)	3.6 (0.7)	3.5 (0.8)	3.2 (0.7)	3.4 (0.5)	3.8 (0.5)	4.2 (0.6)	3.8 (0.4)
Serum total Bilirubin (mg/dl)	0.6 (0.3)	0.5 (0.1)	0.3 (0.1)	0.5 (0.3)	0.6 (0.3)	0.6 (0.3)	0.7 (0.3)
Serum Alkaline Phosphatase (U/L)	77.5 (62)	76 (53.5)	62.5 (30.8)	49 (46)	91 (40.5)	132 (95)	86 (50)
Red blood cell count ($\times 10^{12}/\text{L}$)	3.6 (1.1)	2.8 (0.7)	3 (0.7)	3.2 (0.6)	3.93 (0.8)	3.9 (0.8)	4.5 (0.7)
White blood cell count ($\times 10^9/\text{L}$)	11.3 (92.6)	10.0 (5.5)	8.3 (8.7)	8.3 (3.5)	8.1 (3.2)	7.8 (2.9)	7.3(3.4)
Hemoglobin (g/dl)	11 (3.2)	8.9 (2)	9 (1.8)	9.7 (2.1)	12.3 (3.0)	12.3 (1.9)	12.8 (2.2)
Blood Ureaum	40.3 (38.0)	48 (20.3)	50.1 (22.5)	47.6 (42.0)	30.5 (37.3)	50.8 (45.1)	23 (12)
Prednison daily dose (mg)	25 (20)	120 (62.5)	-	30 (5)	50 (42.5)	20 (28.8)	10 (3.8)
Cyclosporine daily dose (mg)	200 (500)	0 (1300)	-	600 (250)	0 (175)	200 (175)	125 (50)

3 Pharmacokinetics

This section aims to introduce the reader to the principles of pharmacokinetics. In this section, we describe the time course of drug concentration within the body and how that is modeled mathematically. As stated before, the time course of the drug concentration is characterized by four processes: the absorption, distribution, metabolism and excretion of the drug. The speed at which these processes occurs is essential in the description of the overall time course of the drug concentration and is represented by a series of rate processes. Therefore, Section 3.1 introduces the most important rate processes in pharmacokinetics and Section 3.2 discusses the underlying pharmacokinetic processes and their parameters. Section 3.3 elaborates on popular pharmacokinetic models that combine all the underlying processes in order to describe them mathematically. In this respect, we focus on the compartment model which is used in this thesis.

3.1 Rate Processes in Pharmacokinetics

The drug concentration fluctuates over time due to the speed of the underlying pharmacokinetic processes. The speed at which these processes occurs, can be described by multiple rate processes. In general, a drug responds either linearly or exponentially with time, resulting in two different rate processes.

The pharmacokinetic rate processes are mathematically described by *ordinary differential equations (ODE)*. An ODE is an equation that contains an ordinary derivative of unknown functions, which are often involved in the mathematical modeling of real-life phenomena [18]. Essentially, two rate processes are of main importance and occur most often in the pharmacokinetics of a drug: the *zero-order* and the *first-order* rate process. A zero-order rate process, like for instance in the absorption of a drug, means that a drug is absorbed at a constant rate, resulting in:

$$\frac{dX_1(t)}{dt} = -k_0, \quad (3.1)$$

where k_0 is a (zero-order rate) constant and $X_1(t)$ is the amount of the drug remaining to be transferred to the body at time t , the solution is described in Section 3.3.1. In zero-order kinetics, the drug concentration decreases or increases linearly with time. Each rate process has its corresponding *half-life* ($T_{1/2}$), which is defined as the time needed for half of the drug to be depleted (absorbed, distributed or eliminated). The half-life of a zero-order rate process is specified as:

$$T_{1/2} = \frac{X_0}{2k_0}, \quad (3.2)$$

where X_0 is the amount (of a drug) at time = 0.

In first-order kinetics, the rate of reaction is proportional to the amount

remaining to be transferred, resulting in:

$$\frac{dX_1(t)}{dt} = -k_1 X_1(t), \quad (3.3)$$

with k_1 a (first-order rate) constant (for the solution, see Section 3.3.1). In first-order kinetics, the drug concentration decreases or increases exponentially with time. The corresponding half-life is then given by the expression:

$$T_{1/2} = \frac{\ln(2)}{k_1}. \quad (3.4)$$

The majority of drugs follow either zero-order or first-order kinetics. Occasionally, the pharmacokinetics of a drug is described by *second-order* or even *nth-order* ($n > 2$) kinetics. Because these rate processes are exceptional in pharmacokinetics, they will not be further discussed in this thesis.

3.2 ADME

3.2.1 Absorption

The absorption process is defined as the process by which a drug proceeds from its site of administration into the blood stream. Apparently, after an intravenous injection, the drug is administered directly into the blood stream, resulting in the lack of an absorption process. In all other routes of administration, the drug has to cross biological membranes in order to reach the blood stream. For instance, when administered orally, it has to cross the membranes of the gastro-intestinal tract (GIT). The speed at which this process occurs depends on the molecular properties of the drug and the dosage form (a normal tablet will be absorbed more rapidly than a slow release formulation). [19]

The previous section described the two main parameters involved in the absorption process: the *absorption rate constant* (k_a) and corresponding *half-life* ($T_{a,1/2}$). As seen above, k_a is a characteristic of the absorption rate and follows either zero-order or first-order kinetics, described by (3.1) or (3.3) while $X_1(t)$ is the amount remaining to be absorbed at time t with a half-life defined by expression (3.2) or (3.4). Another important parameter in the absorption process is the *bioavailability* (F), which is the proportion of chemically unchanged drug that reaches the systemic blood stream. The term 'chemically unchanged' refers to the early breakdown of a drug, which happens most often with drugs administered orally. There are multiple mechanisms that may cause the breakdown of the active substance of the drug in the GIT (before it reaches the blood stream). For instance, some molecules cannot withstand extreme acidity (when the stomach has a very low pH), or enzymes in the GIT may be responsible for digestion of the molecule (which happens with insulin if it were taken orally). The

final parameter involving the absorption is the *absorption lag time* (T_{lag}), defined as the time delay prior to the start of absorption. The absorption lag time may be influenced by factors such as the stomach emptying process or intestinal motility.

3.2.2 Distribution

Distribution is the process of reversible transfer of a drug to and from the blood stream. When drugs enter the blood circulation, they are distributed throughout the entire body, entering and leaving tissues. The extent of distribution depends on the drug, some drugs remain almost completely in the blood stream, while others remain to a large extent in other tissues. For example, lipid-soluble drugs tend to accumulate in fatty tissues. These fatty tissues slowly release small fractions of the drug back into the blood stream, a process that may continue for days after a single dose of the drug. This process is described by the *distribution rate constant* (k_d) and its corresponding *half-life* ($T_{d,1/2}$), defined in Sections 3.1 and 3.2.1. The *volume of distribution* (V) is the parameter that describes the tendency of a drug to distribute out of the blood into the tissues. V relates the obtained plasma drug concentration to the total amount of drug in the body:

$$\text{Volume of distribution (l)} = \frac{\text{Amount in body (mg)}}{\text{Plasma drug concentration (mg/l)}}. \quad (3.5)$$

It represents the volume of plasma necessary to account for all the drug in the body. When most of the drug remains in tissues, the plasma drug concentration is relatively low compared to the administered dose, in such a case, V tends to be extremely large. For instance, the drug chloroquine (prevents malaria) accumulates in the kidney, liver, lung and spleen, and is strongly bound to melanin containing cells (the eyes and skin), its V is 15.000 l. Suppose 500 mg of chloroquine is administered to a patient, then using expression (3.5), the plasma concentration will probably be near 0.03 mg/l. Then, how much plasma, given that 1 l plasma contains 0.03 mg, is necessary to account for the total drug in the body (500 mg)? The answer to this question is V , i.e. 15.000 l. It represents a hypothetical volume, which provides an intuitive measure of the relationship between drug in plasma and tissue. Note that V can only be determined after the establishment of distribution equilibrium between drug in tissue and that in plasma. [5, 19]

3.2.3 Elimination = Metabolism and Excretion

The elimination process is defined as the irreversible removal of drugs from the body. The speed of this process is characterized by the *elimination rate constant* (k_{el}) and corresponding *half-life* ($T_{el,1/2}$), defined in Sections 3.1 and 3.2.1. Elimination occurs by *excretion* and *metabolism*. Excretion is the

irreversible removal of chemically unchanged drug from the body, which occurs predominantly via the kidneys. Occasionally, drugs are excreted via the bile or in the breath (volatile substances). The major mechanism responsible for elimination of drugs from the body is metabolism, i.e. the conversion of one chemical compound into another. The most common chemical reactions of drug metabolism include oxidation, reduction, hydrolysis and conjugation. In general, the liver is the primary site of drug metabolism, however, occasionally, drugs are metabolized in the kidneys, skin, lungs, blood or gastro-intestinal wall. The elimination mechanism is best described by its parameter *clearance* (CL). CL is a proportionality factor that relates the plasma drug concentration to the rate of elimination:

$$\text{Clearance (l/h)} = \frac{\text{Rate of elimination (mg/h)}}{\text{Plasma drug concentration (mg/l)}}. \quad (3.6)$$

Clearance is the theoretical volume of blood, which is effectively cleared of drug per unit of time. For example, suppose a drug has a CL of 2 l/h, this means that 2 liters of blood are cleared of the drug per hour. When the plasma drug concentration is 10 mg/l, then 20 mg of the drug is cleared per hour. Each elimination process can be described by its own CL , for instance the renal clearance (CL_r) or hepatic clearance (CL_h). The total Clearance (CL_{total}) is the sum of all individual processes. [5]

3.3 Compartment Models

The previous sections covered the underlying pharmacokinetic processes in the time course of plasma drug concentration. These processes need to be combined in order to describe the overall time course of drug concentration within the body. To achieve this, three modeling approaches have been suggested: the *physiological model*, the *compartment model* and the *non-compartment approach*.

In physiology modeling (also known as blood flow or perfusion models), the body is divided into compartments based on anatomical regions (such as the blood, heart and liver). The time course of drug levels in the compartments is calculated using blood flow rates through each compartment in the model. On the other hand, in compartment modeling, the body is divided into compartments but these compartments do not represent realistic, physiological parts of the body but they represent a tissue or group of tissues that have similar blood flow rates and drug affinity. The non-compartment approach does not assume a number of compartments and thereby reduces the number of assumptions necessary to model drug concentration time data. The most common PK parameters are estimated purely based on the plasma concentration levels using area under the curve (AUC) and area under the first moment curve (AUMC) calculations.

This section elaborates on the modeling approach used in this thesis, i.e. compartment modeling. [20]

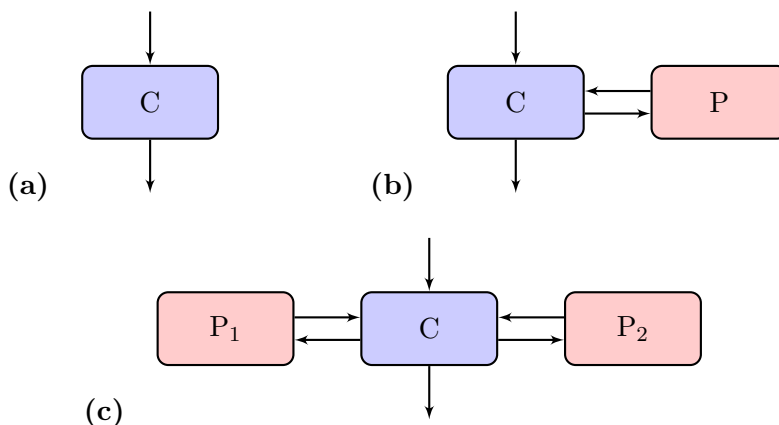


Figure 4: Illustration of (a) one- (b) two- and (c) three-compartment models. With the central compartments (C, blue) and the peripheral compartments (P, pink).

A distinction can be made between one-, two- and three-compartment models, see Figure 4. Within each compartment, the drug is assumed to be uniformly distributed, the compartments are considered to be 'well stirred' and mixing of the drug is assumed to be rapid. This results in a dynamic movement of drugs in and out of the compartments, with an equal probability for each drug molecule to leave or enter a compartment. In a one-compartment model, the body can be seen as one (central) compartment in which the drug is administered into and eliminated from (the body as one giant bucket). It assumes an immediate distribution of the drug throughout the body, resulting in a mono-phasic drug concentration time profile (mono-exponential; see Figure 5(a)). In a two-compartment model, the drug is assumed to distribute between two compartments, a central compartment and a peripheral compartment. Although the compartments do not represent physiological parts of the body, a physiological distinction is assumed to hold. The central compartment is assumed to consist of tissues that are highly perfused (such as the heart, lungs, and kidneys) and the peripheral compartment is assumed to consist less well-perfused tissues (such as the muscles, fat and skin). This results in a bi-phasic drug concentration time profile (bi-exponential, see Figure 5(b)). Finally, in a three-compartment model, the drug is assumed to distribute between three compartments, resulting in a tri-phasic drug concentration time profile (not shown). [5, 20, 21]

Above assumptions make it possible to incorporate the rate processes (described in Section 3.1) into the structure of the compartment model. By combining the ordinary differential equations that represent the rate pro-

cesses (one for each arrow in the compartment model), each compartment can be described mathematically, which is illustrated in Section 3.3.1.

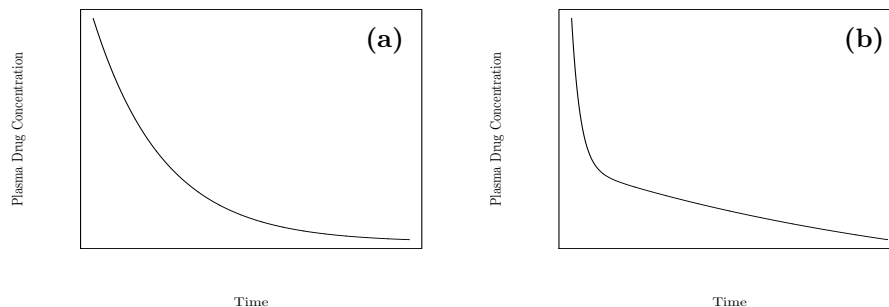


Figure 5: General plasma drug concentration time profiles of (a) one- and (b) two-compartment models with an IV bolus dosing regimen.

3.3.1 Building the Model

Suppose we have pharmacokinetic data of a drug that follows first-order kinetics and was administered with an intravenous bolus dose. We want to describe the drug concentration over time of the central compartment using a two-compartment model. The arrows in the two-compartment model of Figure 4 describe the underlying pharmacokinetic processes. The central compartment of the two-compartment model has three underlying pharmacokinetic processes (the absorption process is omitted because the drug was administered with an intravenous bolus dose): the distribution from the central into the peripheral compartment, the distribution from the peripheral into the central compartment and the elimination from the central compartment. Combining the rates of these processes, results in the following ODE:

$$\frac{dX_1(t)}{dt} = -k_{el}X_1(t) - k_{12}X_1(t) + k_{21}X_2(t), \quad (3.7)$$

where k_{el} is the elimination rate constant, k_{12} is the distribution rate constant representing the rate at which the drug leaves the central compartment, k_{21} is the distribution rate constant representing the rate at which the drug enters the central compartment, $X_1(t)$ is the amount of the drug in the central compartment at time t and $X_2(t)$ is the amount of the drug in the peripheral compartment at time t . Expression (3.7) represents the rate of change of drug concentration versus time. To obtain the expression that describes the drug concentration versus time, we should integrate expression (3.7), which can be done using Laplace transforms [22] resulting in

the following expression:

$$C_p(t) = Ae^{-\alpha t} + Be^{-\beta t}, \quad (3.8)$$

where $C_p(t)$ represents the drug concentration at time t and with A and B defined as:

$$A = \frac{Dose(\alpha - k_{21})}{V(\alpha - \beta)} \quad B = \frac{Dose(k_{21} - \beta)}{V(\alpha - \beta)}, \quad (3.9)$$

where V represents the volume of distribution of the central compartment, $Dose$ represents the administered dose and α and β are defined as:

$$\alpha, \beta = \frac{(\alpha + \beta) \pm \sqrt{(\alpha + \beta)^2 - 4\alpha\beta}}{2}. \quad (3.10)$$

The A , B , α and β terms were derived from the micro-constants k_{12} , k_{21} , k_{el} and V during the integration process. Using the substitutions for the sum and product of α and β ($\alpha + \beta = k_{el} + k_{12} + k_{21}$, $\alpha - \beta = k_{el} - k_{21}$) in expression (3.8) results in expression (3.10). Note that in the numerator of expression (3.10) α is calculated when '+' is used and β is calculated when '-' is used. Thus α is greater than β . All other pharmacokinetic parameters can be expressed in terms of $Dose$ and the micro-constants.

4 Model Estimation Techniques

Before the 1970s, the method for studying pharmacokinetics involved a *two-stage approach* and is often referred to as the *traditional approach*. In the first stage, each of the individual PK parameters is estimated separately through nonlinear regression using the individual's plasma concentration time profile. These estimates are used in the second stage for the calculation of descriptive summary statistics, i.e. mean, variance and covariance, and for establishing a correlation with the patient's characteristics. Several limitations arise from the traditional PK method. For example, in order to obtain reliable PK parameter estimates in the first stage, many appropriately timed blood samples are required (at least three times the number of model parameters), which is often impossible to do in patients from the target population (e.g. AIDS/cancer patients or neonates cannot handle blood loss well). Therefore, traditional PK studies involve often healthy volunteers or a small number of highly selected patients from the target population. Another major limitation of the method is that the subject-specific effects, estimated in the second stage, are likely to be overestimated in absolute value [23–25]. Because of these limitations, the two-stage method does not provide a solid ground for obtaining pharmacokinetic information in order to optimize drug therapy. A turning-point in analysing pharmacokinetic data came when Sheiner *et al.* laid the foundations for *population pharmacokinetic (PPK) modeling* in the 1970s [26, 27]. PPK models involve nonlinear mixed effects models and can be divided into two parts: the *structural* and the *stochastic* part. The structural part estimates individual pharmacokinetic parameters using drug concentration time data (all individuals are processed simultaneously), while the stochastic part describes how the estimates of the PK parameters differ between individuals. A clear definition of PPK is provided by the US Food and Drug Administration (FDA) [28]:

”Population pharmacokinetics is the study of the sources and correlates of variability in drug concentrations among individuals who are the target patient population receiving clinically relevant doses of a drug of interest.”

Major advantages of PPK are that it can handle relatively sparse data, data with unbalanced designs and data obtained during the evaluation of the relationships between dose and efficacy. The nomenclature of this approach is rather deceptive, implying the loss of individual pharmacokinetics, but it does emphasize that interest is focused on the (target) population rather than the individual. [29, 30]

In order to make the analysis of PPK models accessible for a wide range of people studying pharmacokinetics, Sheiner *et al.* developed a software package which contains the algorithms to estimate parameters in nonlinear mixed effects models, i.e. NONMEM [31]. Nowadays, NONMEM is by

far the most popular statistical software for PPK analysis. It uses a joint hierarchical model. A simple form of the first stage of the model is specified as:

$$p(y_{ij}) = f(x_{ij}, \boldsymbol{\phi}_i) + \epsilon_{ij} \quad \epsilon_{ij} \sim N(0, \sigma^2), \quad (4.1)$$

where y_{ij} is the j th observation in individual i . The structural PK model is given by $f(\cdot)$, which is a function of the dosing history x_{ij} and $\boldsymbol{\phi}_i$, the individual-specific PK parameter vector and ϵ_{ij} is the residual variance which is assumed to be normally distributed, i.e. $\epsilon_{ij} \sim N(0, \sigma^2)$.

A simple form of the second stage of the model is specified as:

$$p(\boldsymbol{\phi}_i) = g(\mathbf{z}_i, \boldsymbol{\theta}) + \boldsymbol{\eta}_i \quad \boldsymbol{\eta}_i \sim N(0, \boldsymbol{\Omega}), \quad (4.2)$$

where $g(\cdot)$ is the function of the covariate model, \mathbf{z}_i is the vector of covariates for the i th subject, $\boldsymbol{\theta}$ is the population PK parameter vector and $\boldsymbol{\eta}_i$ is the vector of the individual random effects, with $\boldsymbol{\eta}_i \sim N(0, \boldsymbol{\Omega})$.

NONMEM uses maximum likelihood to estimate the model parameters. In order to obtain estimates of the individual PK parameters $\boldsymbol{\phi}_i$, the marginal likelihood, obtained by integrating out the random effects $\boldsymbol{\eta}_i$, needs to be maximized. However, the resulting integral is very difficult to solve and in order to overcome these intensive numerical integrations, NONMEM uses different approximation techniques, i.e. first-order approximation (FO) [23–25], first-order conditional approximation (FOCE) [32] and Laplacian approximation [33].

The estimation techniques used for the analysis in this thesis involve Bayesian methods, which have the advantage to overcome the integration process involved in estimating the model parameters. Therefore, Bayesian methods can provide exact estimates of the model parameters instead of approximating them. Section 4.1 introduces the reader to Bayesian statistics. In Section 4.2 we describe the interface to the widely used WinBUGS software known as PKBUGS, which was used to determine the PK model parameters in this thesis. In Section 4.3 we describe the implemented Bayesian population pharmacokinetic models.

4.1 Bayesian Methods

In Bayesian statistics, parameters are viewed as random variables. Each parameter involved in a Bayesian model has a distribution attached to it in order to express the uncertainty about its true value. The distribution is known as the *prior* distribution. Prior distributions represent the prior knowledge about the parameter of interest, which is often obtained from historical data (data-based priors). Prior distributions are incorporated in Bayesian analysis using Bayes' Rule. Expression (4.3) gives Bayes' Rule for

continuous parameters:

$$p(\boldsymbol{\theta}|\mathbf{y}) = \frac{L(\boldsymbol{\theta}|\mathbf{y})p(\boldsymbol{\theta})}{p(\mathbf{y})} = \frac{L(\boldsymbol{\theta}|\mathbf{y})p(\boldsymbol{\theta})}{\int L(\boldsymbol{\theta}|\mathbf{y})p(\boldsymbol{\theta})d\boldsymbol{\theta}}, \quad (4.3)$$

where $L(\boldsymbol{\theta}|\mathbf{y})$ is the *likelihood* of the observed data, $p(\boldsymbol{\theta})$ is the distribution of the prior knowledge about the parameter $\boldsymbol{\theta}$, $p(\mathbf{y})$ is the averaged (marginal) likelihood and $p(\boldsymbol{\theta}|\mathbf{y})$ is the resulted *posterior* distribution from which inference is drawn. The relationship between the prior distribution, averaged likelihood and corresponding posterior distribution is illustrated in Figure 6.

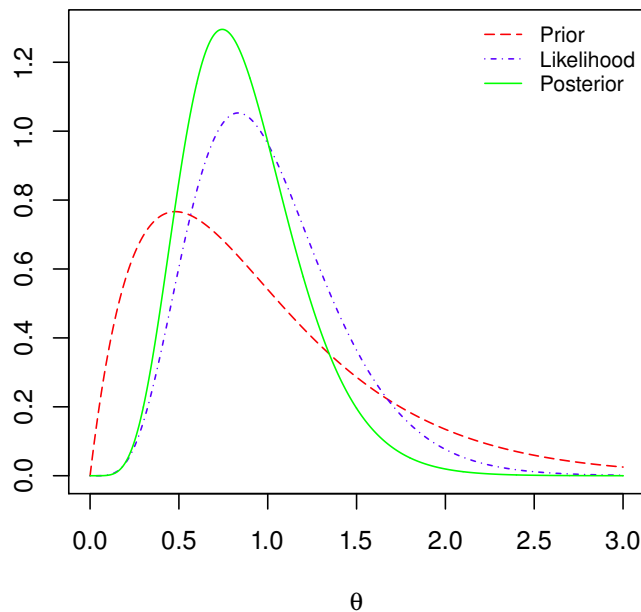


Figure 6: Triplot of a prior (gamma) distribution, (gamma) likelihood of the data and the corresponding (gamma) posterior distribution.

The averaged likelihood is necessary in order for the posterior to be a distribution. By definition, the calculation of the averaged likelihood (and because of that the posterior distribution) involves integration. This integration can become exorbitant, especially when the parameter of interest is high dimensional. For years, the popularity of Bayesian statistics suffered from the impracticable numerical integrations necessary to obtain the posterior distribution. This changed after the introduction of *Markov Chain Monte Carlo (MCMC)* techniques, which resulted in a rise in popularity of Bayesian statistics because it provides a tool to get round the integration process. The most

important and famous MCMC methods include the *Gibbs sampler* [34, 35] and the *Metropolis-Hasting algorithm (MH-algorithm)* [36, 37]. The Gibbs sampler is based on the characteristic that the multivariate distribution is uniquely determined by its conditional distributions. For a two dimensional case, this means that $p(\theta_1, \theta_2 | \mathbf{y})$ is uniquely determined by $p(\theta_1 | \theta_2, \mathbf{y})$ and $p(\theta_2 | \theta_1, \mathbf{y})$. The Gibbs sampler has the following sampling scheme:

- Sample $\theta_1^{(k+1)}$ from $p(\theta_1 | \theta_2^k, \mathbf{y})$,
- Sample $\theta_2^{(k+1)}$ from $p(\theta_2 | \theta_1^{(k+1)}, \mathbf{y})$.

The obtained chain has Markov properties meaning that given θ^k , $\theta^{(k+1)}$ is independent of $\theta^{(k-1)}$, $\theta^{(k-2)}$, etc. It can be proven that sampling from the posterior distribution is achieved following the Gibbs sampling scheme. However, it may take a while for the algorithm to converge and sample from the posterior distribution, therefore an initial part of the chain should be discarded (the *burn-in part*). The MH-algorithm differs from the Gibbs sampler because it does not need the full conditionals but rather uses an instrumental distribution to sample from. The sampled value is then either accepted or rejected. The MH-algorithm samples as follows:

1. Sample a candidate $\tilde{\theta}$ from the proposal density $q(\tilde{\theta} | \theta)$, with $\theta = \theta^k$.
2. The next value $\theta^{(k+1)}$ will be equal to:
 - $\tilde{\theta}$ with probability $\alpha(\theta^k, \tilde{\theta})$ (accept proposal),
 - θ^k otherwise (reject proposal),

with

$$\alpha(\theta^k, \tilde{\theta}) = \min \left(r = \frac{p(\tilde{\theta} | \mathbf{y}) q(\theta^k | \tilde{\theta})}{p(\theta^k | \mathbf{y}) q(\tilde{\theta} | \theta^k)} 1 \right). \quad (4.4)$$

The obtained chain has again Markov properties and it can be proven that the MH-algorithm provides ultimately samples from the posterior distribution. [38]

Multiple Bayesian tools were used in the analysis of this thesis, e.g. to assess and improve convergence, summarize posterior model parameter estimates and check the fit of the model. Below, we discuss these different Bayesian tools, to which we refer later in this thesis.

Over-Relaxation

Numerous techniques have been suggested to improve and accelerate convergence of a Markov chain. *Over-relaxation* is such a technique and is helpful in situations where the elements of the Markov chain are highly positively

correlated. The sampling scheme is adjusted and at step $(k + 1)$ M values $\theta^{k+1,1}, \theta^{k+1,2}, \dots, \theta^{k+1,M}$ are sampled and the current value θ^k is inserted. The value θ^{k+1} is chosen in such a way that it is highly negatively correlated to the current value θ^k , i.e. the $(M + 1)$ values are sorted and given ranks $0, 1, 2, \dots, M$ and θ^{k+1} is taken the value with rank $M - m$ if m is the rank of θ^k . [38, 39]

Assessing Convergence

Reliable and precise posterior summary measures can only be obtained from a converged Markov chain. Therefore, it is important that convergence is assessed before drawing any conclusions from the posterior summary measures. Assessing convergence of a Markov chain involves checking the stationarity of the chain and the accuracy of the posterior summary measures. Multiple techniques have been developed to assess both aspects of the Markov chain, these include graphical and statistical (formal) convergence diagnostics. Graphical diagnostics include, among others, the inspection of the trace plots, which gives an informative impression of the stationarity and mixing rate of the chain for each parameter and the autocorrelation plots, which also show the mixing rate of the chain and the dependency of the chain with its starting position. A short description of the formal diagnostic tests used in this thesis is given below.

Heidelberger and Welch (HW) diagnostic: the HW diagnostic provides tests for the stationarity of the chain and the accuracy of the posterior summary measures. It also provides an estimate of the number of samples that should be discarded as a burn-in sequence. The null hypothesis of convergence uses the Cramer-von-Mises test statistic. If convergence is not rejected, a half-width test is performed by computing the mean and associated $(1-\alpha)100\%$ confidence interval. This test is passed if the half-width of the confidence interval is less than the specified level of accuracy (0.1 as default). [38, 40]

Brooks, Gelman and Ruben (BGR) diagnostic: The above diagnostic is based on single chains, the BGR diagnostic uses multiple chains to test convergence. The test is based on the assumptions that when convergence is reached, the between-chain variability will be relatively small compared to the within-chain variability. The test (the corrected scale reduction factor, CSRf) involves a ratio between the two and is supposed to be passed when the CSRf is smaller than 1.2. [38, 41, 42]

Another tool in the assessment of convergence is the *Effective Sample Size (ESS)*, which estimates the number of independent iterations of the Markov chain. When the elements of a chain experience high autocorrelation, less information is revealed about the posterior distribution of that parameter compared to a chain with independent elements. The ESS measures this amount of information and provides an estimates of the number

of independent Markov samples necessary to give the same precision as the obtained chain.

Posterior Summary Measures

The posterior summary measures reported in this thesis include the mean, standard deviation (sd), *95% credible interval (CI)* and the *MC error*. The 95% CI is the Bayesian analogue of the 95% confidence interval in conventional statistics. In a 95% CI, there's a 95% probability that the mean value of the parameter lies within this interval, while in 95% confidence intervals this probability is either 0 or 1. The MC error is an important tool to assesses the accuracy of the Markov chain. It provides an estimate of the computational error of the mean and is helpful in the assessment of convergence, i.e. establish the MC error (accuracy of the Markov chain) that you wish to attain before considering graphic and formal convergence diagnostics.

Deviance Information Criterion

Another Bayesian tool used in this thesis is the *deviance information criterion (DIC)*, which serves as a model selection criterion. The DIC is a generalization of the Akaike information criterion (AIC). As is the case with AIC, smaller values of DIC indicate a better fitting model. As a rule of thumb, differences of more than 10 definitely favor the model with the lowest DIC, differences between 5 and 10 are substantially favoring the model with the lowest DIC and differences less than 5 indicate that both models have similar fits. For a more extensive background on DIC, the reader is referred to Spiegelhalter *et al.* [43].

The most popular and versatile Bayesian program is WinBUGS. It is the Windows version of the program Bayesian inference Using Gibbs Sampling (BUGS), which is developed in 1989 [44]. The package handles complex Bayesian analyses using Gibbs sampling. We used WinBUGS version 1.3 and 1.4.3 in this thesis. Because the structural part of the population pharmacokinetic model is complex and not straightforward to implement in general statistical software packages (including WinBUGS), an interface was developed for PK/PD modeling within WinBUGS, called PKBUGS. [45] Section 4.2 discusses the basic properties of PKBUGS and appendix A provides a detailed description of analysing PK data using PKBUGS.

4.2 PKBUGS

PKBUGS is an interface for the Bayesian statistical software program WinBUGS that was developed for the analysis of only pharmacokinetic data. The main purpose of PKBUGS is to simplify the specification of PK modeling, which can be done using dialog boxes and menu commands. Currently,

there are two versions of PKBUGS: version 1.1 and 2.0. PKBUGS 1.1 runs on WinBUGS 1.3 and PKBUGS 2.0 on WinBUG 1.4.3. Only PKBUGS 1.1 can be used for the specification of the PK model using menu commands. However, PKBUGS 1.1 generates an equivalent WinBUGS model code which can be runned in WinBUGS 1.4.3 when PKBUGS 2.0 is installed.

As stated before, population pharmacokinetic models can be divided into two parts: the structural part and the stochastic part. PKBUGS 1.1 handles both parts of the population pharmacokinetic model. The first step in specifying the model is data entry, PKBUGS 1.1 recognizes the NONMEM data format and a number of standard data items (like the patient's id, time and the response). Data items that are not recognized by PKBUGS, so called non-standard data items, are assumed to be covariates. Using dialog boxes, the user can regress the covariates against the desired PK parameters. For the specification of the structural part, twenty-eight PK compartment models are implemented, from which the user needs to choose via menu commands. The twenty-eight models comprise one- two- and three-compartment models with the following input characteristics: intravenous (bolus or infusion), zero-order, first-order, zero-order with initial lag time and first-order with initial lag time. This part is called the PK model component, in which the corresponding ordinary differential equations are implemented in PKBUGS and WinBUGS. Besides the specification of the PK model component, the user only needs to specify the priors and a general PK model is generated. The model comprises a general structure of the stochastic part: a three-stage hierarchical model (described in Section 4.3). At this point, the user can either start with the analysis in WinBUGS 1.3 or generate the model code for WinBUGS 1.4.3. The model code is relatively straightforward to modify and incorporate PD components or other extensions. Because the generated model only comprises ISV, in this thesis, the described model is extended with two other levels in its hierarchy, which account for the IOV (described in Section 4.3.1) and the inter-study variability (described in Section 4.3.2). The implemented estimation techniques are the Gibbs sampler and the MH-algorithm. The pharmacokinetic parameters in the PK component are estimated using the MH-algorithm within a Gibbs sampling scheme while all other parameters are estimated using the Gibbs sampler. The use of PKBUGS is further illustrated in Appendix A. [45–48]

4.3 Bayesian Population Pharmacokinetic Models

The general Bayesian population pharmacokinetic model produced by PKBUGS 1.1 and used for the analyses of study 6 (see Section 5.7) includes three hierarchical stages:

Suppose that n_i plasma drug concentrations have been observed for individual i ($i = 1, \dots, K$). Denote the k th observed concentration for individual

i by y_{ij} and corresponding time by t_{ij} . Further, denote the $p \times 1$ vector of pharmacokinetic parameters for individual i by $\boldsymbol{\theta}_i$. At the first stage of the hierarchical model, the form of the probability distribution of each y_{ij} given $\boldsymbol{\theta}_i$ and τ , the inverse of the residual error variance, is specified as:

$$p(y_{ij}|\boldsymbol{\theta}_i, \tau) = N(f_{ij}, \tau^{-1}v_{ij}), \quad \text{for } i = 1, \dots, K, \quad j = 1, \dots, n_i, \quad (4.5)$$

where $y_{ij} \sim N(f_{ij}, \tau^{-1}v_{ij})$ given $\boldsymbol{\theta}_i$ and τ , f_{ij} is the pharmacokinetic model evaluated at time t_{ij} with the individual PK parameters equal to $\boldsymbol{\theta}_i$, i.e. $f(\boldsymbol{\theta}_i, t_{ij})$, and v_{ij} is the residual error structure. Note that other distributions may be chosen instead of a normal distribution, like a lognormal or a student's t-distribution.

At the second stage of the hierarchical model, the following distributional assumptions are made:

$$p(\boldsymbol{\theta}_i|\boldsymbol{\mu}, \boldsymbol{\Omega}^{-1}) = N_p(\boldsymbol{\mu}, \boldsymbol{\Phi}) \quad \text{for } i = 1, \dots, K, \quad (4.6)$$

where $N_p(\cdot)$ denotes the multivariate normal distribution, $\boldsymbol{\mu}$ ($p \times 1$) represents the population pharmacokinetic behavior and $\boldsymbol{\Omega}$ ($p \times p$) is the corresponding variance-covariance matrix representing the ISV.

The third stage of the hierarchical model can be defined by assigning prior densities to the parameters τ , $\boldsymbol{\mu}$ and $\boldsymbol{\Omega}$:

$$p(\tau) = G(\alpha, \beta), \quad (4.7a)$$

$$p(\boldsymbol{\mu}) = N_p(\boldsymbol{\eta}, \mathbf{C}), \quad (4.7b)$$

$$p(\boldsymbol{\Omega}^{-1}) = W_p(\mathbf{R}^{-1}, \rho), \quad (4.7c)$$

where $G(\alpha, \beta)$ represents a gamma distribution with parameters α and β , $\boldsymbol{\eta}$ ($p \times 1$) represent the prior estimates of $\boldsymbol{\mu}$ with variance-covariance matrix \mathbf{C} and W_p denotes a p -dimensional Wishart distribution with mean \mathbf{R}^{-1} ($p \times p$) and degrees of freedom ρ .

4.3.1 Modeling the Inter-Occasion Variability

For simultaneously modeling the ISV and IOV, which was done for the analysis of studies 1 to 5 (see Sections 5.2 to 5.6), another hierarchy is brought into the model:

Suppose that n_{ij} plasma drug concentrations have been observed for individual i on occasion j ($j = 1, \dots, m_i$). Denote the k th observed concentration for individual i on occasion j by y_{ijk} and corresponding time by t_{ijk} and

the $p \times 1$ vector of pharmacokinetic parameters for individual i on the j th occasion by $\boldsymbol{\lambda}_{ij}$. The first stage of the model is specified as:

$$p(y_{ijk}|\boldsymbol{\lambda}_{ij}, \tau) = N(f_{ijk}, \tau^{-1}v_{ijk}),$$

for $i = 1, \dots, K, j = 1, \dots, m_i, k = 1, \dots, n_{ij}$, (4.8)

where $y_{ijk} \sim N(f_{ijk}, \tau^{-1}v_{ijk})$ given $\boldsymbol{\lambda}_{ij}$ and τ , f_{ijk} is the pharmacokinetic model evaluated at time t_{ijk} with the individual PK parameters equal to $\boldsymbol{\lambda}_{ij}$, i.e. $f(\boldsymbol{\lambda}_{ij}, t_{ijk})$, and v_{ijk} is the residual error structure.

The second stage is specified as:

$$p(\boldsymbol{\lambda}_{ij}|\boldsymbol{\theta}_i, \boldsymbol{\Phi}^{-1}) = N_p(\boldsymbol{\theta}_i, \boldsymbol{\Phi}) \quad \text{for } i = 1, \dots, K, \quad j = 1, \dots, m_i, \quad (4.9)$$

where $\boldsymbol{\theta}_i$ ($p \times 1$) represents the mean kinetic behavior of the i th individual and $\boldsymbol{\Phi}$ ($p \times p$) is corresponding variance-covariance matrix representing the IOV.

The third stage of the hierarchical model can be defined by making the following distributional assumptions:

$$p(\boldsymbol{\theta}_i|\boldsymbol{\mu}, \boldsymbol{\Omega}^{-1}) = N_p(\boldsymbol{\mu}, \boldsymbol{\Omega}) \quad \text{for } i = 1, \dots, K, \quad (4.10)$$

where $\boldsymbol{\mu}$ ($p \times 1$) is the mean value of the individual mean parameter vector $\boldsymbol{\theta}_i$ and $\boldsymbol{\Omega}$ ($p \times p$) is the corresponding variance-covariance matrix representing the ISV.

The definition of the hierarchical model is completed by the specification of the fourth stage, in which prior densities are assigned to the parameters τ , $\boldsymbol{\Phi}$, $\boldsymbol{\mu}$ and $\boldsymbol{\Omega}^{-1}$:

$$p(\tau) = G(\alpha, \beta), \quad (4.11a)$$

$$p(\boldsymbol{\Phi}^{-1}) = W_p(\mathbf{G}^{-1}, \gamma), \quad (4.11b)$$

$$p(\boldsymbol{\mu}) = N_p(\boldsymbol{\eta}, \mathbf{C}), \quad (4.11c)$$

$$p(\boldsymbol{\Omega}^{-1}) = W_p(\mathbf{R}^{-1}, \rho). \quad (4.11d)$$

4.3.2 Modeling the Inter-Study Variability

For the meta-analysis, another hierarchy was brought into the model, correcting for the inter-study variability (IS_tV):

Suppose that n_{ijk} plasma drug concentrations have been observed in study i ($i = 1, \dots, K$) for individual j ($j = 1, \dots, s_i$) on occasion k ($k = 1, \dots, m_{ij}$). Denote the l th observed concentration in study i for individual j on occasion k by y_{ijkl} , corresponding time by t_{ijkl} and the $p \times 1$ vector

of pharmacokinetic parameters of study i for individual j on occasion k by λ_{ijk} . The first stage of the model is specified as:

$$p(y_{ijkl}|\lambda_{ijk}, \tau) = N(f_{ijkl}, \tau^{-1}v_{ijkl}), \quad \text{for } i = 1, \dots, K, \\ j = 1, \dots, s_i, \quad k = 1, \dots, m_{ij}, \quad l = 1, \dots, p_{ijk} \quad (4.12)$$

where $y_{ijkl} \sim N(f_{ijkl}, \tau^{-1}v_{ijkl})$ given λ_{ijk} and τ , f_{ijkl} is the pharmacokinetic model evaluated at time t_{ijkl} with the individual PK parameters equal to λ_{ijk} , i.e. $f(\lambda_{ijk}, t_{ijkl})$, and v_{ijkl} is the residual error structure.

The second stage is specified as:

$$p(\lambda_{ijk}|\theta_{ij}, \Phi^{-1}) = N_p(\theta^{(ij)}, \Phi) \quad \text{for } i = 1, \dots, K, \\ j = 1, \dots, s_i, \quad k = 1, \dots, m_{ij}, \quad (4.13)$$

where θ_{ij} ($p \times 1$) represents the mean kinetic behavior in study i of individual j and Φ ($p \times p$) is the corresponding variance-covariance matrix representing the IOV.

The third stage of the hierarchical model can be defined by making the following distributional assumptions:

$$p(\theta_{ij}|\mu_i, \Omega^{-1}) = N_p(\mathbf{Z}_{ij}\mu_i, \Omega) \quad \text{for } i = 1, \dots, K, \quad j = 1, \dots, s_i, \quad (4.14)$$

where μ_i ($p \times K$) is the mean of the individual mean parameter vector θ_{ij} per study, Ω ($p \times p$) is the corresponding variance-covariance matrix representing the ISV and \mathbf{Z}_{ij} is a covariate-effect design matrix for individual j in study i . When no regression analyses is performed, \mathbf{Z}_{ij} can simply be left out the model.

The fourth stage of the hierarchical model is defined by making the following distributional assumptions:

$$p(\mu_i|\kappa, \Upsilon^{-1}) = N_p(\kappa, \Upsilon), \quad (4.15)$$

where κ ($p \times K$) represents the population mean pharmacokinetic behavior and Υ ($p \times p$) is the corresponding variance-covariance matrix representing the IS_tV.

The definition of the hierarchical model for the meta-analysis is completed by the specification of the fifth stage, were prior densities are assigned to the parameters τ , Φ , Ω , κ and Υ :

$$p(\Omega^{-1}) = W_p(\mathbf{R}^{-1}, \rho), \quad (4.16a)$$

$$p(\Phi^{-1}) = W_p(\mathbf{G}^{-1}, \gamma), \quad (4.16b)$$

$$p(\tau) = G(\alpha, \beta), \quad (4.16c)$$

$$p(\boldsymbol{\kappa}) = N_p(\boldsymbol{\eta}, \mathbf{C}), \quad (4.16d)$$

$$p(\boldsymbol{\Upsilon}^{-1}) = W_p(\mathbf{A}^{-1}, \nu), \quad (4.16e)$$

where $\boldsymbol{\eta}$ is the prior estimate of $\boldsymbol{\kappa}$ with variance-covariance matrix \mathbf{C} and $\boldsymbol{\Upsilon}^{-1} \sim W_p(\mathbf{A}^{-1}, \nu)$.

5 Analysis of the Individual Studies

5.1 Introduction

This section describes the results of the pharmacokinetic analysis of each individual study. The goal of these analyses is to explore whether the data of each study is sufficient to provide unique posterior estimates of the model parameters, i.e. the *identifiability* of the likelihood. When the model is overspecified and the likelihood is non-identifiable, strong prior information is indispensable in order to obtain reasonably precise posterior estimates of the model parameters and (some of) these will highly depend on the choice of the priors. Therefore, models with *non-informative*, *fully-informative* and *minimally-informative* priors are fit to the data of each study in order to gain knowledge about the identifiability of the likelihood. The strategies used to obtain non-, fully- and minimally-informative priors and corresponding prior estimates are discussed in Appendix B. In the analysis, emphasis is put on the estimation of the population pharmacokinetic parameters rather than explaining the observed variability by covariates. Subsequently, no regression is performed and the estimates of $\boldsymbol{\mu}$ play a large role in the discussion of the results.

Section 5.1.1 describes the model for estimating the PK parameters per study and the (data-based) priors. The results of the analyses per study are described in Sections 5.2 to 5.7.

5.1.1 The Model

The model that almost always best describes the kinetics of MPA is used to obtain more knowledge about the studies; a two-compartment model with first-order kinetics, lag-time and log MPA concentrations. This model was also used by Van Hest *et al.*. Six pharmacokinetic parameters are estimated using this model: the clearance from the central compartment CL , the inter-compartment clearance Q , the volume of distribution of the central compartment V_1 , the volume of distribution of the peripheral compartment V_2 , the lag-time T_{lag} and the absorption rate constant k_a . The two-compartment model is incorporated in the four-stage hierarchical model described in Section 4.3.1.

Informative priors were obtained from three pharmacokinetic MPA studies [49–51]. In these studies, MPA exposure was measured in renal transplant recipients after different doses of MMF on multiple occasions post-transplantation. Pharmacokinetic parameters were estimated based on a two-compartment model and the statistical analysis was performed using NONMEM. Because the study characteristics correspond to the ones in this study, they are assumed to be suitable as prior knowledge. The studies report estimates of the population pharmacokinetic parameters $\boldsymbol{\eta}$, their variances \boldsymbol{C} , variances for the inter-subject variability \boldsymbol{R} and variances for the

Table 2: Weighted means of η , C , R , G and $1/\sqrt{\tau}$ based on three pharmacokinetic MPA studies [49–51]. For the ISV and IOV, the estimates represent their variances.

Parameter	Weighted mean (var)	ISV	IOV
$\log(CL \text{ (1/h)})$	2.98 (0.6E-2)	0.11	0.02
$\log(Q \text{ (1/h)})$	3.43 (0.4E-2)	0.61	1.60
$\log(V_1 \text{ (l)})$	4.12 (0.01)	0.75	0.42
$\log(V_2 \text{ (l)})$	5.71 (0.01)	14.08	1.00
$\log(T_{lag} \text{ (h)})$	-1.44 (0.3E-3)	0.008	1.00
$\log(k_a \text{ (1/h)})$	1.39 (0.8E-2)	2.11	1.36
Residual error	0.43 (0.9E-4)	-	-

inter-occasion variability G . Weighted means of the relevant parameters were calculated based on the number of participants per study (ranging from 140 to 241 participants). Because WinBUGS estimates the logarithm of the PK parameters, weighted means of the prior variances were obtained using the delta method [52]. Unfortunately, no prior estimates of the inter-occasion variability of Q , V_2 and T_{lag} were found in literature. These prior variances were chosen to be relatively high (coefficients of variation of 100%), in order to reduce the prior influence on these parameters. Table 2 shows the weighted means of the prior estimates.

5.2 Study 1

5.2.1 Model with Non-Informative Prior Information

Based on two chains with over-relaxation, 40,000 iterations leaving out 10,000 burn-in iterations and using non-informative priors, almost no model parameter converged. Note that over-relaxation does not necessarily improve convergence, but it reduces the positive auto-correlation within the Markov chain which was observed for multiple model parameters. The number of iterations was increased to 500,000 but the results were similar. However, it is possible that the model parameters converge after an exorbitant number of iterations, which was not tried due to computational time and memory. Therefore, the below described results are based on 40,000 iterations. Graphical exploration of the posterior distribution in the direction of μ shows stationary chains with rapid mixing for CL and V_1 , see Figure 8 (Appendix C). Trace plots for the other population PK parameter appear 'snake-like', indicating sampling in a dependent matter, which is confirmed by the high autocorrelations for these parameters, the autocorrelations of lag 1 and 50 are 0.44-0.09, 0.88-0.85, 0.75-0.55, 0.98-0.88, 0.81-0.89 and 0.98-0.97 for CL , Q , V_1 , V_2 , T_{lag} and k_a respectively. BGR convergence diagnostics were consulted and provided evidence for non-convergence for

T_{lag} and k_a since the CSRF is higher than 1.2 for these parameters, and no evidence for CL , Q , V_1 and V_2 with CSRF scores of 1.02, 1.07, 1.02 and 1.06 respectively. Because the MC error of Q and V_2 is high and graphical exploration shows evidence for non-convergence, it is assumed that they did not converge. The same trend is observed in the posterior distribution of the individual PK parameters θ , where convergence is only attained for CL , and V_1 . High cross-correlation between the model parameters is associated with slow convergence, however, no high cross-correlation is observed between any of the model parameters. Table 3 shows the posterior summary measures of the model parameters that converged. A higher number of iterations or stronger prior information is necessary for these data to obtain reliable posterior summary measures, the model is probably overspecified and the likelihood is non-identifiable.

5.2.2 Model with Fully-Informative Prior Information

A model with fully-informative priors is fit to these data, see Appendix B.1. All model parameters converged based on both graphical and formal convergence diagnostics. Posterior summary measures are based on two chains with over-relaxation, 15,000 iterations leaving out 2,500 burn-in iterations, see Table 3, and trace plots of μ are shown in Figure 9 (Appendix C). With these, strong informative priors, reasonably precise posterior population PK parameter estimates are obtained, but they depend almost completely on the priors. The estimated posterior means of the logarithm of μ are 2.96, 3.43, 4.15, 5.70, -1.45 and 1.39 for CL , Q , V_1 , V_2 , T_{lag} and k_a respectively with prior point estimates of 2.98, 3.43, 4.12, 5.71, -1.44 and 1.39. With this strong prior information, the posterior PK parameter estimates have been minimally updated by the observed data, this again indicates that the likelihood in the μ subspace is non-identifiable.

5.2.3 Model with Minimally-Informative Prior Information

Finally, a model whereby the least informative prior was specified for each parameter is fit to these data, see Appendix B.1. The resulting minimally-informative priors were still highly informative and the obtained posterior summary measures based on 20,000 iterations leaving out 5,000 burn-in iterations with over-relaxation are shown in Table 3 and trace plots of μ are shown in Figure 10 (Appendix C). Some of the posterior population PK parameters estimates (2.52, 3.44, 4.22, 5.69, -1.47 and 1.37 respectively) have been highly updated by the observed data, however, the prior influence is still very large. The posterior estimates of the IOV and ISV also depend highly on the choice of the priors, which differ a lot between the model with fully-informative and minimally-informative priors. Additionally, the posterior standard deviation of the ISV and IOV is extremely high. Therefore,

no reliable posterior estimates are obtained for the ISV and IOV. It can be concluded that the used model is overspecified for the data of study 1. Only with strong prior information, precise posterior PK parameter estimates are obtained but these depend almost fully on the priors. Therefore, the data are not sufficient to obtain reliable posterior summary measures of all model parameters and the likelihood is non-identifiable. The identifiability problem probably arises because these data contain on average only two occasions per individual and six measurements per occasions, which seems insufficient for estimating the relatively high amount of parameters in this model.

5.3 Study 2

Models with non- fully- and minimally-informative priors were fit to the data of study 2, the posterior summary measures are reported in Table 4 and the trace plots of μ are shown in Figures 11, 12 and 13 (Appendix C). With non-informative priors almost no model parameter converged. Although the number of iterations was increased multiple times with similar results, it could be that the model still needs a higher number of iterations to reach convergence on all model parameters. With fully and minimally-informative priors, all parameters of μ and θ converged based on both graphical and formal diagnostics. However, some parameters of the ISV and IOV did not convergence. Convergence was only attained for these parameters when the degrees of freedom of the corresponding Wishart distributions were increased. When increasing the degrees of freedom of the Wishart distribution, the distribution becomes more informative which resulted in an even higher dependency of all model parameters on the prior information. Therefore, it was chosen to keep the Wishart distribution minimally-informative, see Appendix C. Using minimally-informative priors, the posterior means of μ are 3.11, 3.42, 4.24, 5.71, -1.45 and 1.34 for the logarithm of CL , Q , V_1 , V_2 , T_{lag} and k_a respectively. Most of them depend highly on the prior estimates of 2.98, 3.43, 4.12, 5.71, -1.44 and 1.39 respectively. Because the data of study 2 needs highly informative prior information to obtain precise posterior estimates of the model parameters, and these estimates depend highly on the chosen priors, the model is overspecified and the likelihood is therefore non-identifiable. The data of study 2 contain on average two occasions per individual and 7 measurements per occasion, this is probably not sufficient for estimating the relatively high amount of parameters in this model.

5.4 Study 3

The data of study 3 is the most abundant of all studies, with on average 7 measurements at 9 occasions and a total of 141 individuals. Subsequently, with non-informative priors, almost all model parameters converged. For

Table 3: Posterior summary measures of the model parameters of study 1 with non-informative, fully-informative and minimally-informative priors. Estimates were reported when convergence was attained. For the ISV and IOV, mean (sd) of the variance is given.

Parameter	mean (sd)	Model with non-informative prior				IOV	
		2.5%	median	97.5%	MC error		
$\log(CL)$ (1/h))	2.68 (0.28)	2.09	2.69	3.18	6.20E-3	0.40 (0.25)	1.1 (0.45)
$\log(V_1)$ (l))	4.36 (0.41)	3.54	4.36	5.16	1.81E-2	0.32 (0.34)	0.41 (0.45)
Residual error	0.55 (0.04)	0.47	0.55	0.64	1.05E-3	-	-
Parameter	mean (sd)	Model with fully-informative prior				IOV	
		2.5%	median	97.5%	MC error		
$\log(CL)$ (1/h))	2.96 (0.08)	2.81	2.96	3.11	2.89E-4	1.56 (0.86)	4.40 (1.62)
$\log(Q)$ (1/h))	3.43 (0.06)	3.31	3.43	3.55	4.15E-4	0.64 (0.53)	3.64 (2.16)
$\log(V_1)$ (l))	4.15 (0.12)	3.92	4.15	4.37	1.53E-3	0.45 (0.32)	1.44 (0.68)
$\log(V_2)$ (l))	5.71 (0.09)	5.53	5.71	5.89	8.95E-4	0.73 (0.50)	1.68 (0.91)
$\log(T_{lag})$ (h))	-1.45 (0.02)	-1.48	-1.45	-1.41	4.69E-5	6.61 (4.86)	11.45 (6.92)
$\log(k_a)$ (1/h))	1.39 (0.09)	1.22	1.39	1.56	4.74E-4	6.87 (6.21)	4.75 (4.01)
Residual error	0.46 (0.02)	0.42	0.46	0.50	4.95E-4	-	-
Parameter	mean (sd)	Model with minimally-informative prior				IOV	
		2.5%	median	97.5%	MC error		
$\log(CL)$ (1/h))	2.52 (0.35)	1.81	2.53	3.17	7.11E-3	0.57 (0.30)	1.77 (0.65)
$\log(Q)$ (1/h))	3.44 (0.14)	3.16	3.44	3.71	1.71E-3	0.27 (0.25)	4.34 (3.26)
$\log(V_1)$ (l))	4.22 (0.24)	3.75	4.22	4.68	5.94E-3	0.19 (0.17)	1.45 (0.73)
$\log(V_2)$ (l))	5.69 (0.20)	5.30	5.69	6.07	4.46E-3	0.75 (0.56)	0.74 (0.42)
$\log(T_{lag})$ (h))	-1.47 (0.12)	-1.72	-1.47	-1.23	5.11E-4	7.14 (5.35)	11.4 (5.98)
$\log(k_a)$ (1/h))	1.37 (0.20)	0.98	1.37	1.76	3.99E-3	3.63 (3.57)	0.53 (0.49)
Residual error	0.52 (0.04)	0.45	0.52	0.61	1.06E-3	-	-

Table 4: Posterior summary measures of the model parameters of study 2 with non-informative (based on two chains with 20,000 iterations leaving 5,000 burn-in iterations and over-relaxation), fully-informative (based on two chains with 40,000 iterations leaving 10,000 burn-in iterations and over-relaxation) and minimally-informative priors (based on two chains with 20,000 iterations leaving 5,000 burn-in iterations and over-relaxation). Estimates were reported when convergence was attained. For the ISV and IOV, mean (sd) of the variance is given.

Parameter	Model with non-informative prior information						IOV
	mean (sd)	2.5%	median	97.5%	MC error	ISV	
$\log(CL)$ (1/h))	3.16 (0.08)	2.99	3.16	3.33	2.90E-3	0.14 (0.05)	-
Residual error	0.72 (0.03)	0.66	0.72	0.78	1.83E-3	-	-
Parameter	Model with fully-informative prior information						IOV
	mean (sd)	2.5%	median	97.5%	MC error	ISV	
$\log(CL)$ (1/h))	2.99 (0.07)	2.85	2.99	3.12	4.87E-4	0.47 (0.12)	1.10 (0.18)
$\log(Q)$ (1/h))	3.41 (0.06)	3.29	3.41	3.53	8.15E-4	0.34 (0.17)	0.59 (0.22)
$\log(V_1)$ (l))	4.22 (0.10)	4.03	4.22	4.40	1.93E-3	0.39 (0.19)	0.31 (0.13)
$\log(V_2)$ (l))	5.82 (0.09)	5.64	5.82	5.96	2.27E-3	0.05 (0.04)	0.39 (0.16)
$\log(T_{lag})$ (h))	-1.45 (0.02)	-1.48	-1.45	-1.41	3.06E-5	10.75 (3.71)	6.24 (1.95)
$\log(k_a)$ (1/h))	1.33 (0.09)	1.16	1.33	1.51	2.13E-3	-	-
Residual error	0.60 (0.02)	0.57	0.60	0.64	4.20E-4	-	-
Parameter	Model with minimally-informative prior information						IOV
	mean (sd)	2.5%	median	97.5%	MC error	ISV	
$\log(CL)$ (1/h))	3.11 (0.10)	2.91	3.11	3.29	1.96E-3	0.22 (0.07)	0.52 (0.09)
$\log(Q)$ (1/h))	3.42 (0.06)	3.30	3.42	3.54	1.57E-3	-	0.12 (0.06)
$\log(V_1)$ (l))	4.24 (0.10)	4.05	4.24	4.42	2.50E-3	0.41 (0.22)	0.31 (0.13)
$\log(V_2)$ (l))	5.71 (0.08)	5.54	5.71	5.87	2.77E-3	-	0.37 (0.16)
$\log(T_{lag})$ (h))	-1.45 (0.04)	-1.52	-1.45	-1.37	1.33E-4	10.05 (3.37)	5.87 (1.91)
$\log(k_a)$ (1/h))	1.34 (0.09)	1.17	1.34	1.50	3.85E-3	-	-
Residual error	0.67 (0.02)	0.62	0.67	0.72	7.25E-4	-	-

the estimation of the individual and population PK parameters V_2 and k_a more information is necessary in order to obtain reliable, precise posterior estimates. This resulted in the final, minimally-informative model, where prior information was non-informative for all parameters except for V_2 and k_a . Trace plots of $\boldsymbol{\mu}$ for models with non- fully- and minimally-informative priors are shown in Figures 14, 15 and 16 (Appendix C) and the posterior summary measures of the parameters that converged are reported in Table 5. Because the data of study 3 is abundant, the model is probably not overspecified and the likelihood is expected to be identifiable. However, the estimation of the absorption rate constant k_a is hard, because only a few measurements are taken which make up the absorption process, i.e. measurements before the peak concentration. The volume of distribution of the second compartment V_2 is estimated from the last part of the concentration time curve (when the declining curve changes its rate, see Figure 5b). Although enough measurements make up this part of the curve (on average 4), the variability (ISV and IOV) on these time points is very high resulting in hard to model patterns, see Figures 3e and 3f. This could explain the need for more information in estimating the individual and population PK parameters V_2 and k_a .

5.5 Study 4

Trace plots of $\boldsymbol{\mu}$ for models with non- fully- and minimally-informative priors are shown in Figures 17, 18 and 19 (Appendix C) and the posterior summary measures of the parameters that converged are reported in Table 6. In the analyses with non-informative priors, the uncertainty on all parameters of the model except CL and T_{lag} is huge and these cannot be estimated with non-informative priors, which is probably due to identifiability problems. In the final model, with minimally-informative priors, the posterior means of $\boldsymbol{\mu}$ are 4.19, 3.43, 5.53, 6.00, -2.12 and 0.70 for the logarithm of CL , Q , V_1 , V_2 , T_{lag} and k_a respectively. When comparing these with the prior point estimates (2.98, 3.43, 4.12, 5.71, -1.44 and 1.39 respectively), it is obvious that they have been highly updated by the observed data, except for Q which is the same as its prior estimate. Although strong prior information is necessary, the posterior summary measures do not depend that much on the prior. Therefore, the problems in estimating the model parameters with non-informative priors are probably not a result of a non-identifiable likelihood. Other possible reasons are explained in Section 5.8.

5.6 Study 5

Trace plots of $\boldsymbol{\mu}$ for models with non- fully- and minimally-informative priors are shown in Figures 20, 21 and 22 (Appendix C) and the posterior summary measures of the parameters that converged are reported in Table

Table 5: Posterior summary measures of the model parameters of study 3 with non-informative (based on two chains with 10,000 iterations leaving 2,500 burn-in iterations), fully-informative (based on two chains with 20,000 iterations leaving 7,500 burn-in iterations) and minimally-informative priors (based on two chains with 20,000 iterations leaving 7,500 burn-in iterations). Estimates were reported when convergence was attained. For the ISV and IOV, mean (sd) of the variance is given.

Model with non-informative prior information						
Parameter	mean (sd)	2.5%	median	97.5%	MC error	IOV
$\log(CL)$ (1/h))	3.15 (0.04)	3.08	3.15	3.22	6.95E-4	0.14 (0.02) 0.10 (6.69E-3)
$\log(Q)$ (1/h))	3.55 (0.05)	3.45	3.55	3.65	1.94E-3	0.26 (0.05) 0.18 (0.03)
$\log(V_2)$ (1/h))	5.58 (0.07)	5.44	5.58	5.74	4.79E-3	- -
$\log(T_{lag})$ (1/h))	-1.45 (0.06)	-1.56	-1.45	-1.34	1.90E-3	0.33 (0.05) 0.23 (0.02)
Residual error	0.40 (5.35E-3)	0.39	0.40	0.41	2.18E-4	- -
Model with fully-informative prior information						
Parameter	mean (sd)	2.5%	median	97.5%	MC error	IOV
$\log(CL)$ (1/h))	3.07 (0.04)	2.99	3.07	3.14	8.35E-4	0.20 (0.03) 0.17 (9.11E-3)
$\log(Q)$ (1/h))	3.43 (0.04)	3.35	3.43	3.51	1.27E-3	0.25 (0.04) 0.19 (0.03)
$\log(V_1)$ (1))	3.89 (0.06)	3.76	3.89	4.02	2.40E-3	0.49 (0.08) 0.26 (0.03)
$\log(V_2)$ (1))	5.78 (0.06)	5.65	5.78	5.90	3.12E-3	- 1.13 (0.10)
$\log(T_{lag})$ (h))	-1.44 (0.02)	-1.48	-1.44	-1.41	1.18E-4	1.14 (0.15) 0.21 (0.02)
$\log(k_a)$ (1/h))	0.87 (0.09)	0.71	0.87	1.07	4.03E-3	- 1.82 (0.30)
Residual error	0.40 (5.37E-3)	0.39	0.40	0.41	1.64E-4	- -
Model with minimally-informative prior information						
Parameter	mean (sd)	2.5%	median	97.5%	MC error	IOV
$\log(CL)$ (1/h))	3.10 (0.04)	3.03	3.11	3.18	1.22E-3	0.14 (0.02) 0.09 (5.83E-3)
$\log(Q)$ (1/h))	3.49 (0.06)	3.38	3.48	3.60	1.97E-3	0.24 (0.04) 0.19 (0.03)
$\log(V_1)$ (1))	3.79 (0.08)	3.64	3.79	3.95	3.24E-3	0.47 (0.08) 0.27 (0.03)
$\log(V_2)$ (1))	5.73 (0.07)	5.59	5.73	5.87	4.31E-3	- 1.04 (0.09)
$\log(T_{lag})$ (h))	-1.40 (0.06)	-1.51	-1.40	-1.30	1.06E-3	0.32 (0.04) 0.22 (0.02)
$\log(k_a)$ (1/h))	0.44 (0.11)	0.23	0.44	0.66	5.21E-3	- 1.93 (0.30)
Residual error	0.40 (5.46E-3)	0.39	0.40	0.41	1.76E-4	- -

Table 6: Posterior summary measures of the model parameters of study 4 with non-informative (based on two chains with 20,000 iterations leaving 5,000 burn-in iterations and over-relaxation), fully-informative (based on two chains with 35,000 iterations leaving 15,000 burn-in iterations and over-relaxation) and minimally-informative priors (based on two chains with 50,000 iterations leaving 30,000 burn-in iterations and over-relaxation). Estimates were reported when convergence was attained. For the ISV and IOV, mean (sd) of the variance is given.

Parameter	Model with non-informative prior information						IOV
	mean (sd)	2.5%	median	97.5%	MC error	ISV	
$\log(CL)$ (1/h))	2.92 (0.07)	2.78	2.92	3.06	9.18E-4	0.16 (0.05)	0.20 (0.03)
$\log(T_{lag})$ (1/h))	-1.46 (0.20)	-1.86	-1.46	-1.08	6.03E-3	1.19 (0.34)	0.42 (0.12)
Residual error	0.43 (0.01)	0.41	0.43	0.45	2.67E-4	-	-
Parameter	Model with fully-informative prior information						IOV
	mean (sd)	2.5%	median	97.5%	MC error	ISV	
$\log(CL)$ (1/h))	4.13 (0.09)	3.96	4.13	4.29	1.82E-3	0.27 (0.07)	0.15 (0.02)
$\log(Q)$ (1/h))	3.43 (0.06)	3.30	3.43	3.55	6.80E-4	1.59 (0.68)	0.74 (0.27)
$\log(V_1)$ (1))	5.53 (0.09)	5.36	5.53	5.69	2.53E-3	0.16 (0.05)	0.17 (0.04)
$\log(V_2)$ (1))	6.03 (0.28)	5.49	6.03	6.58	6.77E-3	4.37 (3.03)	4.17 (1.97)
$\log(T_{lag})$ (h))	-2.13 (0.29)	-2.73	-2.12	-1.59	7.14E-3	1.63 (0.72)	2.72 (0.70)
$\log(k_a)$ (1/h))	0.76 (0.26)	0.26	0.76	1.27	7.99E-3	-	4.97 (2.21)
Residual error	0.48 (0.01)	0.46	0.48	0.50	1.50E-4	-	-
Parameter	Model with minimally-informative prior information						IOV
	mean (sd)	2.5%	median	97.5%	MC error	ISV	
$\log(CL)$ (1/h))	4.19 (0.07)	4.05	4.19	4.33	1.74E-4	0.20 (0.05)	0.15 (0.02)
$\log(Q)$ (1/h))	3.43 (0.06)	3.31	3.43	3.56	8.62E-4	1.18 (0.51)	0.74 (0.26)
$\log(V_1)$ (1))	5.53 (0.07)	5.38	5.53	5.67	2.70E-3	0.08 (0.04)	0.16 (0.04)
$\log(V_2)$ (1))	6.00(0.28)	5.45	5.99	6.55	7.12E-3	7.27 (3.78)	3.02 (1.70)
$\log(T_{lag})$ (h))	-2.12 (0.30)	-2.72	-2.11	-1.56	8.10E-3	1.73 (0.75)	2.71 (0.68)
$\log(k_a)$ (1/h))	0.70 (0.25)	0.22	0.70	1.21	8.76E-3	-	5.46 (2.51)
Residual error	0.48 (0.01)	0.46	0.48	0.51	1.88E-4	-	-

7. The data of study 5 is not sufficient to obtain accurate posterior summary measures with non-informative priors. Again this model is probably overspecified and in order to obtain accurate posterior summary measures, strong prior information is necessary. With minimally-informative priors, precise posterior summary statistics on all model parameters were obtained. The posterior means of the logarithm of $\boldsymbol{\mu}$ are 4.28, 3.28, 5.30, 5.80, -2.11 and 0.09 for CL , Q , V_1 , V_2 , T_{lag} and k_a respectively. All estimates have been highly updated by the observed data despite of the strong informative priors. Additionally, the data of study 5 contain high numbers of measurements (mean of 11) on 118 individuals with on average three occasions per individual, therefore, the likelihood is probably identifiable. Other reasons that could lead to the problems in the model with non-informative priors are explained in Section 5.8.

5.7 Study 6

The data of study 6 does not contain measurements of individuals on different occasions, therefore the models for study 6 exclude Φ , see Section 4.3. Trace plots of $\boldsymbol{\mu}$ of models with non- fully- and minimally-informative priors are shown in Figures 23, 24 and 25 (Appendix C) and the posterior summary measures are reported in Table 8. The data of study 6 need strong prior information in order to obtain precise posterior summary measures. With strong prior influence, most of the model parameters depend highly on the prior information. Therefore, it can be concluded that these models are probably overspecified and the likelihood is non-identifiable.

Table 7: Posterior summary measures of the model parameters of study 5 with non-informative (based on two chains with 20,000 iterations leaving 5,000 burn-iterations and over-relaxation), fully-informative (based on two chains with 10,000 iterations leaving 2,500 burn-iterations and over-relaxation) and minimally-informative priors (based on two chains with 30,000 iterations leaving 20,000 burn-iterations and over-relaxation). Estimates were reported when convergence was attained. For the ISV and IOV, mean (sd) of the variance is given.

Parameter	Model with non-informative prior information					
	mean (sd)	2.5% median	97.5% median	MC error	ISV	IOV
$\log(CL \text{ (1/h)})$	4.27 (0.04)	4.20	4.27	4.34	6.82E-4	0.08 (0.02) 0.19 (0.02)
Residual error	0.43 (6.84E-3)	0.42	0.43	0.45	1.83E-4	-
Parameter	Model with fully-informative prior information					
	mean (sd)	2.5% median	97.5% median	MC error	ISV	IOV
$\log(CL \text{ (1/h)})$	4.26 (0.05)	4.17	4.26	4.36	7.84E-4	0.18 (0.03) 0.16 (0.02)
$\log(Q \text{ (1/h)})$	3.22 (0.15)	2.93	3.22	3.51	6.17E-3	1.06 (0.28) 1.79 (0.31)
$\log(V_1 \text{ (1)})$	5.30 (0.07)	5.16	5.30	5.44	2.34E-3	0.09 (0.03) 1.04 (0.09)
$\log(V_2 \text{ (1)})$	5.81 (0.09)	5.63	5.81	5.98	1.70E-3	2.11 (1.14) 3.81 (1.31)
$\log(T_{lag} \text{ (h)})$	-2.12 (0.20)	-2.53	-2.12	-1.74	7.56E-3	1.15 (0.38) 4.70 (0.68)
$\log(k_a \text{ (1/h)})$	0.09 (0.21)	-0.33	0.09	0.52	9.61E-3	2.01 (0.70) 3.96 (0.83)
Residual error	0.43 (6.4E-3)	0.41	0.43	0.44	1.9E-4	-
Parameter	Model with minimally-informative prior information					
	mean (sd)	2.5% median	97.5% median	MC error	ISV	IOV
$\log(CL \text{ (1/h)})$	4.28 (0.05)	4.18	4.28	4.37	7.15E-4	0.18 (0.03) 0.16 (0.02)
$\log(Q \text{ (1/h)})$	3.28 (0.13)	3.03	3.28	3.54	5.03E-3	0.36 (0.15) 1.73 (0.31)
$\log(V_1 \text{ (1)})$	5.30 (0.07)	5.17	5.30	5.44	1.65E-3	0.09 (0.03) 1.03 (0.09)
$\log(V_2 \text{ (1)})$	5.80 (0.09)	5.62	5.80	5.98	1.66E-3	3.03 (1.88) 2.55 (0.89)
$\log(T_{lag} \text{ (h)})$	-2.11 (0.19)	-2.50	-2.11	-1.75	5.51E-3	1.14 (0.36) 4.22 (0.57)
$\log(k_a \text{ (1/h)})$	0.09(0.21)	-0.30	0.09	0.52	7.53E-3	2.04 (0.77) 3.82 (0.80)
Residual error	0.43 (6.62E-3)	0.41	0.43	0.44	1.58E-4	-

Table 8: Posterior summary measures of the model parameters of study 6 with non-informative (based on two chains with 20,000 iterations leaving 5,000 burn-in iterations and over-relaxation), fully-informative (based on two chains with 40,000 iterations leaving 20,000 burn-in iterations and over-relaxation) and minimally-informative priors (based on two chains with 40,000 iterations leaving 20,000 burn-in iterations and over-relaxation). Estimates were reported when convergence was attained. For the ISV and IOV, mean (sd) of the variance is given.

Model with non-informative prior information						
Parameter	mean (sd)	2.5%	median	97.5%	MC error	ISV
$\log(CL)$ (1/h))	2.7 (0.04)	2.63	2.7	2.77	5.36E-4	0.10 (0.02)
$\log(T_{lag})$ (h))	-1.42 (0.14)	-1.72	-1.42	-1.16	3.99E-3	0.89 (0.22)
Residual error	0.37 (0.01)	0.35	0.37	0.40	2.66E-4	-
Model with fully-informative prior information						
Parameter	mean (sd)	2.5%	median	97.5%	MC error	ISV
$\log(CL)$ (1/h))	2.86 (0.05)	2.77	2.86	2.96	4.61E-4	0.22 (0.04)
$\log(Q)$ (1/h))	3.29 (0.06)	3.17	3.29	3.41	9.28E-4	0.60 (0.15)
$\log(V_1)$ (1))	3.90 (0.08)	3.74	3.90	4.06	1.50E-3	0.64 (0.18)
$\log(V_2)$ (1))	5.73 (0.07)	5.60	5.73	5.85	9.85E-4	0.37 (0.27)
$\log(T_{lag})$ (h))	-1.44 (0.02)	-1.48	-1.44	-1.41	4.00E-5	3.24 (0.78)
$\log(k_a)$ (1/h))	1.36 (0.06)	1.23	1.36	1.48	5.75E-4	-
Residual error	0.39 (0.01)	0.37	0.39	0.41	1.51E-4	-
Model with minimally-informative prior information						
Parameter	mean (sd)	2.5%	median	97.5%	MC error	ISV
$\log(CL)$ (1/h))	2.71 (0.04)	2.64	2.71	2.79	9.38E-4	0.07 (0.01)
$\log(Q)$ (1/h))	2.83 (0.09)	2.66	2.82	3.01	3.47E-3	0.20 (0.08)
$\log(V_1)$ (1))	3.46 (0.09)	3.28	3.46	3.65	3.68E-3	0.23 (0.10)
$\log(V_2)$ (1))	5.74 (0.06)	5.61	5.74	5.86	9.78E-4	0.61 (0.27)
$\log(T_{lag})$ (h))	-1.14 (0.11)	-1.36	-1.14	-0.93	2.91E-3	0.64 (0.16)
$\log(k_a)$ (1/h))	1.34 (0.06)	1.22	1.34	1.47	7.12E-4	-
Residual error	0.38 (0.01)	0.35	0.38	0.40	2.92E-4	-

5.8 Conclusion

In the analyses of almost every study, identifiability problems arose when using non-informative priors. These identifiability problems are evident in convergence issues of multiple model parameters. In general, it is hard to investigate the identifiability of the likelihood. Although it was tried to improve convergence by increasing the number of iterations and using over-relaxation, it is unknown if convergence is attained when the number of iterations were increased tremendously. This was not tried due to the impracticability of using such numbers of iterations, i.e. the computational time and memory becomes exorbitant. However, real-life pharmacokinetic data is almost always hard to model. Especially these data, which came from studies which initially did not intent to model the pharmacokinetics of MPA, see Section 2.2. Additionally, the majority of the six studies were multicenter studies (this is unknown for the unpublished studies 1 and 5), and the used models did not correct for this extra variability. In the analysis, individual and population PK parameters V_2 and k_a were particularly hard to converge. For the absorption rate constant k_a this could be explained by the relatively low available observations which make up the absorption process. In general only one measurement was taken between 0 and 12 hours after MMF administration, while the peak concentration appeared around 12 hours. The volume of distribution of the second compartment V_2 is estimated from the last part of the concentration time curve, when the declining curve changes rate. In this last part of the curve, particularly high variability is observed resulting in hard to model patterns, see Figures 2 and 3.

Bayesian methods seem a logical choice for this type of data because prior information can be used to obtain precise estimates of the model parameters despite the identifiability problems. Especially because lots of information about the pharmacokinetics of MPA in renal transplant recipients is available, i.e. historical studies which could serve as prior. Therefore, prior information was increased and estimates of all model parameters were obtained.

As stated before, the used studies had initially other purposes with the data and no publications were found where the pharmacokinetics of MPA were estimated with these data. Therefore, it is unknown if other methods were capable of estimating all model parameters. However, with non-Bayesian software, convergence issues are also common but are harder to notice and easier to look-over.

6 The Individual Participant Data Meta-Analysis

This section describes the analysis of the data of all studies combined, i.e. the meta-analyses. In traditional meta-analysis, only summary data of different studies are combined and analysed, which are often obtained from publications, e.g. modeling the differences in a specific treatment effect of several clinical trials. In this meta-analysis, the complete datasets of the different studies were pooled into one dataset and analyzed simultaneously while correcting for the variability between the studies, resulting in an *individual participant data (IPD) meta-analysis*. The IPD approach improves the quality of the data, the analyses and subsequently the reliability of the results. When combining all data, inference is based on 467 individuals with a median of 8.8 measurements at 3 occasions (Table 1). The used model (Section 4.3.2) corrects for three types of variability, i.e. the inter-subject, inter-occasion and inter-study variability. The inter-study variability (IS_tV) models the heterogeneity between the different studies which arises due to (small) differences in study-design, e.g. different sampling schemes and at different moments after transplantation. In this section we explore whether any of the measured covariates (Table 1) explain the inter-subject variability of MPA exposure. In order to select covariates which may explain this variability, initially, the population and individual pharmacokinetic parameters are estimated based on all data, see Section 6.1. Subsequently, the individual estimates of CL , Q , V_1 , V_2 , T_{lag} and k_a are plotted against the different covariates. When a trend is observed in these graphs, that covariate is introduced in the model and regressed against the PK parameter for which a trend is observed, see Section 6.2. Additionally, this section elaborates about the differences between the used Bayesian methods and the traditional methods, i.e. NONMEM, used by Van Hest *et al.* to analyse these data. In Section 6.3, all results, pitfalls and achievements are summarized and discussed.

6.1 Estimating the Pharmacokinetic Parameters

The model that best describes the pharmacokinetics of MPA and the data is a two-compartment model with first-order kinetics, lag-time and log-MPA concentrations, see Section 5.1.1. This model was incorporated in the five-stage hierarchical model described in Section 4.3.2. The used prior information was based on three historical studies, see Section 5.1.1 and Table 2 [49–51]. Section 6.1.1 further discusses the used prior information and Section 6.1.2 describes the obtained results from the IPD meta-analyses without regressing the individual PK parameter estimates against the covariates.

6.1.1 Prior Information

All prior information was kept close to non-informative, following the guidelines of Lunn *et al.* for obtaining non-informative prior information, see Appendix B. Where Lunn *et al.* [45,46] advises to multiply the prior estimates of the diagonals of \mathbf{R} and \mathbf{G} by the number of diagonals (here 6), in order to obtain non-informative priors. However, multiplying these estimates by 6 resulted in several convergence issues with μ_i and κ . This may indicate that these priors are not completely non-informative or that these priors give too dispersed sampled values which cause the convergence issues. Nevertheless, these priors were optimized in order to attain convergence after a relatively low number of iterations (due to the incredible high computational time with these models), resulting in diagonals of \mathbf{R}^{-1} of [1.46, 2.74, 2.23, 11.83, 19.97, 7.9] and \mathbf{G}^{-1} of [10.83, 12.53, 23.69, 10.00, 1.00, 3.23]. Sampling from these distributions results in relatively high 95% quantiles for the variances of the ISV and IOV indicating that these prior estimates are not non-informative, e.g. the 95% quantiles of the (prior) ISV of CL are [1.77, 21.12]. Because the posterior estimates of the model parameters were not influenced by the changes in the numbers of the diagonals and using non-informative prior information for the ISV and IOV resulted in convergence issues, it was chosen to use these prior estimates.

6.1.2 The Results

Posterior summary measures were obtained based on two chains with 25,000 iterations, leaving out 17,500 burn-in iterations. Based on these iterations, the ESS and MC error were reasonable for all parameters except for k_a , e.g. population PK parameter estimate k_a has a ESS of 20.03 and MC error of 1.50E-2. Increasing the number of iterations or decreasing the burn-in part did not substantially increase the accuracy of k_a . Possible reasons for the issues with respect to parameter k_a are explained below.

Convergence was assessed using graphical and formal (BGR and HW) diagnostics. All individual PK parameters θ attained convergence (based on BGR and HW), all trace plots appear as a horizontal strip with both chains exhibit rapid mixing, Figure 26 (Appendix D) shows trace plots of each PK parameter from a randomly selected individual. Convergence was attained for almost all study PK parameter estimates μ , trace plots of each PK parameter from a randomly selected study is shown in Figure 27 (Appendix D). Trace plots of k_a (for all studies) appear 'snake-like', indicating sampling in a dependent matter, however, based on BGR and HW, convergence was attained for all studies except for study 4. Trace plots for the other PK parameters appear as a horizontal strip and all chains exhibit rapid mixing. This was also the case for population PK parameters κ , see Figure 28 (Appendix D). No reliable and precise posterior summary mea-

tures were obtained for the ISV and IOV of k_a , and for the IS_tV of V_2 and k_a . Further increasing prior information for the ISV and IOV, i.e. increasing the degrees of freedom of corresponding Wishart distributions, may result in convergence for these parameters. However, this would also influence the variances of the other PK parameters from which reliable estimates were obtained. Therefore, it was chosen not to increase prior information in order to maintain the results which are more based on the observed data. Note that the prior information for the IS_tV could not be increased as no prior information is available for this parameter.

As explained before, the estimation of k_a is hard because only a few measurements make up for the absorption process, i.e. the measurements before the peak concentration. Although the data is very abundant, it is not surprisingly that no reliable estimates were obtained on some of the model parameters. As with non-Bayesian methods, e.g. the results of Van Hest *et al.*, also no reliable results were obtained on some of the model parameters (the ISV and IOV of V_2 and T_{lag}), see Section 6.2. Subsequently, the estimate of k_a in the analysis of Van Hest *et al.* is not precise, as the standard deviation is higher than the mean (Section 6.2).

Posterior summary measures of κ , Ω , Φ , Υ and μ are reported in Table 9. The population mean of the MPA clearance from the central compartment CL is 19.49 l/h, with a wide 95% credible interval of [16.28, 23.10] l/h, reflecting high uncertainty about the population mean. This is also the case for the other population PK parameters, i.e. mean of 27.39 l/h with CI [18.73, 40.04] l/h for the inter-compartmental clearance of MPA Q , 35.16 l with CI [29.67, 41.68] l for the central volume of distribution V_1 , 441.42 l with CI [270.43, 780.55] l for the peripheral volume of distribution V_2 , 0.34 h with CI [0.22, 0.53] h for the absorption lag time T_{lag} and 0.44 l/h with CI [0.27, 0.68] l/h for the absorption rate constant k_a . This uncertainty arises due to the high variabilities between and within the studies. Reasonable precise posterior summary estimates of the ISV and IOV were obtained (based on the sd and the, not reported, MC error), but their CIs are wide (also not reported). The estimates of the IS_tV appear neither precise nor reliable, as both the sd and MC error (not reported) are high. The bottom of Table 9 reports the PK parameter estimates of μ , were it becomes evident that high variability is observed between the studies, as the PK parameter estimates vary widely between the studies.

In the next section, it is tried to find covariates responsible for the inter-subject variability. With some of the observed variability explained, more precise estimates of the model parameters may be obtained.

Table 9: Posterior summary measures of the IPD meta-analyses. Estimates were reported when convergence was attained. Top: population PK parameters κ , Ω , Φ and Υ . For the ISV, IOV and IS_tV , mean (sd) of the variance is given. Bottom: PK parameter estimates per study μ_i .

Parameter	mean (sd)	2.5%	median	97.5%	MC error	ISV	IOV	IS_tV
$\log(CL$ (1/h))	2.97 (0.09)	2.79	2.97	3.14	1.18E-3	0.10 (0.01)	0.15 (7.12E-3)	0.04 (0.04)
$\log(Q$ (1/h))	3.31 (0.19)	2.93	3.30	3.69	3.04E-3	0.21 (0.03)	0.26 (0.02)	0.21 (0.22)
$\log(V_1$ (l))	3.56 (0.09)	3.39	3.56	3.73	2.35E-3	0.28 (0.04)	0.33 (0.02)	0.03 (0.04)
$\log(V_2$ (l))	6.09 (0.26)	5.60	6.08	6.66	7.91E-3	0.55 (0.09)	0.47 (0.06)	-
$\log(T_{lag}$ (h))	-1.08 (0.22)	-1.51	-1.09	-0.63	3.56E-3	0.41 (0.05)	0.37 (0.03)	0.27 (0.28)
$\log(k_a$ (1/h))	-0.82 (0.24)	-1.30	-0.81	-0.38	1.50E-2	-	-	-
Residual error	0.40 (3.51E-3)	0.40	0.40	0.41	1.07E-4	-	-	-

Parameter	Study 1		Study 2		Study 3		Study 4		Study 5		Study 6	
	mean (sd)	mean (sd)	mean (sd)	mean (sd)	mean (sd)	mean (sd)	mean (sd)	mean (sd)	mean (sd)	mean (sd)	mean (sd)	mean (sd)
$\log(CL$ (1/h))	2.96 (0.10)	3.13 (0.06)	3.22 (0.03)	2.87 (0.05)	3.22 (0.03)	2.87 (0.05)	2.90 (0.03)	2.77 (0.05)	2.90 (0.03)	2.90 (0.03)	2.77 (0.05)	2.77 (0.05)
$\log(Q$ (1/h))	3.76 (0.20)	3.68 (0.10)	3.48 (0.05)	3.13 (0.09)	3.48 (0.05)	3.13 (0.09)	2.95 (0.06)	2.84 (0.10)	2.95 (0.06)	2.95 (0.06)	2.84 (0.10)	2.84 (0.10)
$\log(V_1$ (l))	3.68 (0.15)	3.65 (0.10)	3.68 (0.06)	3.53 (0.08)	3.68 (0.06)	3.53 (0.08)	3.47 (0.06)	3.38 (0.09)	3.47 (0.06)	3.47 (0.06)	3.38 (0.09)	3.38 (0.09)
$\log(V_2$ (l))	5.50 (0.24)	5.64 (0.14)	5.83 (0.09)	6.38 (0.22)	5.83 (0.09)	6.38 (0.22)	6.54 (0.20)	6.66 (0.28)	6.54 (0.20)	6.54 (0.20)	6.66 (0.28)	6.66 (0.28)
$\log(T_{lag}$ (h))	-0.79 (0.30)	-0.33 (0.15)	-1.45 (0.06)	-1.31 (0.12)	-1.45 (0.06)	-1.31 (0.12)	-1.20 (0.08)	-1.40 (0.12)	-1.20 (0.08)	-1.20 (0.08)	-1.40 (0.12)	-1.40 (0.12)
$\log(k_a$ (1/h))	-0.74 (0.39)	-0.84 (0.35)	-0.72 (0.24)	-	-0.72 (0.24)	-	-0.88 (0.26)	-0.88 (0.32)	-0.88 (0.26)	-0.88 (0.26)	-0.88 (0.32)	-0.88 (0.32)

6.2 The Regression Analyses

6.2.1 The Covariates

Although the complete dataset covered 24 covariates, only 12 covariates were considered for the regression analysis. These 12 include all the covariates without missing values on the first measurement. When including covariates with missing values, multivariate imputations should be included in the model, which is outside the scope of this thesis. Additionally, only the first (baseline) measurement was included in the model instead of time-varying measurements, again because this is outside the scope of this thesis. The 12 covariates that were considered for the regression include race, gender, age, weight, height, hemoglobin, creatinine clearance, use of antacids, use of proton pump inhibitors, use of anti-viral agents, MMF dose and diabetes mellitus. As the relevance for the covariates race, gender, age, weight, height and MMF dose are evident, the biological relevance of the others are discussed below.

Hemoglobin: a protein that contains iron and transports oxygen in red blood cells. Anemia, i.e. a decrease in the number of red blood cells or less than the normal quantity of hemoglobin in the blood, is common after renal transplantation [53].

Creatinine clearance: the volume of blood plasma that is cleared of creatinine (by the kidneys) per unit time. It is a measure of the state of the kidney.

Antacids: neutralizes stomach acidity, treats gastrointestinal complications which are common in renal transplant recipients, e.g. gastrointestinal bleeding or gastroduodenal ulcerations [54].

Proton pump inhibitors (PPI): reduces gastric acid production, treats gastrointestinal complications.

Anti-viral agents: the immune system of renal transplant recipients is highly suppressed to decrease the risk of rejection and viral infections are a significant cause of mortality [55].

Diabetes Mellitus: is a major complication after renal transplantation [56].

Above described covariates were plotted against the individual PK parameter estimates obtained from the analysis described in Section 6.1, see Appendix D. The covariates in graphs with a red asterisk in the right corner were introduced in the first, full regression model.

6.2.2 Prior Information

The same prior information as for the analysis in Section 6.1 was used for κ , Ω , Φ , Υ and τ . The regression coefficients were assigned independent normal priors with mean 0 and variance 10.

6.2.3 Model Selection

After inspection of the covariate regression graphs in Appendix D, 20 covariates were selected and introduced in the first, full model (the ones with a red asterisk in the right corner of the graphs). As most of these showed no effect, i.e. 0 appeared in the CI, the model was gradually reduced by eliminating these covariates from the model. Additionally, the DIC for each model was consulted and decreased each time with at least 5. In this way, the final model was obtained with the covariates creatinine clearance regressed against Q , creatinine clearance and the use of PPI against V_1 and the use of antacids against T_{lag} . Where the DIC of the final model was 17,040, the full model 17,308 and the model without regression (Section 6.1) 17,331. Based on the DIC, the final model provided the best fit.

6.2.4 The Results

Posterior summary measures of the final regression model were obtained based on two chains with 50,000 iterations, leaving out 32,500 burn-in iterations. With this number of iterations, the ESS and MC error is reasonable for all model parameters, but again except k_a , e.g. population PK parameter estimate of k_a has an ESS of 24.99 and MC error of 1.68E-2. Increasing the number of iterations or decreasing the burn-in part resulted in several computer memory issues, which is probably due to the high amount of parameters in the model.

This model took into account the effect of creatinine clearance on Q (coef 1), creatinine clearance (coef 2) and the use of proton pump inhibitors (coef 3) on V_1 and the use of antacids (coef 4) on T_{lag} . Convergence was assessed using graphical and formal (BGR and HW) diagnostics. All individual PK parameters θ reached convergence as well as the three regression coefficients. Convergence issues arose for the study PK parameter estimates μ , where the estimates of k_a from study 1, 2 and 4 did not attain convergence, trace plots are shown in Figure 29. The population PK parameters κ all reached convergence, trace plots shown in Figure 30. But, the ISV for k_a and the IS_tV for V_2 and k_a did not attain convergence. Posterior summary measures of κ , Ω , Φ , Υ , μ and the regression coefficients (coef 1, 2, 3 and 4) are reported in Table 10.

The values of the coefficients reflect the effect of the covariate on the logarithm of the PK parameter, therefore, the effect of creatinine clearance on Q and V_1 seem small. However, both estimates are very precise and it appears that if the creatinine clearance (ml/min) increases with 1 ml/min, it results in a mean decrease of Q with 1,00 l/h and of V_1 with 1,00 l with CI's of [-1.00, -1.00] l/h and [-1.01, -1.00] l respectively. Using proton pump inhibitors results in a mean increase of V_1 by 1.45 l with CI [1.11, 1.92] l. The use of antacids results in a mean decrease of T_{lag} by 1.67 h with CI

[-2.44, -1.13] h.

Compared to the analysis in Section 6.1, the mean population and study PK parameters have changed, especially Q , V_1 and T_{lag} . However, although more variability is explained in this model, the CI of the estimates are still wide. A lot of variability remains unexplained. The estimates of the ISV, IOV and IS_tV are similar to estimates obtained by the analysis in Section 6.1. Although this model provided a better fit than the model without the regression (with DIC of 17,040 and 17,330.5 respectively), this is not reflected in the accuracy of the posterior estimates of the model parameters.

Table 11 reports, again, the posterior estimates of κ , Ω and Φ together with the estimates obtained by van Hest *et al.*. As stated before, these models differ a lot as we did not include covariates with missing values or time-depending covariates but did account for inter-study variability in contrast to van Hest *et al.*. Therefore, we both found different covariates which explain some of the observed variability, see Table 11. Subsequently, the estimates reported in Table 11 differ a lot, even V_2 and k_a which were not regressed against any covariates in both models. These differences reflect the high uncertainty about the true values of the parameters. With Bayesian analysis, more precise estimates are obtained, i.e. the sd is lower for all parameters. However, it would be wrong to conclude that this is a results of the used Bayesian methods because of the differences in both models. The estimates of the IS_tV are very high, and accounting for this variability may result in more precise estimates with non-Bayesian methods. Additionally, except for V_1 and k_a , the estimates of Van Hest *et al.* all lie within the CI of our estimates. Therefore, it is meaningless to conclude anything about the differences in the results of these models.

6.3 Conclusion

The IPD meta-analysis improved the quality of the analysis and the reliability of the results. The obtained results were almost completely based on the observed data, i.e. close to non-informative prior information was used. Some convergence issues arose with the estimation of model parameter k_a , which is a direct result of the observed data, where only a few measurements make up the absorption process. Convergence issues in the variance parameters (ISV, IOV and IS_tV) is probably a result of the very high and hard to model observed variances.

Only 3 of the 12 covariates explained some of the inter-subject variability. However, the regression performed in Section 6.2 has some limitation as no covariates with missing values or time-dependent covariates were considered. Subsequently, it is hard to compare these results with the results obtained by Van Hest *et al.* as they did use both types of covariates in their model. Therefore, we both found different covariates which explain some of the observed variability. Another striking difference between our models is that

Table 10: Posterior summary measures of the IPD meta-analyses with regression. Estimates were reported when convergence was attained. Top: population PK parameters κ , Ω , Φ and Υ . For the ISV, IOV and IS_tV , mean (sd) of the variance is given. Bottom: PK parameter estimates per study μ_i .

Parameter	mean (sd)	2.5%	median	97.5%	MC error	ISV	IOV	IS_tV
$\log(CL (1/h))$	2.98 (0.09)	2.79	2.98	3.16	8.28E-4	0.10 (0.01)	0.14 (7.37E-3)	0.05 (0.05)
$\log(Q (1/h))$	3.40 (0.18)	3.04	3.39	3.76	2.55E-3	0.21 (0.03)	0.26 (0.02)	0.18 (0.18)
$\log(V_1 (l))$	3.67 (0.10)	3.48	3.67	3.87	2.55E-3	0.27 (0.04)	0.33 (0.02)	0.03 (0.05)
$\log(V_2 (l))$	6.07 (0.24)	5.60	6.06	6.57	4.66E-3	0.54 (0.08)	0.44 (0.07)	-
$\log(T_{lag} (h))$	-0.95 (0.32)	-1.56	-0.96	-0.29	3.83E-3	0.42 (0.05)	0.38 (0.04)	0.55 (0.59)
$\log(k_a (1/h))$	-0.91 (0.32)	-1.51	-0.92	-0.27	1.68E-2	0.89 (0.31)	0.36 (0.12)	-
Residual error	0.40 (3.97E-3)	0.40	0.41	0.41	1.40E-4	-	-	-
Coef 1	-2.42E-3 (1.13E-3)	-4.66E-3	-2.40E-3	-2.90E-4	3.77E-5	-	-	-
Coef 2	-2.94E-3 (1.06E-3)	-5.05E-3	-2.93E-3	-9.25E-4	3.49E-5	-	-	-
Coef 3	0.37 (0.14)	0.10	0.37	0.65	3.73E-3	-	-	-
Coef 4	-0.51 (0.20)	-0.89	-0.51	-0.12	8.60E-3	-	-	-

Parameter	Study 1		Study 2		Study 3		Study 4		Study 5		Study 6	
	mean (sd)	mean (sd)	mean (sd)	mean (sd)	mean (sd)	mean (sd)	mean (sd)	mean (sd)	mean (sd)	mean (sd)	mean (sd)	
$\log(CL (1/h))$	2.95 (0.10)	3.18 (0.07)	3.22 (0.03)	3.22 (0.03)	2.87 (0.05)	2.90 (0.04)	2.76 (0.05)					
$\log(Q (1/h))$	3.83 (0.19)	3.73 (0.11)	3.56 (0.06)	3.56 (0.06)	3.22 (0.10)	3.03 (0.07)	2.99 (0.12)					
$\log(V_1 (l))$	3.79 (0.16)	3.80 (0.13)	3.75 (0.07)	3.75 (0.07)	3.62 (0.09)	3.55 (0.08)	3.50 (0.11)					
$\log(V_2 (l))$	5.48 (0.25)	5.70 (0.16)	5.82 (0.09)	5.82 (0.09)	6.33 (0.19)	6.54 (0.19)	6.55 (0.24)					
$\log(T_{lag} (h))$	-0.60 (0.34)	0.17 (0.30)	-1.44 (0.06)	-1.44 (0.06)	-1.28 (0.12)	-1.16 (0.08)	-1.41 (0.12)					
$\log(k_a (1/h))$	-0.81 (0.45)	-0.80 (0.54)	-0.88 (0.29)	-0.88 (0.29)	-	-1.00 (0.32)	-1.02 (0.38)					

Table 11: Posterior summary measures of the population PK parameters κ of the IPD meta-analyses with regression and the estimates of Van Hest *et al.*. For the ISV and IOV, mean (sd) of the variance is given. Estimates of Van Hest *et al.* are obtained using the delta method (as Van Hest *et al.* reported no estimates of the logarithm of the PK parameters). Bayesian estimates: the estimates were obtained taking into account the effect of creatinine clearance on Q , creatinine clearance and the use of proton pump inhibitors on V_1 and the use of antacids on T_{lag} . Estimates van Hest *et al.*: the estimates were obtained taking into account the effect of creatinine clearance, albumin level, cyclosporine pre-dose and hemoglobin level on CL , creatinine clearance, albumin level and the use of antacids on V_1 and cyclosporine dose on k_a .

Parameter	Bayesian IPD meta-analysis			Estimates Van Hest <i>et al.</i>		
	mean (sd)	ISV	IOV	mean (sd)	ISV	IOV
$\log(CL \text{ (1/h)})$	2.98 (0.09)	0.10 (0.01)	0.14 (7.37E-3)	3.14 (0.51)	0.13	0.04
$\log(Q \text{ (1/h)})$	3.40 (0.18)	0.21 (0.03)	0.26 (0.02)	3.53 (1.59)	0.36	0.17
$\log(V_1 \text{ (l)})$	3.67 (0.10)	0.27 (0.04)	0.33 (0.02)	4.23 (1.25)	0.81	0.50
$\log(V_2 \text{ (l)})$	6.07 (0.24)	0.53 (0.08)	0.44 (0.07)	5.70 (1.67)	-	-
$\log(T_{lag} \text{ (h)})$	-0.95 (0.32)	0.42 (0.05)	0.38 (0.04)	-1.43 (0.25)	-	-
$\log(k_a \text{ (1/h)})$	-0.91 (0.32)	0.89 (0.31)	0.36 (0.12)	1.41 (1.53)	1.04	1.38

we account for inter-study variability in contrast to Van Hest *et al.*. This is probably the reason that our estimates of the population PK parameters are more precise as more variability is explained in our model. Although the five different studies have similar study designs, the quite large observed heterogeneity between the studies (estimates of IS_tV) are probably due to the different sampling schemes and different time measurements after transplantation.

The Bayesian population PK parameter estimates κ , differ a lot from the estimates obtained by Van Hest *et al.*, which is probably due to the above described differences between our models. However, despite these differences, it does reveal a high uncertainty about their true values and questions the reliability of both results.

Section 7 summarizes all results, achievements and pitfalls of this thesis and further elaborates on the results of the IPD meta-analysis and the difference between Bayesian and non-Bayesian methods.

7 Conclusion

In general, nonlinear data are hard to model, which is due to the difficult mathematical relationship between the parameters and because most often, a closed form expression for the best-fitting parameter does not exist (in contrast to linear models). The nonlinear models in this thesis include random effects and a high number of model parameters to estimate, two more aspects that are in general hard to model and estimate. Therefore, it is not surprising that many issues and problems arose during the analysis.

Starting with the analysis of the individual studies, where it became evident that most of the studies involved data which were insufficient for estimating all model parameters. This manifested itself in convergence issues and only when strong prior information was used, precise posterior summary measures were obtained. In other words: the models were overspecified and the likelihood was non-identifiable. Not surprising, because all studies initially did not intent to model the pharmacokinetics of MPA. Subsequently, no literature was found on these data where the pharmacokinetics of MPA were estimated. Another limitation was that the studies were multicenter studies and the used models did not correct for this type of variability. However, as Bayesian methods were used, prior information was increased in order to reduce the uncertainty about the true value of the model parameters and precise posterior summary measures were obtained. This is definitely not the most elegant way to overcome identifiability problems. However, emphasis was not put on tackling these identifiability problems but rather on gaining more knowledge about the data of each individual study and providing evidence for performing a meta-analyses.

With all data from the studies combined, i.e. the IPD meta-analysis (with and without regression), the identifiability problems diminished. Subsequently, reliable and sometimes precise posterior summary measures were obtained using close to non-informative prior information. However, convergence issues still arose, especially with the estimation of the absorption rate constant k_a . The data contained too few measurements to make up the absorption process and subsequently, to provide reliable estimates of k_a . Nevertheless, when combining all data, reliable posterior estimates on the mean PK parameters per study were obtained in contrast to the analysis of the individual studies, pointing out the advantages of a meta-analysis.

For the regression, no covariates with missing values and time-depending covariates were considered. These limitations resulted in differences between our found covariates which explain some of the observed variability and the ones found by Van Hest *et al.*. Because of these limitations, it is meaningless to conclude anything about these, found differences. Another difference between our models is that we account for inter-study variability in contrast to Van Hest *et al.*. This source of variability was quite high and therefore necessary to correct for. This was probably the reason that our estimates

of the population PK parameters were more precise than the obtained estimates by Van Hest *et al.*. As a result of all these differences between our models, the population PK parameter estimates differed a lot, revealing a high uncertainty about their true values. In order to make more statements about these differences, it would be necessary to include multivariate imputations in the model to introduce the covariates with missing values and to explain the IOV, i.e. include time-dependent covariates. Additionally, another source of variability could be included which corrects for the multicenters in each individual study and the dataset could be expanded with more studies on MPA in order to obtain more reliable estimates on the population PK parameters.

Pharmacokinetic data is in general hard to collect because a relatively high number of blood samples within a short time span are required to describe the overall time course of drug concentration within the body. Especially when a drug experiences high IOV and ISV and when its pharmacokinetics are best described by a two- (or higher) compartment model, lots of measurements on different individuals and occasions are necessary to obtain reliable estimates on the model parameters. However, because the pharmacokinetics of each drug on the market has been thoroughly investigated during its clinical trials, there is already highly reliable information available on its pharmacokinetics. Therefore, the contribution of prior information, i.e. the use of Bayesian methods, is a huge advantage in the analysis of pharmacokinetic data and may even be crucial for obtaining reliable estimates from smaller studies.

As WinBugs is the most popular Bayesian program, it is an obvious choice for performing Bayesian population pharmacokinetic analysis. However, as these models can be complex and much less straightforward to implement, it would be inaccessible to perform these models in WinBUGS without the existence of PKBUGS. Extensive knowledge of pharmacokinetics, Bayesian statistics and the BUGS language is required in order to implement PK models in WinBUGS. Although it is straightforward to specify and analyse a relatively simple PK model (the one described in Section 4.3 and the example in Appendix A) in PKBUGS, when the model becomes more difficult, e.g. with the introduction of IOV and/or IS_tV , it becomes less straightforward. In these cases, the model needs to be printed by PKBUGS and the user should adjust the structural and stochastic part of the model code. Additionally, PKBUGS rearranges the data in specific formats and these should be adjusted as well. In order to make these adjustments, knowledge of the BUGS language and a good understanding of the generated model by PKBUGS is a prerequisite. Another limitation of PKBUGS is that the structural part of the PK model is 'hidden' within the PK model component (see Section 4.2), consequently, the ODE's that describe the time course of drug concentration within the body (from which the PK parameters are estimated) are not accessible to the user. This limits

the freedom to make adjustments in this part of the model (for instance to fix the value of a parameter within the ODE's or assign a distribution to it). Although in PKBUGS 2.0 it is possible to implement the ODE's manually, this requires extensive knowledge of the Pascal programming language. Despite of these limitations, PKBUGS makes the analysis of Bayesian population pharmacokinetic models accessible to a wide range of people and under the right circumstances it provides a good alternative to other pharmacokinetic programs.

A PKBUGS Analyses

This section describes a basic Bayesian PPK analysis using PKBUGS. It aims to make the reader more familiar with PKBUGS. For a more extensive use of PKBUGS, the reader is advised to read Sections 3, 4 and 5 of the PKBUGS User Manual [45].

A.1 The Data

Pharmacokinetic data of the anti-asthmatic drug theophylline is used in this analysis. The data originates from a study by Dr. Robert Upton [57–59]. Twelve subjects were given one oral dose of theophylline and serum concentrations were measured at 11 time points over the next 25 hours. The data can be obtained via R (library datasets):

```
> data(Theoph)
> head(Theoph)

  Subject  Wt Dose Time  conc
1        1 79.6 4.02 0.00  0.74
2        1 79.6 4.02 0.25  2.84
3        1 79.6 4.02 0.57  6.57
4        1 79.6 4.02 1.12 10.50
5        1 79.6 4.02 2.02  9.66
6        1 79.6 4.02 3.82  8.58
```

Where Subject is the id of the subjects (1 to 12), Wt is the weight (kg), Dose is the administered dose (mg/kg), Time is the time of the measurements (h) and conc is the measured theophylline concentration (mg/l).

The data should be adapted in order for PKBUGS to recognize it:

```
> head(Theoph)

  id  Wt  amt time  dv evid
1  1 79.6 4.02 0.00  0.74   1
2  1 79.6 4.02 0.25  2.84   0
3  1 79.6 4.02 0.57  6.57   0
4  1 79.6 4.02 1.12 10.50   0
5  1 79.6 4.02 2.02  9.66   0
6  1 79.6 4.02 3.82  8.58   0
```

Note that some of the column names are changed (into the names of the data items PKBUGS recognizes) and an extra variable is created called evid. Evid stands for event identification and summarizes what type of event each observation belongs to. In this study we only have observations (evid = 0) and dose events (evid = 1).

A.2 Model Specifications

When the data is copied to WinBUGS (version 1.3 with PKBUGS 1.1), it is important that the data and the data items (column names) are saved in separate documents. To load the data items in PKBUGS, make sure that the data item document is the top window and select **Load item names** from the PKBugs menu. In the status bar, the message 'items names loaded' appears. The same procedure can be followed for loading the data, by selecting **Load data** and the message '*data loaded*' appears. The next step is to define the model, which can be done by selecting **Define model** from the PKBugs menu. For this example, we select a one-compartment model with normal residuals, then click on **Check model**. In the status bar the message 'model ok' appears. The priors can be specified by clicking **Priors** in the PKBugs menu. In this example we choose as priors, the estimated PK parameters from a PK study on theophylline by Hussain *et al.* (with weight = 70 kg and age = 40 yrs) [60]; CL : 2.44 l/h with an ISV of 38.7%, V : 39.9 l with an ISV of 40%. The box covariates shows the covariates in this data set (weight and log weight), which can be selected for each PK parameter. In this example we will not select any, and just estimate the PK parameters. When the priors are specified, click on **Done**. Then the message '*priors ok*' appears. A new window automatically opens with the above specified priors. These can be loaded in PKBUGS by choosing **Load priors** from the PKBugs menu and the message '*priors loaded*' appears in the status bar. Again, a new window opens with the initial values for the population parameters which PKBUGS automatically generates. Click on **Load inits (pop)** in the PKBugs menu to load these initial values into PKBUGS. Then the message '*initial values for population parameters loaded*' appears. A new window has opened showing the initial values for theta and these can be loaded in PKBUGS by clicking on **Load inits (theta)** from the PKBugs menu, the status bar shows the message '*initial values for theta loaded*'. Now we can choose between two options, **Export model** or **Compile** both in the PKBugs menu. If we export the model, PKBUGS generates an equivalent WinBUGS code for the specified model, which runs in WinBUGS version 1.4.3 with PKBUGS 2.0. This option can be used when it is necessary to modify the standard three-stage hierarchical model. This is not necessary for the analysis in this example and we choose to compile the model.

A.3 The Analysis

The actual analysis does not differ from any other analysis in WinBUGS. PKBUGS automatically opens the trace plots for the most important parameters. Where $\mu[1]$ is the logarithm of CL , $\mu[2]$ is the logarithm of V , $\Omega[1, 1]$ is the inter-subject variability of CL , $\Omega[2, 2]$ is the inter-subject variability of V and σ is the standard deviation of the residual

error. Other nodes that may be of interest can be specified in the **Sample Monitor Tool**, we choose *model*, which represents the pharmacokinetic model evaluated at time *t* and *theta*, which represent the individual PK parameter estimates. Since the individual PK parameters are estimated using the MH-algorithm, we would also like to monitor its acceptance rates by choosing **Metropolis monitor** in the PKBugs menu. When all nodes are specified we can start sampling, we run 100,000 iterations.

A.4 The results

The autocorrelation function (ACF) shown in Figure 7 tells us that most of the sampled values are basically uncorrelated. The initial sampled values for *CL*, show some correlation, therefore, 5,000 iterations are discarded. The density plots shown in Figure 7 tell us that all of the sampled values are generated from a unimodal distribution. Based on these plots and the trace plots (not shown), convergence seems to be attained. Nevertheless, these plots are not a proof that the posterior is sampled appropriately and convergence diagnostics are needed. Using the BOA package in R, convergence diagnostics were consulted. Heidelberger & Welch confirmed that convergence is attained (output not shown). The Bayesian summary measures shown in Figure 7 show that the posterior mean of *CL* is 0.02 l/h with 95% credible interval [0.01, 0.03], *V* is 0.70 l [0.58,0.85], ISV of *CL* 0.15 [0.03, 0.47], ISV of *V* 0.09 [0.03, 0.22] and sigma (standard deviation of residual error) 2.10 [1.84, 2.42]. Sigma appears to be large, perhaps a two-compartment model provides a better fit for this data.

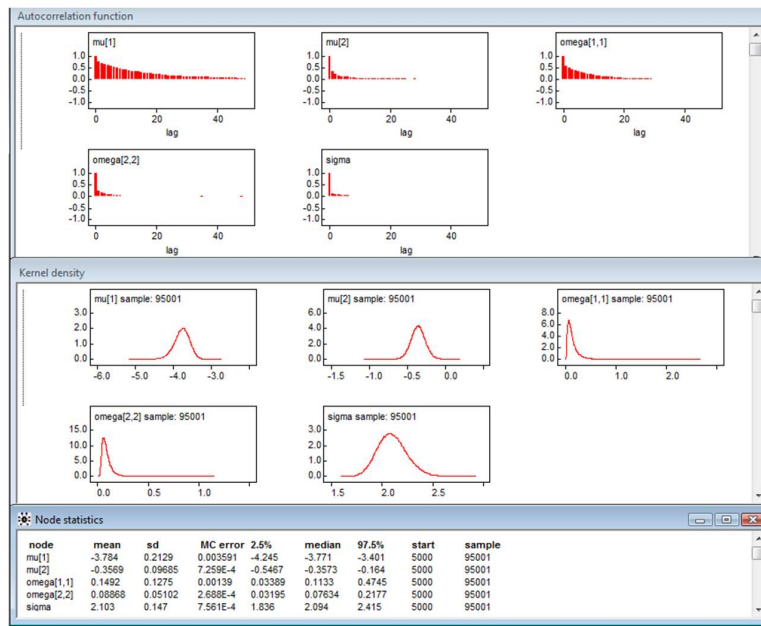


Figure 7: Theophylline PPK analysis: diagnostic and output screens obtained from Winbugs. Top: ACF after 100000 iterations. Middle: smooth density plot after 95000 iterations (5000 burn-in). Bottom: Bayesian summary measures based on 95000 iterations (5000 burn-in)

B Priors per Study

This section describes the prior estimates of the models per study (Section 5). Strategies for obtaining non-, fully- and minimally-informative priors are described below.

Non-informative priors were chosen following the guidelines of Lunn *et al.* [45,46] and were the same for all studies. Where the population pharmacokinetic parameters $\boldsymbol{\eta}$ were set to the prior point estimates reported in Table 2, their variances, the diagonals of \mathbf{C} , were set to 10,000, the prior point estimates of the variances for the inter-occasion and inter-subject variability, the diagonals of \mathbf{R} and \mathbf{G} , were multiplied by 6 (the number of PK parameters in the model) with degrees of freedom, ρ and γ , equals 6 and the prior mean of the precision (τ) was set to 1 with a prior variance of 1000.

The starting position for obtaining fully-informative priors includes the specification of the prior estimates reported in Table 2 (for $\boldsymbol{\eta}$, the diagonals of \mathbf{C} , \mathbf{R} and \mathbf{G}). With these informative priors, the data of the studies still need more information for estimating all model parameters, i.e. convergence is not attained for all model parameters and the likelihood is non-identifiable. In order to make the model identifiable, the precisions of the model parameters that did not converge were increased by multiplying it by a scalar until convergence was attained (for \mathbf{C} , \mathbf{R} or \mathbf{G}). Whereby the degrees of freedom of the Wishart distributions for the ISV and IOV, ρ and γ , were kept at 6 or increased to 7 (when convergence was not reached). A minimally-informative Wishart distribution is obtained by choosing the degrees of freedom in this way, i.e. $\boldsymbol{\Omega}^{-1} \sim W_p(\mathbf{R}^{-1}, \rho)$, where \mathbf{R} is a $p \times p$ positive definite matrix and ρ is chosen as p or $p + 1$ [38]. Prior information on the ISV and IOV highly influenced the individual and subsequently population PK parameter estimates. For instance, when the degrees of freedom of the Wishart distribution were further increased (making the distribution highly informative by setting them equal to the number of individuals who participated in the prior studies), the estimates of $\boldsymbol{\theta}$ and $\boldsymbol{\mu}$ became almost identical to the prior estimates. In order to gain as much information as possible from the data of the studies, convergence could not always be reached for the parameters $\boldsymbol{\Omega}$ and $\boldsymbol{\Phi}$ and emphasis was put on the convergence of $\boldsymbol{\mu}$ and $\boldsymbol{\theta}$.

With fully-informative priors, prior information is at least equal to the prior estimates reported in Table 2. However, some parameters may not need such informative priors. Therefore, minimally-informative priors were obtained by further optimizing the prior estimates of the variances. This was done in such a way that convergence was attained for previously mentioned parameters with the least informative priors, i.e. it was tried to reduce the prior information on each parameter while remaining convergence.

The resulting multiplication factors for \mathbf{C} , \mathbf{R} and \mathbf{G} for fully- and minimally-informative priors per study are reported in Sections B.1 to B.6.

B.1 Study 1

Fully-informative priors: the prior point estimates of the diagonals of \mathbf{C} were multiplied by 1, \mathbf{R} were multiplied by [1, 1, 1, 0.03, 0.003, 0.03] for Q , V_1 , V_2 , T_{lag} and k_a respectively, with ρ equals 7, \mathbf{G} were multiplied by [1, 0.03, 0.3, 0.1, 0.015, 0.05] with γ equals 7 and the mean of τ was set to the prior point estimate with variance 0.3.

Minimally-informative priors: the prior point estimates of the diagonals of \mathbf{C} were multiplied by [1000000, 5, 5, 5, 50, 5], \mathbf{R} were multiplied by [3, 3, 3, 0.03, 5, 0.068] and ρ was set to 7, \mathbf{G} were multiplied by [3, 0.03, 0.3, 0.3, 0.015, 0.5], γ was set to 7 and the posterior mean of τ was set to 1 with variance 1000.

B.2 Study 2

Fully-informative priors: the prior point estimates of the diagonals of \mathbf{C} were multiplied by 1, \mathbf{R} were multiplied by [1, 1, 1, 0.5, 1, 1], with ρ equals 7, \mathbf{G} were multiplied by [1, 0.1, 1, 0.5, 0.0167, 1] with γ equals 7 and the mean of τ was set to the prior point estimate with variance 0.3.

Minimally-informative priors: the prior point estimates of the diagonals of \mathbf{C} were multiplied by [2000000, 1, 1, 1, 5, 1], \mathbf{R} were multiplied by [3, 1, 1, 0.5, 1, 11] and ρ was set to 7, \mathbf{G} were multiplied by [3, 1, 1, 0.5, 0.02, 2], γ was set to 7 and the posterior mean of τ was set to 1 with variance 1000.

B.3 Study 3

Fully-informative priors: the prior point estimates of the diagonals of \mathbf{C} were multiplied by [1, 1, 1, 1, 1, 1], \mathbf{R} were multiplied by [1, 1, 1, 1, 1, 1], with ρ equals 7, \mathbf{G} were multiplied by [1, 1, 1, 0.05, 1, 0.007] with γ equals 7 and the mean of τ was set to the prior point estimate with variance 0.3.

Minimally-informative priors: the prior point estimates of the diagonals of \mathbf{C} were multiplied by [1.61E6, 2.42E6, 6.99E5, 3.79, 3.17E7, 2.49], \mathbf{R} were multiplied by [6, 6, 6, 6, 6, 0.6] and ρ was set to 7, \mathbf{G} were multiplied by [6, 6, 6, 0.01, 6, 0.007], γ was set to 7 and the posterior mean of τ was set to 1 with variance 1000.

B.4 Study 4

Fully-informative priors: the prior point estimates of the diagonals of \mathbf{C} were multiplied by [1, 1, 1, 1, 1, 1], \mathbf{R} were multiplied by [1, 1, 1, 0.047, 1, 0.322], with ρ equals 7, \mathbf{G} were multiplied by [1, 0.1, 0.737, 0.04, 0.0154, 0.0118] with γ equals 7 and the mean of τ was set to the prior point estimate with variance 0.3.

Minimally-informative priors: the prior point estimates of the diagonals of \mathbf{C} were multiplied by [10, 1, 10, 10, 1000, 10], \mathbf{R} were multiplied by [10,

10, 10, 0.047, 10, 1] and ρ was set to 7, \mathbf{G} were multiplied by [10, 0.1, 1, 0.05, 0.0199, 0.0124], γ was set to 7 and the posterior mean of τ was set to 1 with variance 1000.

B.5 Study 5

Fully-informative priors: the prior point estimates of the diagonals of \mathbf{C} were multiplied by [100, 10, 10, 1, 1000, 10], \mathbf{R} were multiplied by [1, 0.1, 0.1, 0.01, 10, 0.02], with ρ equals 7, \mathbf{G} were multiplied by [10, 0.01, 0.01, 0.025, 0.003, 0.005] with γ equals 7 and the mean of τ was set to the prior point estimate with variance 0.3.

Minimally-informative priors: the prior point estimates of the diagonals of \mathbf{C} were multiplied by [100000, 10, 10, 1, 1000, 10], \mathbf{R} were multiplied by [1, 1, 1, 0.01, 10, 0.02] and ρ was set to 7, \mathbf{G} were multiplied by [10, 0.01, 0.01, 0.05, 0.003, 0.005], γ was set to 7 and the posterior mean of τ was set to 1 with variance 1000.

B.6 Study 6

Fully-informative priors: the prior point estimates of the diagonals of \mathbf{C} were multiplied by [1, 1, 1, 0.5, 1, 0.5], \mathbf{R} were multiplied by [1, 1, 0.5, 0.05, 1, 0.08], with ρ equals 7 and the mean of τ was set to the prior point estimate with variance 0.3.

Minimally-informative priors: the prior point estimates of the diagonals of \mathbf{C} were multiplied by [1614435, 10, 1000, 0.5, 31738623, 0.5], \mathbf{R} were multiplied by [100, 10, 10, 0.05, 10, 0.08], ρ was set to 7 and the posterior mean of τ was set to 1 with variance 1000.

C Trace Plots of the Analysis per Study

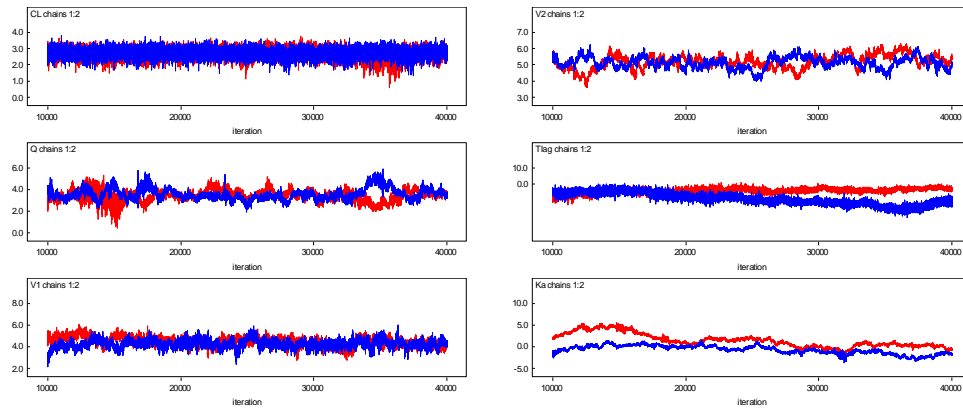


Figure 8: Study 1: trace plots of the population PK parameters obtained from WinBUGS based on two overdispersed starting positions using non-informative priors.

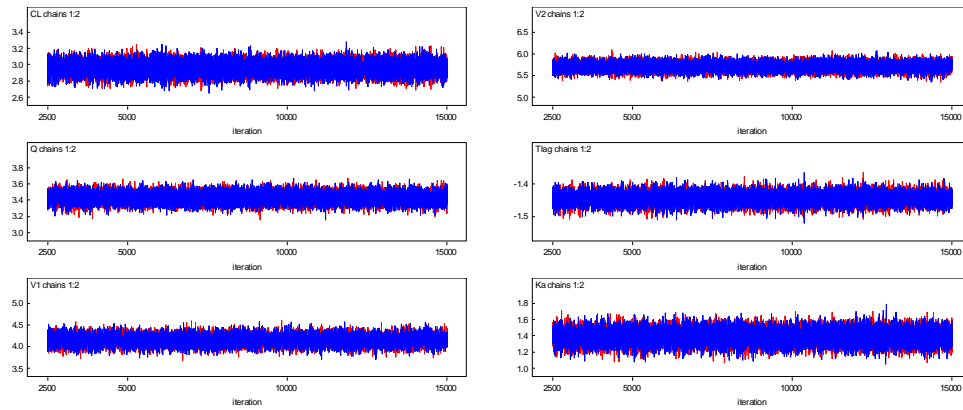


Figure 9: Study 1: trace plots of the population PK parameters obtained from WinBUGS based on two overdispersed starting positions using fully-informative priors.

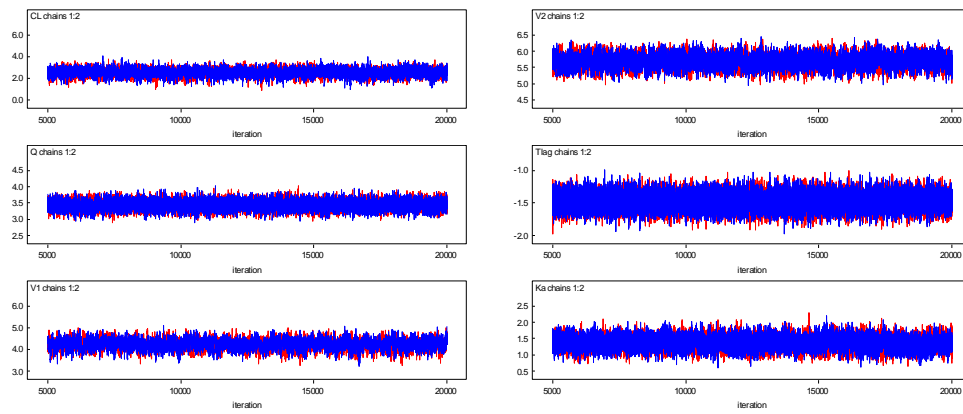


Figure 10: Study 1: trace plots of the population PK parameters obtained from WinBUGS based on two overdispersed starting positions using minimally-informative priors.

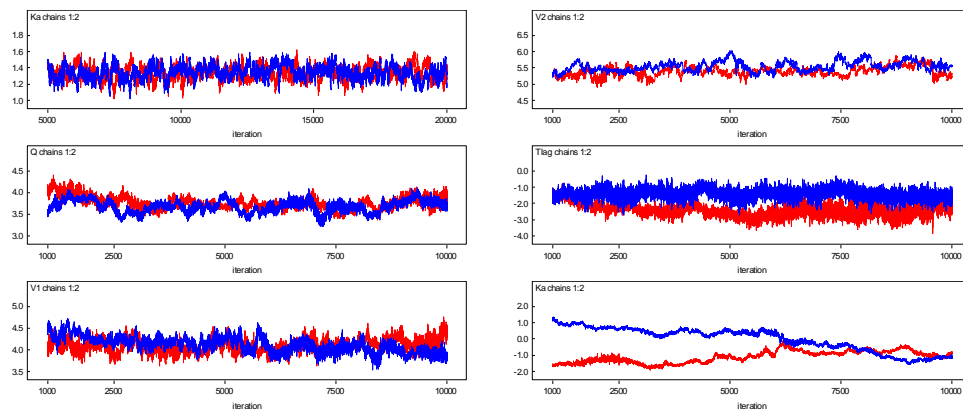


Figure 11: Study 2: trace plots of the population PK parameters obtained from WinBUGS based on two overdispersed starting positions using non-informative priors.

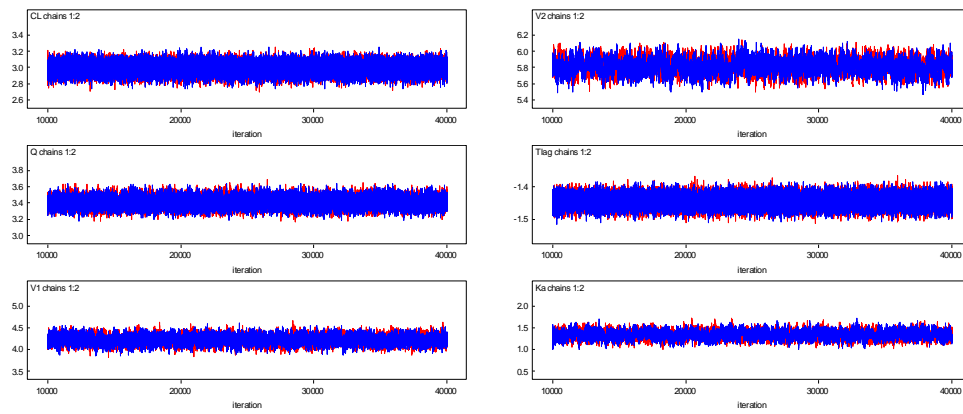


Figure 12: Study 2: trace plots of the population PK parameters obtained from WinBUGS based on two overdispersed starting positions using fully-informative priors.

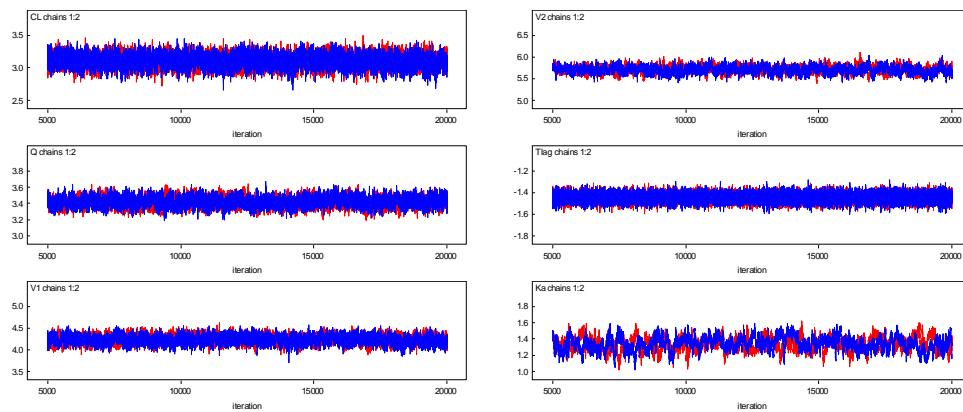


Figure 13: Study 2: trace plots of the population PK parameters obtained from WinBUGS based on two overdispersed starting positions using minimally-informative priors.

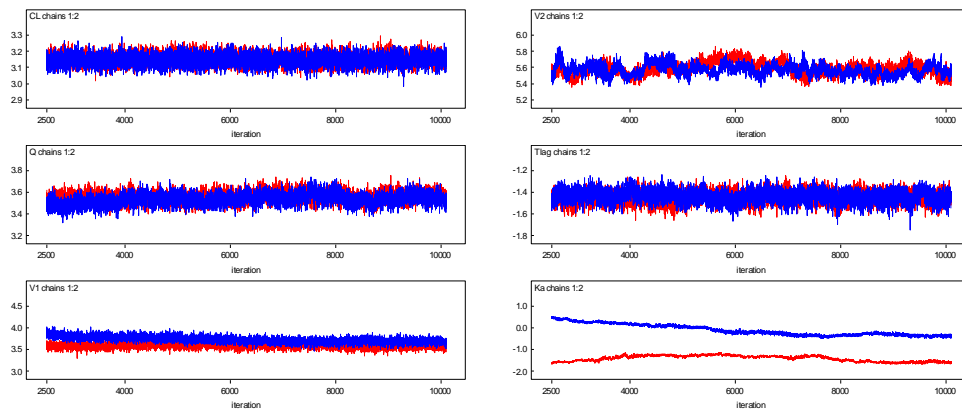


Figure 14: Study 3: trace plots of the population PK parameters obtained from WinBUGS based on two overdispersed starting positions using non-informative priors.

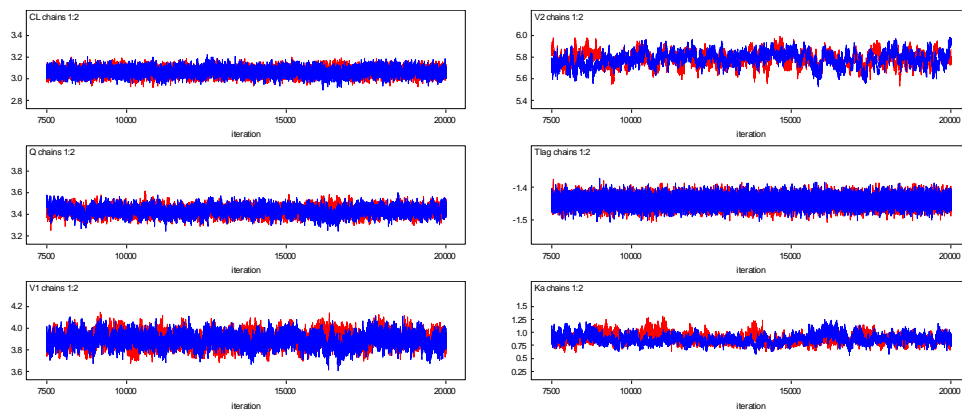


Figure 15: Study 3: trace plots of the population PK parameters obtained from WinBUGS based on two overdispersed starting positions using fully-informative priors.

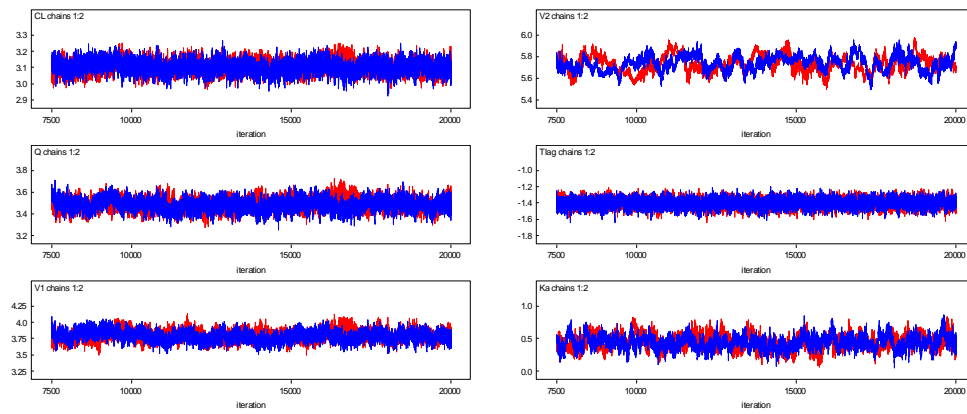


Figure 16: Study 3: trace plots of the population PK parameters obtained from WinBUGS based on two overdispersed starting positions using minimally-informative priors.

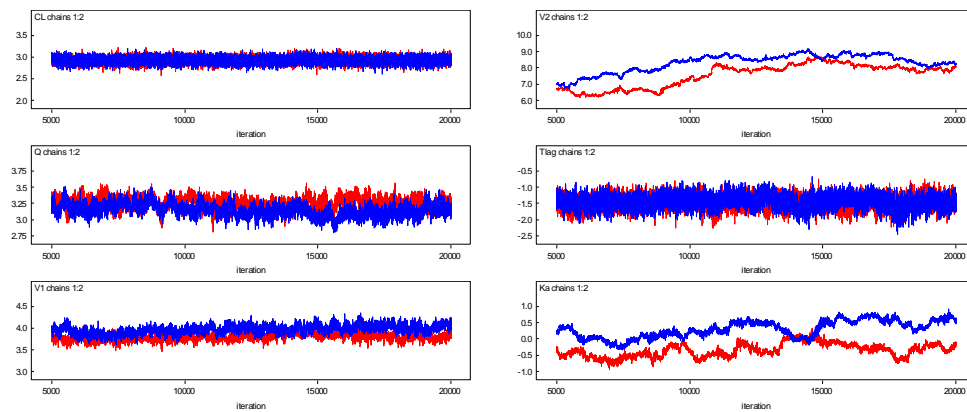


Figure 17: Study 4: trace plots of the population PK parameters obtained from WinBUGS based on two overdispersed starting positions using non-informative priors.

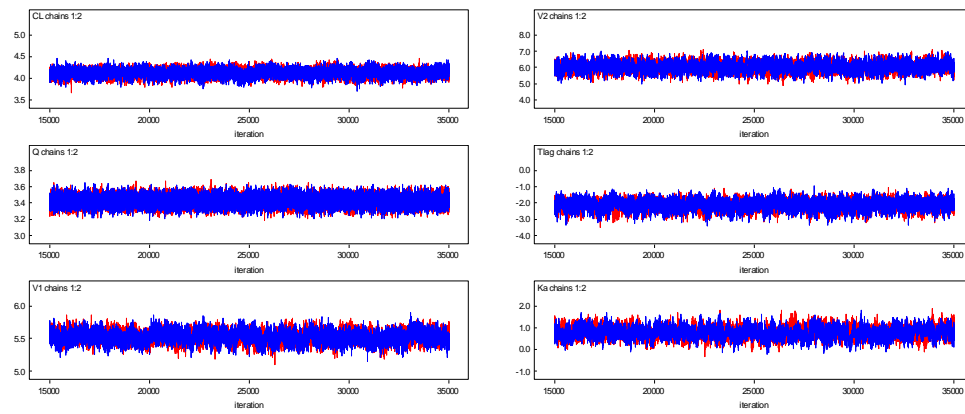


Figure 18: Study 4: trace plots of the population PK parameters obtained from WinBUGS based on two overdispersed starting positions using fully-informative priors.

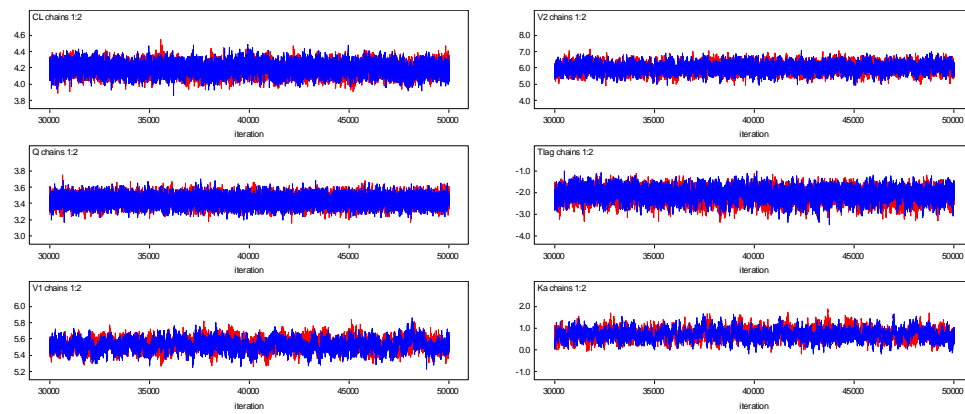


Figure 19: Study 4: trace plots of the population PK parameters obtained from WinBUGS based on two overdispersed starting positions using minimally-informative priors.

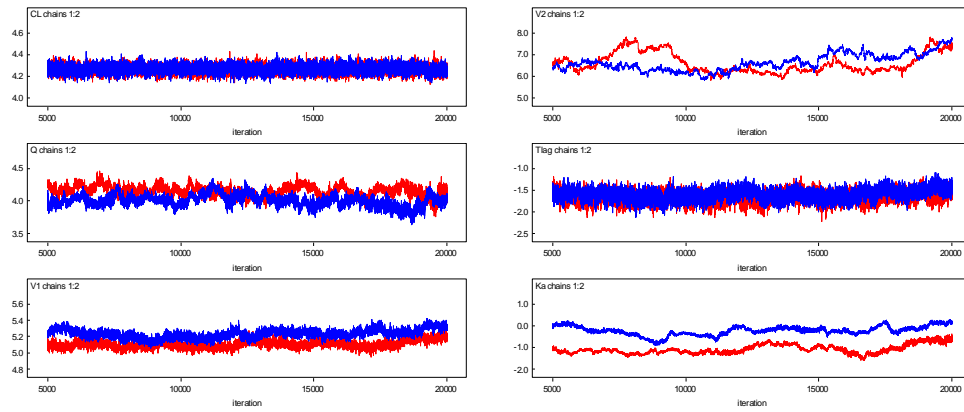


Figure 20: Study 5: trace plots of the population PK parameters obtained from WinBUGS based on two overdispersed starting positions using non-informative priors.

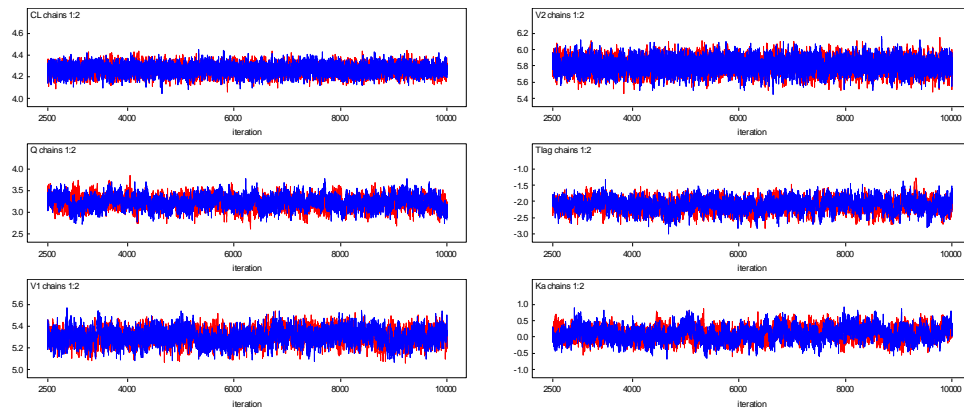


Figure 21: Study 5: trace plots of the population PK parameters obtained from WinBUGS based on two overdispersed starting positions using fully-informative priors.

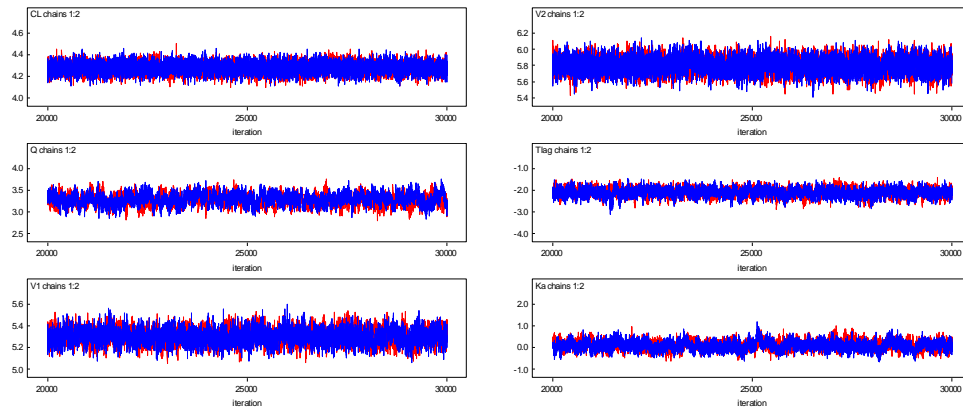


Figure 22: Study 5: trace plots of the population PK parameters obtained from WinBUGS based on two overdispersed starting positions using minimally-informative priors.

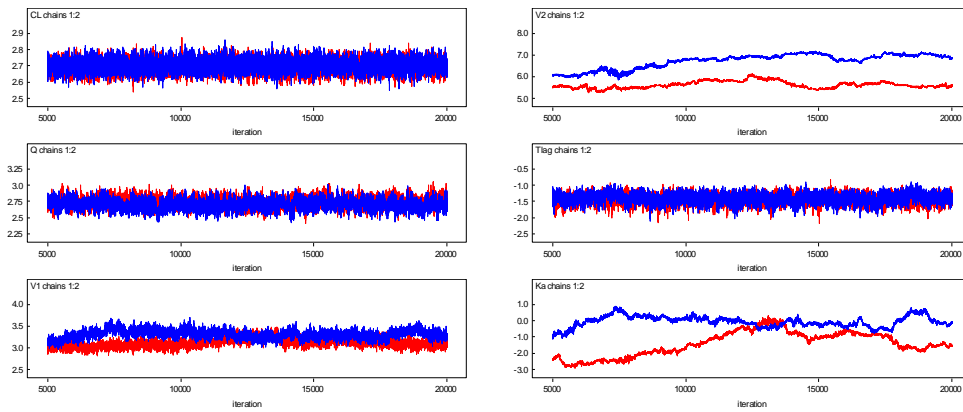


Figure 23: Study 6: trace plots of the population PK parameters obtained from WinBUGS based on two overdispersed starting positions using non-informative priors.

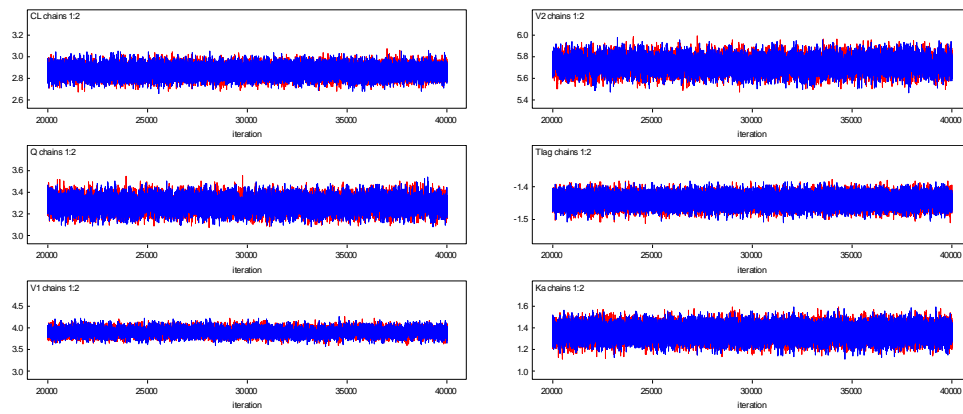


Figure 24: Study 6: trace plots of the population PK parameters obtained from WinBUGS based on two overdispersed starting positions using fully-informative priors.

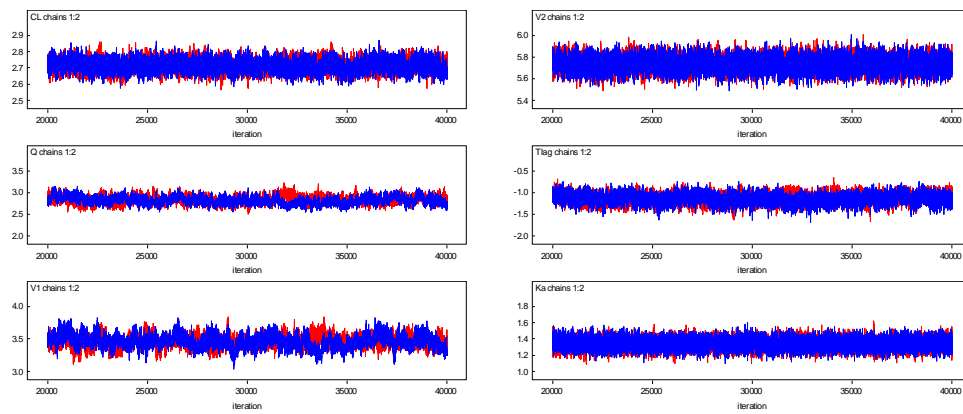


Figure 25: Study 6: trace plots of the population PK parameters obtained from WinBUGS based on two overdispersed starting positions using minimally-informative priors.

D Trace Plots of the Meta-Analysis

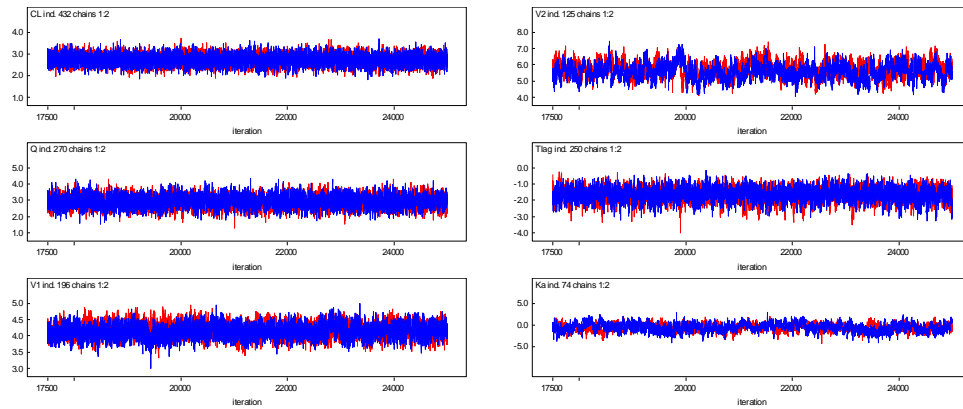


Figure 26: Trace plots of the individual PK parameters obtained from WinBUGS based on two overdispersed starting positions using non-informative priors. Individuals were randomly selected

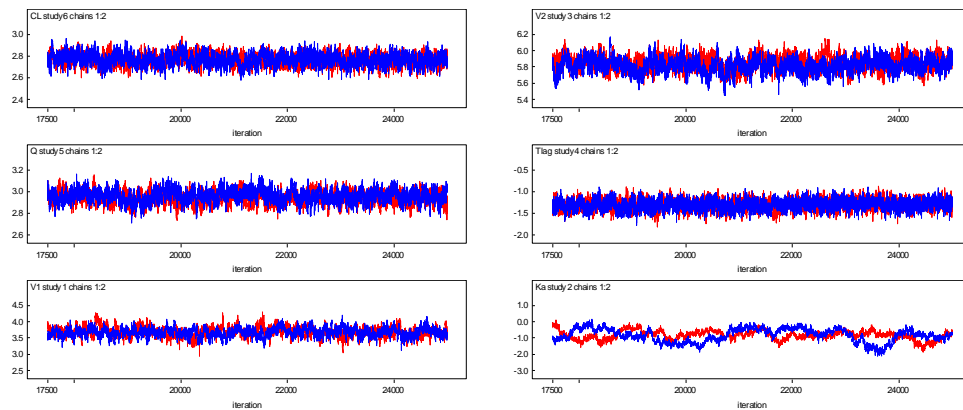


Figure 27: Trace plots of the study PK parameters obtained from WinBUGS based on two overdispersed starting positions using non-informative priors. The studies were randomly selected

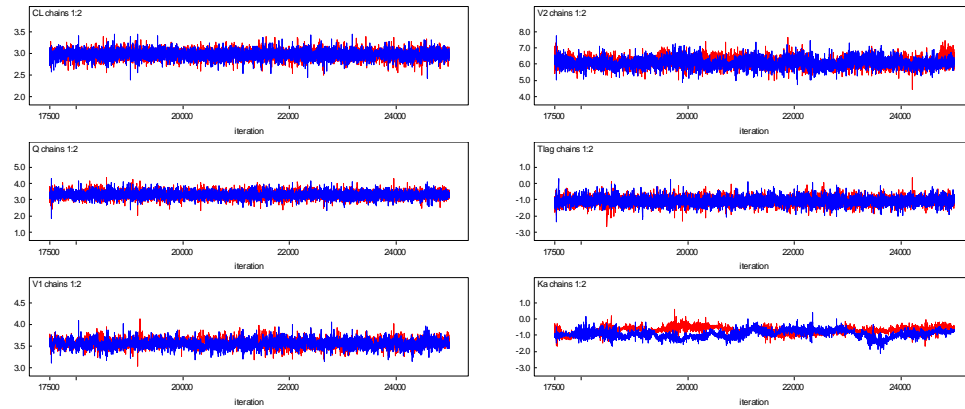


Figure 28: Trace plots of the population PK parameters obtained from WinBUGS based on two overdistributed starting positions using non-informative priors.

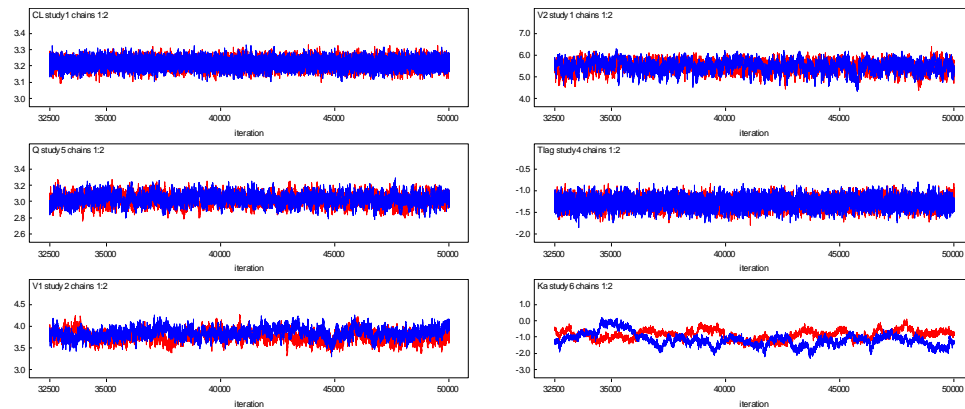


Figure 29: Regression analyses: trace plots of the study PK parameters obtained from WinBUGS based on two overdistributed starting positions using non-informative priors. The studies were randomly selected

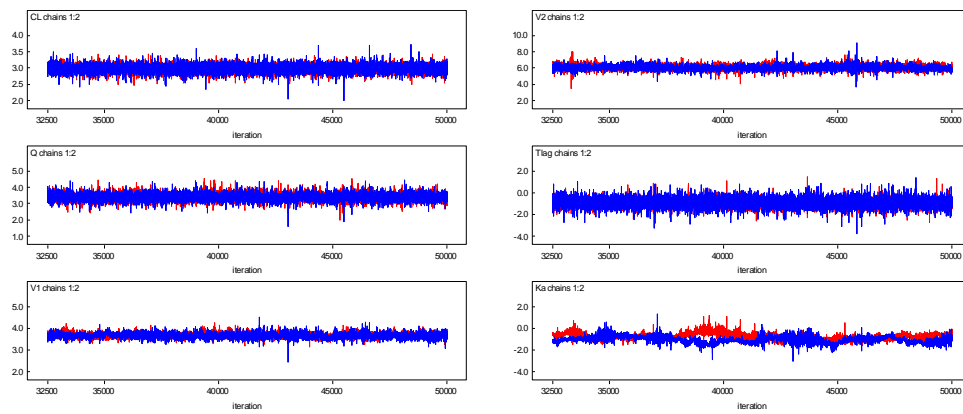


Figure 30: Regression analyses: trace plots of the population PK parameters obtained from WinBUGS based on two overdispersed starting positions using non-informative priors.

E Regression Plots of the Covariates

E.1 Gender

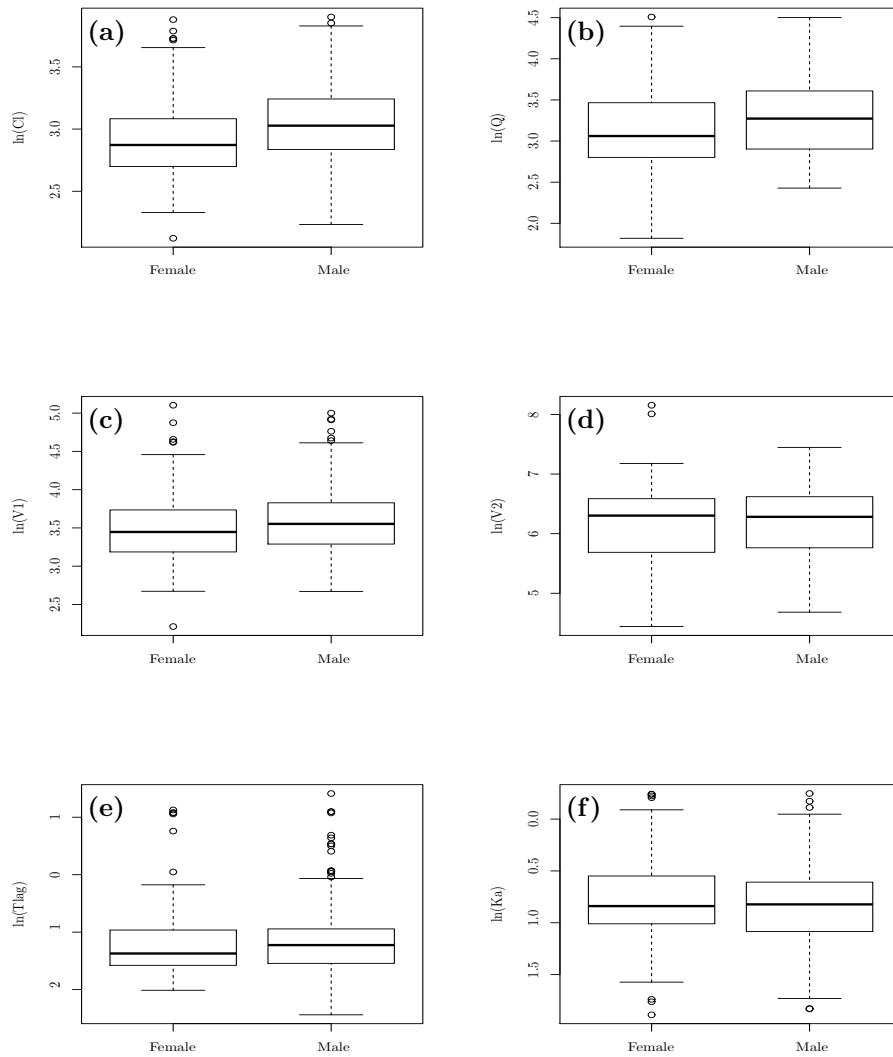


Figure 31: Individual effect of gender on the pharmacokinetic parameters (a) clearance, (b) inter-compartmental clearance, (c) volume of distribution of the central compartment, (d) volume of distribution of the peripheral compartment, (e) absorption lag-time and (f) absorption rate constant.

E.2 Race

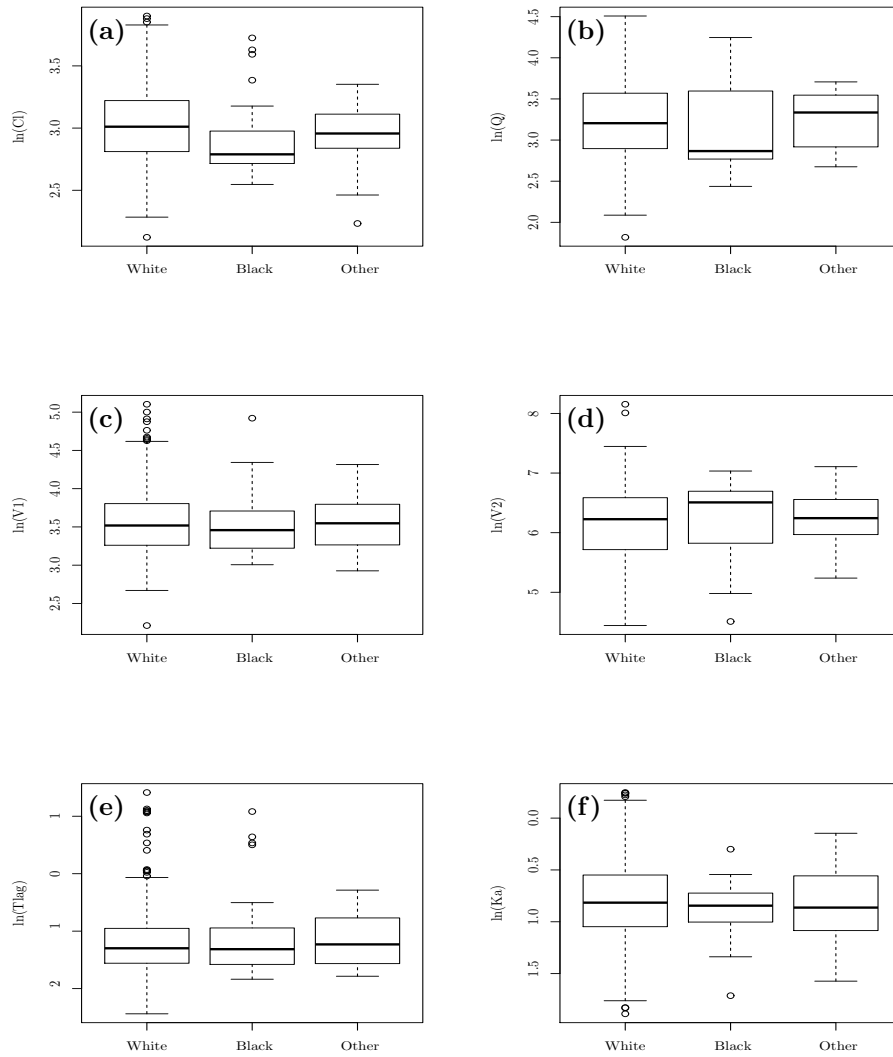


Figure 32: Individual effect of race on the pharmacokinetic parameters (a) clearance, (b) inter-compartmental clearance, (c) volume of distribution of the central compartment, (d) volume of distribution of the peripheral compartment, (e) absorption lag-time and (f) absorption rate constant.

E.3 Weight

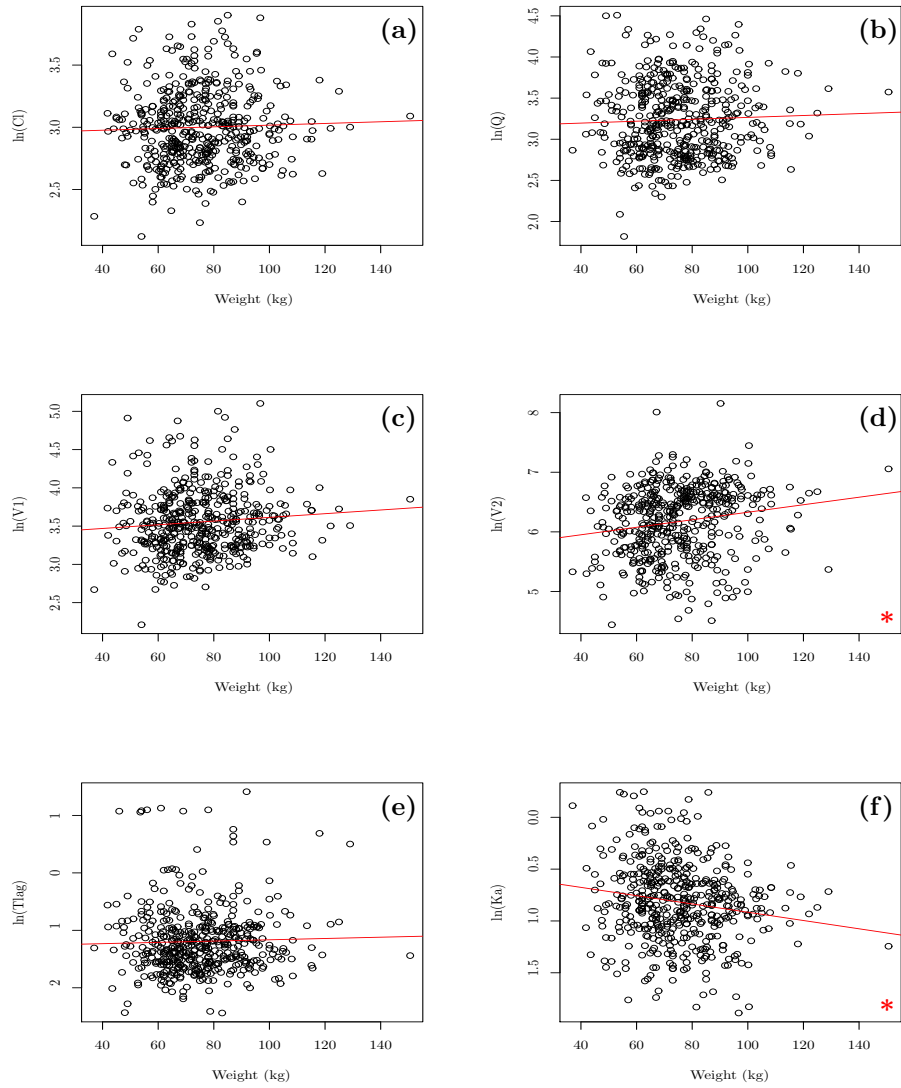


Figure 33: Individual effect of weight on the pharmacokinetic parameters (a) clearance, (b) inter-compartmental clearance, (c) volume of distribution of the central compartment, (d) volume of distribution of the peripheral compartment, (e) absorption lag-time and (f) absorption rate constant. * Introduced in the full regression model.

E.4 Height

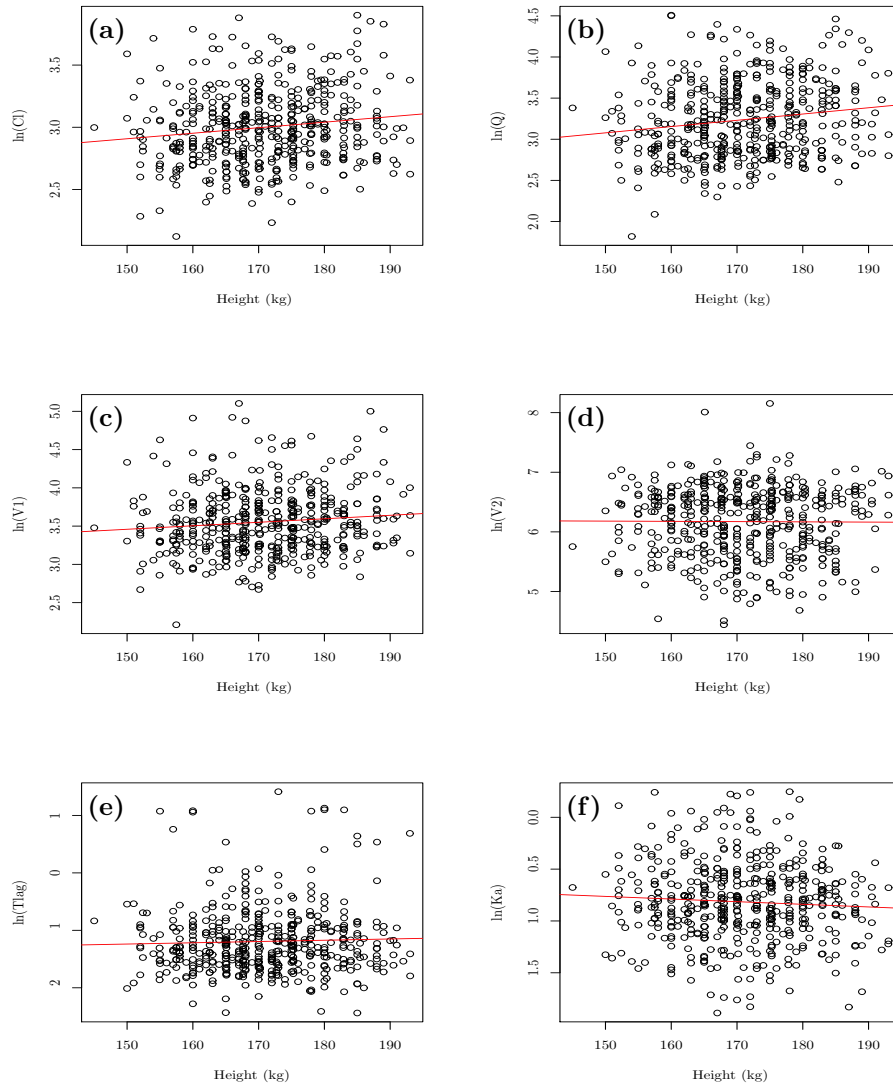


Figure 34: Individual effect of height on the pharmacokinetic parameters (a) clearance, (b) inter-compartmental clearance, (c) volume of distribution of the central compartment, (d) volume of distribution of the peripheral compartment, (e) absorption lag-time and (f) absorption rate constant.

E.5 Age

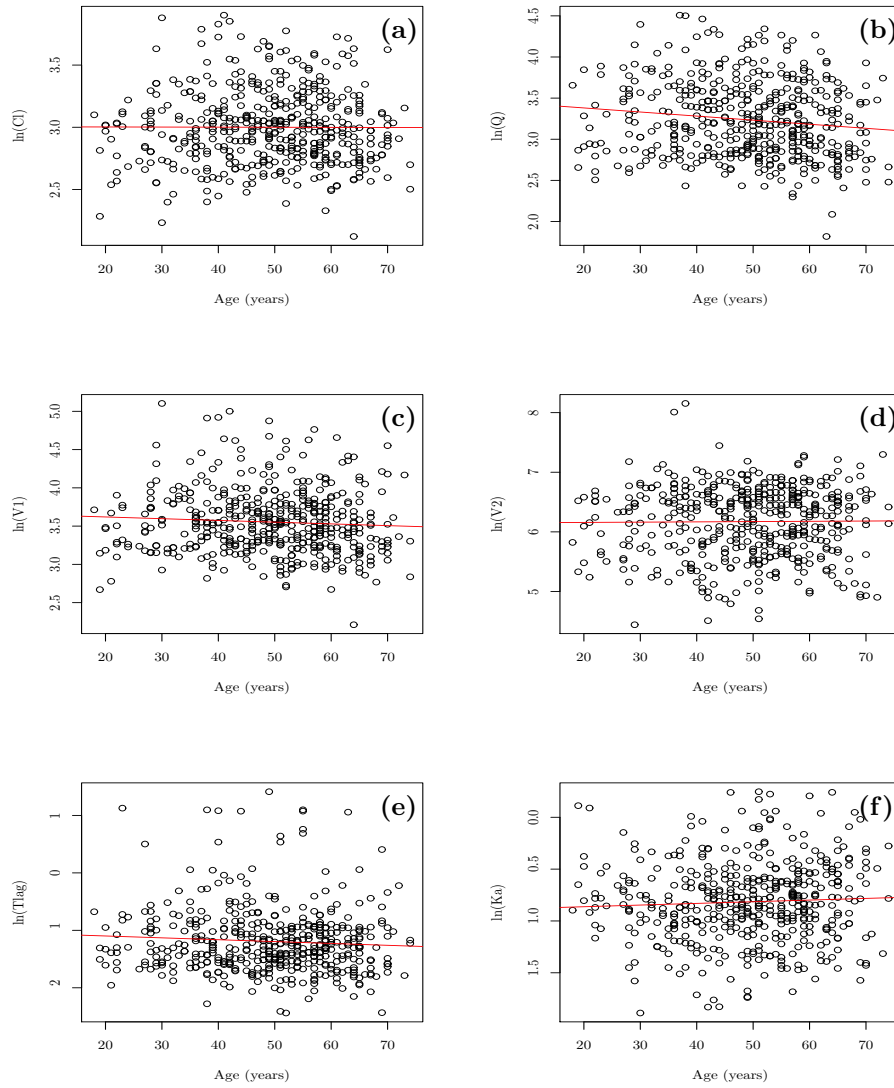


Figure 35: Individual effect of age on the pharmacokinetic parameters (a) clearance, (b) inter-compartmental clearance, (c) volume of distribution of the central compartment, (d) volume of distribution of the peripheral compartment, (e) absorption lag-time and (f) absorption rate constant.

E.6 MMF dose

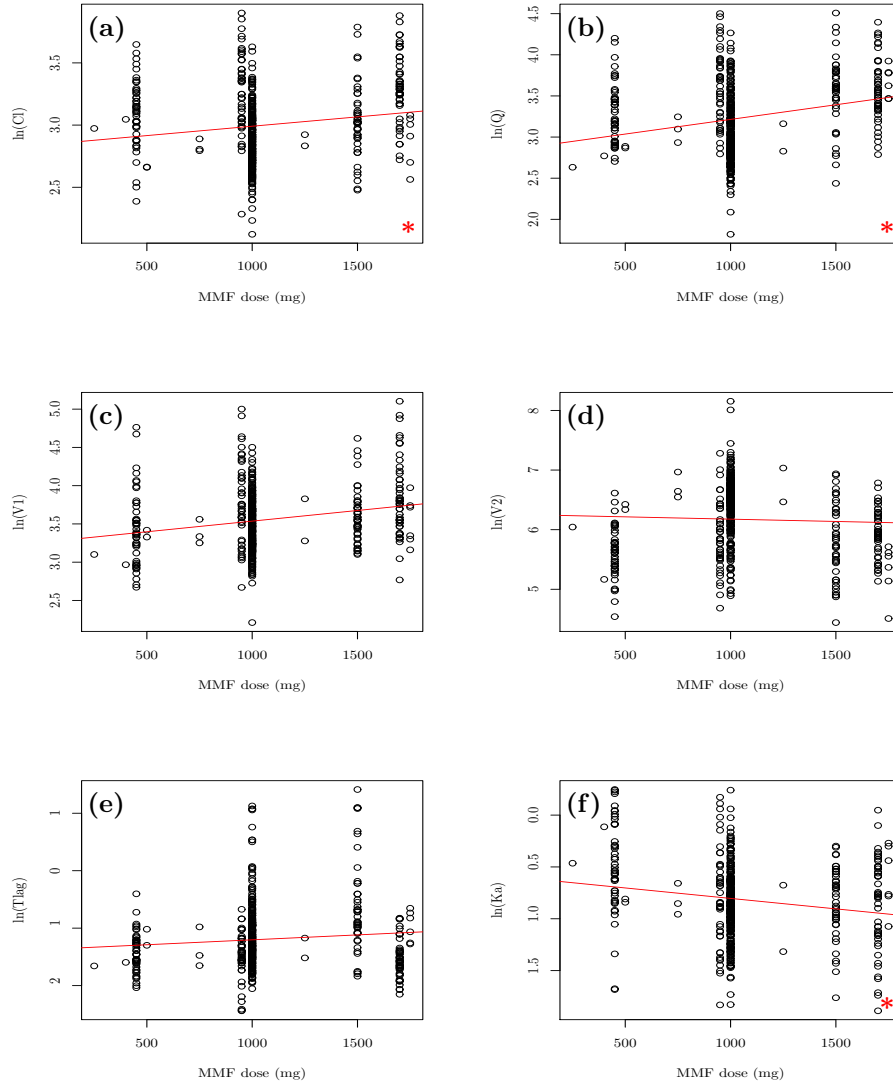


Figure 36: Individual effect of the MMF dose on the pharmacokinetic parameters (a) clearance, (b) inter-compartmental clearance, (c) volume of distribution of the central compartment, (d) volume of distribution of the peripheral compartment, (e) absorption lag-time and (f) absorption rate constant. * Introduced in the full regression model.

E.7 Diabetes Mellitus

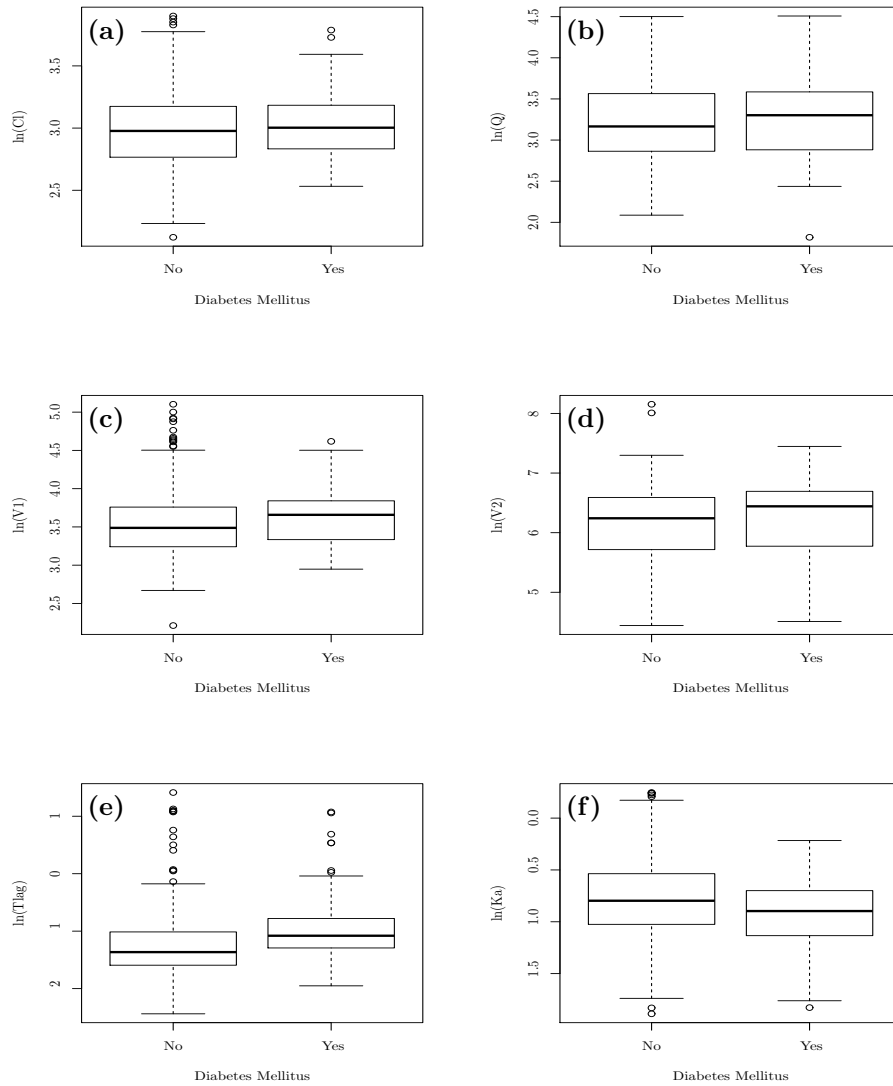


Figure 37: Individual effect of Diabetes Mellitus on the pharmacokinetic parameters (a) clearance, (b) inter-compartmental clearance, (c) volume of distribution of the central compartment, (d) volume of distribution of the peripheral compartment, (e) absorption lag-time and (f) absorption rate constant.

E.8 Creatinine Clearance using Cockcroft & Gault

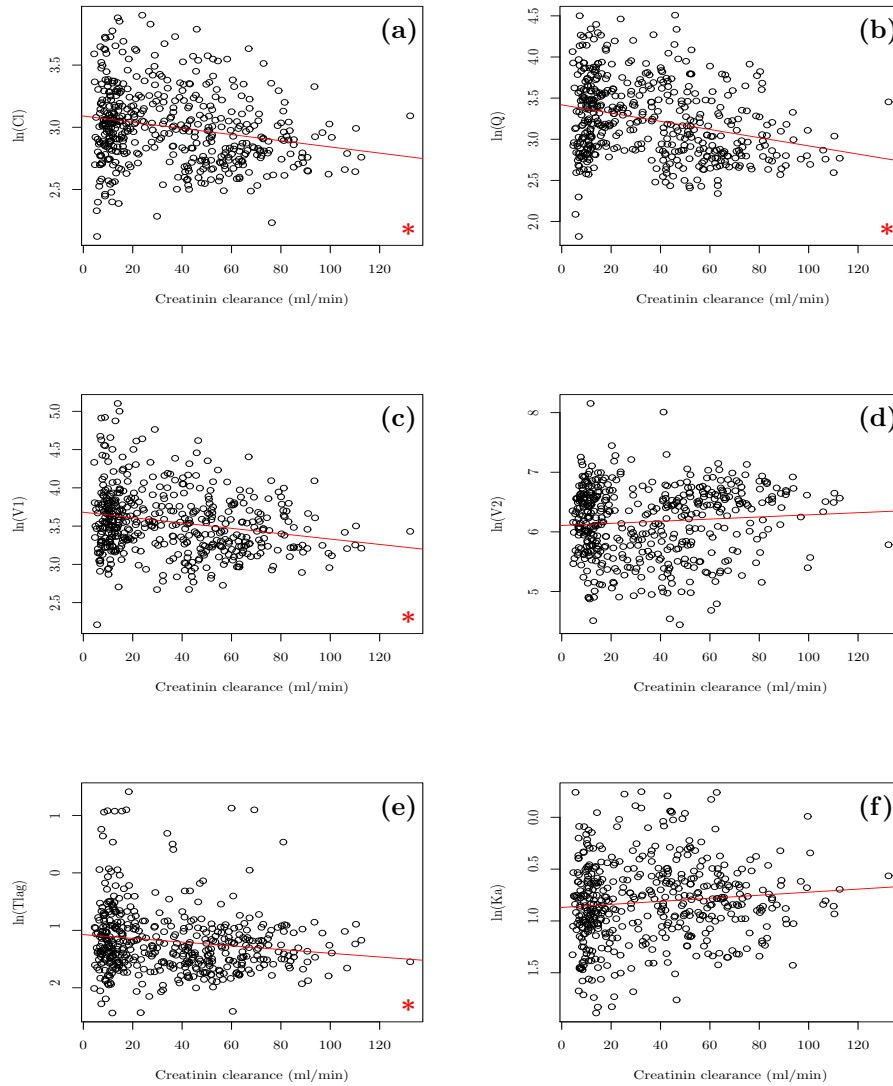


Figure 38: Individual effect of creatinine clearance on the pharmacokinetic parameters (a) clearance, (b) inter-compartmental clearance, (c) volume of distribution of the central compartment, (d) volume of distribution of the peripheral compartment, (e) absorption lag-time and (f) absorption rate constant. * Introduced in the full regression model.

E.9 Hemoglobin

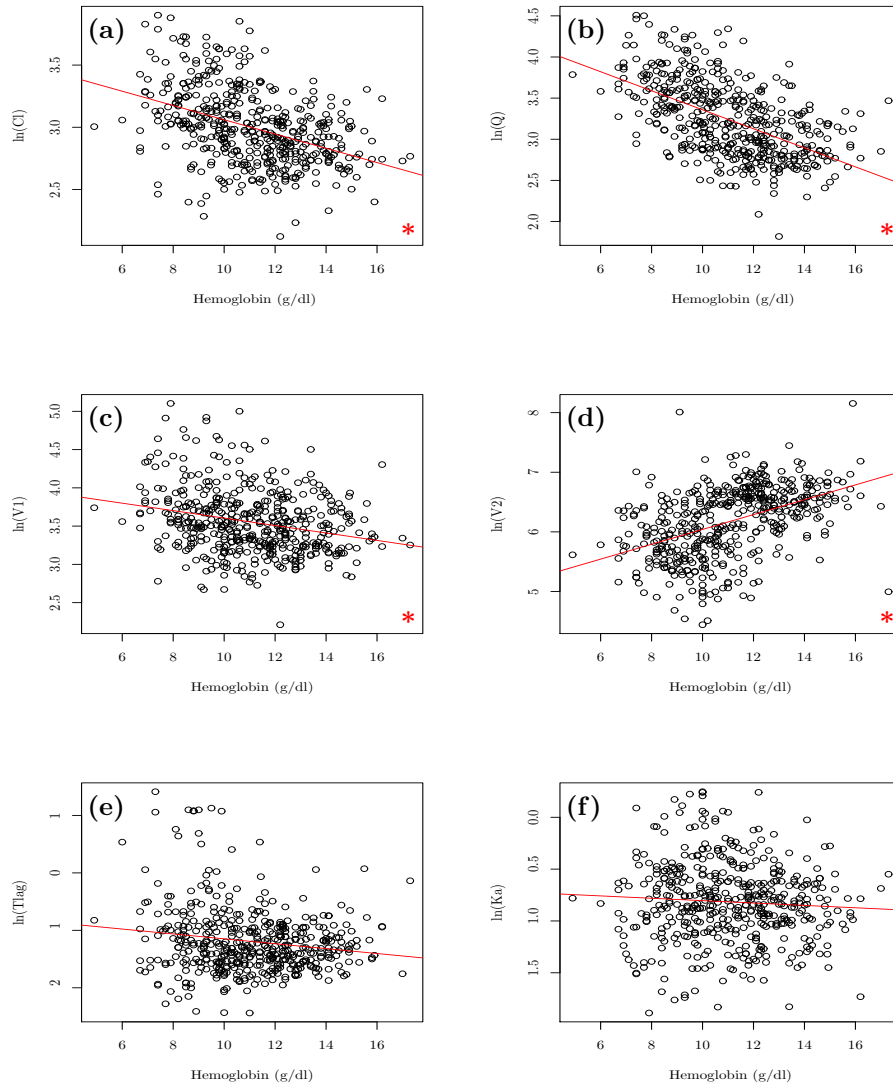


Figure 39: Individual effect of Hemoglobin on the pharmacokinetic parameters (a) clearance, (b) inter-compartmental clearance, (c) volume of distribution of the central compartment, (d) volume of distribution of the peripheral compartment, (e) absorption lag-time and (f) absorption rate constant. * Introduced in the full regression model.

E.10 The use of Antacids

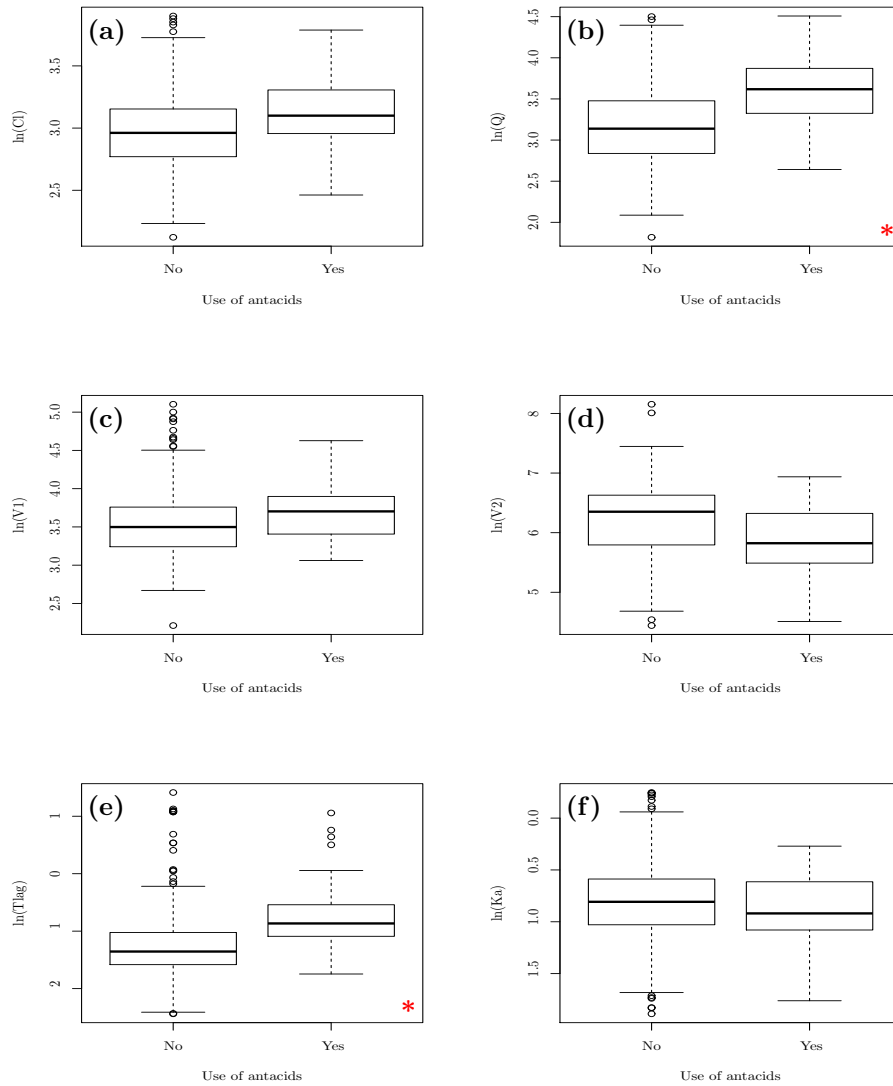


Figure 40: Individual effect of the use of antacids on the pharmacokinetic parameters (a) clearance, (b) inter-compartmental clearance, (c) volume of distribution of the central compartment, (d) volume of distribution of the peripheral compartment, (e) absorption lag-time and (f) absorption rate constant. * Introduced in the full regression model.

E.11 The use of Proton Pump Inhibitors

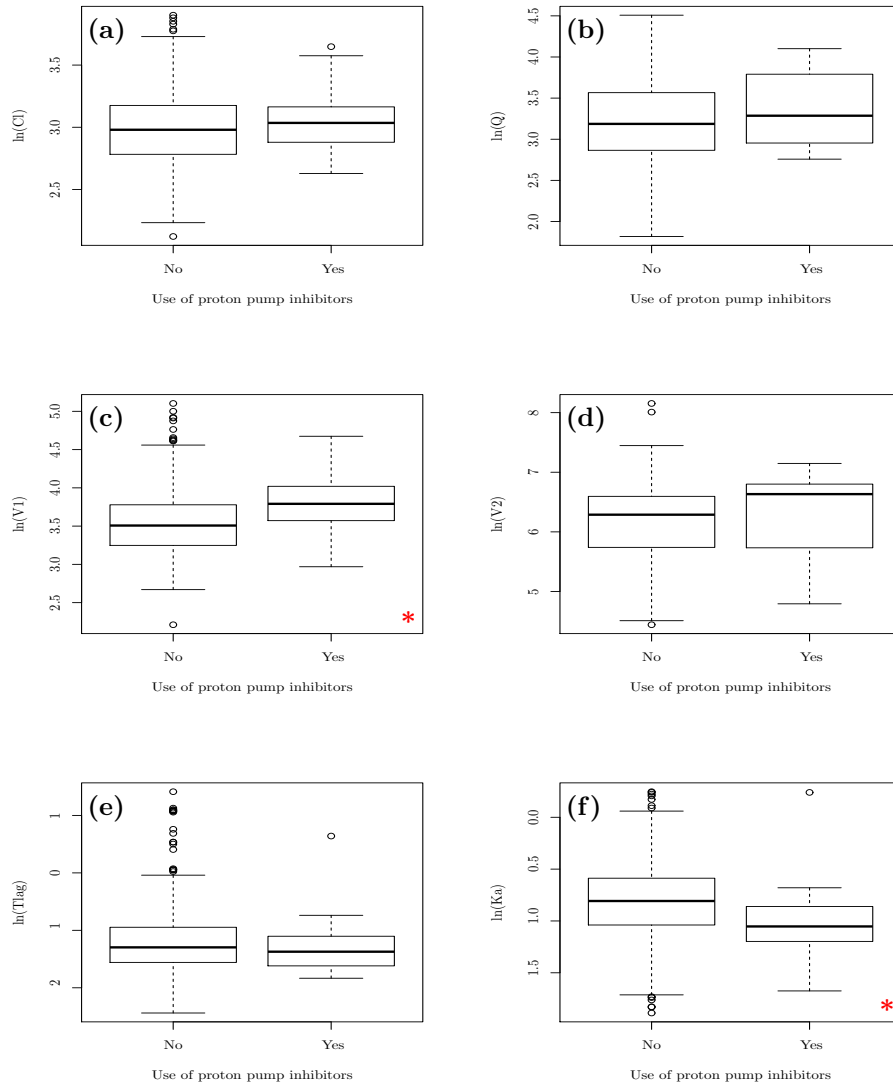


Figure 41: Individual effect of the use of proton pump inhibitors on the pharmacokinetic parameters (a) clearance, (b) inter-compartmental clearance, (c) volume of distribution of the central compartment, (d) volume of distribution of the peripheral compartment, (e) absorption lag-time and (f) absorption rate constant. * Introduced in the full regression model.

E.12 The use of Anti-Viral Agents

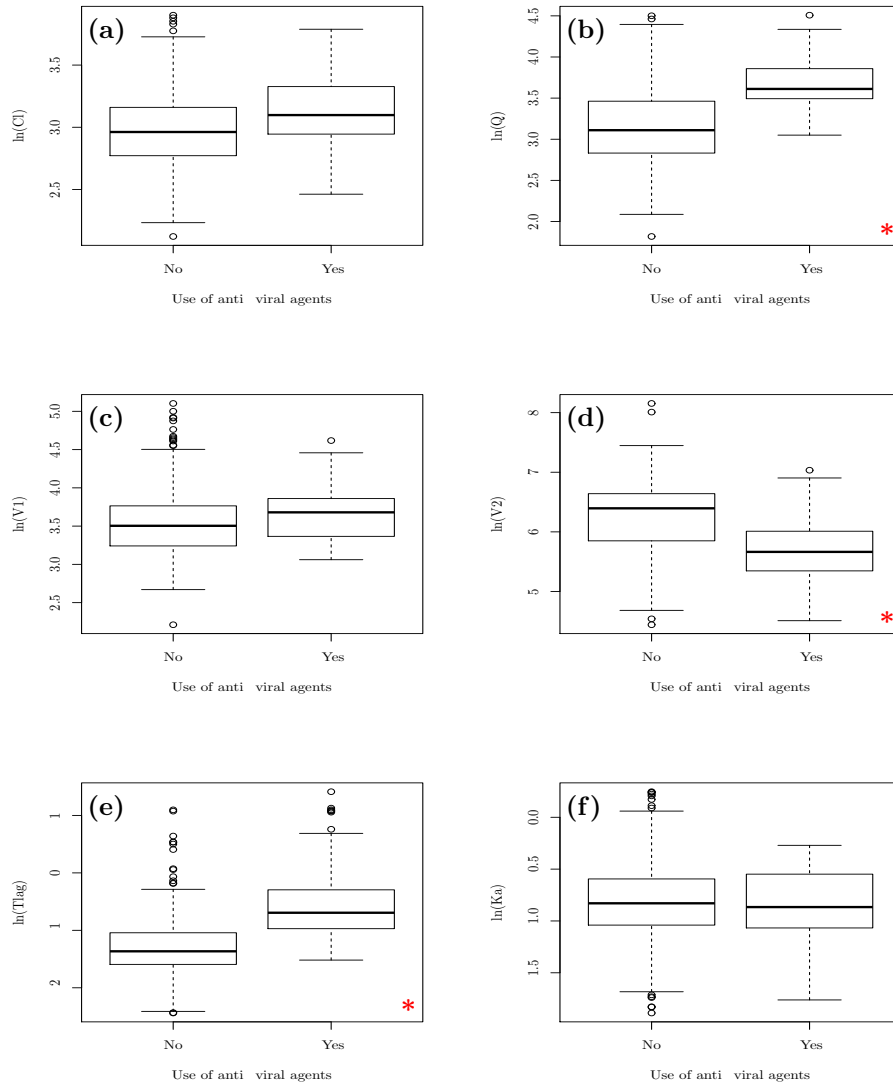


Figure 42: Individual effect of the use of anti-viral agents on the pharmacokinetic parameters (a) clearance, (b) inter-compartmental clearance, (c) volume of distribution of the central compartment, (d) volume of distribution of the peripheral compartment, (e) absorption lag-time and (f) absorption rate constant. * Introduced in the full regression model.

Bibliography

- [1] C Adams and V Brantner. Estimating the cost of new drug development: Is it really \$802 million? *Health Affairs*, 25:420–428, 2006.
- [2] JR Mitchell and MG Horning. *Drug Metabolism and Drug Toxicity*. Raven Press, 1984.
- [3] R Panchagnula and NS Thomas. Biopharmaceutics and pharmacokinetics in drug research. *International Journal of Pharmaceutics*, 201:131–150, 2000.
- [4] J DiPiro, W Spruill, W Wade, and R Blouin. *Concepts in Clinical Pharmacokinetics*. American Society of Health-System Pharmacists, 2005.
- [5] M Rowland and TN Tozer. *Clinical Pharmacokinetics Concepts and Applications*. Lippincott Williams & Wilkins, 1995.
- [6] P Halloran, T Mathew, S Tomlanovich, C Groth, L Hooftman, and C Barker. Mycophenolate mofetil in renal allograft recipients: A pooled efficacy analysis of three randomized, double-blind, clinical studies in prevention of rejection. *Transplantation*, 63:39–47, 1997.
- [7] LM Shaw, M Korecka, R Venkataramanan, L Goldberg, R Bloom, and KL Brayman. Mycophenolic acid pharmacodynamics and pharmacokinetics provide a basis for rational monitoring strategies. *American Journal of Transplantation*, 3:534–542, 2003.
- [8] LM Shaw, M Korecka, S Aradhye, R Grossman, L Bayer, C Innes, A Cucciarra, C Barker, A Naji, and A Nicholls. Mycophenolic acid area under the curve values in African American and Caucasian renal transplant patients are comparable. *Journal of Clinical Pharmacology*, 40:624–633, 2000.
- [9] DA Hesselink, RM Van Hest, RA Mathot, F Bonthuis, W Weimar, RW de Bruin, and T Van Gelder. Cyclosporine interacts with mycophenolic acid by inhibiting the multi-drug resistance-associated protein 2. *American Journal of Transplantation*, 5:987–994, 2005.
- [10] RM Van Hest, T Van Gelder, AG Vulto, and RA Mathot. Population pharmacokinetics of mycophenolic acid in renal transplant recipients. *Clinical Pharmacokinetics*, 44:1083–1096, 2005.
- [11] HW Sollinger. Mycophenolate mofetil for the prevention of acute rejection in primary cadaveric renal allograft recipients. *Transplantation*, 60:225–232, 1995.

- [12] The US Renal Transplant Mycophenolate Mofetil Study Group. Mycophenolate mofetil for the prevention of acute rejection in primary cadaveric kidney transplants: Status of the MYC 1866 study at 1 year. *Transplantation Proceedings*, 29:348–349, 1997.
- [13] T Van Gelder, LB Hilbrands, Y Vanrenterghem, W Weimar, JW de Fijter, JP Squifflet, Rj Hene, GA Verpooten, MT Navarro, MD Hale, and AJ Nicholls. A randomized double blind, multicenter plasma concentration study of oral mycophenolate mofetil for the prevention of acute rejection after kidney transplantation. *Transplantation*, 68:261–266, 1999.
- [14] MD Hale, AJ Nicholls, RE Bullingham, R Hene, A Hoitsma, JP Squifflet, W Weimar, Y Vanregenterghem, FJ Van de Woude, and GA Verpooten. The pharmacokinetic-pharmacodynamic relationship for mycophenolate mofetil in renal transplantations. *Clinical Pharmacology and Therapeutics*, 64:672–683, 1998.
- [15] H Ekberg, J Grinyo, B Nashan, Y Vanrenterghem, F Vincenti, E Calleja, C Nasmyth-Miller, and M Truman. The use of daclizumab and mycophenolate mofetil in combination with corticosteroids and cyclosporine (low dose versus low dose followed by withdrawal) to optimize renal function in recipients of renal allografts. *Transplantation*, 78(2):458–459, 2004. On behalf of the CAESAR study group.
- [16] MD Pescovitz, A Guasch, R Gaston, P Rajagopalan, S Tomlanovich, S Weinstein, GL Bumgardner, L Melton, P Sanwald Ducray, L Banken, J Hall, and BX Boutouyrie. Equivalent pharmacokinetics of mycophenolate mofetil in African-American and Caucasian male and female stable renal allograft recipients. *American Journal of Transplantation*, 3:1581–1586, 2003.
- [17] RM Van Hest, RA Mathot, MD Pescovitz, R Gordon, RD Mamelok, and T Van Gelder. Explaining variability in mycophenolic acid exposure to optimize mycophenolate mofetil dosing: A population pharmacokinetic meta-analysis of mycophenolic acid in renal transplant recipients. *Journal of the American Society of Nephrology*, 17:871–880, 2006.
- [18] CH Edwards and DE Penney. *Calculus*. Prentice Hall, 6 edition, 2002.
- [19] P Rowe. *Pharmacokinetics*. Ventus Publishing ApS, 2012.
- [20] L Lunney. Pharmacokinetics: A look at different approaches. Master's thesis, North Carolina State University, 1992.
- [21] S Dhillon and A Kostrzewski. *Clinical Pharmacokinetics*. Pharmaceutical Press, 2006.

- [22] G Doetsch. *Introduction to the Theory and Application of the Laplace Transformation*. Springer-Verlag, 1974.
- [23] LB Sheiner and SL Beal. Evaluation of methods for estimating population pharmacokinetic parameters. I. Michaelis-Menten model; routine clinical pharmacokinetic data. *Journal of Pharmacokinetics and Biopharmaceutics*, 8:553–571, 1980.
- [24] LB Sheiner and SL Beal. Evaluation of methods for estimating population pharmacokinetic parameters. II. Biexponential model; experimental pharmacokinetic data. *Journal of Pharmacokinetics and Biopharmaceutics*, 9:635–651, 1981.
- [25] LB Sheiner and SL Beal. Evaluation of methods for estimating population pharmacokinetic parameters. III. Monoexponential model: routine clinical pharmacokinetic data. *Journal of Pharmacokinetics and Biopharmaceutics*, 11:303–319, 1983.
- [26] LB Sheiner, B Rosenberg, and KL Melmon. Modelling of individual pharmacokinetics for computer-aided drug dosage. *Computers and Biomedical Research*, 5:411–459, 1972.
- [27] LB Sheiner, B Rosenberg, and VV Marathe. Estimation of population characteristics of pharmacokinetic parameters from routine clinical data. *Journal of Pharmacokinetics and Biopharmaceutics*, 5:445–479, 1977.
- [28] Population Pharmacokinetic Working Group. Guidance for industry population pharmacokinetics. Technical report, Center for Drug Evaluation and Research (CDER) and The Food and Drug Administration (FDA), 1999.
- [29] L Aarons. Population pharmacokinetics: Theory and practice. *British Journal of Clinical Pharmacology*, 32:669–670, 1991.
- [30] ME Burton, LM Shaw, JJ Schentag, and WE Evans. *Applied Pharmacokinetics and Pharmacodynamics*. Lippincot Williams & Wilkins, 2006.
- [31] LB Sheiner and SL Beal. The NONMEM system. *The American Statistician*, 34:118–119, 1980.
- [32] MJ Lindstrom and DM Bates. Nonlinear mixed effects models for repeated measures data. *Biometrics*, 46:673–687, 1990.
- [33] R Wolfinger. Laplace’s approximation for nonlinear mixed models. *Biometrika*, 80:791–795, 1993.

- [34] AE Gelfand and AFM Smith. Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, 85:398–409, 1990.
- [35] S Geman and D Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6:721–741, 1984.
- [36] N Metropolis, AW Rosenbluth, MN Rosenbluth, AH Teller, and E Teller. Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, 21:1087–1092, 1953.
- [37] WK Hasting. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57:97–109, 1970.
- [38] E Lesaffre and AB Lawson. *Bayesian Biostatistics (Statistics in Practice)*. John Willey & Sons, 2012.
- [39] R Neal. Suppressing random walks in Markov chain Monte Carlo using ordered overrelaxation. Technical report, Technical Report 9508, 1995.
- [40] P Heidelberger and Welch P. Simulation run length control in the presence of an initial transient. *Operations Research*, 31:1109–1144, 1983.
- [41] A Gelman and DB Rubin. Inference from iterative simulation using multiple sequences. *Statistical Science*, 7:457–511, 1992.
- [42] S Brooks and A Gelman. General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, 7:434–455, 1998.
- [43] DJ Spiegelhalter, NG Best, BP Carlin, and A Van Der Linde. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(4):583–639, 2002.
- [44] W Gilks, A Thomas, and D Spiegelhalter. A language and program for complex Bayesian modelling. *The Statistician*, 43:169–177, 1994.
- [45] DJ Lunn, J Wakefield, A Thomas, N Best, and D Spiegelhalter. *PKBUGS User Guide*, 1999.
- [46] DJ Lunn, N Best, A Thomas, J Wakefield, and D Spiegelhalter. Bayesian analysis of population PK/PD models: General concepts and software. *Journal of Pharmacokinetics and Pharmacodynamics*, 29(3):271–307, 2002.

- [47] DJ Lunn and LJ Aarons. Markov chain Monte Carlo techniques for studying interoccasion and intersubject variability: Application to pharmacokinetic data. *Journal of Applied Statistics*, 46(1):73–91, 1997.
- [48] DJ Lunn and LJ Aarons. The pharmacokinetics of saquinavir: A Markov chain Monte Carlo population analysis. *Journal of Pharmacokinetics and Biopharmaceutics*, 26(1):47–74, 1998.
- [49] RM Van Hest, T Van Gelder, AG Vulto, and RAA Mathot. Population pharmacokinetics of mycophenolic acid in renal transplant recipients. *Clinical Pharmacokinetics*, 44:1083–1096, 2005.
- [50] BCM De Winter, T Van Gelder, P Glander, D Cattaneo, H Tedesco-Süva, I Neumann, L Hilhrands, RM Van Hest, MD Pescovitz, K Budde, and RAA Mathot. Population pharmacokinetics of mycophenolic acid. *Clinical Pharmacokinetics*, 47:827–838, 2008.
- [51] BCM De Winter, RAA Mathot, F Sombogaard, AG Vulto, and T Van Gelder. Nonlinear relationship between mycophenolate mofetil dose and mycophenolic acid exposure: Implications for therapeutic drug monitoring. *Clinical Journal of the American Society of Nephrology*, 6:656–663, 2011.
- [52] G Casella and RL Berger. *Statistical Inference*. Duxbury, 2002.
- [53] Y Vanrenterghem, C Ponticelli, JM Morales, D Abramowicz, K Baboolal, B Eklund, V Kliem, C Legendre, Sarmiento ALM, and F Vincenti. Prevalence and management of anemia in renal transplant recipients: A European survey. *American Journal of Transplantation*, 3(7):835–845, 2003.
- [54] C Ponticelli and P Passerini. Gastrointestinal complications in renal transplant recipients. *Transplant International: Official Journal of the European Society of Organ Transplantation*, 18(6):643–650, 2005.
- [55] BC Weikert and EA Blumberg. Viral infection after renal transplantation: Surveillance and management. *Clinical Journal of the American Society of Nephrology*, 3:S76–S86, 2008.
- [56] BL Kasiske, JJ Snyder, D Gilbertson, and AJ Matas. Diabetes mellitus after kidney transplantation in the United States. *American Journal of Transplantation*, 3(2):178–185, 2003.
- [57] AJ Boeckmann, LB Sheiner, and SL Beal. *NONMEM Users Guide: Part V*. University of California, 1994.
- [58] M Davidian and DM Giltinan. *Nonlinear Models for Repeated Measurement Data*. Chapman & Hall, 1995.

- [59] JC Pinheiro and DM Bates. *Mixed-Effects Models in S and S-PLUS*. Springer, 2000.
- [60] M Hussain, N Fazal, S Nichani, G Lawson, RK Firmin, and DR Upton. Population pharmacokinetics of theophylline during paediatric extracorporeal membrane oxygenation. *British Journal of Clinical Pharmacology*, 55:23–31, 2003.