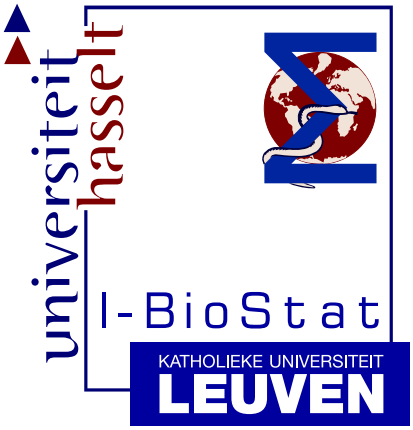
Non-linear mixed-effects modeling for complex biopharmaceutical data

Proefschrift voorgelegd tot het behalen van de graad van
Doctor in de Wetenschappen, richting Wiskunde
te verdedigen door

TOM JACOBS

Promotor: Prof. dr. Geert Molenberghs



Dankwoord

When I finished the biostat program at the Limburgs Universitair Centrum back in 1999, I didn't expect to return six years later to what is now called the Hasselt University. Even at the start of this mission, I was not fully confident that I would be successful; All the statisticians with a PhD that I knew at that time were bright students with excellent results and a thorough understanding of the broad field of statistics. Clearly, I didn't fit in this picture in 1999. Having spent six years in industry, I finally had a good understanding of what I would call basic modeling concepts and pharmacology, with a special thank you to Heidi De Smedt. When I saw the announcement for a position at the Hasselt University at the end of 2004, I knew that I had to apply for it; If not, I would regret it for the rest of my life. However, would I manage to survive these four years successfully?

Fortunately this wasn't a one-man's job and I received valuable input of many people. It is impossible to list all of them, but I would like to mention some explicitly. In the first place there is of course Geert Molenberghs. Geert, you always gave me so many new ideas during our meetings that I could continue for several months with these views from a new angle to solve the problem. I am still astonished about how fast you manage to help me out with any problem I had or paper I sent for review. Even a crash of a hard disk didn't stop you to send me your comments from South America.

Enthusiasm and knowledge are, however, insufficient to start a PhD. Funding is at least as crucial. Luc Bijnens, your combined scientific curiosity and management skills not only ensured the funding, but you also challenged me to think outside the box. Further, you forced me to focus on scientific results by submitting abstracts for conferences without having the draft paper yet. Without this, I would probably still work on my second draft paper. Also, it was you who got me into contact with Adrian Dunne. Adrian, thank you for teaching me the concepts of convolution-based IVIVC modeling and helping me with this research. You always managed to find

the weaknesses and strengths in my work and helped tremendously to improve the structure of my papers.

Of course, I didn't spent my days at Janssen Pharmaceutica playing only petanque, I also had many interesting discussions about statistics, pharmacology and various other not research-related topics such as Monty Pyton and the importance of the number 42. Although listing all these people would be a chapter of its own, I would like to mention Filip De Ridder, Roel Straetemans, Bart Michiels, Sarah Rusch, An Vermeulen, Stefaan Rossenu, Erik Mannaert, Achiel Van Peer, Adriaan Bouwknecht and the IVIVC working group for their valuable contribution to turn my visits to Beerse into both a scientific and inter-personal success.

I also enjoyed working with the people of Censtat. The transition from industry back to university was strange, but smoothed by the many discussions we had among the "Janssen boys", in particular with Jan and Kristien. On the other hand, the dullness of the former teaching room C101B was compensated excellently by the presence of the roommates. Thank you Tetyana, Philippe and Auguste for all the funny discussions and weird talks over the last four years.

Als laatste wil ik de voor mij belangrijkste mensen bedanken. Hanne en Maarten, bedankt dat jullie me er elke dag aan herinnerden dat het uiteindelijk maar werk is. Door met jullie te spelen en voor jullie te zorgen kon ik elke dag alles eens van me af zetten, het in perspectief plaatsen en even afstand nemen. Debby, je hebt me van in het begin gesteund, zowel als ik weer eens gefrustreerd was als een programma raar deed, of een reviewer in mijn ogen weer eens de bal volledig mis sloeg, maar ook als ik je weer eens alleen thuis achter liet met de kindjes als ik op congres was. Bedankt om er altijd voor me te zijn.

Tom Jacobs
Diepenbeek, 20 May 2009

Contents

List of Tables

List of Figures

1

Introduction

The (human) body is one of the most particular and complex systems to our knowledge. It is in the human nature to investigate, understand, and mimic the world surrounding him. As such, a substantial amount of research is attributed to understanding the human body. One part of this study consists of how the human body handles substances administered one way or another. This type of study is known as pharmacokinetics. It is one of the corner stones of current pharmaceutical research and has wide applications from dose finding, drug interactions, formulation development, bioequivalence up to many advanced pharmaceutical modelling. Its key elements is briefly introduced in Section 1.1. The main part of this dissertation handles about drug formulations with a specific administration, i.e., it is controlled, altered, or delayed. Section 1.2 introduces briefly the concepts and pitfalls of the analysis of such controlled-release formulations using *In Vitro-In Vivo Correlation* (IVIVC) models, which is a technique that relates an *in-vitro* property of the formulation to the *in-vivo* behavior. Section 1.3 gives a short overview of the contents of the dissertation.

1.1 Pharmacokinetics

During drug development and after registration of the drug candidate, patients are exposed to drug products, in order to the benefit of the patients' condition. As the

administration of such drug substances has an impact on the body, it is important to understand their behavior over time. Two time evolutions can be considered: the time evolution of the systemic exposure itself is called pharmacokinetics, i.e., it studies what the body does to the drug substance over time (Gibaldi and Perrier, 1982). This is in contrast to pharmacodynamics, where one studies what a drug substance does to the body over time Gabrielson and Weiner (2000). The latter consists of studying any beneficial or adverse effect, such as blood pressure changes, receptor binding (Kenakin, 1997), etc. Although the human body is very complex, it is important to translate these processes into mathematical and statistical models to allow for a better understanding and to be able to predict. For a more thorough introduction to pharmacokinetics, I refer to Gabrielson and Weiner (2000), Wagner (1975), and Gibaldi and Perrier (1982).

In the case of an orally administered drug formulation, the capsule or tablet dissolves and becomes available for the body. Four processes can be identified from this point onwards and pharmacokinetics can be considered as the study of this so-called "ADME" of a drug substance (Eddershaw, Beresford and Bayliss, 2000). This abbreviation means:

- absorption: how does the body transfer the extravascular administration of the drug substance into the bloodstream?
- distribution: how is the drug substance distributed throughout the entire body?
- metabolism: is the drug substance broken down to more soluble molecules that are easier to eliminate? This process involves the liver enzymes.
- excretion: how is the drug substance eliminated from the body? The main elimination path is via the kidneys for most compounds.

One can obtain a good understanding by repeated plasma sampling and analyzing the plasma concentrations as a time profile. In general, such a time profile is characterized by a steep increase corresponding primarily to the absorption phase, followed by the gradual decay mainly due to metabolism and elimination.

One can classify the analysis of pharmacokinetic concentration-time profiles roughly in two classes, a non-compartmental and a compartmental analysis. A non-compartmental analysis is most frequently used in standard analyzes. It summarizes the concentration-time profile to a number of summary measures, such as an Area Under the Curve (AUC), maximum concentration ($C_{\max}$), time to maximum concentration ($t_{\max}$), total apparent body clearance (CL_{F}), etc (Rowland and Tucker, 1980).

In a compartmental analysis, one assumes that the plasma concentration-time profile behaves according to a mathematical model (Gabrielson and Weiner, 2000). These parametric models describe the heterogeneity and the rates with which the drug substance is changing over time in the body. Not only is the drug substance eliminated from the body at a certain rate, it is also distributed within the body at a certain rate. If the rate from one part of the body back to the bloodstream differs from the other transfer rates, the body behaves in a heterogeneous way. Mathematically, different parts of the body with the same transfer rates are grouped into compartments. Such compartments are mathematical representations of the heterogeneity of the drug substance in the (human) body, such as blood versus tissue or the blood brain barrier. Typically, the change of drug substance from one compartment to another is characterized by a subject specific transfer rate constant.

As the plasma concentration-time profile evolves over time, it is natural to describe this change. Mathematically, change can be expressed as a differential equation (Wagner, 1975). Therefore, the concentration-time profile can be described by a differential equation. Following equation represents a one-compartment model, i.e., the human body behaves homogeneously, as an example (Gabrielson and Weiner, 2000).

$$\frac{dC}{dt}(t) = \frac{D}{V} I(t) - kC(t), \tag{1.1}$$

where $C(t)$ is the plasma concentration, D the administrated dose, V the volume of distribution, k the elimination rate constant, and $I(t)$ the input function depending on the mode of administration. Typical input functions $I(t)$ are 0 for an IV-bolus injection, and $\exp(-k_a t)$ for a first order absorption as for a tablet. Other input functions can be found in Gabrielson and Weiner (2000).

The use of the compartmental analysis is not only to describe and to understand, but more important to predict the drug exposure under different circumstances, such as in the case of multiple dosing, drug-drug-interactions, and special populations like hepatic impaired patients. Further, pharmacokinetic and pharmacodynamic data can then be combined to understand the dose-response relationship of clinical symptoms and benefits. All these models and applications share the same goal: to predict the drug exposure and, hopefully, the efficacy and toxicity related to it.

Pharmacokinetics (PK) and pharmacodynamics (PD) cannot be considered separately and Figure 1.1, taken from Gabrielson and Weiner (2000), shows how the different parts of pharmacokinetics and pharmacodynamics are entangled.

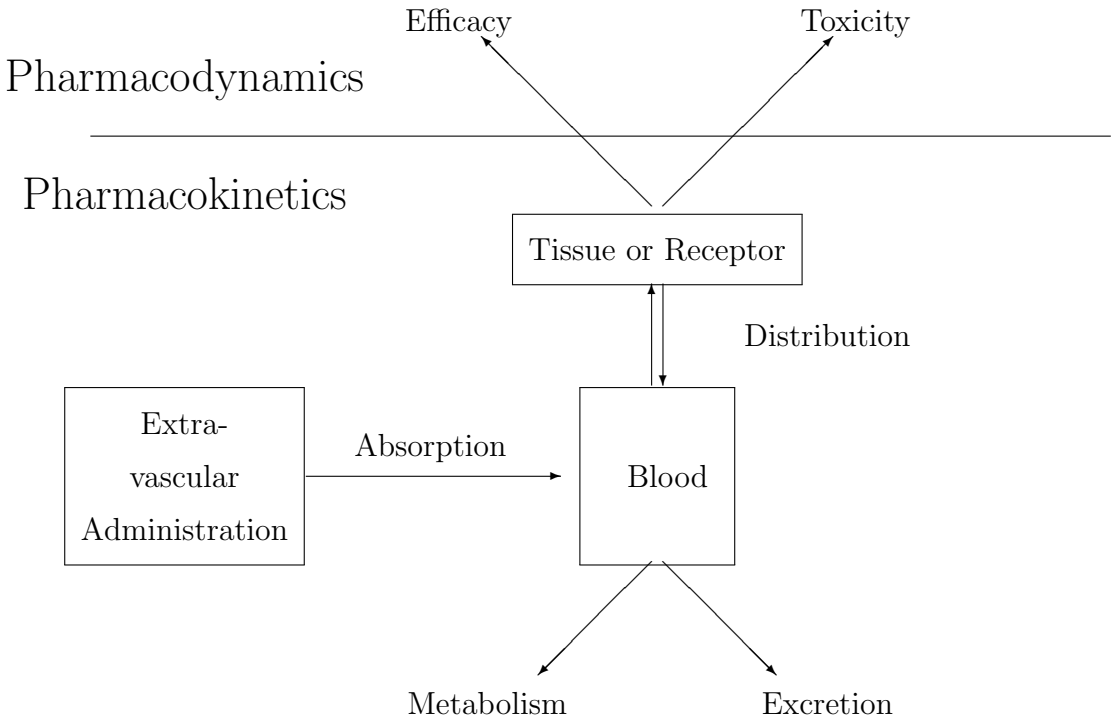


Figure 1.1: A schematic representation of the relation between pharmacokinetics and pharmacodynamics

1.2 In-vitro – In-vivo Correlation Modeling

From this point onwards the focus will be on a very specific form of pharmacokinetic analysis, called In-vitro – In-vivo Correlation modeling (IVIVC). Whereas in the previous section any drug product was considered, IVIVC modeling is limited to the case where the administration of an existing drug product is altered, spread over time, or even delayed (Emami, 2006). Such a formulation of a drug substance is typically referred to as a controlled-release formulation or a slow-acting formulation. Examples of such drug formulations are depot formulations, where the drug product is released gradually over time.

An IVIVC is defined as

> "A predictive mathematical model describing the relationship between an *in-vitro* property of an extended-release dosage form (usually the rate or extent of drug dissolution or release) and a relevant *in-vivo* response, e.g., plasma drug concentration or amount of drug absorbed."

in the FDA guideline (CDER 1997). This means that the release mechanism over time is tested in a laboratory setting, typically using a rotating basket or paddle. These are two of the most commonly adapted apparatus for the testing (Ju and Liaw, 1997). The obtained release time profiles are sometimes referred to as dissolution time profiles to indicate that it often represents the drug substance going into solution. A mathematical or statistical model is sought to relate these release profiles with the *in-vivo* systemic drug exposure such as the plasma concentration-time profiles in patients.

In FDA guideline (CDER 1997), one makes the distinction between four types of IVIVC correlation.

- Level A: It represents a point-to-point relationship between an *in-vitro* dissolution and the *in-vivo* release, i.e., using hierarchical modeling. The observed and model predicted plasma concentrations are compared directly. This means that individual plasma concentration-time profiles can be predicted with sufficient accuracy.

- Level B: In this case, the IVIVC the mean *in-vitro* dissolution and the mean *in-vivo* release profile are linked, typically using a marginal model. This means that only an average profile is predicted.

- Level C: The IVIVC establishes a single point relationship with a dissolution parameter. This can be a percentage dissolved at a specified time point versus a

non-parametric pharmacokinetic parameter. Whereas before the entire plasma concentration time curve is predicted for level A and B, the focus is here on the prediction of a summary parameter.

- Multiple Level C: As the name suggests, it relates some pharmacokinetic parameters to the amount of drug dissolved at several time points of the dissolution profile.

Only level A IVIVC models will be considered in this dissertation because of its clear superior predictive properties.

As the name and the definition state, the *in-vitro* properties of the drug are translated to the *in-vivo* setting. Therefore, it is very valuable for the pharmaceutical researchers: it enables to replace some *in-vivo* tests by its *in-vitro* equivalent. As such, it is a resource saving technique: *in-vitro* testing does not involve living organisms, is less expensive and often a less time consuming approach to demonstrate the required effects. For these reasons, it is frequently applied in formulation development and scale-up and post-approval changes (SUPAC). Therefore, it comes with no surprise that pharmaceutical researchers invest in developing an IVIVC model for their drug products. To illustrate its importance, the following is an incomplete list of drug products for which an IVIVC can be found in literature; acetaminophen (Dalton *et al.,* 2001), buspirone hydrochloride (Takka *et al.,* 2003), carbamazepine (Veng-Pedersen *et al.,* 2000), diltiazem (Sirisuth *et al.,* 2002), divalproex (Dutta *et al.,* 2005), doxorubicin (Chueng *et al.,* 2004), levosimendan (Kortejarvi *et al.,* 2006), metformin (Balan *et al.,* 2001), metoprolol (Mahayni *et al.,* 2000), (Sirisuth and Eddington, 2000), montelukast (Okumu *et al.,* 2008), ...

Some drug products have a fast elimination. There are two possibilities to ensure that the patient retains an efficacious exposure: give a very high dose once a day, or a frequent administration scheme. Both are however unacceptable. The high dose once a day may lead to unacceptable high drug exposures shortly after administration and thus may induce unwanted adverse effects, whereas a frequent administration scheme may not be practical for patients. Also compliance issues play its role here. However, if such a frequent administration of the drug product would be altered to a single administration with a release at multiple time points, the drug product may be preferred by patients and, therefore, correspond to a medical need. The use of a controlled-release formulation might also enable the development of a drug product with suboptimal pharmacokinetic properties. Rather than creating a multitude of controlled-release formulations, an optimal drug exposure is first determined. Then one derives the *in-vivo* release time profile required for the corresponding plasma

concentration-time profile. If an IVIVC is already established at this point, the optimal *in-vitro* properties can be determined directly. On the other hand, if no IVIVC model is established yet, a plausible relation between the *in-vitro* dissolution time profile and the *in-vivo* release time profile is chosen. As such, one can construct a limited number of controlled-release formulations to be tested in man. Without the use of such an IVIVC, a larger number of formulations would be required to ensure the identification of one leading to an optimal drug exposure. Thus a smaller number of drug formulations are tested in a clinical trial, which reduces the sample size and thus saves time, and resources.

The application of IVIVC is however not restricted to formulation development. The technique has also its use after marketing of the drug product. During development, drug product is produced according to certain procedures to obtain a relative small amount of drug product. However, these procedures may have to be altered to allow for large scale production. By changing these production procedures, one might inadvertent modify the release of the drug product and as such obtain a drug formulation with a different efficacy/safety profile. The impact of these scale-up changes to the production procedures can be assessed *in-silico*: One assesses the *in-vitro* dissolution time profile of the newly produced drug product and translates the impact of the production procedures to the potential effect on the *in-vivo* drug exposure. Here, one assumes that plasma concentration-time profiles can be used as a surrogate for the clinical effect.

Therefore, IVIVC is an inexpensive technique that can replace bioequivalence testing between batches and other SUPAC-related modifications, but also allows for the optimization of the drug exposure during drug development (Emami, 2006). However, not all drug products allow for the development of a controlled-release drug formulation.

The biopharmaceutics classification system (BCS) groups drug products in four classes according to the solubility and permeability of the drug product, as these govern the rate and extend of drug absorption for solid oral dosage forms (Amidon, G.L., Lennernas, H., Shah, V.P., and Crison, J.R. 1995). Permeability is determined as the ability of the drug molecule to permeate through the intestine into the systemic circulation. One can roughly group drug products according having good or bad solubility and good or bad permeability properties. BCS is considered as a guideline for determining when an IVIVC can be expected.

For drug products from BCS-1, the limiting factor is the gastric emptying. If the dissolution of the controlled-release formulation is slower than the gastric emptying, it is expected that a level A IVIVC can be established unless the permeability is site

Table 1.1: The biopharmaceutics classification system of drug products is based on the solubility and permeability of the drug product

	Good Solubility	Poor Solubility
Good Permeability	BCS-1	BCS-2
Poor Permeability	BCS-3	BCS-4

dependent.

The solubility and hence the dissolution of BCS-2 drugs is the main constraint for absorption. This can for example be caused by insufficient fluid in the intestines to allow for complete dissolution of the dose. In such cases, high doses cannot attain complete dissolution. For smaller doses, an IVIVC is expected depending the *in-vitro* properties.

The permeability and as such the absorption is the blocking factor for BCS-3, and no IVIVC would be expected. BCS-4 drug products have both absorption and solubility challenges and are not expected to form a good oral controlled-release formulation. No IVIVC is expected.

It is important to note that the BCS system is for oral dosage forms. The above does not apply to other administration modes such as depot formulations. A more extensive description of BCS, and applications of IVIVC in general, can be found in Emami (2006).

1.3 Thesis Overview

In what follows is an overview of the different chapters of the dissertation.

Two case studies are introduced in **Chapter 2**. The complexity and particularities of the data is described to familiarize with the setting of IVIVC modeling.

In **Chapter 3** the philosophy and concepts of IVIVC modeling are introduced. The IVIVC models in literature and traditional pharmacokinetic models are at first sight very different. The hidden relation between traditional pharmacokinetic models and IVIVC models is illustrated for a special case. This is then followed by a more formal description of current convolution-based IVIVC modeling. The chapter also contains methods for IVIVC model evaluation and practical aspects of IVIVC modeling.

IVIVC modeling traditionally restricts to homogeneous drug formulations, i.e., the drug unit is filled with one type of drug formulation. As such, all the drug product is released in a controlled way. A new model is developed in **Chapter 4** to cope with a heterogeneous drug formulation, i.e., where one part of the drug substance is released immediately and the remaining part is released gradually over time. The IVIVC model in this chapter uses a one-stage approach, i.e., all *in-vitro* and *in-vivo* data is modelled simultaneously. However, the convolution-based IVIVC modeling as introduced by O'Hara *et al.* (2001) fits first the immediate-release plasma concentration-time profile for each subject and subsequently imputes the obtained parameter estimates in the second stage of the model. In **Chapter 5**, the impact of such a two-stage modeling approach is assessed in comparison to a one-stage modeling.

One important part of modeling is model diagnostics. In the IVIVC setting, the model diagnostics are traditionally restricted to residual analysis and inspection of the Average Percentage Prediction Error ($\%PE$). We show in **Chapter 6** that this method is appropriate but insufficient to detect potential outlying subject's plasma concentration-time profiles. Local influence (Cook, 1986) is introduced as a suitable alternative approach. Other alternative techniques are based on the leave-one-out principle, see for example Sadray, Jonsson, and Karlsson (1999).

Now that the concepts of IVIVC modeling are clarified and illustrated, the relevance of IVIVC modeling to the research and development is illustrated in **Chapter 7**. In this chapter, a new application of IVIVC modeling is presented. The IVIVC model is used in combination with a PK/PD model to understand the clinical impact of changes in the *in-vitro* dissolution profile. As such, one can determine clinically relevant dissolution specifications. This expands the ideas of Hayes *et al.* (2004) from a bioequivalence setting towards a clinical setting.

One of the more traditional applications of IVIVC modeling is to prove that a new batch of the drug product is bioequivalent to the reference batch. This reference batch is typically the phase III batch. One claims that the efficacy and safety is maintained between the batches in case the *in-vivo* test shows bioequivalence or the *in-vitro* dissolution time-profile complies to the dissolution specifications. As such, the IVIVC model is a surrogate for the bioequivalence testing. However, bioequivalence testing (Schuirmann, 1987, Boddy *et al.,* 1995) often does not take into account the dose response relationship of the clinical effects. In **Chapter 8**, we extend bioequivalence testing to incorporate the dose response relationship via the therapeutic window. In contrast to current methods, our method complies to the FDA guideline (CDER 2003) by explicitly preventing therapeutic failure or unwanted side effects.

Up to this point, plasma concentration-time profiles are considered. One often

assumes that this is a good surrogate for the clinical effects of the drug product. Further, the physiology of the human body is taken into account to explain clinical effects. However, plasma concentrations are not always available. In **Chapter 9**, we demonstrate that plasma concentrations are not always required to fit a model that contains a physiological interpretation to the parameter estimates. Such models are referred to as K-PD models to stress the fact that only the pharmacodynamic response is measured without the plasma concentration-time profile (Jacqmin *et al.,* 2007). The potential bias of this model is illustrated in an example.

Finally, in **Chapter 10**, some concluding remarks regarding the different chapters are offered.

2

Case Studies

The novel methodology, models and concepts in the dissertation will be illustrated with some case studies. As the particularities of the data gave rise to reconsidering the current methodology, it is worth spending some time on these examples before discussing the methodology.

2.1 Galantamine

The acetylcholinesterase inhibitor Galantamine is used for the treatment of Alzheimer's disease (Lilienfeld 2002). Galantamine formulations currently on the market are tablets, a syrup and extended-release capsules.

Within the population of subjects with Alzheimer's disease, the duration of drug exposure can sometimes be too short to guarantee sufficient protection for a certain time period due to poor compliance. Therefore, a controlled-release formulation of Galantamine was developed in an attempt to optimize drug exposure. Whereas an immediate-release formulation dissolves instantaneously and the drug product is immediately available in the gut, an extended-release formulation releases the drug product slowly over time allowing the body to absorb the drug product gradually. The controlled-release formulation under investigation here consisted of the extended- and immediate-release components combined in the same pellet as 2 layers (ratio CR/IR: 3/1) separated by a rate-controlling membrane containing 5-12%

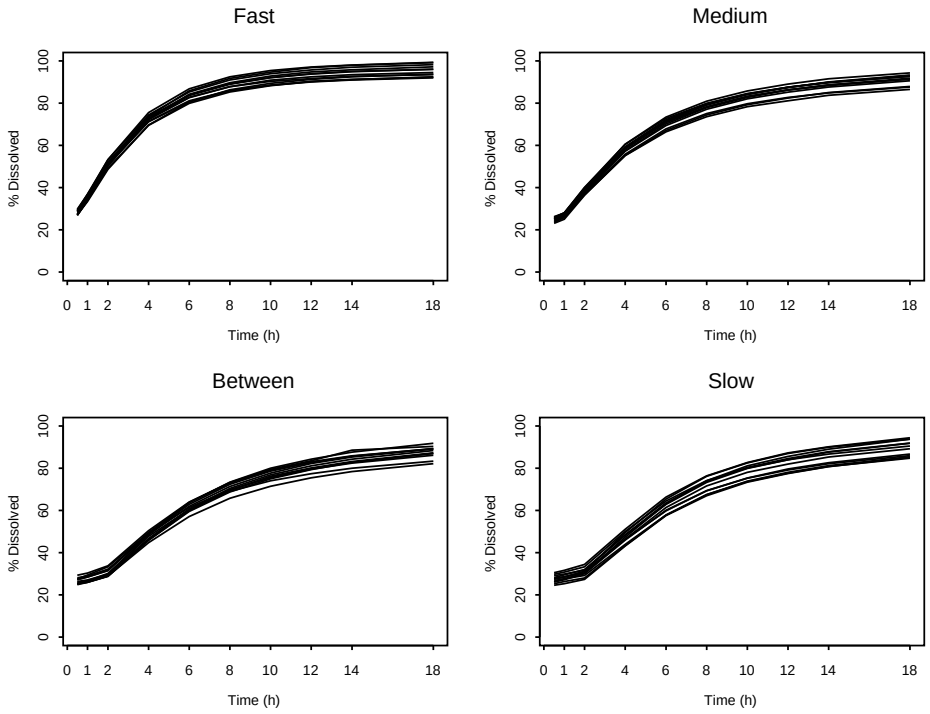


Figure 2.1: Galantamine: *In-Vitro* dissolution curves of twelve individual capsules of the four controlled-release Galantamine formulations

ethylcellulose/hydroxypropyl-methylcellulose (EC/HPMC; ratio: 75/25). The relatively high water solubility (3.3 g/100 $m\ell$ water, pH=5.2) and absolute oral bioavailability (88.5%) of Galantamine are pharmaceutical characteristics indicative of a drug whose controlled-release formulation is a good candidate for IVIVC exploration.

Four different controlled-release formulations were studied (slow, fast, between and medium), however for the sake of simplicity the focus is only on one controlled-release formulation (the slow one). For each controlled-release formulation, twelve dissolution curves were assessed *in-vitro*. The dissolution data were generated using an USP apparatus 2 - paddle with 50 rpm (s.e. 2 rpm) speed of shaft rotation. The dissolution medium used was a volume of 900 $m\ell$ of 0.050 M phosphate buffer at pH 6.5. The percentage dissolution was registered between 0.5 and 18 hours, as shown in Figure 2.1 for a controlled-release formulations.

In a clinical trial, seventeen healthy subjects were first administered to the immediate-

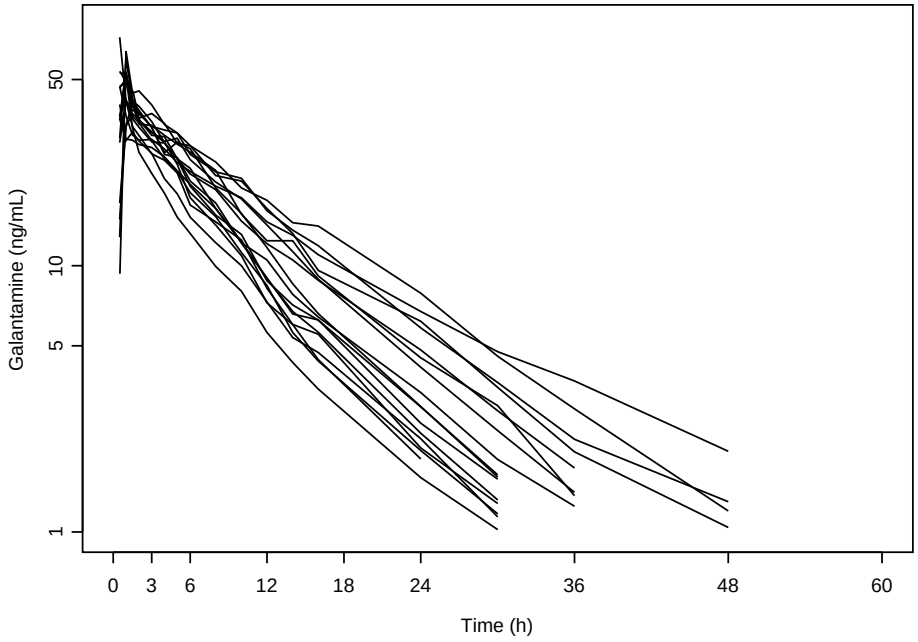


Figure 2.2: Galantamine: Individual *in-vivo* plasma concentrations for the immediate-release formulation of Galantamine.

release formulation and then randomized according to a four period latin square design. Treatments were the four controlled-release formulations (slow, fast, between and medium) of Galantamine. One subject dropped out after the immediate-release period. He did not receive the controlled-release formulations and was included as such in the analysis. To demonstrate our methodology in chapter 4, only one of the four controlled-release formulation, the slow one, is included in this analysis. A venous blood sample was taken for the measurement of Galantamine plasma concentrations at prespecified time points during the study, from pre-dose (0 hour) until 60 hours post-dose for the immediate-release formulation, and up to 72 hours post-dose for the controlled-release formulations.

The immediate-release plasma concentration-time data are shown in Figure 2.2, while the plasma concentration-time data for the controlled-release formulations are presented in Figure 2.3. In the former, maximal plasma concentrations were reached

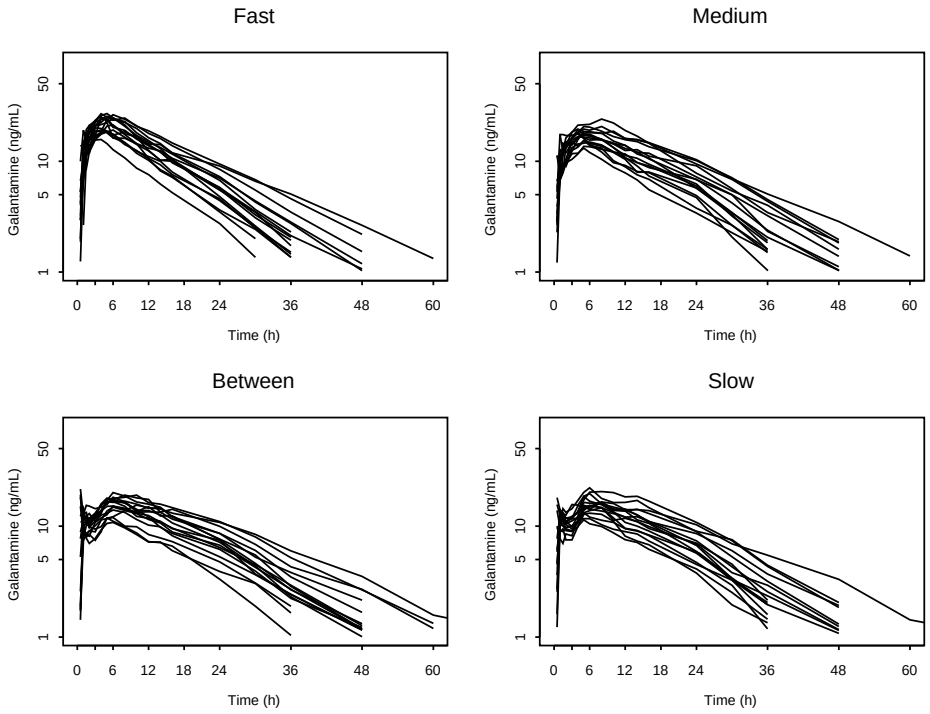


Figure 2.3: Galantamine: Individual *in-vivo* plasma concentrations for the four controlled-release formulations of Galantamine.

faster and were higher, but they decreased rapidly. In the latter, a bimodal profile was present: one steep peak was present after 30 minutes followed by a second smoother peak 6 hours after intake. In addition, the decrease of plasma concentration was slower after the second peak.

The advantage of combining the extended- and immediate-release formulation lies in this bimodal profile. The goal of the extended-release part is to ensure that patients remain in the effective plasma concentration range from 3-4 until 24 hours, and therefore it is hoped that the patients remain protected for the full 24 hours. The extended-release fraction on its own would not reach the therapeutic window quickly enough; levels would remain too low during the first 3 hours post-dose. Therefore, a loading dose consisting of an immediate-release fraction, is added. Hence, patients remain protected for the full 24 hours.

The current IVIVC methodology is restricted to homogeneous formulations. The

heterogeneity of the capsules resulted in a modification of the convolution-based IVIVC model as described in Chapter 4.

2.2 Long Acting Injectable Antipsychotic Agent

This second case study is a long acting injectable formulation for intramuscular administration for an antipsychotic agent with potent dopamine-D2 antagonistic properties. The IVIVC model will be combined with a PK/PD model to give a clinical interpretation to the *in-vitro* dissolution specifications. This novel concept is explained in Chapter 7. The data is also used to illustrate the local influence technique in Chapter 6.

The long acting injectable formulation consists of microspheres of a biologically degradable polymer matrix in which the active compound is embedded. These microspheres are to be reconstituted and consequently the obtained aqueous suspension can be injected in the gluteal or deltoid muscle. Gradual degradation of the polymer at the site of injection ensures a slow but steady release of the drug product over a period of several weeks. In single dose studies, the plasma concentrations of the active antipsychotic fraction showed release of a minor fraction of the active compound within 8 to 24 hours following the intramuscular injections of the long acting injectable. A gradual release of the main fraction of the long acting injectable started 2-3 weeks following injection, peaked at about 5 weeks and lasted until 7 weeks after the injection.

The controlled-release plasma concentration-time profiles used in our analysis were obtained from an open-label phase I trial in 54 patients. Plasma levels were measured for 12 weeks. The elimination transfer rates were estimated from the unit impulse response data (IV bolus formulation). These were obtained from another open 3-treatment cross-over bioavailability trial in 12 healthy subjects. For the controlled-release formulation, *in-vitro* release curves were obtained. The averaged percentage release was registered up to 43 days. The *in-vitro* specifications were determined as releasing 50% of the drug product within days 26–35, and 80% should be released before day 41. Figure 2.4 shows the controlled-release plasma concentration time profiles after the administration of a single dose. The majority of the subjects have a small peak shortly after administration followed by a peak after four to seven weeks. Four patients have a marked release during the first week in contrast to the other patients.

The unit impulse response data (IV-bolus formulation) was obtained from a differ-

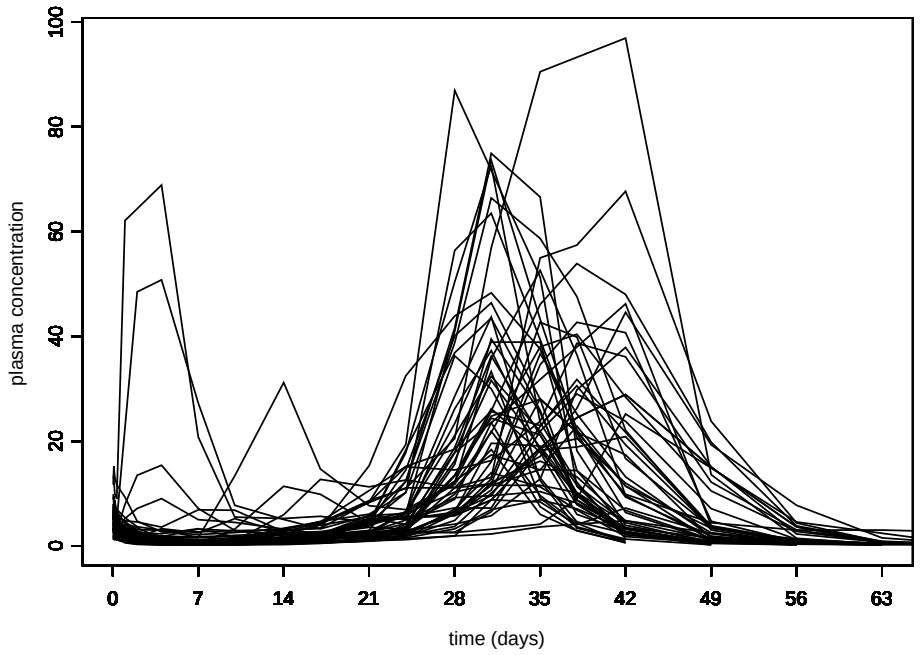


Figure 2.4: Long Acting Injectable: The controlled-release plasma concentration-time profiles following the single administration of the long acting injectable intramuscular formulation of the antipsychotic agent.

ent, open 3-treatment cross-over bioavailability trial in 12 healthy subjects. Pharmacokinetic parameters were derived from this small study under the assumption that the population of both studies are comparable. The data is shown in Figure 2.5.

For each controlled-release formulation, twelve dissolution curves were assessed *in-vitro*. The percentage dissolution was registered up to 43 days, as shown in Figure 2.6 for the controlled-release formulation. The graph clearly demonstrates that the drug product is released after several weeks unlike more conventional controlled-release drug formulations.

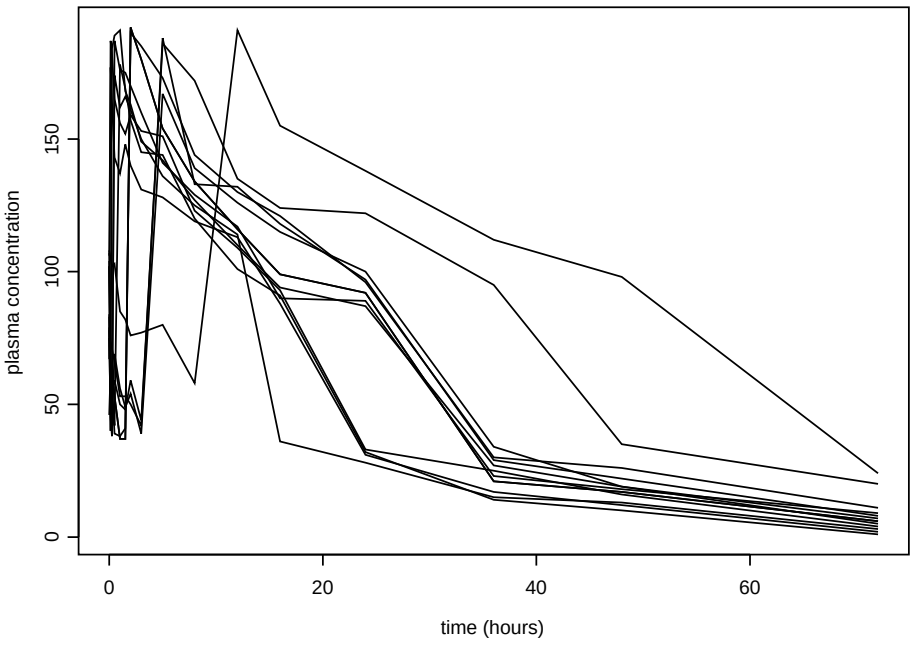


Figure 2.5: Long Acting Injectable: The immediate-release plasma concentration data used to establish the IVIVC model.

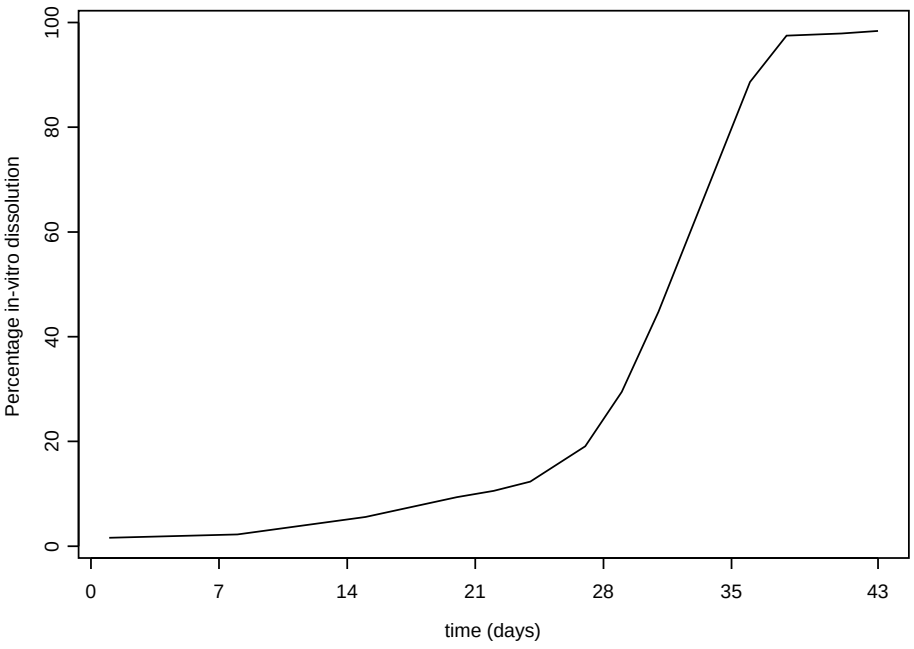


Figure 2.6: Long Acting Injectable: The averaged *in-vitro* dissolution time profile of the long acting injectable intramuscular formulation of the antipsychotic agent.

3

Methodology

As can be seen from the case studies and the definition, IVIVC modeling requires nonlinear hierarchical modeling approaches. Further more, it necessitates the combination of several responses, i.e., the *in-vitro* dissolution, the immediate-release and controlled-release plasma concentration-time profiles. This chapter focuses on the methodology needed for such an analysis. Before postulating complicated equations and model statements, the general concept of IVIVC modeling is introduced using an intuitive approach in Section 3.1. In this section, the IVIVC models will be translated to the standard PK-setting. The convolution-based models as introduced by Dunne *et al.* (1999) and O'Hara *et al.* (2001) are explained more formally in Section 3.2. The goodness of the model fit is quantified using the average percentage prediction error ($\%PE$). This metric is introduced in Section 3.3. Some pitfalls of this measure are listed. Finally, some modeling tricks are described in Section 3.4.

The following notation will be used in this dissertation. $Y_{i1\ell}(t)$ stands for the measured dissolution for the *in-vitro* data of drug-unit i at time t, $Y_{i2k\ell}(t)$ for the measured *in-vivo* plasma concentration of subject k after administration of drug-unit i of formulation ℓ at time t. The index 1 denotes the *in-vitro* data, while 2 will be used for *in-vivo*, i is the statistical unit representing the drug-unit for *in-vitro*; k denotes the subject; ℓ denotes the formulation. The immediate-release formulation will be denoted with δ instead of ℓ, owing to its special status in IVIVC modeling and to emphasize that the underlying probability density function of the release mechanism

follows, in this case, Dirac's Delta distribution. F will denote the actual dissolution/release fraction, c stands for the actual plasma concentration profile, $c_{2k\delta}(t)$ is the actual immediate-release plasma concentration-time profile, also referred to as the unit impulse response. This is traditionally, but not necessarily, based on a compartmental model.

3.1 Bridging PK and IVIVC

Before embarking on equations and models, the concept of controlled-release formulations is introduced based on a theoretical example. Further, it is noted that current literature on IVIVC modeling and the existing pharmacokinetics literature do not correspond. Therefore, the typical IVIVC models are bridged towards the traditional PK setting.

Gillespie and Veng-Pedersen (1985) proved that the controlled-release plasma concentration-time profile at time t, denoted by $c_{i2k\ell}(t)$, can be derived as the convolution of the unit impulse response $c_{2k\delta}(t)$, which is typically an IV bolus plasma concentrations, and the *in-vivo* release mechanism $F'_{i2k\ell}$,

$$c_{i2k\ell}(t) = D \int_0^t c_{2k\delta}(t-\tau) F'_{i2k\ell}(\tau) d\tau. \tag{3.1}$$

To have an intuitive understanding of this formula, it is important to understand how plasma concentrations accumulate after repeated administration. It is based on the principle of superposition within pharmacokinetics, i.e., the assumption that each mechanism acts independently of each other and there is linear kinetics. This is a crucial assumption in the IVIVC setting.

The third column in Figure 3.1 shows the repeated plasma concentration-time profile in the case of a one-compartment bolus injection (middle column). For the example chosen here, the plasma concentration-time profile decreases very rapidly. To ensure an effective drug exposure for 24 hours, a very high single dose should be administered as can be seen in the top row. However, the high plasma concentrations shortly after administration would potentially lead to toxicity. B.i.d. administration seems not to improve the situation as can be seen in the second row. The plasma concentrations of the second administration can be calculated easily as the addition of the equation at the time of administration to the single dose profile.

The administration every three hours in the lowest row seems more appropriate: a steady state condition is reached after the fifth administration. Additional administrations of the same dose with the same time interval would not change the pattern.

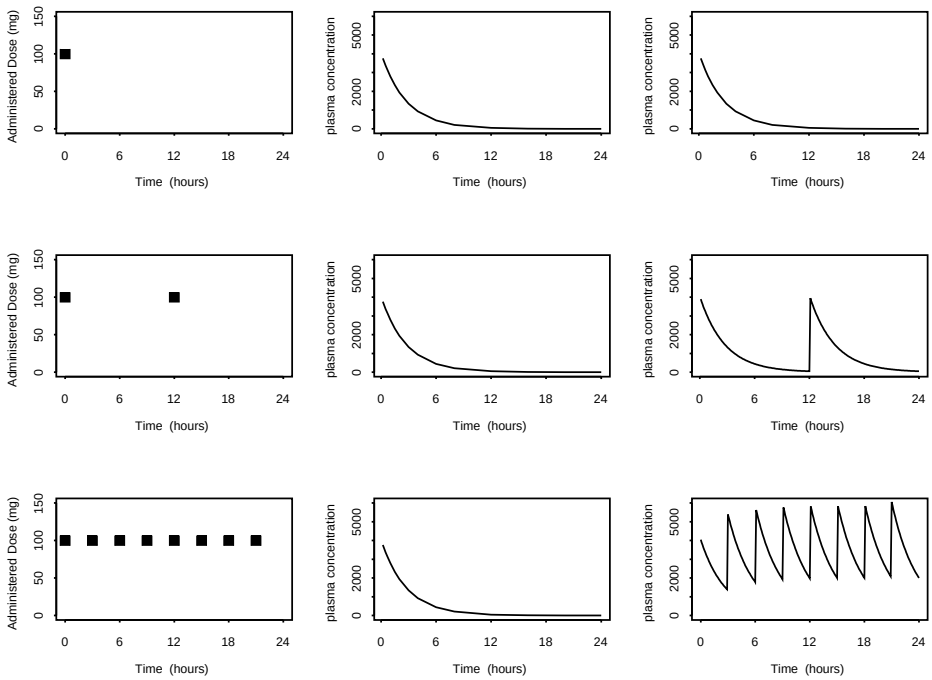


Figure 3.1: Plasma concentration-time profiles in the case of single, b.i.d. and repeated administration over time. The first column represents the administration time and dose, the second corresponds to the single dose plasma concentration, and the last column is the resulting plasma concentration-time profile.

Again the plasma concentrations from the next administration can be calculated as the iterative summation of the bolus equation for each administration.

However, it is not in the patients interest to have a drug product that has to be administered every three hours. Therefore, an option to save the drug product is the creation of a controlled-release formulation. Figure 3.2 shows the administration of 100 mg of the product spread over several administrations and dose levels. Again, the plasma concentration-time profile is the summation over the individual administrations. Whereas the first two rows represent a discrete administration, the last row corresponds to a continuous, controlled release over time. It is easy to understand that if the 100 mg is spread even more frequently over time that the plasma concentration-time profile becomes smoother. The mathematical consequence is that

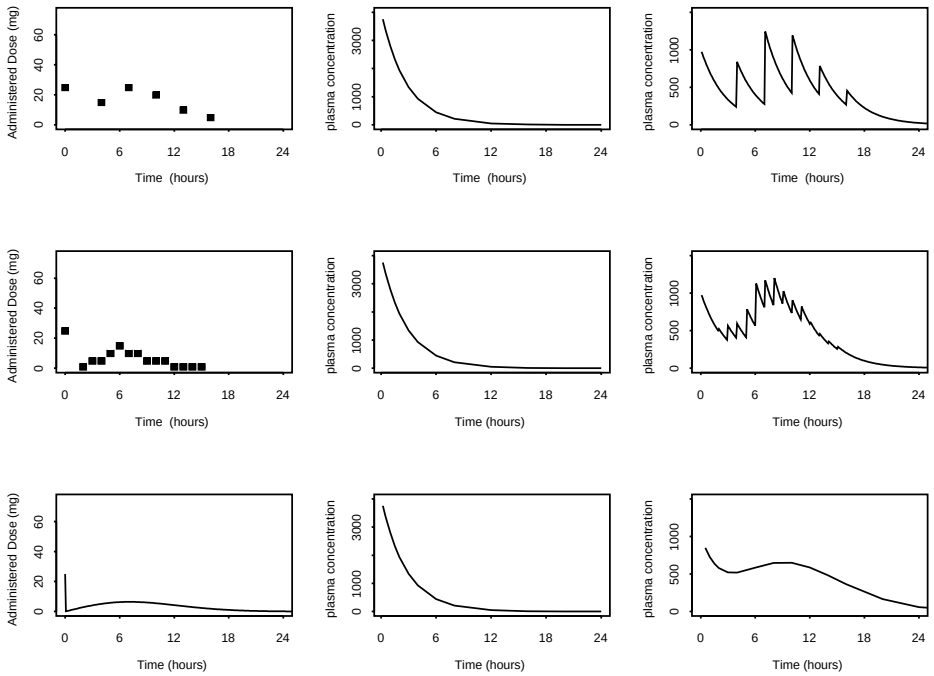


Figure 3.2: Plasma concentration-time profiles in the case of very discrete, more frequent and continuous administration over time.

the weighted summation of the bolus plasma concentration-time profile converges towards the integral found in (3.1).

The strong asset is the smoothness of the profile, the low maximal drug concentration and effective plasma concentrations for a much longer duration than in the case of the 100 mg bolus injection as seen in the middle column of the above figures.

One of the difficulties with IVIVC modeling is that (3.1) does not correspond to the typical derivative equations as found in pharmacokinetic modeling. To understand the relation between both forms of notation, the specific case of a one-compartment model is chosen for illustration purposes only.

Following equation as found on p. 302 in Khuri (1993) will be applied: If

$$G(x_2) = \int_{\lambda(x_2)}^{\theta(x_2)} f(x_1, x_2) dx_1,$$

then

$$\frac{dG}{dx_2} = \int_{\lambda(x_2)}^{\theta(x_2)} \frac{\partial f(x_1, x_2)}{\partial x_2} dx_1 + \theta'(x_2) f(\theta(x_2), x_2) - \lambda'(x_2) f(\lambda(x_2), x_2).$$

This reduces our model to

$$\begin{aligned}
\frac{dc_{i2k\ell}}{dt}(t) &= \frac{d}{dt} D \int_0^t c_{2k\delta}(t-\tau) F'_{i2k\ell}(\tau) d\tau \\
&= D \int_0^t \frac{\partial}{\partial t} (c_{2k\delta}(t-\tau) F'_{i2k\ell}(\tau)) d\tau + D c_{2k\delta}(0) F'_{i2k\ell}(t) \\
&= \frac{D}{V} F'_{i2k\ell}(t) + D \int_0^t F'_{i2k\ell}(\tau) \frac{\partial}{\partial t} c_{2k\delta}(t-\tau) d\tau \\
&= \frac{D}{V} F'_{i2k\ell}(t) - kD \int_0^t c_{2k\delta}(t-\tau) F'_{i2k\ell}(\tau) d\tau \\
&= \frac{D}{V} F'_{i2k\ell}(t) - k c_{i2k\ell}(t).
\end{aligned}$$

The obtained equation corresponds to (1.1), where $F'_{i2k\ell}(t)$ is the input function $I(t)$ representing the *in-vivo* release mechanism and V is the theoretic volume over which the dose D is spread. Even though the above derivation is for the specific case of a one-compartmental model, a similar derivation can be performed for higher order compartmental models. However, in that case, the amount of drug exposure is modeled rather than the plasma concentration. This amount of drug exposure is easy to understand: it is the total amount of drug product within a compartment and corresponds to the product of the drug concentrations and the volume of the corresponding compartment.

3.2 Convolution-based Models

Gillespie and Veng-Pedersen (1985) were the first to state that a controlled-release drug-product can be predicted based on the *in-vitro* dissolution data of the drug-unit and the *in-vivo* plasma concentration-time profile derived from the immediate-release formulation. They used the convolution product of these quantities. For decades, this method was applied using a deconvolution-based method. Deconvolution-based methods are, however, out of the scope of this dissertation. O'Hara *et al.* (2001) introduced a convolution-based method, which will be the method under study in this thesis.

An IVIVC model typically consists of three types of data: (1) the *in-vitro* dissolution time profile $Y_{i1\ell}(t) = F_{i1\ell}(t) + \varepsilon_{i1\ell}$, with $\varepsilon_{i1\ell} \sim N(0, \sigma_0^2)$, (2) the immediate-release plasma concentration profile $Y_{2k\delta}(t) = c_{2k\delta}(t) + \varepsilon_{2k}$ with $\varepsilon_{2k\delta} \sim LN(0, \sigma_1^2)$,

and (3) the controlled-release plasma concentration profile $Y_{i2k\ell}(t) = c_{i2k\ell}(t) + \varepsilon_{i2k\ell}$ with $\varepsilon_{i2k\ell} \sim LN(0, \sigma_2^2)$.

O'Hara *et al.* (2001) proposed to use a two-stage modeling approach. The first stage consists of modeling the immediate-release plasma concentration-time profile for each subject separately. Typically, a compartmental model is used for this unit impulse response. Then, the parameter estimates are incorporated in the second-stage modeling, where the *in-vitro* dissolution data and the controlled-release plasma concentrations time profile are modelled simultaneously. Dunne *et al.* (2005) and Gaynor *et al.* (2006) showed that this method is more robust and superior to the deconvolution-based methods.

As mentioned before, Gillespie and Veng-Pedersen (1985) proved that the controlled-release plasma concentration-time profile at time t, denoted by $c_{i2k\ell}(t)$, can be derived as the convolution of the unit impulse response $c_{i2k\delta}$ and the *in-vivo* release mechanism $F'_{i2k\ell}$,

$$\begin{aligned} Y_{i2k\ell}(t) &= D \int_0^t c_{2k\delta}(t-\tau) F'_{i2k\ell}(\tau) d\tau + \varepsilon_{i2k\ell}, \\ \varepsilon_{i2k\ell} &\sim LN(0, \sigma_2^2), \end{aligned}$$

where the *in-vitro* dissolution data $F_{i1\ell}$ is often modeled using a Weibull type of model (Comets and Mentré, 2001). However, the *in-vivo* release is required in the convolution model, rather than the *in-vitro* dissolution.

Therefore, Dunne *et al.* (1999) proposed to consider the *in-vitro* dissolution curve $F_{i1\ell}$ and the *in-vivo* release curve $F_{i2k\ell}$ both as the cumulative distribution function of the stochastic process representing the *in-vitro* dissolution of the molecule into solution and the *in-vivo* release, respectively. The quantity $F'_{i2k\ell}$, as part of the convolution model, stands for the density function corresponding to the underlying *in-vivo* mechanism of the molecule release into the solution. However, this latter quantity is unobserved. Dunne *et al.* (1999) proposed to use an hypothetical link between these two cumulative distribution functions. The most stringent relation

$$F_{i2k\ell}(t) = F_{i1\ell}(t) \tag{3.2}$$

corresponds to superimposable release mechanisms *in-vitro* and *in-vivo*. From a physiological point of view, this is a serious restriction as there is no a priori reason that the *in-vitro* setting mimics perfectly the *in-vivo* release. A more liberal relation relies on the concept of proportionality, such as the proportional odds, proportional hazards, and the proportional reversed hazards model.

$$g\left(F_{i2k\ell}\right)(t) = \alpha + g\left\{F_{i1\ell}\left(t\right)\right\} \tag{3.3}$$

Based on this interpretation, O'Hara (2001) incorporated following more liberal IVIVC link between the in-vitro and in-vivo release in the case of one formulation:

$$F_{i2k\ell}(t) = g^{-1}\left\{\theta_0 + \theta_1 t + s_{ik} + u_i + g[F_{i1\ell}(t)]\right\}, \tag{3.4}$$

with $u_i \sim N(0, \sigma_u^2)$ representing the dosage unit variability as found in the *in vitro* dissolution time profile; $s_{ik} \sim N(0, \sigma_s^2)$ the combined effect of both the subject and dosage unit variability. $g(\cdot)$ is either the logit, log-log, or the complementary log-log link function, although other link functions could be considered too. In this formula, the parameters θ_0, θ_1, and s_{ik} capture the influence of the gastrointestinal tract, whereas u_i and $g[F_{1\ell}(t)]$ correspond to the dissolution. Typically, the variability associated to *in-vitro* dissolution testing is very small, implying that u_i will be omitted in the simulations. The term s_{ik} will therefore correspond to a random intercept associated to the fixed effect θ_0 with variability $\sigma_{\theta_0}^2$.

3.3 Goodness-of-fit

Regulatory guidances (FIP 1996, CDER 1997) proposed that the adequacy of IVIVC model fitting is determined by the average absolute percent prediction error ($\%PE$), calculated both for $C_{\max}$ and AUC separately. This was defined as the mean of

$$\left|\frac{x_{obs,k} - x_{pred,k}}{x_{obs,k}}\right| \times 100, \tag{3.5}$$

where $x_{.,k}$ is the Area Under the Curve to the last measurable observation (AUC_{last}) or the maximal concentration ($C_{\max}$) of the empirical Bayes estimates and the observed concentrations for subject k.

The $\%PE$ of the IVIVC model fitting is classified into three categories (CDER 1997):

- $\%PE \leq 10\%$ establishes the predictability of the IVIVC if the $\%PE$ for each formulation does not exceed 15%;
- $10\% \leq \%PE \leq 20\%$ indicates inconclusive predictability and the need for further study using additional data sets;
- $\%PE > 20\%$ indicates inadequate predictability.

In the case of an IVIVC model with a fourth formulation, the model is created using only three formulations and the fourth one will serve as an external validation. Then a similar evaluation of the $\%PE$ occurs only using the external validation data.

Unfortunately it is not specified in the guidelines how to calculate the $\%PE$. Here are some possible ways to calculate it:

- geometric version: Calculate the $\%PE$ per subject and formulation. Calculate the average of the log-transformed quantities and back-transform it.
- arithmetic version: Calculate the $\%PE$ per subject and formulation and subsequently calculate the average.
- marginalized version: Marginalize the model by integrating out the random effects, subsequently calculated a marginalized AUC_{last} and $C_{\max}$. The observed plasma concentration are averaged over the subjects and the $\%PE$ is calculated by comparing the marginalized and the averaged observed values.
- median version: Set the random effects equal to zero to obtain a form of median subject profile. This is then compared with the AUC_{last} and $C_{\max}$ using the median of the observed plasma concentrations per time point.
- average version: Calculate the average of the observed and the average of the predicted observations. Use these to obtain the observed and predicted AUC_{last} and $C_{\max}$.

However, the statistical appropriateness of the different methods is questionable for the last four methods: The latter three are easy to rule out as they do not agree with the hierarchical modeling philosophy used in IVIVC modeling. Given the skewed distribution of the subject $\%PE$, the logarithm of the ratios approximates better normality. Therefore, only the geometric version is statistically appropriate and will be used in the rest of this thesis.

The $\%PE$ measure does not incorporate the measurement's uncertainty. Our first proposal for modification, relative to the existing literature, is to use the upper limit of the one-sided 95% confidence interval rather than the average as a criterion. It yields the range of the population mean with 95% certainty. This is clearly conservative when compared to the existing criterion. It ensures that most profiles have an adequate fit, whereas if the focus is on the mean, one might encourage overfitting of the data, in the sense that some profiles might fit extremely well, while the model fails to predict the profile for others.

Calculating the $\%PE$ on the original dataset, which is used for parameter estimation, is known to be kind for the model. Alternative approaches would be jackknife or leave-one-subject-out approach. Also the recent normalized prediction distribution errors (Brendel et al, 2006) is an option. However, given the long run-time of the

IVIVC models, bootstrapping methods or similar jeopardize the practical implementation of these methods.

Following simulated example illustrates the danger of judging the model fit solely on the $\%PE$. Assume following dissolution curve:

$$F_{i1\ell}(t) = 100\left(1 - \exp\left\{-\exp\left[\phi_1\left(t - \phi 2\right)\right]\right\}\right), \tag{3.6}$$

with $\phi_1 \sim N(0.14, 0.01^2)$, $\phi_2 = 36$, and a residual standard deviation of 0.5. The immediate-release plasma concentration-time profile is defined as

$$c_{2k\delta}(t) = \frac{D}{V}\exp(-k_e t), \tag{3.7}$$

where $D = 100$ is the dose, $\log(k_e) \sim N(-1, 0.42^2)$, $\log(V) = -3.7$, and a residual standard deviation of 0.01. For the IVIVC part of the model, $\theta_0 \sim N(1, 4^2)$ and $\theta_1 = 0$ was chosen. When fitting the data, the *in-vitro* release will be considered as the perfect predictor for the *in-vivo* release, i.e., the parameters θ_0 and θ_1 are fixed to 0. As this is an artificial example to illustrate the weakness of the criterion, this does not restrict the result. The model fit is depicted in figure 3.3. It is clear that the peak of the plasma concentration profile is systematically predicted too early or too late. Nevertheless, the geometric mean $\%PE$ is 9.87% and 8.23% for AUC and $C_{\max}$, respectively, indicating an adequate fit.

3.4 Practical Aspects

Pharmacokinetic models are highly non-linear models. Non-linear mixed-effects modeling exhibits a number of additional complexities compared to linear and generalized linear mixed-effects modeling, such as starting values for the iterative maximum likelihood optimization.

However, IVIVC modeling has an additional complexity. The controlled-release plasma concentrations are modelled as the convolution product of the immediate-release plasma concentration profile and the *in-vivo* release of the drug product. Typically, this convolution integral is too complicated to find an analytical solution. Therefore, the integral was approximated as a finite sum over small time intervals.

$$\begin{aligned} c_{i2k\ell}(t) &= \int_0^t c_{2k\delta}(t-\tau)F'_{i2k\ell}(\tau)d\tau + \varepsilon_2, \\ &\cong \sum_{t_i}^{t} \frac{\left[c_{2k\delta}(t-t_i)F'_{i2k\ell}(t_i) + c_{2k\delta}(t-t_{i-1})F'_{i2k\ell}(t_{i-1})\right]}{t_i - t_{i-1}} + \varepsilon_{i2k\ell}, \\ \varepsilon_{i2k\ell} &\sim N(0, \sigma_2^2). \end{aligned}$$

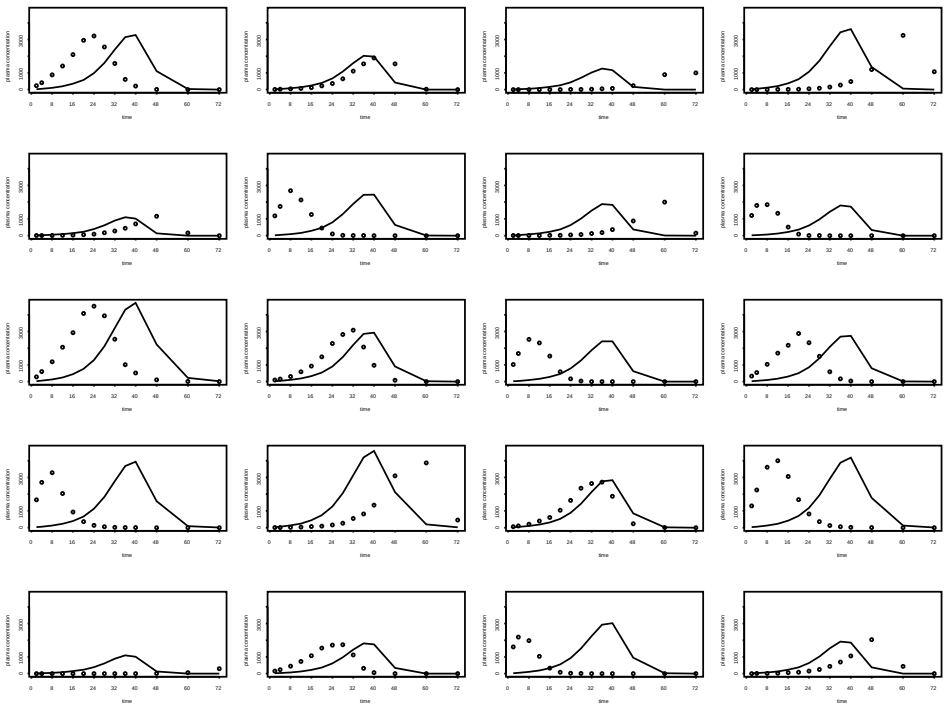


Figure 3.3: A simulated example of a poor model fit despite acceptable $\%PE$ for both AUC and C_{max}.

It is important not only to cope with the traditional complexities of non-linear mixed-effects modeling, but also to take into account the influence of the time intervals of this finite summation. Therefore, one should consider the use of different time intervals for the convolution model and its effect on the parameter estimates. This means that models with decreasing time intervals should be fitted until parameter estimates stabilize.

Throughout the thesis, maximum likelihood estimation of the set of models was implemented in the SAS procedure NLMIXED (version 9.1). Model convergence was obtained using the first-order integration method of Beal and Sheiner (1982).

4

Combined Models for Data From IVIVC Experiments

Contrary to other published results (Modi *et al.* 2000, Veng-Pedersen *et al.* 2000) a heterogeneous formulation is modelled in this chapter. As described in Chapter 2, the galantamine formulation contains an extended-release part overencapsulated with immediate-release material, and will be referred to as controlled-release capsule in the remainder of the chapter. For illustration purposes, the analysis is limited to the slowest release formulation. The formulation index ℓ is left out to simplify the notation in this chapter. The present chapter is based on Jacobs *et al.* (2008).

4.1 Methodology

First, four types of models used for describing the *in-vitro* dissolution curves are introduced in Section 4.1.1. Then a novel *in-vivo* convolution-based IVIVC methodology coping with heterogeneous drug products is described in Section 4.1.2; This model is fitted on the galantamine data in Section 4.1.3.

4.1.1 *In-vitro* Dissolution Models

The *in-vitro* dissolution profile is often described by a Weibull function (Comets and Mentré, 2001). Besides the Weibull function, also a simpler exponential function and a more complex Gompertz *in-vitro* dissolution model will be evaluated.

The simplest model for the *in-vitro* dissolution profile is given by the exponential model:

$$\begin{aligned} Y_{i1}(t) &= F_{i1}(t) + \varepsilon_{i1}, \quad \varepsilon_{i1} \sim N(0, \sigma_0^2), \\ F_{i1}(t) &= \phi_1 \left\{1 - \exp[-(t - \phi_3)\phi_{2i}]\right\}, \end{aligned} \tag{4.1}$$

with $0 < \phi_1 \leq 1$, $\phi_{2i} \sim logN(\phi_2, \sigma_{\phi_2}^2)$, and ϕ_{2i} the capsule-specific scale parameter and ϕ_3 a lag time. This model has a steep increase in the beginning and converges slowly to the asymptotic maximal dissolution, ϕ_1.

The following extension of the exponential model copes with the heterogeneity of the formulation via the ϕ_4-parameter.

$$\begin{aligned} Y_{i1}(t) &= F_{i1}(t) + \varepsilon_{i1}, \quad \varepsilon_{i1} \sim N(0, \sigma_0^2), \\ F_{i1}(t) &= \phi_4 + (\phi_1 - \phi_4) \left\{1 - \exp[-(t - \phi_3)\phi_{2i}]\right\}, \end{aligned} \tag{4.2}$$

where ϕ_4 captures an initial jump followed by the previous version of the exponential model.

The previous models, however, lack the capability to fit a sigmoidal curvature. Therefore, the traditional Weibull function with the initial jump ϕ_4 is proposed to check for the improvement under these conditions, see model 4.3. The parameter ϕ_2 has the same interpretation as for the exponential model, whereas ϕ_3 determines the shape.

$$\begin{aligned} Y_{i1}(t) &= F_{i1}(t) + \varepsilon_{i1}, \quad \varepsilon_{i1} \sim N(0, \sigma_0^2), \\ F_{i1}(t) &= \phi_4 + (\phi_1 - \phi_4) \left\{1 - \exp[-(t\phi_2)^{\phi_{3i}}]\right\}. \end{aligned} \tag{4.3}$$

The dissolution profiles in Figure 2.1 contain both an asymmetrical S-shaped curvature and an initial jump. The Gompertz curve has the first property, but it has its short curvature at the end. The following modification of the Gompertz function (Lindsey 1997) will serve to model this feature and to challenge the performance of the Weibull model:

$$\begin{aligned} Y_{i1}(t) &= F_{i1}(t) + \varepsilon_{i1}, \quad \varepsilon_{i1} \sim N(0, \sigma_0^2) \\ F_{i1}(t) &= \phi_4 + [\phi_{1i} - \phi_4] \exp\left\{-\exp[-\phi_2(t - \phi_{3i})]\right\}, \end{aligned} \tag{4.4}$$

where ϕ_4 represents the initial jump. The coefficient $\phi_{1i} \sim N(\phi_1, \sigma^2_{\phi_1})$ corresponds to the asymptotic maximum dissolution. ϕ_{3i} represents a capsule specific lag-time, ϕ_2 is the scale parameter and is related to the inverse half-life of the curve.

4.1.2 Combination Models

All published models are limited to homogeneous formulations. A naive approach would be to ignore the heterogeneity of the formulation and fit the traditional model mentioned in Section 3.2 for a homogeneous formulation. However, in case of a heterogenous formulation of both immediate and extended-release material, the cumulative distribution function does not start at 0 but rather at a quantity approximately similar to the proportion of immediate-release material within the mixture. The inclusion of this initial jump alters the density function $F'_{i2k}(\tau)$. Therefore, we propose in this section a new model that takes this heterogeneity into account.

For the convolution model a similar derivation is possible. Recall from Section 2.1 that the capsules represent a heterogeneous formulation, consisting in part of immediate-release and in part of extended-release. As a result, two different underlying dissolution processes can be expected to be present.

The principle of superposition within pharmacokinetics, i.e., the assumption that each mechanism acts independently of each other and there is linear kinetics, means that the plasma concentration-time profile of the controlled-release formulation can be described as a weighted combination of each of the drug product plasma concentration-time profiles. This is a valid assumption (Piotrovsky *et al.* 2003). Based on this principle, one part of the profile corresponds to the immediate-release drug product within the formulation, the other one corresponds to the extended-release drug product. Clearly, these considerations imply a specific form for the model to be considered. The plasma concentration time profile corresponding to the immediate-release drug product can be considered as identical to the one observed from the immediate-release formulation whereas the latter follows the convolution model as described in Section 3.2. Therefore, the following new model is proposed:

$$\begin{aligned} Y_{i2k}(t) &= \phi_4 D c_{2k\delta}(t) + [\phi_1 - \phi_4]\, D \int_0^t c_{2k\delta}(t-\tau) F'_{i2k}(\tau) d\tau + \varepsilon_{i2k}, \qquad (4.5) \\ \varepsilon_{i2k} &\sim LN(0, \sigma_2^2), \end{aligned}$$

where ϕ_4 is the weight corresponding to the quantity of immediate-release drug product within the formulation and D represents the dose. This corresponds to the initial jump observed in the *in-vitro* models.

Furthermore, the *in-vivo* release $F_{i2k}(t)$ model is slightly modified compared to the proposal of O'Hara *et al.* (2001):

$$F_{i2k}(t) = g^{-1}\left\{\theta_0 + \theta_1 t + g\left[F_{i1}(t)\right]\right\}, \tag{4.6}$$

where the index i stands for the capsule i dependent variability of the *in-vitro* dissolution F_{i1}. As this is unobserved for the capsule administered to the subject k, this is indirectly included via the subject level. Thus, the random effects are included at the *in-vitro* level of the model rather than as a random intercept. This corresponds to the underlying source of variation. Further, the gastro-intestinal subject level s_{ik} was removed from the IVIVC, because the inclusion of an additional random intercept, next to the presence of the random effect in the *in-vitro* part of the equation, could jeopardize convergence or lead to very long run-times. Additionally, one can only estimate these inter-subject gastro-intestinal differences when multiple formulations per subject are analyzed to enable the dissociation between subject- and drug-unit-driven variability.

Although it is not in the traditional sense, but this newly proposed combination model can be considered as a mixture distribution at two levels. A first mixture is situated in the *in-vitro* dissolution: it represents a mixture cumulative distribution function of a step function for the immediate-release material of the formulation on one hand and a second cumulative distribution function such as the weibull distribution for the slow release drug product on the other hand. The mixture is also present at a second level: it is a combination of two log-normally distributed processes for the immediate-release plasma concentration time-profile on one hand and the convolution based profile for the slow release product on the other hand. Again, the weight ϕ_4 comes in in order to attribute the ratio of the two distinct underlying release processes.

4.1.3 Model Fitting

The models for the immediate-release plasma levels and the *in-vitro* dissolution were initially fitted separately to obtain good starting values for fitting the IVIVC model. The immediate-release pharmacokinetics of Galantamine are known to follow a two-compartmental model (Piotrovsky *et al.* 2003). This was based on population mod-

eling of several studies in elderly patients:

$$
\begin{aligned}
Y_{2k\delta}(t) &= c_{2k\delta}(t) + \varepsilon_{2k\delta}, \\
c_{2k\delta}(t) &= \frac{k_a}{V_F}\left[\frac{(k_{21}-\alpha_k)e^{-\alpha_k(t-t_{lag})}}{(k_a-\alpha_k)(\beta_k-\alpha_k)} + \frac{(k_{21}-\beta_k)e^{-\beta_k(t-t_{lag})}}{(k_a-\beta_k)(\alpha_k-\beta_k)}\right. \\
&\quad \left. + \frac{(k_{21}-k_a)e^{-k_a(t-t_{lag})}}{(\alpha_k-k_a)(\beta_k-k_a)}\right], \\
\varepsilon_{2k\delta} &\sim LN(0,\sigma_1^2), \qquad (4.7) \\
\begin{bmatrix} \alpha_k \\ \beta_k \end{bmatrix} &\sim N\left(\begin{bmatrix} \alpha \\ \beta \end{bmatrix}, \begin{bmatrix} \sigma_\alpha^2 & c_{\alpha\beta}\sigma_\alpha\sigma_\beta \\ c_{\alpha\beta}\sigma_\alpha\sigma_\beta & \sigma_\beta^2 \end{bmatrix}\right).
\end{aligned}
$$

In this model, V_F is the apparent volume of distribution, t_{lag} is a lag-time, k_a is the absorption coefficient, k_{21}, α, and β are transfer rate constants. The best fit to the data was attained by choosing random effects on α and β, with associated variabilities σ_α^2 and σ_β^2. This was based on visual inspection of the fit of the individual profiles as well as by comparison of the likelihood functions. The absorption rate constant k_a could not be estimated by fitting the immediate-release formulation alone because too few samples were taken during the absorption phase shortly after drug intake and k_a had to be fixed in the immediate-release model.

O'Hara *et al.* (2001) fits first the immediate-release profile per subject and then in a second stage fits the convolution. In a second stage he fits then the convolution and the IVIVC using the empirical Bayes' estimates of the immediate-release plasma concentration time profile. A possible drawback of such a two-stage modeling approach is that this might lead to biased results (Verbeke and Molenberghs 2000). Our proposal is to fit all the submodels simultaneously to tackle this potential source of bias. The two methods are formally compared in Chapter 5. The possible impact is discussed further in Section 4.3 as well as in Chapter 5.

4.2 Results

As mentioned in Section 4.1.1, the modified Gompertz function fitted the data well. Random effects were added to ϕ_1 and ϕ_3 since, as seen in Figure 2.1, the asymptotic maximum dissolution was drug-unit dependent. The random effect on the lag time improved the fit further. Conventional model selection tools, such as the likelihood ratio and the Akaike Information Criterion (AIC), were used.

A system of sub-models is proposed for the IVIVC modeling consisting of the combination of models 4.4, 4.5, 4.6, and 4.7. Contrary to O'Hara *et al.* (2001) are all

four submodels fitted simultaneously. This allows exchange of information between models. Whereas the absorption rate constant k_a could not be estimated for the immediate-release model alone, owing to insufficient early sampling points, the pooling of information about common parameters from different data sources did allow for estimation of k_a for model 4. In models 1–3, $\log(k_a) = -3$ had to be fixed to allow convergence. Additionally, $\log(V_F) = -4.64$ had to be fixed in models 1–2. Some simplifications were done for the system of sub-models compared to the separate submodels: (i) No random effect was used on the α-component of the two-compartmental model because inclusion of this random effect made the model diverge (non-positive hessian matrix). While this seems a disadvantage, one ought not to forget that the models are highly non-linear and rather complex. Therefore, one should keep the complexity of the random effects structure within reasonable limits; (ii) The dissolution random effects ϕ_{i1} and ϕ_{3i} were forced to be independent otherwise estimation of the correlation could not be established: many observations are needed to accurately estimate correlations between random effects.

The following models were fitted: Model (1)–(2) using the exponential dissolution and logit link as a convolution and a combination model, Model (3) as a combination model with a Weibull dissolution and logit link function, and Model (4) as a combination model with a Gompertz dissolution and logit link. Model (2)–(4) are the newly proposed models to cope with the heterogeneity of the data. Estimates of the model parameters can be found in Table 4.1 for the four different dissolution curves.

Table 4.1: Parameter estimates (95% confidence interval) for Models 1–4 using a one-stage convolution-based approach.

	Model 1	Model 2	Model 3	Model 4
	Est.	Est.	Est.	Est.
Parameter	(95% CI)	(95% CI)	(95% CI)	(95% CI)
Immediate-release				
k_a (hr^{-1})				0.026
				(0.004 ; 0.163)
V_F (L)			0.00356	0.0018
			(0.00305 ; 0.00417)	(0.0003 ; 0.0121)
k_{21} (hr^{-1})	0.43	0.48	0.067	0.033
	(0.26 ; 0.73)	(0.28 ; 0.83)	(0.063 ; 0.071)	(0.005 ; 0.205)
α (hr^{-1})	0.37	0.41	2.54	2.57
	(0.23 ; 0.62)	(0.25 ; 0.69)	(2.13 ; 3.03)	(2.12 ; 3.11)
β (hr^{-1})	11.5	11.5	0.14	0.14
	(10.9 ; 12.1)	(10.9 ; 12.1)	(0.123 ; 0.16)	(0.12 ; 0.16)
σ_β	0.06	0.05	0.24	0.21
	(0.04 ; 0.09)	(0.03 ; 0.08)	(0.14 ; 0.33)	(0.13 ; 0.29)
σ_1	0.29	0.29	0.19	0.19
	(0.26 ; 0.32)	(0.27 ; 0.32)	(0.18 ; 0.21)	(0.17 ; 0.21)

Table 4.1: Parameter estimates (95% confidence interval) for Models 1–4. (continued)

	Model 1	Model 2	Model 3	Model 4
	Est.	Est.	Est.	Est.
Parameter	(95% CI)	(95% CI)	(95% CI)	(95% CI)
In Vitro Dissolution				
ϕ_1	1.02	0.91	0.91	0.88
	(0.98 ; 1.05)	(0.88 ; 0.94)	(0.89 ; 0.93)	(0.86 ; 0.90)
ϕ_2	0.12	0.31	0.14	0.29
	(0.10 ; 0.13)	(0.28 ; 0.34)	(0.13 ; 0.15)	(0.28 ; 0.31)
ϕ_3	-1.74	0.72	1.39	4.07
	(-1.98 ; -1.50)	(0.46 ; 0.97)	(1.28 ; 1.50)	(3.83 ; 4.30)
ϕ_4		0.27	0.23	0.22
		(0.26 ; 0.28)	(0.21 ; 0.25)	(0.20 ; 0.24)
σ_0	2.95	2.33	3.26	1.33
	(2.54 ; 3.37)	(1.99 ; 2.66)	(2.82 ; 3.71)	(1.13 ; 1.53)
σ_{ϕ_1}				0.03
				(0.019 ; 0.04)
σ_{ϕ_2}	0.012	0.23		

Table 4.1: Parameter estimates (95% confidence interval) for Models 1–4. (continued)

	Model 1	Model 2	Model 3	Model 4
	Est.	Est.	Est.	Est.
Parameter	(95% CI)	(95% CI)	(95% CI)	(95% CI)
	(0.004 ; 0.020)	(0.16 ; 0.31)		
σ_{ϕ_3}			0.001	0.21
			(0.001 ; 0.100)	(0.10 ; 0.33
$\rho_{\phi_1\phi_3}$				0.74
				(0.32 ; 1.00)
Controlled-release				
θ_0	4.37	2.41	0.01	1.43
	(4.19 ; 4.56)	(2.19 ; 2.63)	(-0.22 ; 0.25)	(1.23 ; 1.63)
θ_1	-0.78	-0.17	-0.08	0.09
	(-0.93 ; -0.63)	(-0.33 ; -0.01)	(-0.10 ; -0.06)	(0.06 ; 0.12)
σ_2	0.56	0.53	0.68	0.66
	(0.54 ; 0.59)	(0.50 ; 0.55)	(0.64 ; 0.71)	(0.63 ; 0.70)

Table 4.2: Model fit based on the criterion of average absolute percent prediction error and its 90% confidence interval for Models 1–4.

Model	$C_{\max}$	AUC_{last}	AUC_{0-4}
1	82.95 (81.70 ; 84.22)	75.27 (73.19 ; 77.40)	76.92 (75.68 ; 78.18)
2	82.94 (81.83 ; 84.06)	74.06 (73.01 ; 75.13)	77.09 (75.82 ; 78.39)
3	8.06 (5.24 ; 12.40)	7.32 (5.13 ; 10.45)	8.46 (5.93 ; 12.08)
4	9.21 (5.54 ; 15.31)	4.67 (3.20 ; 6.83)	7.08 (4.34 ; 11.56)

The fit based on the different dissolution models was formally compared by Akaike's Information Criterion (AIC): 937.8, 880.3, 668.3, and 502.4 for model 1–4, respectively. The model prediction of the controlled-release plasma concentration of a randomly chosen subject for the different models is depicted in Figure 4.1. The fit of the Gompertz odds model was judged based on visual inspection of the empirical Bayes estimates versus the observed controlled-release profiles on the one hand and the average absolute percent prediction error on the other hand. Figure 4.3-4.5 contains the observed as well as the model predictions of the *in-vitro* dissolution, the immediate and the controlled-release plasma concentrations time profiles for the Gompertz model.

The $\%PE$ of the different models can be found in Table 4.2. The first two columns correspond to the $C_{\max}$ and AUC_{last} as requested in the regulatory guidelines (FIP 1996, CDER 1997), the last represents AUC_{0-4} and is an indication of the model fit for the data up to 4 hours. The model with the exponential dissolution does not fit the data well. The addition of the combination model to the exponential dissolution does not improve the model. The *in-vitro* exponential mixture dissolution model however misses the S-shape as observed in the data, see Figure 2.1. Therefore, the model is extended to the Weibull model and the Gompertz model as a combination model. The $\%PE$ indicates a significant improvement of the model fit for models 3 and 4.

The parameter estimates for the immediate-release part of the model show a permutation: values for α in models 1 and 2 have the same magnitude, and hence physiological meaning as β for models 3 and 4. The residual variance is lower for the latter models (standard deviation of 19 instead of 29). Some of the inadequacy of the first model is demonstrated in the asymptotic maximal dissolution parameter ϕ_1. In theory, all material should dissolve. Even though dissolution is not fully complete after 18 hours, the *in-vitro* release profile contradicts this. The half-life of the disso-

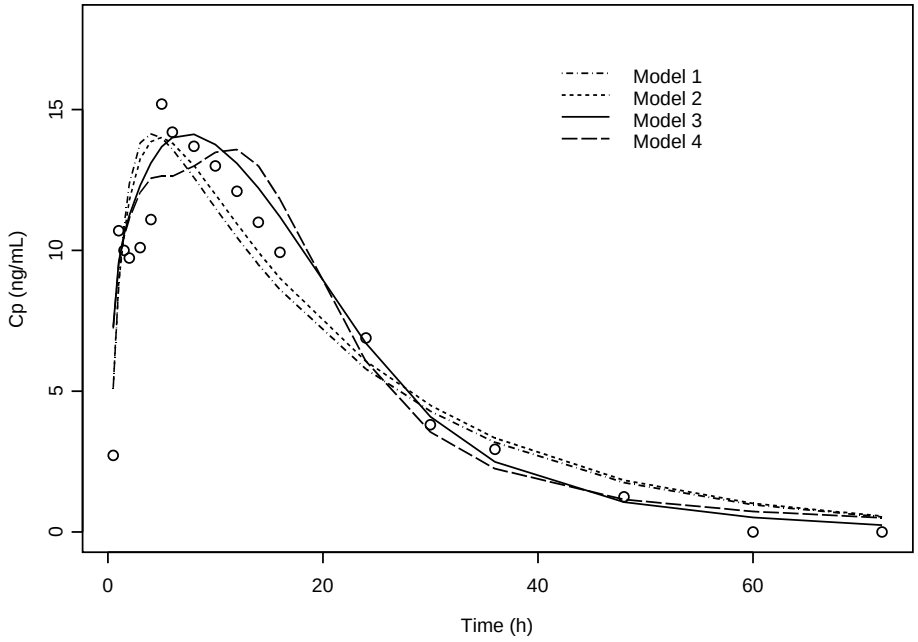


Figure 4.1: Observed and predicted controlled-release Galantamine concentrations of one randomly chosen subject for the different models.

lution ϕ_2 is estimated to be 0.1 for models 1 and 3 whereas models 2 and 4 produce an estimate of 0.3. This 3-fold difference might indicate that also model 3 is not free of issues. The estimated proportion immediate-release formulation ϕ_4 is close to the known formulation heterogeneity ratio of 0.25 for all 3 models. The parameters θ_0 and θ_1 have no physiological meaning and differences between the models are difficult to interpret.

4.3 Discussion

A model with clear improvements over the standard IVIVC models at two levels is presented: It allows the fitting of formulations containing both extended- and immediate-release material and it is a true one-stage analysis method. We employed

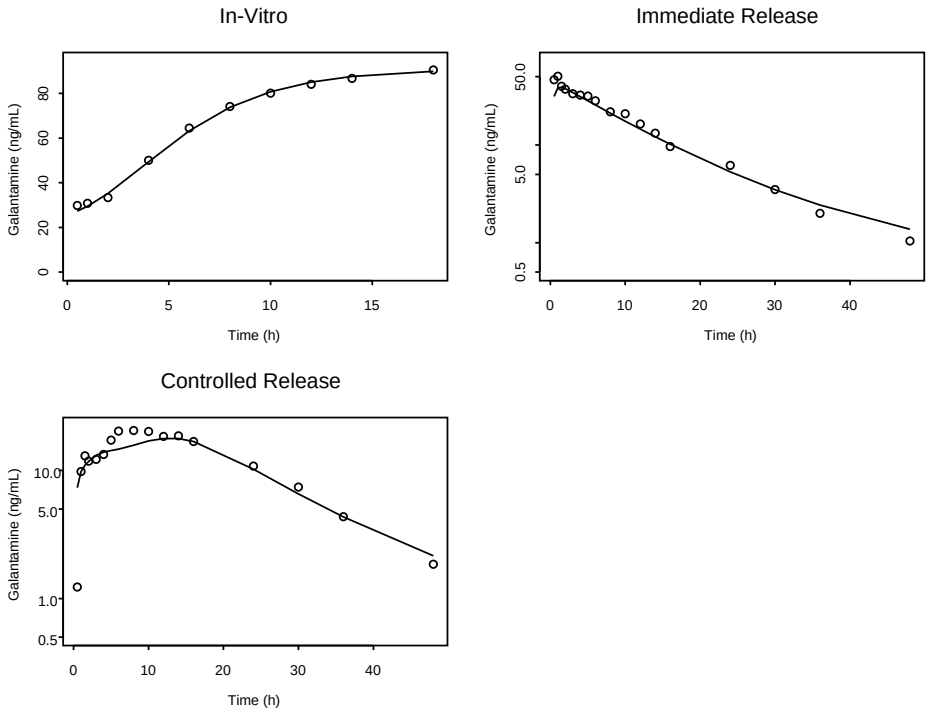


Figure 4.2: Observed and predicted *in-vitro* as well as *in-vivo* controlled and immediate-release Galantamine concentrations time profile for a randomly chosen subject and capsule. Predictions are based on the Gompertz odds model.

the SAS procedure Nlmixed rather than the standardly used NONMEM package.

All publications up to now have been limited to homogeneous formulations. In this chapter, we extend the convolution-based method for a heterogeneous formulation of both an immediate- and an extended-release drug product by a combination model. Four different models were evaluated during the model building: The first model used the convolution with the Exponential dissolution model and the logit link function. The average percent error $\%PE$ of both $C_{\max}$ and AUC_{last} remained well above the 10% criterion as in the guidelines (FIP 1996, CDER 1997), see Table 4.2, indicating an inadequate model fit.

The model was therefore extended to the combination model with the logit link function because this used the underlying heterogeneous structure of the capsules and fitted a bimodal profile. The use of this combination imposes, however, no restrictions

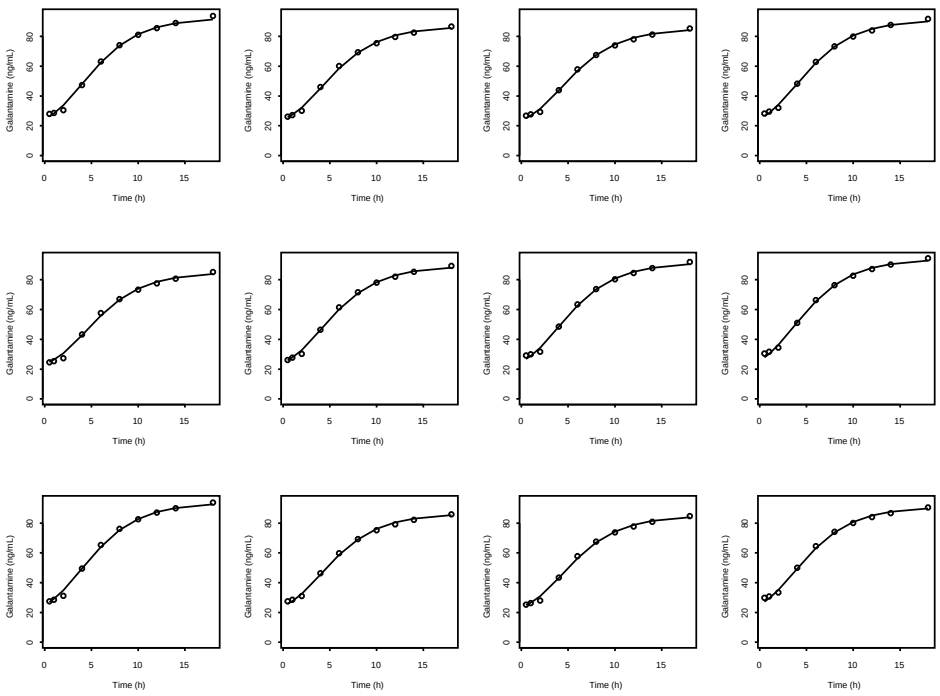


Figure 4.3: Observed and predicted in-vitro dissolution per drug-unit for the Gompertz odds model.

on the model. It relies on the principle of superposition within pharmacokinetics, i.e., the assumption that each mechanism acts independently of each other and there is linear kinetics. The metabolism of the drug remains unchanged during the drug product release and only depends on the amount of drug product released. The standard convolution model itself assumes already the superposition principle and linear kinetics. The $\%PE$ for $C_{\max}$ of this model remained in the same order of magnitude. The fit of the *in-vitro* dissolution data indicated however that further refining was required: the exponential model has a steep incline for the first hours and converges to its asymptotic limit, whereas the *in-vitro* dissolution data showed an asymmetric S-shaped curve. The model was finetuned with the use of the Weibull and the Gompertz model for the *in-vitro* dissolution in combination with the logit link function. This lead to a substantial decrease in $\%PE$. Although the $\%PE$ was well below 10% for all parameters, only the Gompertz model had the upper limit of

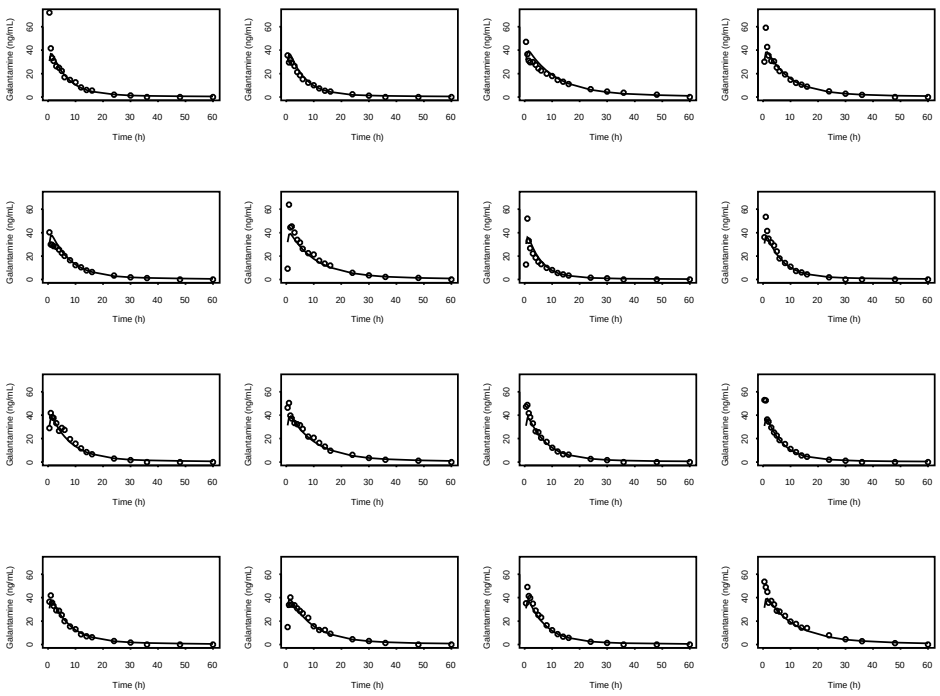


Figure 4.4: Observed and predicted immediate-release Galantamine concentrations per subject for the Gompertz odds model.

the 90% confidence interval below 10% for AUC_{last}. The Gompertz model can be considered as the superior model given this better prediction of the overall exposure AUC_{last} and the ability to estimate all parameters of the model without fixing any of them.

This illustrates that traditional models should not be used in case of formulations consisting of both immediate-release and slow-release drug product. The risk of overfitting is limited given that the construction of the model is based on the formulation properties and the clearly bimodal profiles.

Models 3 and 4 meet the regulatory specifications on the point estimate, see Table 4.2. Whereas the guidelines (FIP 1996, CDER 1997) focus only on the mean $\%PE$ being less than 10% to conclude IVIVC predictability, this does not take into account the possible variability of the prediction. Even though the average might be less than 10%, a large variability of the individual $\%PE$ might indicate that some

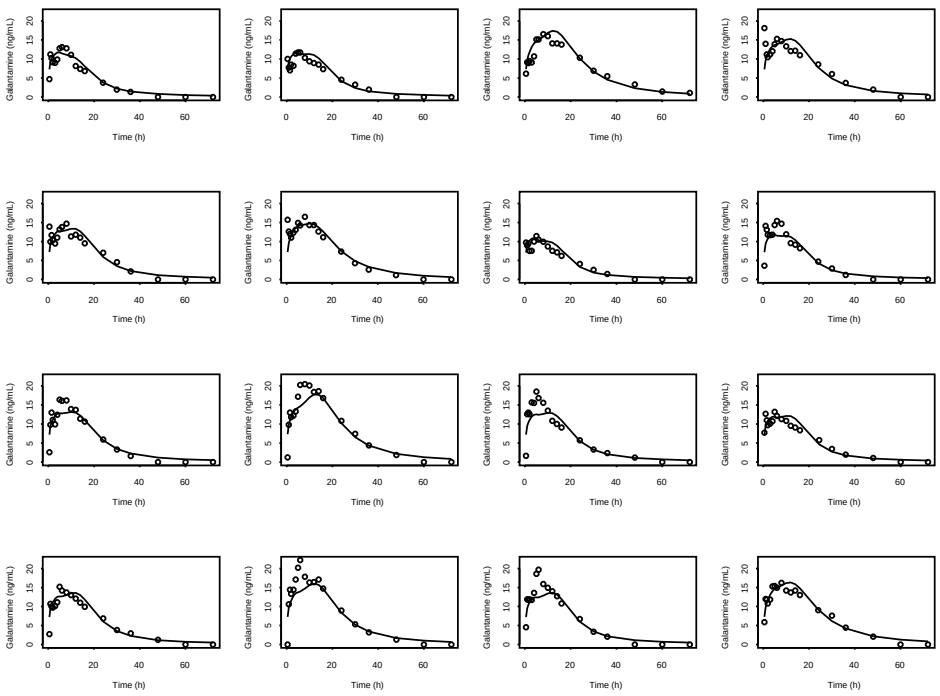


Figure 4.5: Observed and predicted controlled-release Galantamine concentrations per subject for the Gompertz odds model.

subject's controlled-release profile is poorly estimated. Therefore, one should rather use the non-inferiority philosophy and look at the upper limit of the 90% confidence interval.

A second, more fundamental change to the convolution method of O'Hara (O'Hara *et al.* 2001) is that all models are fitted simultaneously, whereas O'Hara's method first fits the immediate-release profile per subject and then fits in a second stage the convolution and the IVIVC using the empirical Bayes' estimates of the immediate-release plasma concentration time profile. A possible drawback of such a two-stage modeling approach is that this might lead to biased results (Verbeke and Molenberghs, 2000). In the first stage, the immediate-release plasma concentration time profile is reduced to a couple of summary statistics and residual error is ignored. In the second stage, these estimates are used as if they are error-free. Hence, the possible error of these coefficients will be reallocated to the remaining coefficients and as such

introduce possibly bias. Fitting everything at once however does not ignore the error in the individual compartmental pharmacokinetic parameters. On the contrary, it allows a pooling of information about common parameters of the immediate- and the extended-release model. This might lead to more accurate parameter estimation like for example the k_a in the case study. A formal comparison of the two approaches is given in Chapter 5. The advantage of the two-stage approach is that the parameter space is split. As a result, the two-stage approach is more flexible, in the sense of adding random effect, and model convergence is easier and faster.

It is not clear, based on these data, whether a modified sampling would have allowed for better estimation. Intuitively, additional early sampling might enhance estimation of the parameter k_a, but this was not formally established. On the other hand, only a limited number of blood samples can be taken for ethical and practical reasons. Therefore, the current sampling scheme is arguably the best feasible one: additional early samples might jeopardize later sampling.

In conclusion, a novel one-stage methodology was proposed as well as a combination model to cope with heterogeneous formulations in IVIVC testing. Based on the case study it was shown superior to the traditional model.

5

One- versus Two-stage IVIVC Modeling

An IVIVC model combines the *in-vitro* dissolution properties with the *in-vivo* immediate-release plasma concentration-time profile to accurately, i.e., the estimation is unbiased, and precisely, i.e., with low uncertainty, predict the *in-vivo* controlled-release plasma concentration time profile. Once such an IVIVC model is obtained, it can be used for the development of new formulations, as well as to assess the impact of manufacturing procedures, and even to detect batch differences. Therefore, an accurate estimation of the area under the plasma concentration time curve (AUC) and maximal concentration ($C_{\max}$) is indispensable.

Gaynor, Dunne, and Davis (2007) demonstrated that the current standard method, based on deconvolution (FDA CDER 1997), results in a potentially biased prediction of both AUC and C_{max}, whereas this is less pronounced for the two-stage convolution technique.

The convolution method, as introduced by O'Hara *et al.*, (2001) fits the immediate-release plasma concentration time profile first per subject. Then, in a second stage, the convolution model as well as the *in-vitro* dissolution model is fitted simultaneously. However, as the convolution model combines the immediate-release plasma concentration time profile and the *in-vivo* release of the drug substance, this convolution

model shares a number of parameters with the immediate-release plasma concentration model as well as with the *in-vitro* dissolution model via the hypothetical IVIVC link function. This implies that information on parameters of these submodels is contained within the controlled-release plasma concentration time profile. However, the model as introduced by O'Hara *et al.* (2001) does not use this shared information. Therefore, it is of interest to compare the behavior of the model as introduced by O'Hara *et al.* (2001) versus the one-stage modeling approach, where all three submodels are fitted simultaneously (Jacobs *et al.* 2008). For the latter, a non-linear mixed-effects model (Gabrielson and Weiner, 2000) is used for the immediate-release plasma concentration time profile, rather than an approach based on fitting a separate model to each subject.

Gaynor, Dunne, and Davis (2007) demonstrated the unbiased prediction of the summary measures AUC and $C_{\max}$ for the two-stage convolution method. Although this suffices for most applications, the accurate time profile prediction required in, for example, novel formulation development, cannot be taken for granted based on this result. Therefore, a simulation study was set up to compare the parameter estimates with the values used for data generation. The two-stage convolution-based method (O'Hara *et al.* 2001) was compared with the novel one-stage convolution-based method (Jacobs *et al.* 2008). This chapter is based on Jacobs *et al.* (2009) and is organized as follows. The set-up of the simulation study to investigate the difference between a one- and a two-stage approach is described in Section 5.1. The analysis of the simulation study can be found in Section 5.2.

5.1 Design of Simulation Study

As already clear from the previous chapters, there are two possible convolution-based approaches. The two-stage modeling of O'Hara *et al.* (2001) as introduced in chapter 3 will be compared to the one-stage modeling discussed in chapter 4 by way of a simulation study. For simplicity, attention is restricted to a single formulation.

An overview table with the parameter values used for the simulation can be found in Table 5.1. Two types of *in-vitro* dissolution models are used to ensure a representative simulation setting. The first dissolution model is chosen to have a rapid release, i.e., from the start onwards, substantial amounts of drug product are released over time. A good candidate for such a model is the exponential dissolution model, taking the form:

$$F_{i1\ell}(t) = \phi_1 \left\{1 - \exp\left[-\left(t - \phi_3\right)\phi_{2i}\right]\right\}, \tag{5.1}$$

Table 5.1: These parameter values were used to simulate the *in vitro* data as well as the *in vivo* controlled-release and immediate-release plasma concentration profiles for the different settings and levels of variability.

	Exponential 1	Exponential 2	Gompertz 1	Gompertz 2
D	100	100	100	100
ϕ_1	0.95	0.95	0.95	0.95
ϕ_2	-1	-1	0.1	0.1
ϕ_3	-0.5	-0.5	-0.5	-0.5
σ_{ϕ_2}	0.04	0.04	0.01	0.01
σ_0	0.5	0.5	0.5	0.01
$\log(k_e)$	-1	-1	-1	-1
$\log(V)$	-3.7	-3.7	-3.7	-3.7
σ_{k_e}	0.42	0.05	0.05	0.05
σ_1	0.05	0.01	0.05	0.01
θ_0	1	2	3	4
θ_1	0.2	0.2	0.2	0.2
σ_{θ_0}	0.2	0.2	0.2	0.2
σ_2	0.05 ; 0.1 ; 0.15	0.05 ; 0.1 ; 0.15	0.05 ; 0.1 ; 0.15	0.05 ; 0.1 ; 0.15
	0.2 ; 0.25 ; 0.3	0.2	0.2 ; 0.25 ; 0.3	0.2

where $\phi_1 = 0.95$ is the asymptotic maximal dissolution, $\phi_{2i} \sim N(p_2, \sigma^2_{\phi_2})$ is the drug-unit-specific scale parameter, $\phi_3 = -0.5$ the dissolution lag-time, and $\varepsilon_{i1\ell} \sim N(0, \sigma_0^2)$. Values of -3 and -1 for p_2 were used and labelled "exponential 1" and "exponential 2," respectively.

The second dissolution model has a built-in delay, i.e., a small amount of drug product is released in the beginning of the *in-vitro* dissolution testing, and the majority of the drug product is released only after some time. A Gompertz dissolution model is then well suited:

$$F_{i1\ell}(t) = \phi_1 \left(1 - \exp\left\{-\exp\left[\phi_{2i}\left(t - \phi_3\right)\right]\right\}\right), \tag{5.2}$$

where, again $\phi_1 = 0.95$, is the asymptotic maximal dissolution, $\phi_{2i} \sim N(0.1, \sigma^2_{\phi_2})$ is the drug-unit-specific scale parameter, $\phi_3 = 30$ the dissolution lag-time, $\varepsilon_{i1\ell} \sim$

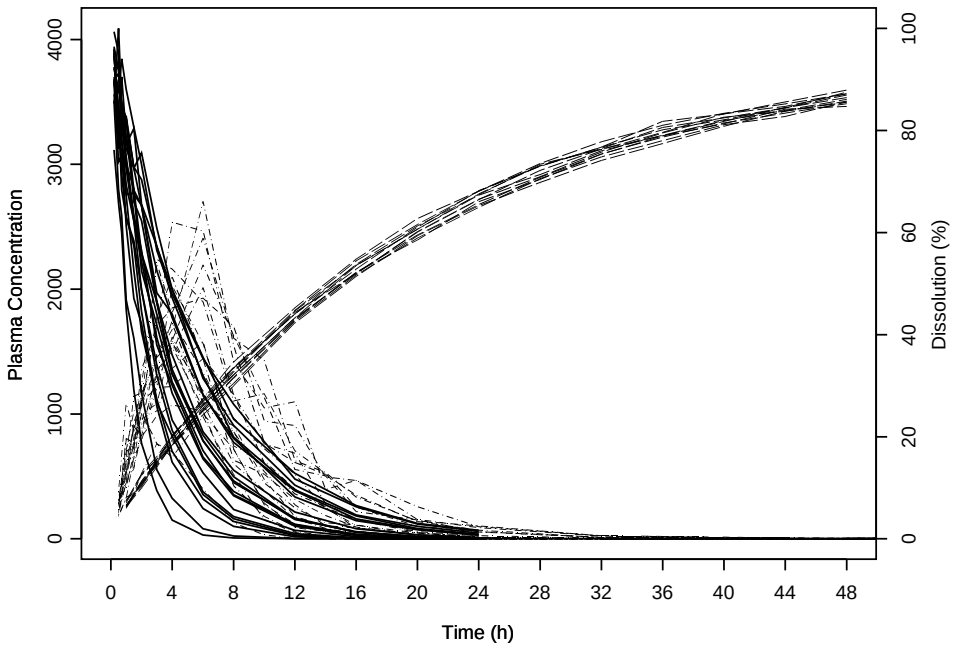


Figure 5.1: Simulation of an IVIVC study using one controlled-release formulation (– . –), an exponential dissolution curve (– –) and an IV bolus as the immediate-release plasma concentration time profile (—).

$N(0, \sigma_0^2)$, and $\sigma_{\phi_2}^2 = 0.01$.

For the immediate-release plasma concentrations $c_{2k\delta}(t)$, an intravenous bolus one-compartmental model (Gabrielson and Weiner, 2000) will be used:

$$c_{2k\delta}(t) = \frac{D}{V} \exp(-k_e t), \tag{5.3}$$

where $D = 100$ is the administered dose, $\log(k_e) \sim (-1, \sigma_{k_e}^2)$ is the subject dependent elimination rate constant, $V = 0.27$ is the volume of distribution, and $\varepsilon_{2k\delta} \sim LN(0, \sigma_1^2)$. Further, $\sigma_1^2 = 0.05(0.01)$ for the first (second) exponential setting. The standard deviation σ_{k_e} of the random effect was set to 0.42, 0.05, 0.05, and 0.05 for the first and second exponential and Gompertz settings, respectively.

Examples of the simulation data can be found in figure 5.1 and 5.2 for an exponential and Gompertz setting. The poor choice of both θ_0 and θ_1 to be positive leads to

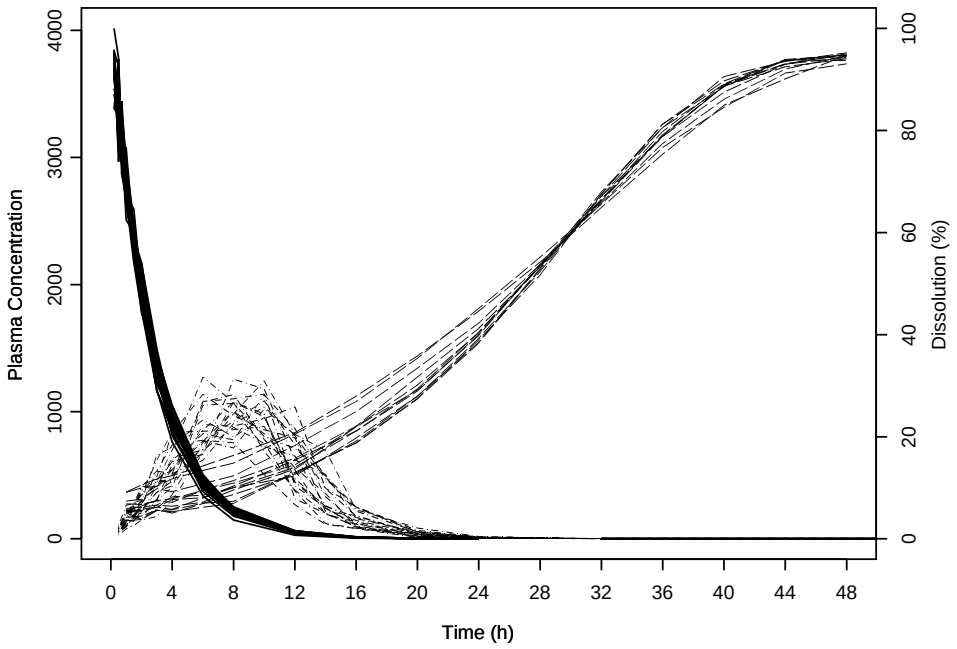


Figure 5.2: Simulation of an IVIVC study using one controlled-release formulation (– . –), a gompertz dissolution curve (– –) and an IV bolus as the immediate-release plasma concentration time profile (—).

observed plasma concentrations for the first 24 hours only. In reality, the parameter estimate of θ_0 is typically negative. However, the choice of the parameters does not jeopardize the results of this simulation study.

In the two-stage method (O'Hara *et al.* 2001) the immediate-release plasma concentration time profile are first modelled separately per subject. In the one-stage method, a non-linear mixed-effects submodel (Molenberghs and Verbeke 2005, Pinheiro and Bates 2000) is implemented.

The parameter estimates obtained by both methods in the simulation were divided by the original values used for data generation and subsequently averaged. Values close to 1 indicate more unbiased estimation, values further from 1 points to the presence of bias. At the same time, precision of the estimation is assessed by means of standard deviation.

Maximum likelihood estimation of the set of models was implemented in the SAS procedure NLMIXED (version 9.1). Model convergence was monitored using the first-order integration method of Beal and Sheiner (1982). Gaussian quadrature might have yielded better numerical accuracy (Molenberghs and Verbeke, 2005), but the very long run times prohibited the use of this method in the simulation setting. A very small simulation setting was additionally performed using adaptive gaussian quadrature to ensure that the results are not restricted to the first-order method. Only three quadrature points were used to allow a feasible runtime.

A large number of different simulation settings are studied to ensure the representation of the results. For each setting, 50 virtual studies were generated and analyzed. These studies can be grouped into four main categories.

- Exponential 1: high variable clearance and high residual variance for $c_{2k\delta}(t)$, a fast exponential *in-vitro* dissolution curve, six different variances for the slow-release plasma concentrations.

- Exponential 2: low variable clearance and small residual variance for $c_{2k\delta}(t)$, a slow exponential *in-vitro* dissolution curve, four different variances for the slow-release plasma concentrations.

- Gompertz 1: high variable clearance and high residual variance for $c_{2k\delta}(t)$, a Gompertz *in-vitro* dissolution curve, six different variances for the slow-release plasma concentrations.

- Gompertz 2: low variable clearance and low residual variance for $c_{2k\delta}(t)$, a Gompertz *in-vitro* dissolution curve, four different variances for the slow-release plasma concentrations.

5.2 Results

The accuracy and precision of one- versus two-stage modeling is assessed via a simulation study. For a number of settings in our simulation study, the Hessian of the model was not positive definite, a typical issue in non-linear mixed-effects modeling (Pinheiro and Bates 2000, Davidian and Giltinan 1995). Therefore, these studies were excluded from further analysis (see Table 5.2). The fairness of comparison is not jeopardized because only studies are included where both the one- and two-stage model converged. It is of interest to note that almost all of the Hessian matrix problems are due to the two-stage models (Tables 5.3 and 5.4).

Table 5.2: Number of studies used to compare the one-stage and two-stage approaches for the different settings and levels of variability. '.' denotes not investigated settings.

σ_2	Exponential 1	Exponential 2	Gompertz 1	Gompertz 2
0.05	28	42	35	36
0.10	27	43	35	32
0.15	25	41	34	34
0.20	34	47	41	34
0.25	28	.	44	.
0.30	27	.	40	.

Table 5.3: Number of studies with a positive Hessian matrix using a one-stage model for the different settings and levels of variability.

σ_2	Exponential 1	Exponential 2	Gompertz 1	Gompertz 2
0.05	50	50	49	49
0.1	48	50	49	49
0.15	50	50	50	49
0.2	50	50	50	50
0.25	50	.	50	.
0.3	50	.	50	.

Figures 5.3 and 5.4 and Table 5.5 and 5.6 contain the relative accuracy of the parameter estimates as well as the precision of both modeling methods for the four groups of simulation settings. Values closer to 1 indicate less bias, whereas large deviations from 1 indicate that the averaged parameter estimate within the setting deviates considerable from the target parameter value.

The simulation study showed systematically that parameter estimates were less accurate for the two-stage approach compared to the one-stage approach. The parameter estimation of the dissolution parameters ϕ_1, ϕ_2, and ϕ_3 can be considered as very accurate for both methods for most settings. The fixed effects θ_0 and θ_1, on the contrary, exhibit considerable inaccuracy for the two-stage method, and little to no bias for the one-stage method. At first sight, θ_1 is estimated with less accuracy

Table 5.4: Number of studies with a positive Hessian matrix using a two-stage model for the different settings and levels of variability.

σ_2	Exponential 1	Exponential 2	Gompertz 1	Gompertz 2
0.05	28	42	35	37.00
0.10	29	43	35	32.00
0.15	25	41	34	34.00
0.20	34	47	41	34.00
0.25	28	.	44	.
0.30	27	.	40	.

than θ_0. However, the parameter is not more sensitive to bias. The inaccuracy of θ_0 is spread over both the fixed effect and the inflated random effects variability. Therefore, it gives the false impression of being more accurately estimated than θ_1. The residual standard deviation σ_2 seems most problematic. It was systematically overestimated for all four settings in the case of the two-stage approach: 4.118, 2.943, 5.109, and 6.697 versus 1.098, 1.017, 0.999, and 109.8 for the first and second exponential and Gompertz, in the case of two- versus one-stage modeling, respectively. Also, the random-effects standard deviations σ_{ϕ_2} and σ_{θ_0} are substantially overestimated for the two-stage modeling approach.

Not only the behavior in terms of bias is inferior for the two-stage approach, also its precision is worse for most parameter estimates: the standard deviations on the original scale are systematically larger than for the one-stage method for almost all parameters. Again, this is most prominent for the residual standard deviations and random-effects standard deviations, whereas the difference in precision is negligible for the fixed effects of the dissolution part of the model.

A very small simulation of 100 studies using Gaussian quadrature confirmed these results. Also in this case, only 24 out of 100 studies converged for the two-stage method versus 95 out of 100 for the one-stage method. The bias of the estimates is however smaller using gaussian quadrature.

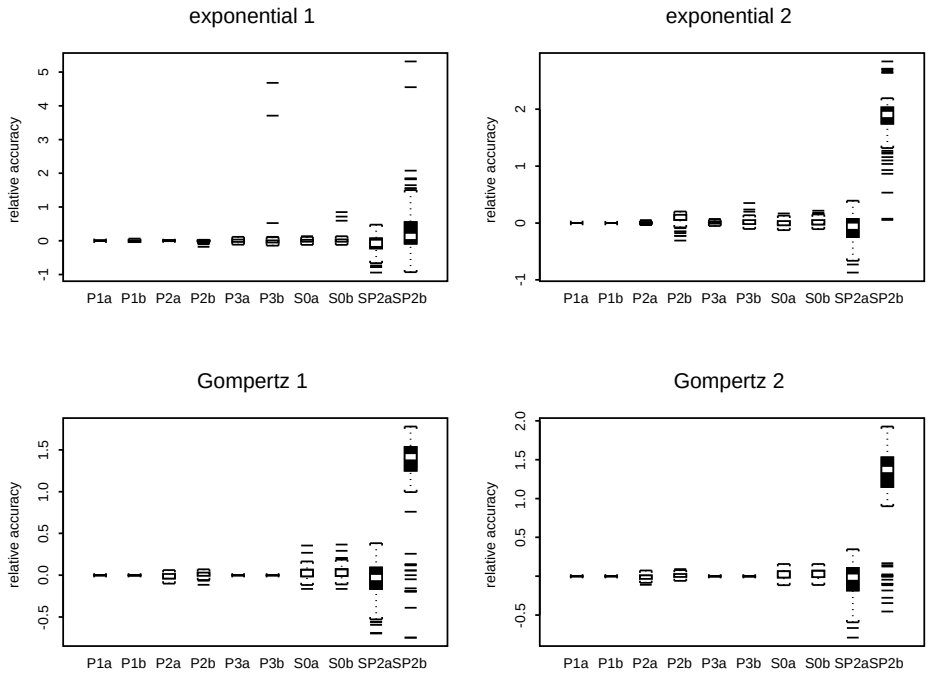


Figure 5.3: Boxplot of the log-transformed parameter estimates for the dissolution parameters for the different settings. In the legend, a and b indicates the one- or two-stage estimation, P1, P2, P3, S0, and SP2 are ϕ_1, ϕ_2, ϕ_3, σ_0, and σ_{ϕ_2}, respectively.

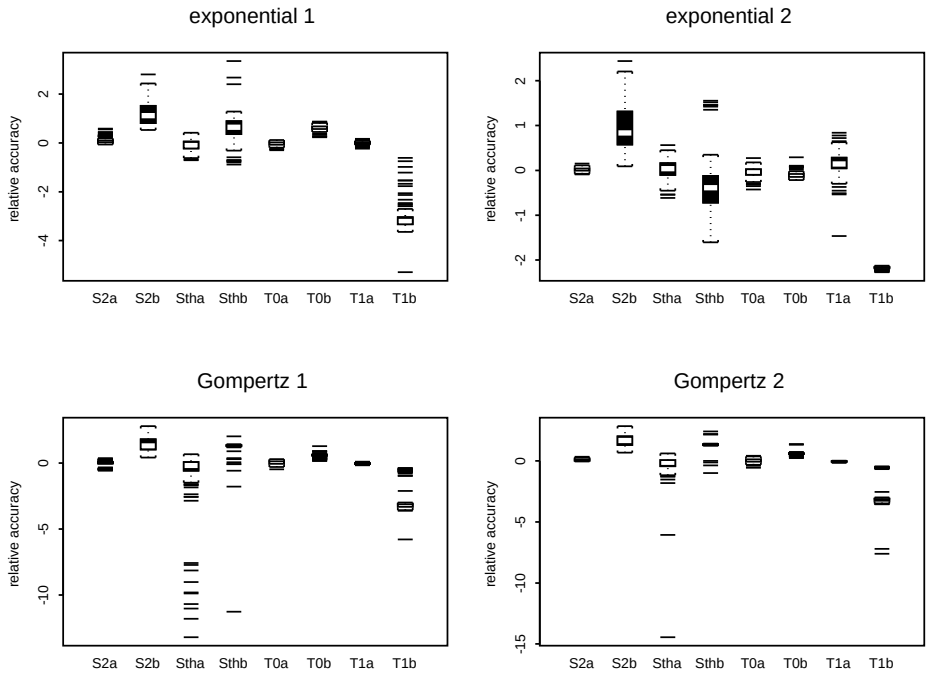


Figure 5.4: Boxplot of the log-transformed parameter estimates for the IVIVC parameters for the different settings. In the legend, a and b indicates the one- or two-stage estimation, S2, Sth, T0, and T1 are σ_2, σ_{θ_0}, θ_0, and θ_1, respectively.

Table 5.5: Relative bias and absolute efficiency of the parameter estimates after the simulation of two- versus one-stage IVIVC modeling using an exponential *in-vitro* dissolution function.

		2-stage		1-stage	
		relative mean	std	relative mean	std
exponential 1 ($N = 169$)	ϕ_1	0.998	0.0066	1.000	0.0021
	ϕ_2	0.991	0.0200	1.000	0.0044
	ϕ_3	1.855	8.7735	0.999	0.0482
	σ_{ϕ_2}	2.770	17.339	0.935	0.2457
	σ_0	1.023	0.1515	1.002	0.0486
	θ_0	1.868	0.2566	0.974	0.0687
	θ_1	0.056	0.0651	1.002	0.0545
	σ_{θ_0}	2.034	2.6012	0.924	0.1966
	σ_2	4.118	2.8185	1.098	0.1408
exponential 2 ($N = 173$)	ϕ_1	1.000	0.0006	1.000	0.0005
	ϕ_2	1.090	0.0867	1.001	0.0144
	ϕ_3	1.022	0.0632	1.005	0.0242
	σ_{ϕ_2}	6.364	3.1339	0.935	0.2164
	σ_0	1.018	0.0592	1.001	0.0525
	θ_0	0.900	0.0628	0.959	0.0969
	θ_1	0.114	0.0028	1.192	0.2843
	σ_{θ_0}	0.689	0.8804	0.986	0.3721
	σ_2	2.943	1.9414	1.017	0.0405

Table 5.6: Relative bias and absolute efficiency of the parameter estimates after the simulation of two- versus one-stage IVIVC modeling using a Gompertz *in-vitro* dissolution function.

		2-stage		1-stage	
		relative mean	std	relative mean	std
Gompertz 1 ($N = 229$)	ϕ_1	0.999	0.0012	0.999	0.0012
	ϕ_2	1.011	0.0286	0.989	0.0351
	ϕ_3	0.999	0.0011	1.000	0.0011
	σ_{ϕ_2}	2.082	3.4603	0.938	0.2959
	σ_0	1.036	0.0761	1.028	0.0714
	θ_0	1.817	0.1931	1.047	0.1411
	θ_1	0.095	0.1642	0.964	0.0401
	σ_{θ_0}	3.174	1.7929	0.762	0.4320
	σ_2	5.109	3.5562	0.999	0.1838
Gompertz 2 ($N = 136$)	ϕ_1	0.999	0.0012	0.999	0.0012
	ϕ_2	1.009	0.0278	0.988	0.0334
	ϕ_3	1.000	0.0012	1.000	0.0012
	σ_{ϕ_2}	1.691	3.6415	0.977	0.1986
	σ_0	1.031	0.0599	1.025	0.0592
	θ_0	1.833	0.3368	1.045	0.1470
	θ_1	0.107	0.1799	0.961	0.0292
	σ_{θ_0}	3.154	2.1457	0.843	0.3813
	σ_2	6.697	3.9426	1.098	0.1046

5.3 Discussion

Up to now, deconvolution-based methods are standard practice in IVIVC modeling. However, Gaynor, Dunne, and Davis (2007) demonstrated that the two-stage convolution-based method (O'Hara *et al.* 2001) is superior for the prediction of the overall (AUC) and peak ($C_{\max}$) exposure of controlled-release drug products. Although this is an excellent result for most applications, in certain circumstances, such as formulation development, the interest of the IVIVC model predictions is not in the accurate prediction of drug exposure measures AUC and $C_{\max}$, but rather in the accurate prediction of the individual plasma concentration time profiles, which corresponds to the estimation of the parameter estimates. To get a better understanding of the accuracy and precision of the two-stage convolution-based method, its estimation properties are compared to the more recent one-stage convolution-based method of Jacobs *et al.* (2008) in a simulation study.

The numerical performance of both models is very different. Indeed, using the same starting values, the two-stage method leads more frequently to a non-positive Hessian matrix than for the one-stage model. Not surprisingly, also the run-times are strikingly different: the partitioning of the parameter space allows for a much faster convergence for the two-stage method. The increase over run-time mainly depends on the complexity of the random-effects structure, an important disadvantage of the one-stage method in the case of complex submodels.

Comparison of the accuracy of the parameter estimates indicates that the one-stage method outperforms the two-stage method. In particular, the estimation of the residual and random-effects standard deviations are worrisome for the two-stage method due to the serious overestimation. The fixed-effects estimation in the convolution submodel exhibits a similar inaccuracy. The fixed effects of the *in-vitro* dissolution submodel have acceptable accuracy.

From both methodological considerations and simulations, it is clear that the submodels share a number of parameters as well as information on parameters, i.e., the controlled-release plasma concentrations time profiles also contain information on the parameters of the *in-vitro* dissolution model as well as of the parameters of the immediate-release plasma concentration time profile. Therefore, it comes as no surprise that also the precision of the one-stage method is better than the two-stage method.

One should be aware that the first-order integration method of Beal and Sheiner (1982) was used rather than the more accurate Gaussian quadrature. The long run-times of the latter prevented the implementation in a large simulation study. A very

small simulation study indicated that its estimation was better than in the case of the first-order method, but the discrepancy between the one- and the two-stage method remained. Also the impact of the approximation of the convolution model 3.1 as a finite sum might have have introduced bias.

Does this mean that one should abandon not only the deconvolution-based method, but also the two-stage convolution-based method of O'Hara *et al.* (2001)? There is no straightforward answer. Indeed, as long as interest in and application of the IVIVC model is restricted to the estimation of overall and peak exposure estimation, Gaynor, Dunne, and Davis (2007) demonstrated that the two-stage method performs well. The shorter run-times are a very convenient additional feature. However, if the application of the IVIVC model should predict accurately the individual time concentration profile, such as in formulation development, then the more accurate parameter estimations of the one-stage method would lead to more reliable simulated plasma concentration time profiles, and, therefore, it seems to be the appropriate method in that case.

6

Outlying Subject Detection in In-vitro – In-vivo Correlation Models

Having introduced the concepts and an example of IVIVC modeling in the previous chapters, the focus of this chapter will be on the model diagnostics, i.e., how well does the model behave and are there outlying observations that might influence the estimation. As already mentioned, the accurate prediction of the drug exposure plays a crucial role in IVIVC modeling. It is conventionally assessed with the average absolute percentage prediction error, $\%PE$ (CDER 1997). However, special attention is required to particular observations that have a strong influence on the estimation of the IVIVC model, as these might represent perfectly normal data, but also may represent problematic data such as outliers, indicate inadequate data sampling, or might be due to an incorrect model (Verbeke and Molenberghs, 2000).

Several diagnostic tools were developed to detect outlying observations. For the linear regression model, residual analysis, leverage, Cook's distance, and local influence are standard techniques. However, not all of these methods retain validity in the case of nonlinear models (St-Laurent and Cook, 1996). Local influence (Cook 1986) assesses the impact of small perturbations of the likelihood function. There-

fore, this method retains validity for nonlinear mixed effects models. Local influence is, however, not always a calibrated metric and measures the relative local influence compared to the other observations. Poon and Poon (1999) extended the method using the conformal normal curvature. This ensures a calibrated metric falling within the unit interval.

Although local influence is a likelihood-based and therefore generally valid method, its use is frequently restricted to (generalized) linear (mixed) models (Verbeke and Molenberghs, 2000, Thomas and Cook, 1990). Some papers applied local influence on other types of data and models, such as nonlinear structural equation models (Lee and Tang 2004), incomplete multivariate and longitudinal data (Jansen et al. 2003), survival data (Escobar and Meeker, 1992), growth data (Shi and Ojeda 2004). However, no paper focuses on local influence for nonlinear mixed effects models in pharmacokinetics. The diagnostic tools used in pharmacokinetics often are restricted to the goodness-of-fit of the model, such as residual analysis (Gabrielson and Weiner, 2001). It must be said that some outlier detection methods have been proposed, but these are based on a leave-one-out principle, e.g., Sadray, Jonsson, and Karlsson (1999).

Whereas $\%PE$ is an overall indicator of the goodness-of-fit recommended by regulatory authorities (CDER 1997), one could also consider the $\%PE$ evaluated per subject, i.e., how well does the hierarchical model predict the AUC and $C_{\max}$ for each subject separately. This quantity assesses how well the model fit the time-profile for the particular subject. In the current paper, we explore the relation between both $\%PE$ and local influence as methods for detecting outlying subjects using a real life example and a simulation study.

The chapter is based on Jacobs *et al.* (2009) and is organized as follows. Local influence is introduced in Section 6.1. The set up of the simulation study is described in Section 6.2. The relation between both outlier detection methods is found in Section 6.4.

The likelihood in this chapter was implemented using the Laplace approximation as described in Pinheiro and Bates (2000).

6.1 Local Influence

Failing to model the IVIVC relation might be due to a number of reasons. Even under the assumption of a correctly specified model, it might occur that some of the data have an unusually high influence on the model estimation and influence the results of the analysis. Such data might represent valid data, but might also signal model

inadequacy, inappropriate data, or insufficient data sampling. Therefore, a thorough study of such potential outlying and influential observations is deemed necessary.

Cook (1986) introduced the concept of local influence. It comprises of assessing the impact of a set of observations on the likelihood function, which is defined as

$$L(\Theta) = \sum_{n=1}^{N} L_n(\Theta), \tag{6.1}$$

where Θ assembles all parameters, N is the number of observations, and $L_n(\Theta)$ is the likelihood contribution of the n^{th} observation. Likelihood-based methods consist of maximizing the quantity $-2\log[L(\Theta)]$. Local influence studies the impact of minor perturbations to this function. If the impact is large, then the perturbation is highly influential. In this case, the data might represent for example an outlying observation. The impact of perturbations to subject k can be assessed via

$$L(\Theta \mid \omega) = \sum_{n=1}^{N} w_n L_n(\Theta), \tag{6.2}$$

where w_n is a weight, 1 if the observation belongs to subject k, 0 else. Then, the local influence of the subject k on the likelihood $L(\Theta \mid \omega)$ can be derived to take the form (Verbeke and Molenberghs, 2000)

$$C_k = 2 \mid \Delta_k' \ddot{L}^{-1} \Delta_k \mid, \tag{6.3}$$

where Δ is defined as the vector of second derivatives of $L(\Theta \mid \omega)$ with respect to w_n and all components of Θ and evaluated in the original estimated maximum log-likelihood. Δ_k is the k^{th} column. $\ddot{L}$ represent the estimated Hessian matrix of the log-likelihood.

However, local influence as defined by Cook (1986) has one strong disadvantage: it is not calibrated, and can take any value (Poon and Poon 1999). Hence, the local influence measurement of a particular subject should be interpreted relative to the local influence of the other subjects. Poon and Poon (1999) introduced the conformal normal curvature approach of local influence, which incorporates a calibration factor:

$$B_k = \frac{\mid \Delta_k' \ddot{L}^{-1} \Delta_k \mid}{\sqrt{tr[(\Delta' \ddot{L}^{-1} \Delta)^2]}}, \tag{6.4}$$

where tr is the trace of a matrix. This calibration ensures that B_k falls within the unity interval. Values above twice the average B_k require further scrutinizing (Verbeke and Molenberghs, 2000).

In our setting, the interest is in detecting influential subjects. Therefore, the w_n are set to 1 for a particular subject and to zero for all others.

The derivation Δ_k of the likelihood to the parameters of the model was performed numerically, i.e.,

$$\frac{L(\Theta + h_\Theta \mid \omega) - L(\Theta \mid \omega)}{h_\Theta}, \tag{6.5}$$

where h_Θ was set to 10^{-7} for each parameter, and each element of the column Δ_k represents the derivative to a different model parameter evaluated for the subject k.

$\%PE$, as introduced in Section 3.3, is one of the traditional model diagnostic utilities. It is of interest to verify how much local influence and $\%PE$ overlap each other as diagnostic tools. Therefore, calculating the $\%PE$ for each subject separately could also be considered as a measurement of the goodness-of-fit of the model for the corresponding subject, hence be used as a method to detect potential outliers. On the other had, one might also consider the $\%PE$ for each subject as a form of aggregated residual analysis ignoring the time component of the data.

6.2 Simulation Set up

A simulation study was set up to study the relation between local influence as defined by Poon and Poon (1999) and $\%PE$. An exponential *in-vitro* dissolution time profile with lag-time ϕ_3 is chosen.

$$F_{i1\ell}(t) = 100\phi_1 \left\{1 - \exp\left[-\left(t - \phi_3\right)\exp\left(\phi_{2i}\right)\right]\right\} + \epsilon_{i1}, \tag{6.6}$$

where $\epsilon_{i1} \sim N(0, \sigma_0^2)$ and $\phi_{2i} \sim N(0, \sigma_{\phi_2}^2)$. ϕ_1 corresponds to the asymptotic maximum amount of drug product release *in-vitro*. The unit impulse response $c_{2k\delta}(t)$ consisted of a one-compartmental IV-bolus injection.

$$c_{2k\delta}(t) = \frac{D}{V}\exp(-k_e t), \tag{6.7}$$

where D represents the dose, V is the apparent volume of distribution, k_e is the subject-dependent elimination rate, and residual error has a log-normal distribution. The hypothetical link function between the *in-vitro* dissolution and the *in-vivo* release was set to

$$F_{i2k\ell}(t) = g^{-1}\left\{\theta_0 + \theta_1 t + s_{ik} + g\left[F_{i1\ell}\left(t\right)\right]\right\},$$

where θ_0, θ_1, and s_{ik} are parameters correcting for the subject-dependent differences between the *in-vivo* release and the *in-vitro* dissolution.

A perturbation was added to one of the three parameters (ϕ_1, θ_0, and θ_1) for the first subject of each simulated study to explore the impact of an outlying subject on

the likelihood estimation. The chosen parameter was multiplied with a factor 2.5, 5, 7.5, or 10. The factors are chosen to generate mildly to extreme outlying profiles. As such, the simulation also explores the power of the method to detect outlying profiles. Per parameter and perturbation factor, 500 studies were generated with 20 subjects. The quantities B_k, $\%PE_{AUC}$, and $\%PE_{C_{\max}}$ were calculated for the obtained model fits.

The correspondence between the obtained measurements is subsequently studied to assess the degree of one method confirming the other. The verification of such correspondence can be done using reliability theory. The concept was introduced by Fleiss (1986). For a more general approach of the concept, we refer to Vangeneugden et al. (2004). In our case, the correspondence reduces to the correlation between the metrics.

6.3 The Case Study

Our case study is a long acting injectable formulation for intramuscular administration. The data is described in Section 2.2. An IVIVC model is established for one batch only. The impact of each subject on the likelihood function is assessed and potential outlying subjects are identified. A more thorough understanding of the data allows for an improved model estimation. An IVIVC for the formulation is set up. However, the degree of metabolism of compound is subject-dependent. Therefore, the metabolic ratio M is included in both the clearance and the IVIVC link function to correct for poor versus extensive metabolizing subjects,

$$F_{i2k\ell}(t) = g^{-1}\left\{\theta_0 + \theta_{01}M + \theta_1 t + \theta_{11}Mt + g\left[F_{i1\ell}(t)\right]\right\}.$$

Note that the random effect s_{ik} was removed from the model, because it was not statistically significant. The controlled-release plasma concentration-time profile data was modelled as a mixture of a slow-release and a small immediate-release process, because a small amount of immediate release of drug product was observed in all patients. This was due to specific steps in the manufacturing process of the formulation.

6.4 Results

The quality of the IVIVC modeling is investigated using residual analyses and $\%PE$ following the FDA guidelines. Although the FDA guidelines do not mention diagnostics for outlier detection, it is worth to verify whether the quantity $\%PE$ would be a suitable diagnostic tool, in particular for the assessment of outlier detection. A

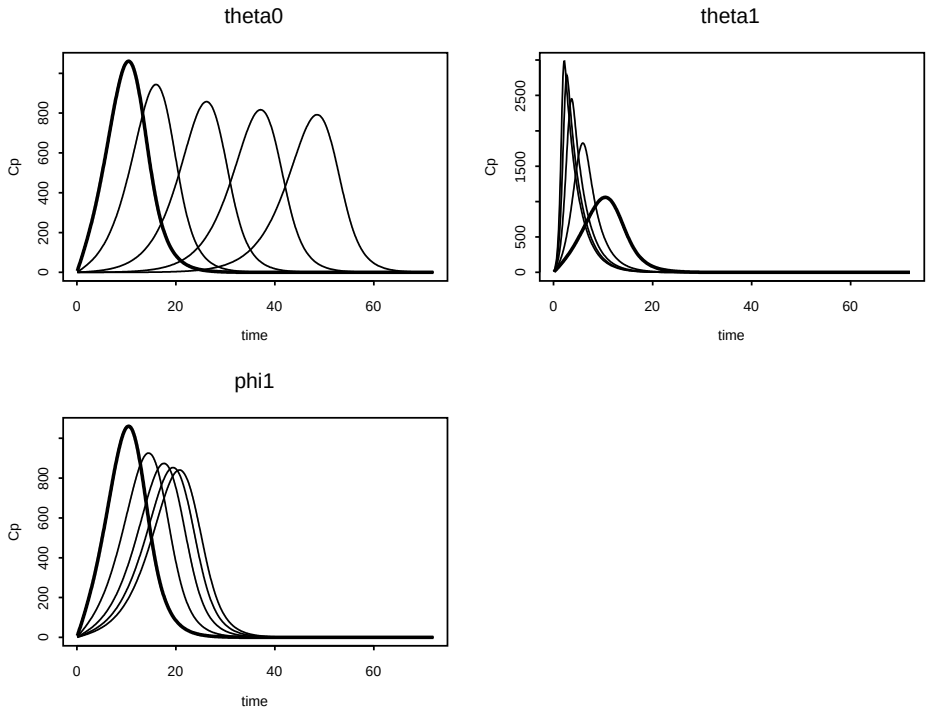


Figure 6.1: The unperturbed (bold) and perturbed controlled-release profiles with increasing degrees of perturbation (2.5,5,7.5,10) used in the simulation study.

simulation was set up to explore a potential relation between $\%PE_{AUC_k}$ and local influence B_k and the power of both methods to detect subjects with outlying plasma concentration-time profiles.

Figure 6.1 shows the standard profile as well as the perturbed profiles for one subject. The unperturbed simulated profiles are shown in bold. A perturbation is added to the model parameters for one subject leading to the other profiles. Residual and random-effects variances are set to zero for graphical purposes only. Table 6.1 contains the number of studies with an appropriate model convergence and positive-definite Hessian matrix. This number is low despite the use of a grid of starting values. However, the focus of the chapter is on the outlier detection rather than fitting of the simulated data. The $\%PE_{AUC}$ of the perturbed subjects are depicted in Figure 6.2 for the different perturbation scenarios. From this graph, it is clear that $\%PE_{AUC}$ is capable of detecting outlying pharmacokinetic time profiles for parameters θ_0 and

Table 6.1: Number of simulated studies with an appropriate model convergence as a function of the perturbation factor.

	0	2.5	5	7.5	10
ϕ_1	324	322	330	334	344
θ_0		328	329	319	281
θ_1		351	322	340	311

Table 6.2: The correlation between B_k and $\log(\%PE_{AUC_k})$ indicate an absence of relation for each of the perturbation scenarios.

	0	2.5	5	7.5	10
ϕ_1	0.11	0.09	0.12	0.12	0.13
θ_0		0.11	0.15	0.17	0.17
θ_1		0.10	0.08	0.09	0.11

to a smaller extent also for θ_1, but not for ϕ_1, i.e., the boxplots show a large and small increase of $\%PE_{AUC}$ with increasing perturbation for θ_0 and θ_1, respectively. Figure 6.3 suggests that $\%PE_{C_{\max}}$ fails to detect outlying subjects. On the other hand, Figure 6.4 depicts the boxplots of the local influence metric B_k of the simulation study as a function of the degree of perturbation. Again, the figure is restricted to the perturbed subject of each simulated trial. No impact of the perturbation on the local influence is observed for the dissolution parameter ϕ_1, whereas a clear increase of B_k is obtained for the IVIVC parameters θ_0 and θ_1. This is a reassuring, though surprising result, i.e., it is possible to detect outlying subjects using local influence in the case of IVIVC modeling; however, only if the perturbation occurs on the *in-vivo* release parameters θ_0 and θ_1, and not when the perturbation occurs at ϕ_1, which corresponds to the asymptotic cumulative percentage of product released *in-vitro*. The insensitiveness of the method for the perturbation at ϕ_1 is striking; Even in the extreme case of ϕ_1 reduced by a factor ten, both methods are not able to detect the outlying behavior, although the impact of the perturbation on ϕ_1 is very small in Figure 6.1.

Bearing in mind this result, one might wonder whether both methods confirm

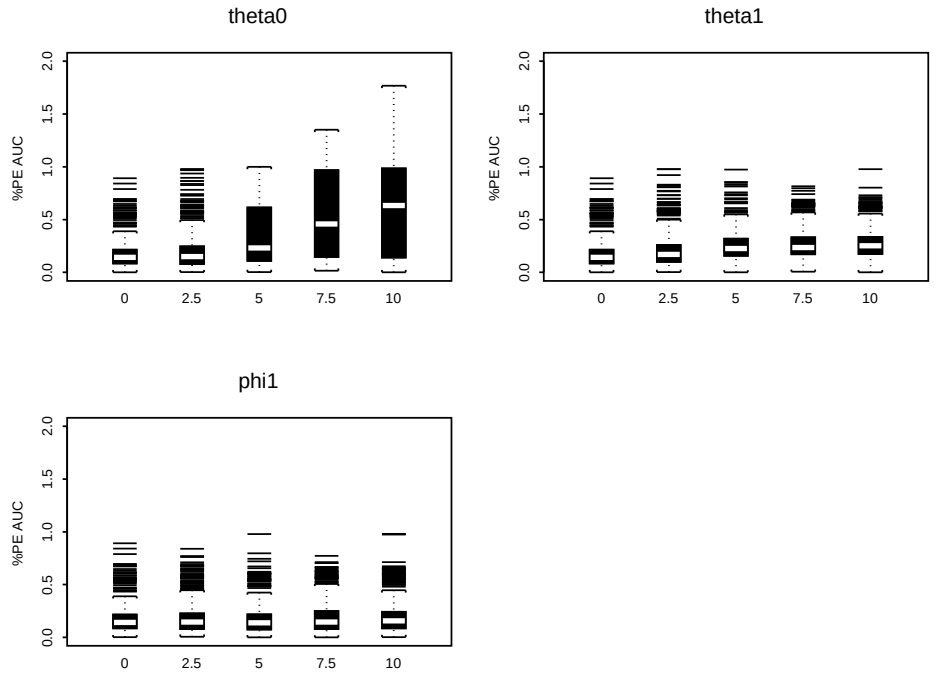


Figure 6.2: Boxplots of $\%PE_{AUC}$ of the perturbed subject as a function of the perturbation factor. For each perturbation level the number of subjects was 20 (19 unperturbed, 1 perturbed) and 500 replicates were generated.

each other, thus whether there is a one-to-one relation between $\%PE_{AUC_k}$ and B_k. Therefore, the local influence measurements B_k and $\%PE_{AUC_k}$ are depicted per perturbation scenario in Figure 6.5. The figure suggests that only a weak relation between B_k and $\%PE_{AUC_k}$ can be expected for the different scenarios. A more formal analysis of the simulation was carried out with a correspondence analysis (Vangeneugden *et al.,* 2004). Table 6.2 contains the correspondence $\text{Corr}(B_k, \%PE_{AUC_k})$ for the different perturbation settings. This formally confirms that the methods share little to no information on outlier detection, despite the fact that both are capable to identify outliers. Therefore, the focus on residual analysis and $\%PE_{AUC_k}$, which is a special case of the former, can be considered as appropriate but insufficient and the data can be explored from a different angle using local influence. The relation between B_k and $\%PE_{C_{\max_k}}$ is tabulated in Table 6.3. Also in this case, the correspondence is weak

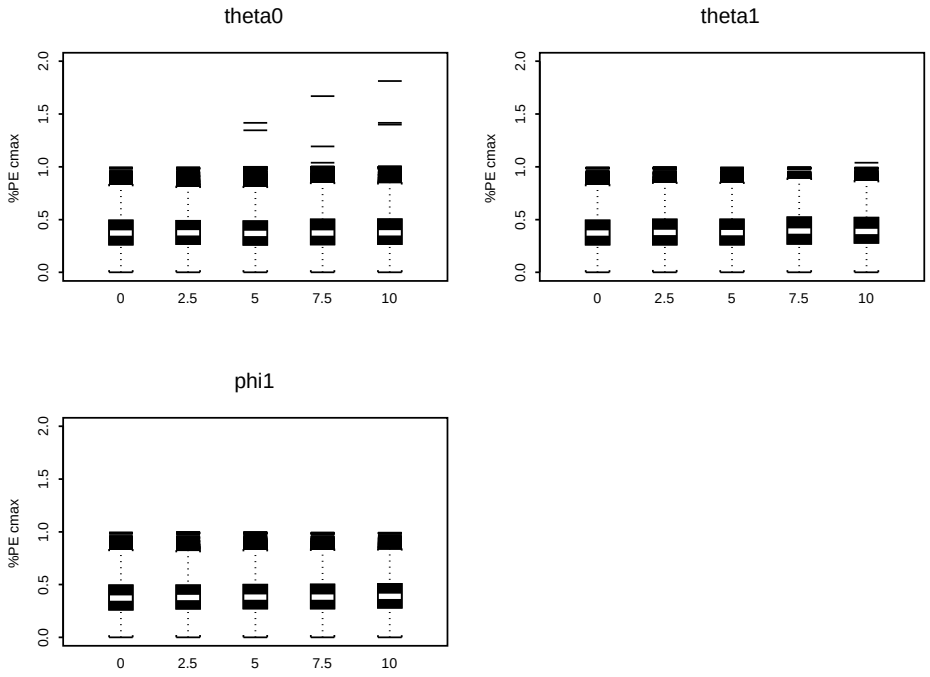


Figure 6.3: Boxplots of $\%PE_{C_{\max}}$ of the perturbed subject as a function of the perturbation factor. For each perturbation level the number of subjects was 20 (19 unperturbed, 1 perturbed) and 500 replicates were generated.

(25-30%). The scatterplot between B_k and $\%PE_{C_{\max_k}}$ show a similar picture as in Figure 6.5.

After the introduction of the local influence method and the exploration of its properties, it is worth submitting the method to a stress test to find potential issues or weaknesses of the technique. Our case study has some specific characteristics which pose some challenges to IVIVC modeling and thus provide a good opportunity to explore the behavior of model diagnostics under circumstances where they matter most; (1) an imbalanced study design in the sense that immediate-release and controlled-release plasma concentrations are not measured in the same patients, (2) unusual profiles with the majority of the plasma samples close to the detection limit, (3) only few samples measured near the peak plasma concentrations, and (4) the specific manufacturing process of the formulation leads to a small amount of immediate

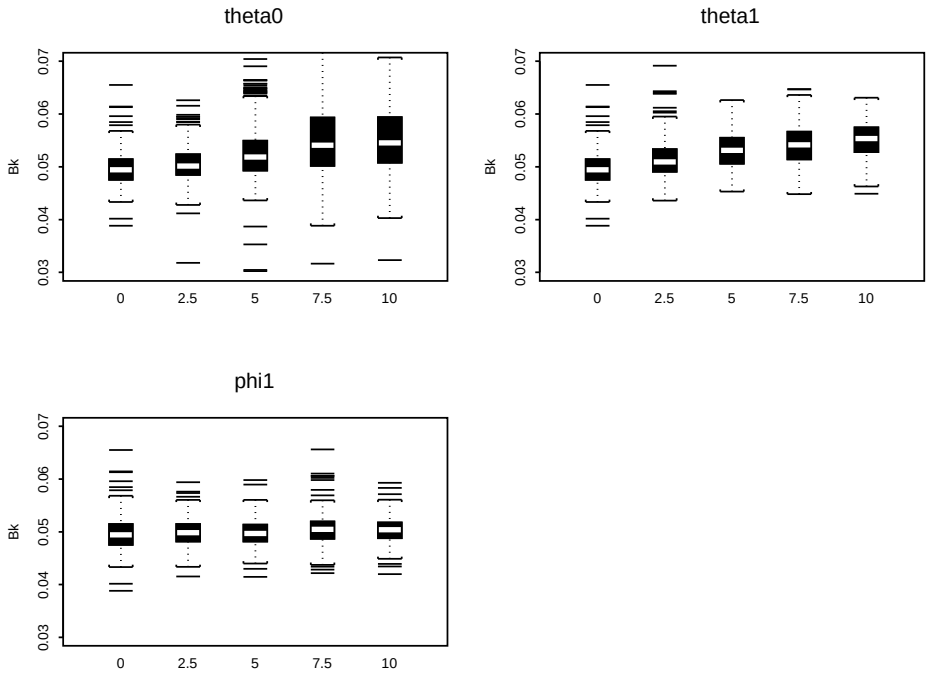


Figure 6.4: Boxplots of B_k of the perturbed subject as a function of the perturbation factor. For each perturbation level the number of subjects was 20 (19 unperturbed, 1 perturbed) and 500 replicates were generated.

release of the drug product.

The goodness-of-fit of the model for the case study was 20% and 15% for $\%PE_{AUC}$ and $\%PE_{C_{\max}}$, respectively. Figure 6.6 shows the local influence of the different patients. Six profiles had a local influence of more than twice the average of B_k and are worth further scrutinizing. The model fit of the nine patients with the highest influence metric are depicted in Figure 6.7. It turns out that the four patients with the highest influence have a model fit for which the estimated peak plasma concentration is estimated too early or too late. Whereas at least seven detectable plasma levels are sampled for more traditional pharmacokinetic profiles, the profiles in this case study have only 2 to 5 plasma samples which contribute to the peak plasma concentration estimation. Thus, these most extreme local influence values indicate that a different plasma sampling scheme with more sampling during the main release period of the

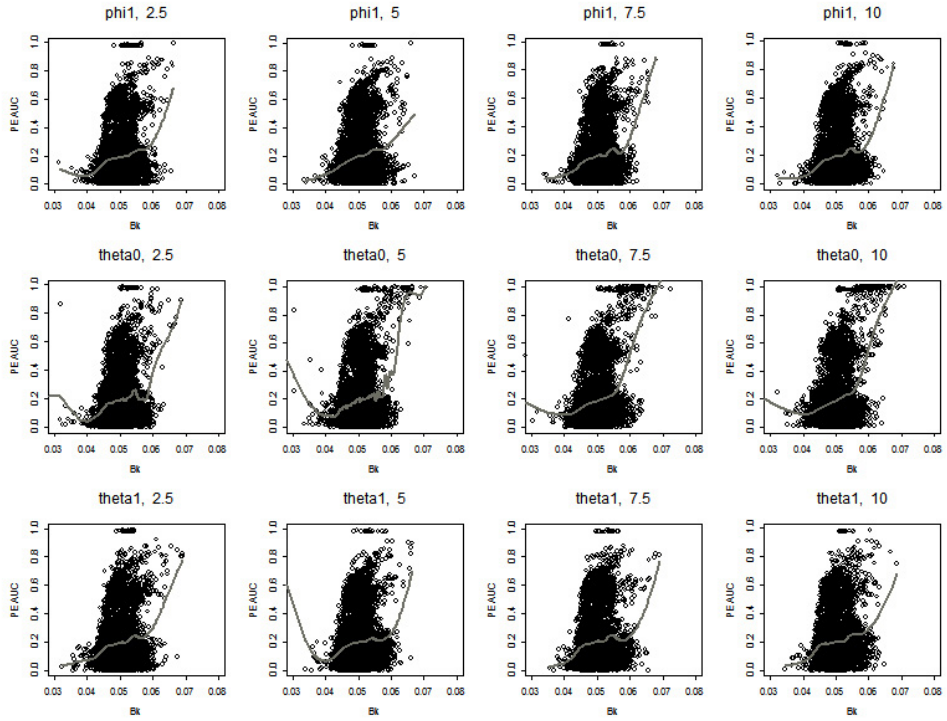


Figure 6.5: Explorative relation between influential subjects B_k and $\%PE_{AUC_k}$ in a simulation study with different degrees of perturbations implemented for each parameter estimate. The grey line depicts a smoothed relation between the two metrics.

product ought to be considered. However, also some other patients in Figures 6.8 – 6.12 with smaller influence metrics have predicted peak plasma concentrations not coinciding with the ones observed. In these cases, less plasma samples were taken near the peak and, therefore, less observations were predicted erroneously. As the local influence of the subject is an aggregate measure over the entire profile, few outlying samples are insufficient to yield the entire profile as influential. As a result, the local influence is smaller for these patients.

No major deviation was observed for patient 30068 (Figure 6.7) despite the high values for local influence. On the contrary, patient 30024 has the fifth most extreme local influence. He has an aberrant profile, i.e., high plasma concentrations during the first week after administration. This is due to an unintentional early release. It is of interest to note that the profile of the patient having the 8^{th} highest (30010)

Table 6.3: The correlation between B_k and $\log(\%PE_{C\max_k})$ indicate an absence of relation for each of the perturbation scenarios.

	0	2.5	5	7.5	10
ϕ_1	0.28	0.27	0.27	0.28	0.27
θ_0		0.25	0.22	0.24	0.20
θ_1		0.29	0.31	0.29	0.33

local influence also exhibits an elevated release in the first week. The third patient (30032), showing such an early release out of a total of 54 patients, does not have a high influence metric (0.012), but in this case, the early peak is restricted to two sampling points, hence the method might discard the patient as having an erroneous sample rather than an aberrant profile. These three patients can be considered as real outliers because the release of the drug product is not immediately after injection like for the other patients, but seems to take several hours or even some days. Therefore, the model was refitted after exclusion of these patients. However, the $\%PE$ remains unaltered for both AUC and $C_{\max}$ as can be expected from the absence of correlation between $\%PE$ and local influence.

6.5 Discussion

The emphasis of statistical model diagnostic techniques in the case of nonlinear mixed-effects modeling traditionally restricts to residual analyses. In the specific case of pharmacokinetics, the exclusion of aberrant observations is typically subject to the pharmacokinetisist's subjective judgment. The reason for this is that only few outlier detection methods are available in standard software or described in literature, to our knowledge. These are based on the leave-one-out principle, see for example Sadray, Jonsson, and Karlsson (1999). In this paper, we have shown that local influence is a powerful tool for the detection of outliers not only in linear and generalized linear (mixed effects) models, but that it also can be used to detect complex outlying pharmacokinetic time profiles. The focus has been on a specific type of pharmacokinetic models, i.e., IVIVC models.

Regulatory guidelines state that the model fit of IVIVC models should comply with $\%PE < 10\%$. This $\%PE$ can be considered as an aggregated residual measurement

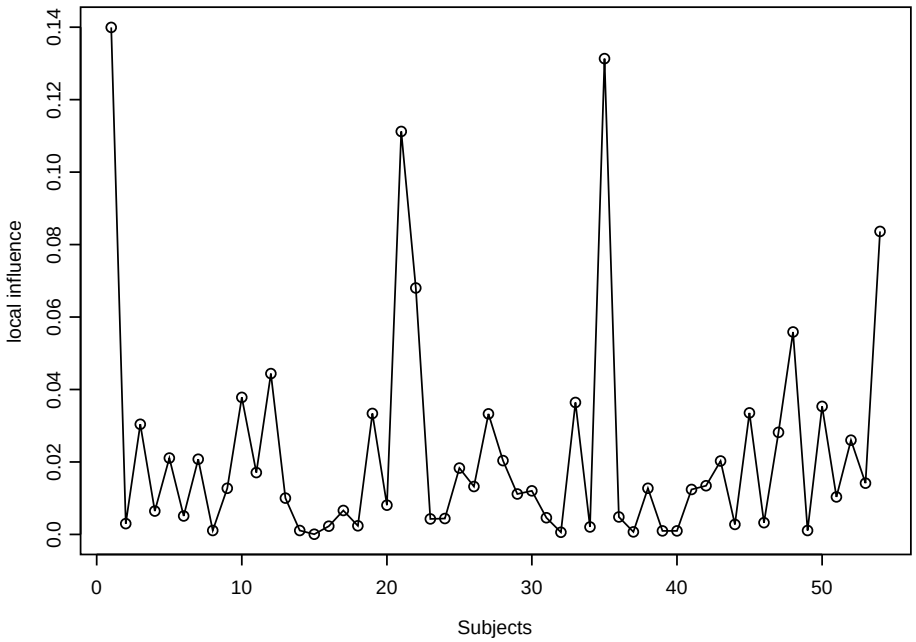


Figure 6.6: Influence of the subjects on the likelihood estimation for a long acting injectable formulation.

that ignores the time aspect of the data. An increase of $\%PE_{AUC}$ is found for the perturbated profiles in the simulation. This standard method detects the deviation and, therefore, can be used to detect outliers. On the contrary, $\%PE_{C_{\max}}$ turns out to be useless to detect outliers. The local influence method is also applied in the same simulation. Again, a perturbation dependent increase is found by using local influence. This proves that local influence is also capable to detect outlying subject profiles. It is worth noting that the correlation between both methods is weak. This suggests that both methods detect different aspects of the outlying profiles. Therefore, the focus on $\%PE_{AUC}$ as per guideline can be considered as appropriate but insufficient. The calculation of the local influence of a subject's plasma concentration-time profile sheds some new light on the data and focuses on the problem from a different angle.

However, the case study clearly points out one of the limitations of both techniques. Both $\%PE_{AUC}$ and local influence are an aggregate measurement of the fit of

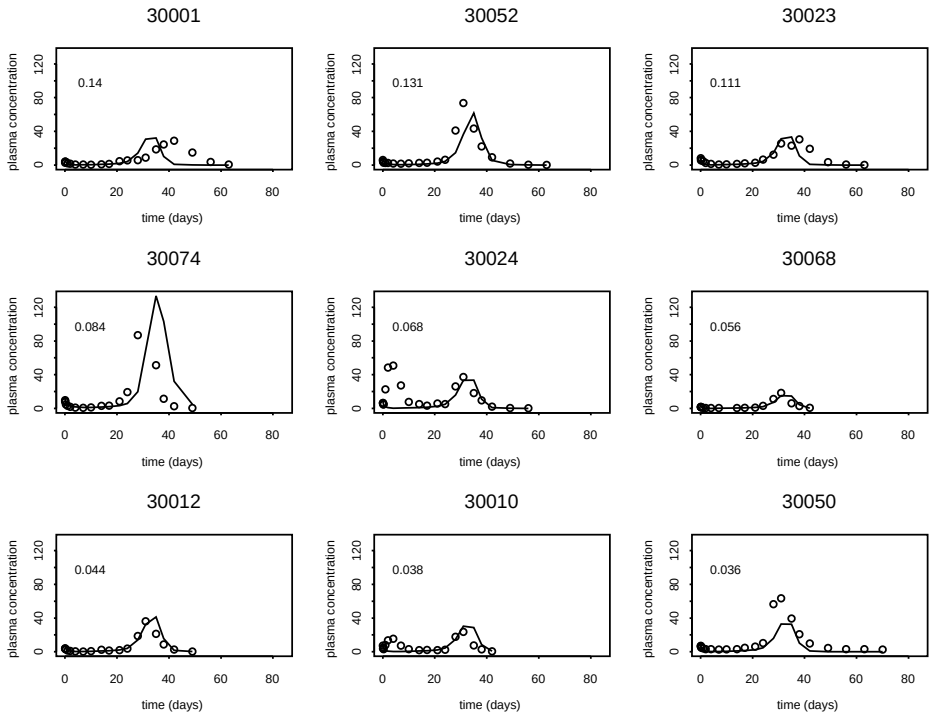


Figure 6.7: Model fit for the nine subjects with the largest influence on the likelihood estimation for a long acting injectable formulation. The influence is added in the upper left corner for each subject.

the entire profile. The methods work well for traditional pharmacokinetic profiles, but their performance might decrease for cases of controlled-release formulations where the majority of the many plasma concentrations is close to the detection level and only few samples describe the peak plasma concentration. However, local influence does take the time aspect into account via the likelihood approach, in contrast to $\%PE$ based on AUC and $C_{\max}$, which ignores the time aspect. As a result, local influence is capable of detecting a change in the time of the peak plasma concentration $t_{\max}$, which $\%PE$ based on AUC cannot. However, if very few plasma samples are taken near the peak plasma concentration as in the case study, the misfit of the $C_{\max}$ is not detected with local influence, but would be found with the $\%PE$ method. Therefore, it is our recommendation not to restrict the goodness-of-fit to either of the techniques, but rather to consider both as complementary due to their focus on different aspects

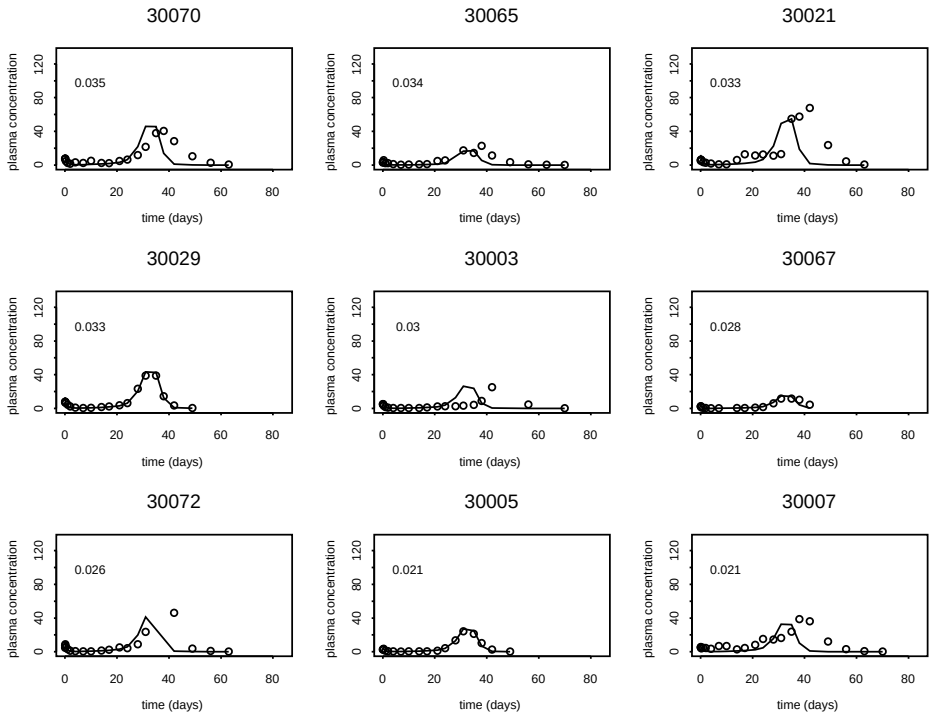


Figure 6.8: Model fit for the nine subjects with the largest influence on the likelihood estimation for a long acting injectable formulation. The influence is added in the upper left corner for each subject (continued).

of data modeling.

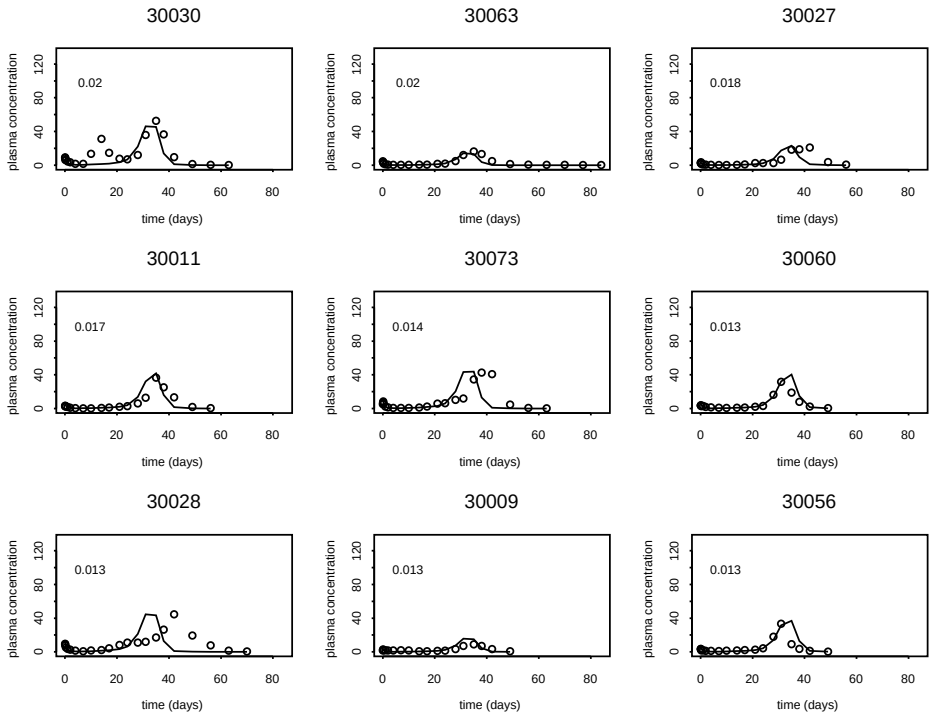


Figure 6.9: Model fit for the nine subjects with the largest influence on the likelihood estimation for a long acting injectable formulation. The influence is added in the upper left corner for each subject (continued).

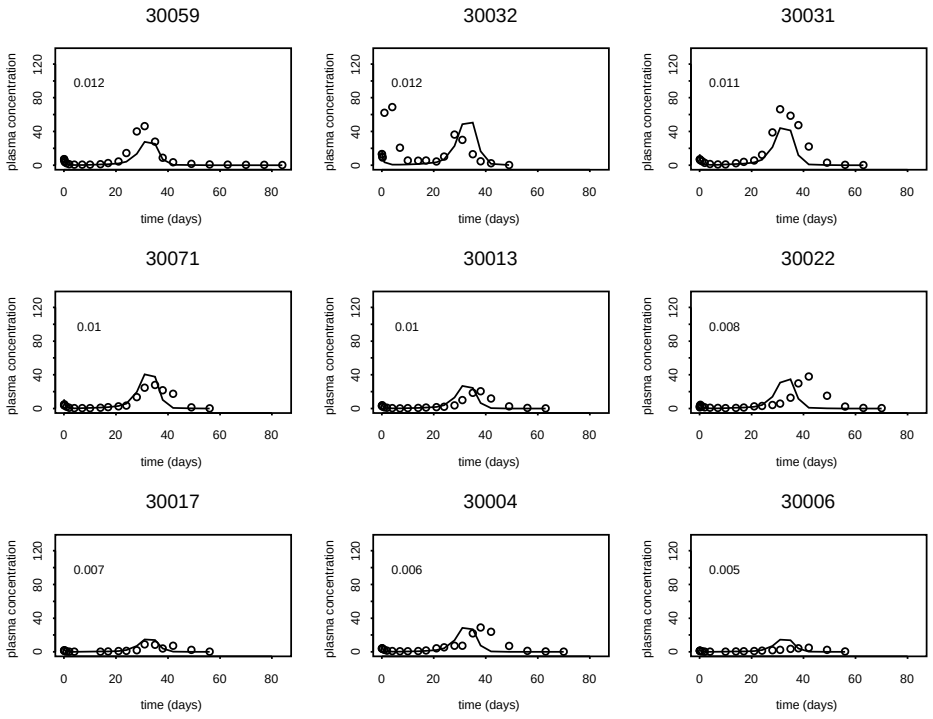


Figure 6.10: Model fit for the nine subjects with the largest influence on the likelihood estimation for a long acting injectable formulation. The influence is added in the upper left corner for each subject (continued).

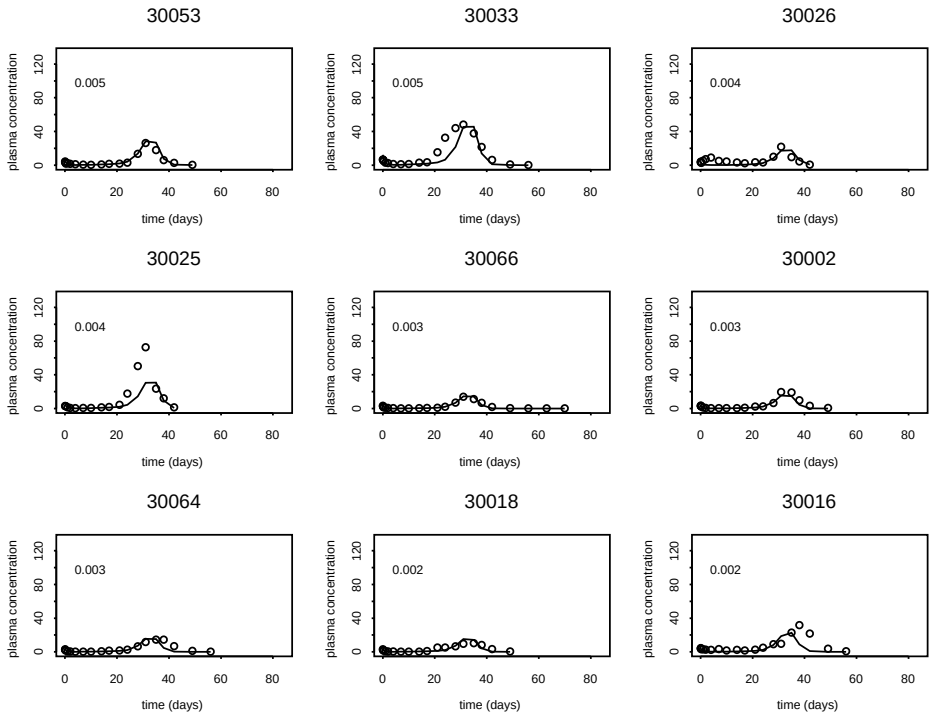


Figure 6.11: Model fit for the nine subjects with the largest influence on the likelihood estimation for a long acting injectable formulation. The influence is added in the upper left corner for each subject (continued).

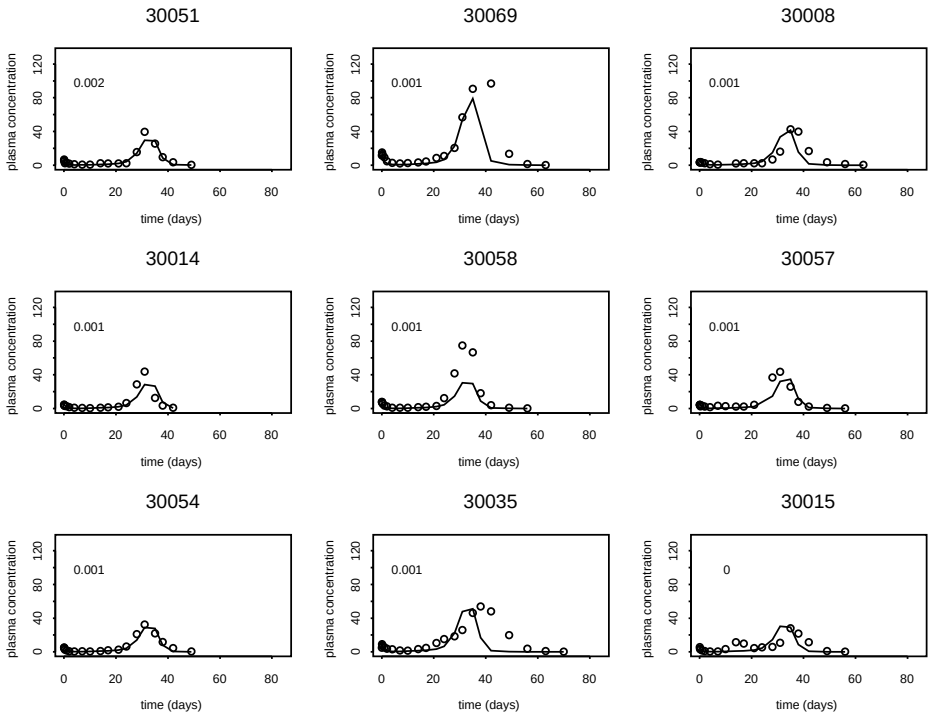


Figure 6.12: Model fit for the nine subjects with the largest influence on the likelihood estimation for a long acting injectable formulation. The influence is added in the upper left corner for each subject (continued).

7

Can *In-vitro* Dissolution Specifications be Determined by its Clinical Relevance?

In-vitro tests play a vital role in the pharmaceutical industry. They avoid the use of resource-intensive clinical trials by performing less resource consuming, well-controlled laboratory tests instead. One of these tests is the *in-vitro* dissolution test for controlled-release formulations. It assesses the *in-vitro* release over time of the medicinal product. Setting dissolution specifications of such controlled-release formulations corresponds to defining as the number of release percentages that must be attained after a prespecified time interval under controlled conditions, such as pH, dosage form, and excipient. Such dissolution specifications ought to retain the *in-vivo* drug exposure bioequivalent, i.e., within 20% of the reference exposure, and as such the therapeutic effect should be unaltered. Hayes *et al.* (2003) described that these dissolution specifications can be optimized as a function of the bioequivalence criteria.

We study the clinical relevance of these dissolution specifications by formulating an answer to the question as to whether these limits guarantee that the clinical effects observed remain unchanged? If specifications are too liberal, this might lead to an increase of unwanted effects or the loss of efficacy. If, on the contrary, the speci-

fications are too conservative, too many acceptable batches of drug product would unnecessarily be discarded for alleged insufficient quality.

In this chapter, we present a case study in which we determine the relation between the *in-vitro* release-time profile and the clinical outcome for the controlled-release formulation. We will use this relation to translate *in-vitro* changes into changes of the clinical response. As such, clinically relevant *in-vitro* release specifications can be determined. It ought to be noted that it is not our intent to show that a perfect relation exists for our case study. Rather, the principal idea of our paper is that one can apply the concept to any other drug product, under certain conditions.

Instead of translating changes of the *in-vitro* dissolution time profile directly into clinical effects, we use a pharmacokinetic/pharmacodynamic (PK/PD) relationship to relate the controlled-release plasma concentration-time profile to the receptor binding. For our case study, a consensus about the clinical interpretation of the receptor binding exists (Kapur *et al.* 1999, 2000), reducing our research quest to finding a relation between the *in-vitro* dissolution time profile and the controlled-release plasma concentration-time profile. *In Vitro-In Vivo Correlation* (IVIVC) modeling combines the *in-vivo* drug release with the immediate-release plasma concentration-time profile to predict the controlled-release plasma concentration-time profile. This is established by imposing a posited relationship between the *in-vitro* dissolution and the unobserved *in-vivo* release and verification of its predictive capacity.

The combination of the above models will enable us to relate D_2 receptor binding with *in-vitro* release properties of the drug formulation and, therefore, indirectly to associate a clinical interpretation to the dissolution specifications, i.e., the increased hazard to occurrence of extrapyramidal symptoms with D_2 receptor occupancy exceeding 80%. The lack of efficacy is associated with a D_2 receptor binding below 65%.

The chapter is based on Jacobs *et al.* (2009) and organized as follows. The case study, motivating this research, is described in Section 7.1. The set up of the simulation study is described in Section 7.2. The model fits as well as the simulations for establishing the clinically relevant dissolution specifications are the subject of Sections 7.3.1 and 7.3.2, respectively.

7.1 The Case Study

Again, the data from long acting injectable formulation for intramuscular administration of an antipsychotic agent with potent dopamine-D2 antagonistic properties were used to relate *in-vitro* changes to *in-vivo* effects. The averaged percentage *in-vitro*

release was registered up to 43 days. The *in-vitro* specifications were determined as releasing 50% of the drug product within days 26–35, and 80% should be released before day 41.

A PK/PD relationship between plasma concentrations and target D_2 receptor binding was established (Nyberg *et al.* 1995, Remington *et al.* 1998, and 2006, and Kapur *et al.* 1999). All of these studies indicate that 80% receptor occupancy is a threshold above which a higher incidence of extrapyramidal symptoms (EPS) is observed. Beyond this point, a steep increase of this side effect is observed, resulting in 30% of the subjects passing this threshold to develop EPS. On the other hand, a receptor binding of 65% is required to achieve clinical benefit (Kapur *et al.* 1999, 2000).

7.2 Simulation Study

Investigations of *in-vitro* release-time profiles traditionally focus on two aspects: when are 50% and 80% of the drug product released? The rationale for this is the consensus that the time of 50% release is considered to correspond to the time the maximum plasma concentration ($t_{\max}$) is attained. The difference between the time of 80% and 50% release on the other hand is considered to indicate the release rate. Intuitively, one can expect higher peak plasma concentrations ($C_{\max}$) in case of a faster release than for a shallow release mechanism. Therefore, two simulations are required; the $t_{\max}$ simulation will study the impact of changes to the time of 50% release while maintaining the rate of release fixed, the $C_{\max}$ simulation assesses the impact of altering the rate of release while fixing the time of 50% release.

Both simulations have a similar set up to mimic the effects of switching from the reference batch to a novel one, i.e., the reference batch is administered five times, followed by three times the test batch. In order to mimic reality, parameter estimates of the analysis of actual data were used. A two-week administration interval was used. For each simulation, a number of *in-vitro* release-time profiles is generated. Subsequently, the IVIVC model is combined with the PK/PD model to translate the changes of the *in-vitro* release-time profile into changes of the clinical outcome. The plasma concentration-time profiles of two hundred patients are generated for each release profile and the number of times that the clinical threshold is passed is counted. The plasma concentration-time profiles were generated using the established IVIVC model. The inter-individual variability of the clearance as observed in the case study is included in the simulation.

The $t_{\max}$ simulation investigates the impact of altering the time of 50% release

(t_{50}) while fixing the release rate $F'_{i2k}(t_{50}, \lambda, \kappa)$,

$$\begin{aligned} F'_{i2k\ell}(t_{50}, \lambda, \kappa) &= p_{1\ell}\frac{\kappa}{\lambda}\left[\left(\frac{t_{50}}{\lambda}\right)^{(\kappa-1)}\right]\exp\left[-\left(\frac{t_{50}}{\lambda}\right)^{\kappa}\right] \\ &\quad +(1-p_{1\ell})\frac{\kappa_{2\ell}}{\lambda_{2\ell}}\left[\left(\frac{t_{50}}{\lambda_{2\ell}}\right)^{(\kappa_{2\ell}-1)}\right]\exp\left[-\left(\frac{t_{50}}{\lambda_{2\ell}}\right)^{\kappa_{2\ell}}\right]. \quad (7.1) \end{aligned}$$

The parameters $\kappa_{2\ell}$ and $\lambda_{2\ell}$ determine the Weibull function corresponding to the limited release after 10 to 20 days and will be fixed. The following set of equations is then solved for κ and λ to alter the main release such that the t_{50} varies from day 10 to 59, while the overall slope at t_{50} remains unaltered.

$$\begin{aligned} 0.50 &= p_{1\ell}\left\{1-\exp\left[-\left(\frac{t_{50}}{\lambda}\right)^{\kappa}\right]\right\}+(1-p_{1\ell})\left\{1-\exp\left[-\left(\frac{t_{50}}{\lambda_{2\ell}}\right)^{\kappa_{2\ell}}\right]\right\}, \\ F'_{i2k\ell}(t_{50}, \lambda, \kappa) &= F'_{i2k\ell}(t_{50_{\text{ref}}}, \lambda_{1\ell}, \kappa_{1\ell}). \quad (7.2) \end{aligned}$$

Here, $t_{50_{\text{ref}}} = 31.63$ is defined as the time 50% has dissolved for the reference batch. The set of equations was solved numerically.

The $C_{\max}$ simulation investigates the impact of the release rate $F'_{i2k\ell}(t_{50}, \lambda, \kappa)$ when t_{50} is fixed to $t_{50_{\text{ref}}}$. Therefore, λ and κ should satisfy following set of equations:

$$\begin{aligned} F_{i2k\ell}(t_{50_{\text{ref}}}, \lambda, \kappa) &= F_{i2k\ell}(t_{50_{\text{ref}}}, \lambda_{1\ell}, \kappa_{1\ell}), \quad (7.3) \\ F'_{i2k\ell}(t_{50_{\text{ref}}}, \lambda, \kappa) &= F'. \quad (7.4) \end{aligned}$$

Again, the set of equations was solved numerically. In the simulation, $\log(F')$ varies from -4.0 to -1.5, by steps of 0.05. The reference slope $F' = 0.096$.

For both simulations, it is important to minimize the number of incorrect decisions for each batch of the product, i.e., the batch being compliant to the dissolution specifications but with clinically significant effects; otherwise it would be violating the dissolution specifications while having no clinical changes. This means we want to maximize the probability of a correct decision:

$$P(\text{correct}) = P(C \mid D)P(D) + P(C' \mid D')P(D'), \quad (7.5)$$

where $P(C \mid D)$ corresponds to the probability that no clinical changes (C) are observed given that the batch complies with the dissolution specification (D). $P(C' \mid D')$ is the probability that clinically significant effects are observed for a batch with its dissolution profile outside the specifications. The above probabilities can be calculated based on the simulations and the optimal dissolution specifications can be obtained.

It is important to note that all simulated patients receive a fixed dose of 50 mg. For many patients, such a dose would be too high and lower doses would be administered

to reduce the risk of safety issues while maintaining the clinical benefits. As such, the incidence of passing the safety threshold does not correspond to clinical reality. However, the importance of the simulation lies in the change of passing the safety threshold rather than the absolute number. This applies to both of the simulations.

7.3 Results

7.3.1 IVIVC model

A two-stage IVIVC hierarchical model is fitted to the data set. Attention is restricted to the active moiety. In the first stage, the immediate-release plasma concentration-time profile as well as the *in-vitro* release-time profile are modelled. The estimates obtained are then used in the second stage to model the controlled-release plasma concentration-time profile from the commercial batch, further denoted as the reference batch, according to convolution model (3.1). From the graphs of the individual fits in the previous chapter, it can be seen that the model exhibits a good fit for the majority of the patients. For some patients, the time of the peak is estimated either too early or too late. A few other subjects had a plasma concentration peak higher (lower) than estimated, suggesting that these patients had a lower (higher) elimination rate than observed from the immediate-release plasma concentration-time profile. Further, all plasma concentration-time profiles share a small peak shortly after administration. The manufacturing of a depot formulation requires the overencapsuling of the particles. At the end of this process, the remaining free product is washed away with ethanol. However, a very small, variable quantity ($\sim 1\%$), close to the overencapsuled product, is not washed off. This free product is released immediately after injection. Therefore, the controlled-release plasma concentration time-profile data are modelled as the mixture of a controlled-release and a very small amount of immediate-release drug product (Jacobs *et al.* 2008). The fit is quantified using the average percent prediction error ($\%PE$) and its 95% confidence interval for both AUC and $C_{\max}$. It is estimated to be 20 and 15%, respectively.

The proposed IVIVC model has some weaknesses: the fit is not perfect for each subject's plasma concentration-time profile. A number of explanations can be postulated. First, the immediate-release data originate from a different study, i.e., patients different from the ones who received the controlled-release plasma concentrations; hence one might question whether the estimation of the clearance and its associated variability is representative for the clearance of the patients from the controlled release study. Second, the formulation itself is a long acting injectable formulation with the main release after several weeks, i.e., it has many observations below the detection

limit for the first weeks. The likelihood-based model fitting encounters the problem that it aims at optimizing the fit of the model at all time points. However, the logarithmic transformation of the plasma concentration is fitted in accordance with the distribution of the data. Owing to the over-representation of values close to zero and the logarithmic distribution, the algorithm could experience difficulties to optimize the fit of the plasma concentration peaks. This problem is absent in traditional PK-analyses. The sampling issue for the case study is described formally in Jacobs *et al.* (2009).

Despite the fit of the IVIVC model failing the acceptance criteria of to the FDA guideline (CDER 1997), the model is retained; the key message of the current paper is not the establishment of an IVIVC model for our case study, but rather to capitalize on the concept of attributing a clinical interpretation to changes of the release-time profile and as such also to the release specifications.

7.3.2 Simulation Results

The designs of the two simulations are described in Section 7.2. Both simulations mimic the situation where patients at steady state conditions switch from an old to a new batch of product.

The release profiles in the $t_{\max}$ simulation are depicted in Figure 7.1(a). Figure 7.2 contains a selection of the corresponding simulated plasma concentration-time profiles. As expected, the effect of changing t_{50} has a dramatic impact on the plasma concentration-time profiles when the time of 50% release differs substantially from the one of the reference batch, i.e., t_{50} equal to day 32 versus much earlier such as t_{50} before day 25. The extreme case of t_{50} at day 15 results in a coincidence of the plasma concentration peaks from the last administration from the reference batch and the first administration from the new batch. In this case, the coincidence of the release of the fifth and sixth administration leads to plasma concentrations far beyond the safety threshold. The more the release profile approaches the reference batch and t_{50} increases towards day 32, the more the plasma concentration-time profiles stemming from the fifth and sixth administration disentangle, i.e., the more the plasma concentrations exhibit two peaks. This can be seen in Figure 7.2, where the plasma concentration-time profiles of the fifth and sixth administration coincide for t_{50} up to day 20. The plasma concentration-time profile still coincides partially for t_{50} at day 25. For t_{50} from day 30 onwards, the profiles clearly separate. Having simulated the plasma concentration-time profiles for the batches with different values for t_{50}, one can now calculate the probability of a correct classification as a function of the

in-vitro dissolution specifications. Figure 7.1(b) depicts this probability as a function of the lower and upper *in-vitro* limit. From this figure, it turns out that the highest probability of a correct classification is obtained for $27 < t_{50} < 59$.

The *in-vitro* release profiles used in the $C_{\max}$ simulation are depicted in Figure 7.1(c). The reference slope at t_{50} of the reference batch is 0.096. The plasma concentration-time profiles for a selection of the different release batches are depicted in Figure 7.3. It is clear from the simulated plasma concentration-time profiles that steeper (slope=0.135) slopes for the release profiles result in higher plasma concentration peaks, whereas more shallow (slope=0.027) *in-vitro* release profiles yield lower plasma concentration-time profiles. This is not surprising: the peak plasma concentration is higher if the release is limited to a very short time frame compared to when the release would be spread over time. Again, one can calculate the probability of a correct classification as a function of the choice of the *in-vitro* dissolution specification for the slope F' at t_{50}. This is depicted in Figure 7.1(d). The optimal dissolution specification for the slope turns out to be $-3.4 < \log(F') < -2.3$.

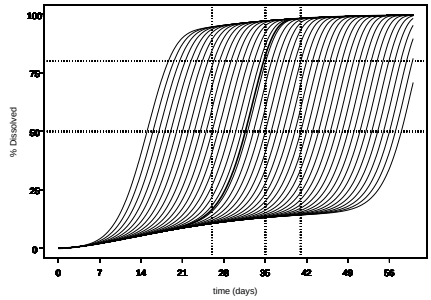
7.4 Discussion

Quality control plays an important role in the manufacturing process of medicines. In the case of a controlled-release drug formulation, an *in-vitro* release-time profile and the associated release specifications is one of the tests to ensure that the drug product retains the proven clinical behavior within limits. However, this clinical behavior is not incorporated at the time the release specifications are set, apart from the bioequivalence criteria, which claim that a difference of 20% in exposure is not clinically relevant. In this paper, we have described a way to include this information and applied it to a real life manufacturing study.

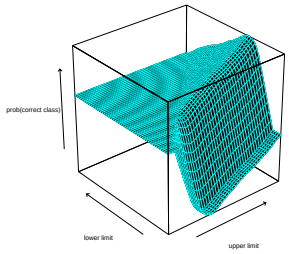
The link of the clinical behavior to the release specification is done by establishing a relationship between changes of the release-time profile and the clinical effects. An IVIVC model is fitted to relate the release-time profile to the plasma concentration-time profile in a first step. The second step consists of a PK/PD model, which relates the controlled-release plasma concentration-time profile to the target receptor binding. Occupancy levels at these target receptor have an established clinical interpretation (Kapur *et al.* 1999, 2000). Combining the above models allows for simulated plasma concentration-time profiles for a different number of batches. The specification of the dissolution limits reduces to an optimalization problem: to maximize the probability of a correct decision, i.e., if the *in-vitro* dissolution profile is compliant, then the clinical effects should remain unaltered, and if the *in-vitro* violate the specifications, then this ought to translate in clinically significant changes.

IVIVC modeling combines an immediate-release plasma concentration-time profile with an *in-vivo* release profile to predict the controlled-release plasma concentration-time profile. Owing to the design of the study, we decided to fit the immediate-release plasma concentrations and the *in-vitro* release profile in a first stage. The obtained estimates were then used to fit the controlled-release plasma concentration-time profile. The rationale behind this is that immediate- and controlled-release plasma concentrations were measured in different studies with different patients. Therefore, no exchange of information within the same patient is possible and the advantage of a simultaneous fit disappears. This is in contrast to O'Hara *et al.* (2001), who also use a two-stage method but only the immediate-release plasma concentrations are fitted at the first stage, and Jacobs *et al.* (2008), who use a one-stage approach.

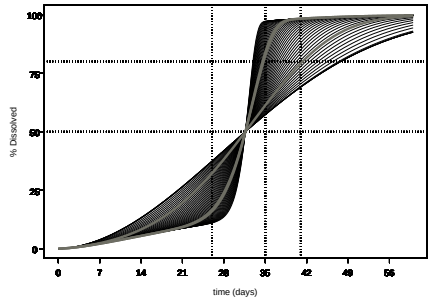
Two simulation studies were set up to study the criteria at the time of 50% and 80% release, t_{50} and t_{80} respectively. t_{50} is a marker for $t_{\max}$, whereas $t_{80} - t_{50}$ is a marker for the release rate and therefore an indication of $C_{\max}$. The $t_{\max}$ simulation demonstrates that for a novel batch with an early release, e.g., at day 15, the



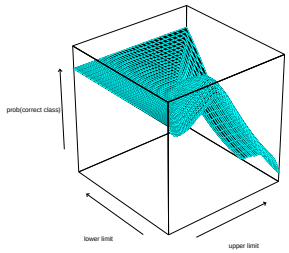
(a) in-vitro time point simulation



(b) prob. correct classification, time point simulation



(c) in-vitro time interval simulation



(d) prob. correct classification, time interval simulation

Figure 7.1: Panels (a) and (c) represent the *in-vitro* dissolution curves utilized in the $t_{\max}$ and $C_{\max}$ simulation, respectively. The dotted lines represent the prespecified dissolution specifications. The left optimal specification found in the $t_{\max}$ coincides with the clinical batch, whereas the right optimal specification corresponds to the specification that 80% should be released before day 41. Panels on the right represent the probability of a correct classification as a function of the lower and upper limit of the *in-vitro* dissolution specification for the (b) $t_{\max}$ and (d) $C_{\max}$ simulation, respectively.

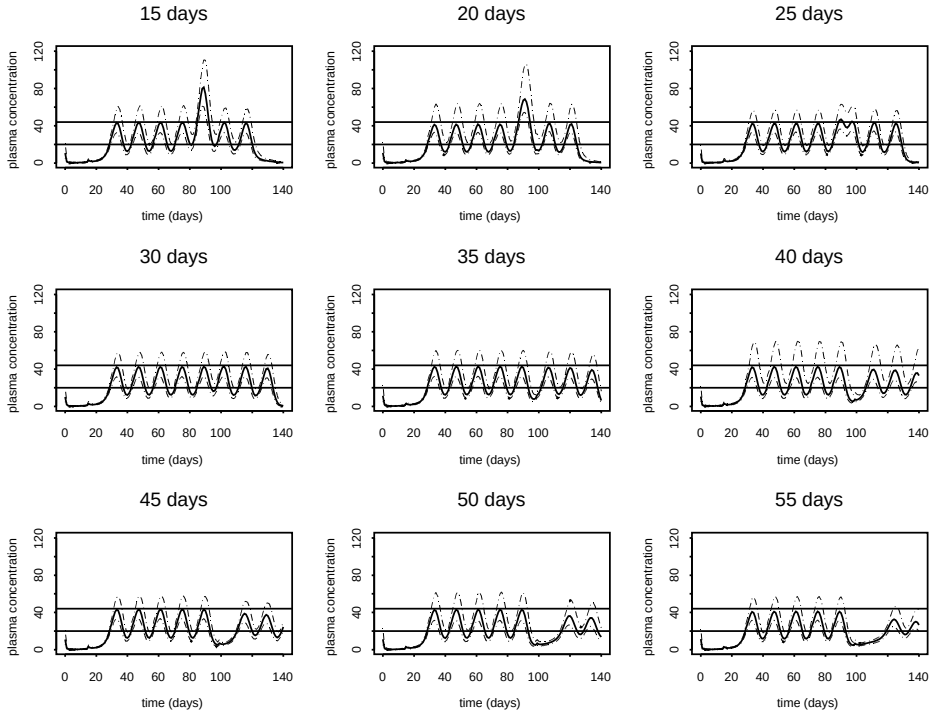


Figure 7.2: Median, 5^{th} and 95^{th} quantile of the simulated controlled-release plasma concentration-time profiles from the $t_{\max}$ simulation for shifted times of 50% release and constant release slopes. The horizontal lines represent the plasma concentrations corresponding to the efficacy and safety threshold as obtained from the PK/PD model.

plasma concentration peak tends to coincide with the plasma concentration peak of the previous administration from the reference batch. The simulation indicates the appropriateness of the release specification for t_{50} to avoid an increase of the safety hazard. The release rate is investigated in the $C_{\max}$ simulation. It was prespecified that 80% should be released before day 41. Based on the simulation, the current dissolution specification might require some fine-tuning, because it allows for a drug release within a short time interval. In such a case, the incidence of exceeding the safety threshold might become an issue. Therefore, it would be appropriate to add a restriction to the release rate in the form of a restriction on the slope of the release function. The $C_{\max}$ simulation indicated that restricting the release rate F' at t_{50} to

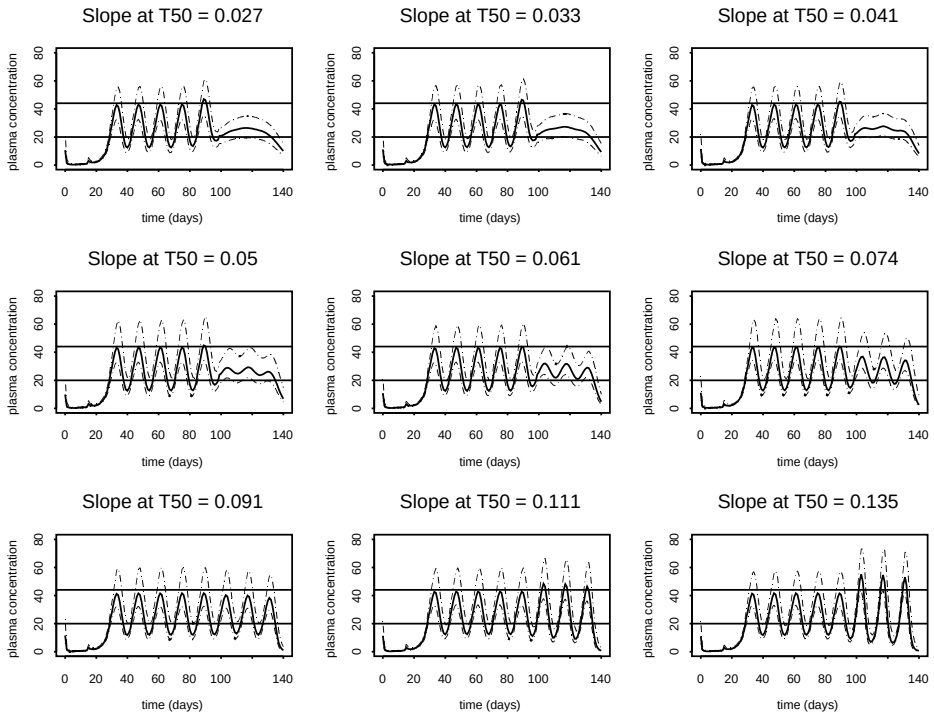


Figure 7.3: Median, 5^{th} and 95^{th} quantile of the simulated controlled-release plasma concentration-time profiles from the $C_{\max}$ simulation for altered release rates while maintaining the time of 50% release fixed. The horizontal lines represent the plasma concentrations corresponding to the efficacy and safety threshold as obtained from the PK/PD model.

the interval -3.4 to -2.3 is optimal. This means that indeed 80% of the drug product should be released before day 41, but not faster than the current phase 3 batch.

Although other drug products do not target the same receptor, the knowledge of the method of action exists for many compounds. This can be used to set up a similar relation between the *in-vitro* tests and the clinical outcome. As such, it is possible to obtain more relevant release specifications that guarantee the clinical response of the novel batches. The idea behind such a technique is that the release specifications one obtains from the simulations ensure that the new batch will retain the same efficacy and safety profile as the reference batch, hence it is in the patients' interest. On the other hand, it might be that in some cases the release specifications are prespecified

too conservative, hence good batches are unnecessarily discarded, so the company gains using the approach as exemplified here. In the latter case, less base product is required, so the method also has an environmental benefit. Further, the method does not restrict to a research setting, but also allows for specification modifications based on novel post-marketing information. A prerequisite for applying the concept, is that during the different stages of drug development the appropriate studies are done to collect the data needed to built the different models.

8

Incorporating Therapeutic Window in Bioequivalence Acceptance Limits

The main topic of the previous chapters was on IVIVC modeling. As explained in Section 1.2, one of the main applications after the set up of an IVIVC model is related to SUPAC, i.e., ensuring that a new batch of the drug product will lead to the same *in-vivo* drug exposure. However, pharmaceutical companies would also like to test for equality of drug exposure before IVIVC models are established. This is then performed using an *in-vivo* bioequivalence study. This is typically a cross-over study where healthy volunteers, patients or animals receive both the reference and test drug product in a random order. The bioequivalence testing procedure and methodology quantifies the maximal deviation of the exposure from the test versus the reference drug product. This maximal deviation is uniform for all drug products. Our criticism is that this equivalence testing ought to be considered in perspective; can one request the same criterion for a drug product for which minor changes of the exposure lead to major clinical changes compared to a drug formulation for which altering the exposure with factor two would not have a clinical impact? Therefore, the current methodology is extended to differentiate according to the therapeutic window

(Jacobs *et al.* 2008). Part of the history and the philosophy of bioequivalence testing is explained in Section 8.1. The methodology is extended in Section 8.2, followed by the performance of the method as evaluated through simulations in Section 8.3. Finally the method is applied to three known examples of narrow index drugs (theophylline, digoxin, and phenytoin) in Section 8.4.

8.1 Philosophy and Rationale of Bioequivalence Testing.

Bioequivalence studies are important in drug development to prove that two drug products give similar *in-vivo* exposure, and therefore that the safety and efficacy profile is not altered and therapeutic equivalence can be claimed. Bioavailability and bioequivalence studies are performed to evaluate differences in drug products, for example research versus market tablets, various batches, or production sites. At the same time, those techniques are also used for evaluating food effects, drug-drug interactions, and comparing administration routes.

Schuirmann (1987) laid the foundations of modern bioequivalence testing. He proposed to perform two one-sided tests, to test the hypothesis that the ratio of the key pharmacokinetic parameters AUC and $C_{\max}$ is contained within a prespecified range, which usually is 80–125%. At the end of the twentieth century, average bioequivalence as proposed by Schuirmann was questioned because it only focusses on whether the average exposure of the study population is equivalent (Anderson and Hauck, 1990, Scheiner 1992). In the typical situation where drugs are on the market, each patient should maintain the same exposure independent of his choice. This led to the concept of individual bioequivalence (Anderson and Hauck, 1990), also known as switchability. Owing to the complexity of the technique and its favoring of highly variable drug products (Hsuan 2000), individual bioequivalence has not been used extensively to date.

There are two situations in which the traditional approach with a fixed acceptance range is not optimal: first the one of highly variable drug products, and secondly narrow index drugs, i.e., drugs where comparatively small differences in dose or concentration lead to dose-and concentration-dependent, serious therapeutic failures and/or serious adverse drug reactions.

An area of discussion is the bioequivalence assessment of highly variable drug products, i.e., products with a within-subject variability of more than 30%. Authorities acknowledge that the large sample sizes for trials with such drug products

cannot always be ethically justified (FDA 2003, CPMP EMEA 2006). The simplest correction for highly variable drug products, is by extending the acceptance limits from 80–125% to 75–133% (CPMP EMEA 2001). Boddy (1995) proposed to modify the limits for highly variable drug products according to a predefined estimate of the within-subject variability of the reference drug product. The disadvantage of the 30% threshold is a discontinuity at that threshold: For example, it is possible that in a given study, a within-subject variability of 29% is observed and no modification of the limits is applied, while if the variability was slightly more than 30%, adaptation of the acceptance limits could have yielded to a different conclusion.

Karalis et al. (2004, 2005) modified the idea of extending the bioequivalence limits. Whereas Boddy et al. (1995) categorize drug substances according to a within-subject variability of less versus more than 30%, Karalis expands the bioequivalence limits in a continuous fashion as a function of the within-subject variability. However, expanding the acceptance limits increases the risk of false positives, i.e., falsely concluding two drug products to be bioequivalent. Therefore, Karalis proposed to incorporate the observed geometric mean ratio of the pharmacokinetic parameters AUC and $C_{\max}$ in the acceptance limits: the further the geometric mean ratio deviates from equality, the more conservative the acceptance range becomes.

As suggested by the FDA guidance, the therapeutic window should be taken into account instead of performing an automatic extension of the acceptance ranges:

> "Where the test product generates plasma concentrations that are substantially above those of the reference product, the regulatory concern is not therapeutic failure, but the adequacy of the safety database from the test product. Where the test product has plasma concentrations that are substantially below those of the reference product, the regulatory concern becomes therapeutic efficacy. When the variability of the test product rises, the regulatory concern relates to both safety and efficacy, because it may suggest that the test product does not perform as well as the reference product, and the test product may be too variable to be clinically useful." (FDA 2003)

The aim of the research in this paper is to present further approaches in bioequivalence acceptance taking into account the therapeutic window as suggested by the guidelines (FDA 2003). The proposed bioequivalence limits in this paper consider the position of the therapeutic dose with respect to the lowest effective dose (LED) and the maximum tolerated dose (MTD). A dose close to the LED and/or the MTD may require more stringent limits ensuring exposure remains within the therapeutic

window.

8.2 Methodology

Let us first introduce some notation: U and L are the upper and lower acceptance limit, α the traditional limit (125%), β the extended limit (143%), Ψ the geometric mean ratio, σ_w the within-subject standard deviation, γ, δ, and θ are rate constants. D is the therapeutic dose, which usually corresponds more or less to the administered dose, however, the phenytoin example is an example where the administered dose is lower than the therapeutic dose.

A first approach to adapt the bioequivalence limits for studies with highly variable drug substances was introduced by Boddy (1995). His method maintains the original method and acceptance ranges proposed by Schuirmann (1987) for drug substances with a low variability, i.e., $\%CV < 30\%$. For drug substances with a higher variability, the acceptance ranges are rescaled using the within-subject variability, with the 90% confidence interval of the difference on the logarithmic scale satisfying the criterion:

$$\mid \mu_T - \mu_R \mid \leq \vartheta \sigma_w, \tag{8.1}$$

where the left side of the expression is the treatment difference on the logarithmic scale, σ_w the within-subject standard deviation, and usually $\vartheta = 1$. However, this approach has one important weakness: there is a discontinuation in the acceptance ranges at a within-subject variability of 30%. To illustrate this, imagine a study with an observed $\%CV$ of 29%. To conclude bioequivalence the 90% confidence interval on the logaritmic scale should fall within the interval (-0.223; 0.223). However, if samples had been slightly less accurately analysed, leading to a $\%CV$ of 31%, then the 90% confidence interval should fall within the interval (-0.294; 0.294), by using relation (8.6) between σ_w and $\%CV$. In this case, the study with the the higher accuracy will fail to show bioequivalence, whereas bioequivalence is concluded for the less precise study but with the same observed geometric mean ratio.

Karalis *et. al.* (2005) tried to overcome the problems associated with the discontinuity and proposed three types of bioequivalence limits depending on the geometric mean ratio and at the same time rescaling according to the within-subject variability in a continuous manner. In this paper Weibull type limits will be used to further refine the proposed approach of Karalis:

$$U = \alpha + (5 - 4\Psi)(\beta - \alpha)\left\{1 - e^{-(\gamma\sigma_w)^2}\right\}, \tag{8.2}$$

with γ a constant to regulate the expansion of the acceptance limit, the lower acceptance limit L is $1/U$. For low within-subject variability and $\Psi = 1$, the upper limit remains approximatively α, whereas for large variability and $\Psi = 1$, the upper limit approximates β. With $\Psi = \alpha$, the upper limit is fixed to α, regardless of the variability.

A more general formulation is

$$U = \alpha + 5\left(1 - \frac{1}{\alpha}\Psi\right)(\beta - \alpha)\left\{1 - e^{-(\gamma\sigma_w)^2}\right\} 1_{\Psi \leq \alpha}, \tag{8.3}$$

with $L = 1/U$ as before. Using $\alpha = 125\%$ in the above equation simplifies to (8.2). $1_{\Psi \leq \alpha}$ is added to indicate explicitly that Ψ should fall within the acceptance range. As the focus is not on the choice of γ, it will be fixed in the rest of the paper to a value of 3. This restricts by no means the results of the paper and is mainly chosen based on the simulations from Karalis (2005) to ensure that the acceptance ranges remain close to the standard 80–125% for small variabilities.

In this paper, the expansion of the acceptance range will not only depend on the within-subject variability, but will also depend on the therapeutic window. Therefore, a second correction factor, which represents a similar sigmoidal function of the therapeutic window is added.

$$\begin{aligned} U &= \alpha + 5\left(1 - \frac{1}{\alpha}\Psi\right)(\beta - \alpha)\left\{1 - e^{-(\gamma\sigma_w)^2}\right\}\left\{1 - e^{-\left(\delta\frac{MTD}{D}\right)^2}\right\} 1_{\Psi \leq \alpha}, \\ L &= \frac{1}{\alpha + 5\left(1 - \frac{1}{\alpha}\Psi\right)(\beta - \alpha)\left\{1 - e^{-(\gamma\sigma_w)^2}\right\}\left\{1 - e^{-\left(\delta\frac{D}{LED}\right)^2}\right\} 1_{\Psi \geq \alpha}}, \end{aligned} \tag{8.4}$$

where D is the therapeutic dose, and δ is a constant for the rate of change. The therapeutic window is defined as the ratio D/LED and the ratio MTD/D. Note the asymmetric character of the acceptance limits: the lower limit depends on the distance between the dose and the LED, whereas the upper limit depends on the distance between the dose and the MTD.

A more conservative approach can be applied for narrow-index drugs. The concern has been introduced in the Canadian guideline (Ministry of Health Canada, 2006) that for certain drugs the 80–125% acceptance range would be too liberal. Therefore, the standard 125% limit, which is used as a starting point in the current approach, can be modified in a similar way. This renders the resulting acceptance ranges even more narrow in case of narrow-index drugs. As a result, the following type of bioequivalence

acceptance range is introduced:

$$\begin{aligned}
U &= \alpha\text{"} + 5\left(1 - \frac{1}{\alpha\text{"}}\Psi\right)(\beta - \alpha\text{"})\left\{1 - e^{-(\gamma\sigma_w)^2}\right\}\left\{1 - e^{-\left(\delta\frac{MTD}{D}\right)^2}\right\}1_{\Psi \le \alpha\text{"}}, \\
L &= \frac{1}{\alpha' + 5\left(1 - \frac{1}{\alpha'}\Psi\right)(\beta - \alpha')\left\{1 - e^{-(\gamma\sigma_w)^2}\right\}\left\{1 - e^{-\left(\delta\frac{D}{LED}\right)^2}\right\}1_{\Psi \ge \alpha'}}, \\
\alpha' &= 1 + (\alpha - 1)\left\{1 - e^{-\left(\theta\left(1 + \frac{D}{LED}\right)\right)^2}\right\}, \\
\alpha\text{"} &= 1 + (\alpha - 1)\left\{1 - e^{-\left(\theta\left(1 + \frac{MTD}{D}\right)\right)^2}\right\}, \qquad (8.5)
\end{aligned}$$

where, as before, δ, γ, and θ are rate constants.

A new study protocol would consist of fixing the rate constants γ, δ, and θ, based upon interaction with the regulatory authorities, stating the therapeutic window LED and MTD of the drug product, and the anticipated therapeutic dose, i.e., the dose most frequently used by patients. The sample size calculation can then be performed using simulation-based methodology.

8.3 Simulation Study

The proposed bioequivalence acceptance ranges (8.5) depend on the therapeutic window as well as on the within-subject variability. These parameters, as well as the influence of the choice of the parameters δ, θ, and γ, are explored through simulations.

In the first simulation run, the acceptance ranges are calculated using (8.5) to explore their behavior for different values of θ, δ, the ratios MTD/D and D/LED, and the within-subject variability with $\Psi = 1$. The within-subject variability is presented as a coefficient of variation ($\%CV$) in line with pharmacokinetics practice. It is linked to the within-subject variability, as follows:

$$\sigma_w = \sqrt{\ln(1 + \%CV^2)}. \qquad (8.6)$$

Figure 8.1 shows the new acceptance range for different choices of δ and θ as a function of the ratio MTD/D for the upper limit and D/LED for the lower limit. It shows that, for each choice of δ and θ, the upper or the lower acceptance limit is reduced when the tested dose approaches the boundary of the therapeutic window, i.e., when MTD/D or D/LED approach unity. For doses far from the boundary of the therapeutic window, the ratios MTD/D and D/LED are larger and the acceptance ranges broaden. This is a conservative approach to ensure patients maintain a safe and efficacious exposure. With θ increasing from 0.1 to 1, the slopes of the acceptance

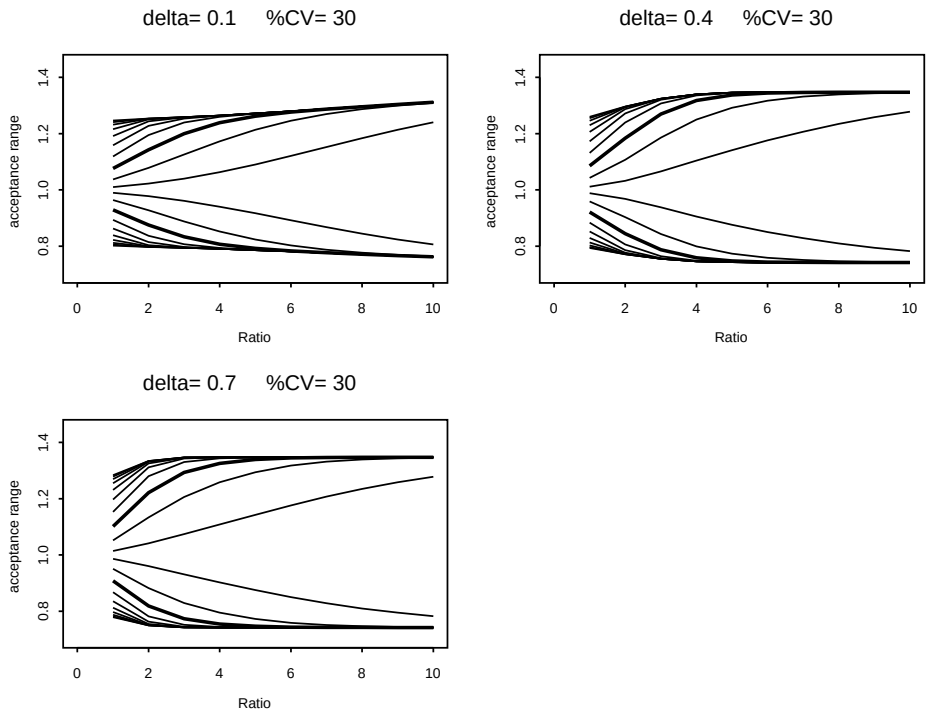


Figure 8.1: Illustration of the influence of therapeutic window by varying θ from 0.1 (middle) to 1 (outside) on the newly proposed bioequivalence acceptance range for different δ. The tick line represents the case $\theta = 0.3$. For the upper limit, the ratio in the x-axis represents MTD/D whereas D/LED for the lower limit. A %CV of 30% was assumed.

ranges become steeper near the therapeutic borders. A value of 0.3 for θ seems reasonable: the resulting shallow slope protects patients by imposing strict acceptance limits close to the borders of the therapeutic window. For a higher value of θ, the influence of the ratios MTD/D and D/LED vanishes and they may not be sufficiently conservative.

Whereas the parameter θ regulates the shrinkage of the acceptance range with respect to the therapeutic window, the parameter δ determines the expansion of the limits as a function of the therapeutic window and the within-subject variability. It basically means that for a highly variable drug product with a dose near the boundaries of the therapeutic window, the expansion of the acceptance limits is smaller

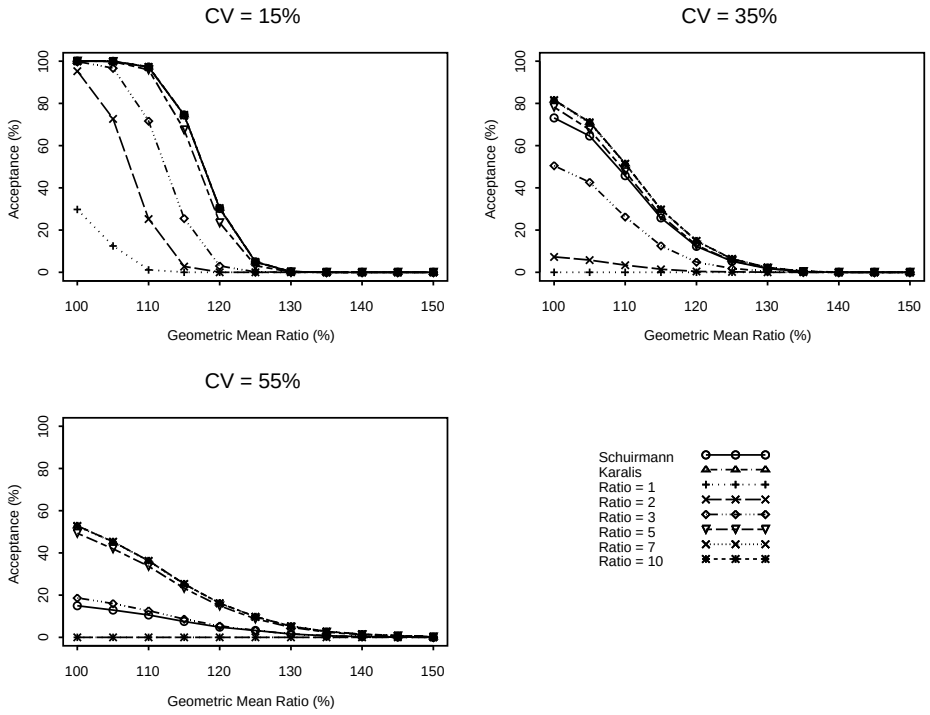


Figure 8.2: Influence of the within-subject variability on the acceptance (%) of bioequivalence trials using Schuirmann's method, Karalis and our new proposal with $MTD/D = D/LED$ from 1 to 10. The sample size is fixed to 36 subjects.

than the ones proposed by Karalis et al. (2005). For a dose far from the therapeutic boundary, the acceptance ranges behave similar to the ones in the above article.

A small value for δ, e.g., 0.1, penalizes the acceptance ranges in a very conservative way, whereas values ranging over 0.7–1 are too liberal and impose little restriction (Figure 8.1). Therefore, an intermediate value of 0.4 for δ seems reasonable.

Although not demonstrated in the figures, the approach of Karalis et al. (2005) is maintained and extended: the acceptance ranges depend on the within-subject variability of the drug products and gradually expand from 80–125% to 70–143%, as a function of the within-subject variability.

Whereas the previous calculations mainly illustrated the general concepts of the new approach to acceptance limits, the ensuing set of simulations was performed to compare it to existing methods (Figure 8.2). One thousand two-treatment, two-

period cross-over studies with 36 subjects were simulated per condition, defined by the within-subject variability ($\%CV$ of 15%, 35%, and 55%), and the true geometric mean ratio Ψ (100% to 150%). θ was fixed at 0.3, and δ set to 0.4. The simulation was simplified in a first step by setting MTD/D equal to D/LED. These ratios varied from 1 to 10 in the simulation, to cover a broad spectrum of therapeutic windows. This resulted in 33,000 trials simulated for a sample size of 36 patients. Simulations were performed using SAS 9.1 and analyzed with procedure MIXED.

Our acceptance limits and those obtained by the method of Karalis et al. coincided for 35%CV for a MTD/D ratio from 7 onwards. For a narrow-index drug, e.g., a MTD/D of 3 or less, the acceptance rate is strongly decreased due to desired shrinkage of the acceptance limits. For low-variability drugs (15% CV), the methods are essentially equivalent to the Schuirmann method, but for highly variable drug products (55%CV), the methods give clear differences. The proposed method behaves as liberal as the Karalis' method for drugs with a broad therapeutic window, and more conservative than Schuirmann for narrow-index drugs. The simulations also indicate that only a very small amount of the studies with Ψ superior to 125% conclude bioequivalence as would be expected based on (8.5). For any Ψ, the acceptance rate increases with the therapeutic window. Further exploration of the effect of the sample size and changes of the within-subject variability in Figures 8.3 and 8.4 confirmed the previous conclusions.

Table 8.1 contains a summary of the above simulations for the specific case of $\Psi = 125\%$, i.e. the point from where onwards bioequivalence is rejected in classic bioequivalence testing. It represents the proportion of simulated trials where bioequivalence is concluded, whereas in fact the two products are bio-inequivalent. This corresponds to the type-1 error for the Schuirmann method. These values are larger than for Schuirmann, but correspond to the method of Karalis for large MTD/D, but they decrease well below the Schuirmann error rate when the dose approaches the MTD. Therefore, the new acceptance limits are conservative when it is in the patients interest. This illustrates well the strength of the method.

Figures 8.5 – 8.7 represent the same simulations as before, but now for a situation where the dose is closer to the LED than the MTD, where D/LED varies from 1 to 10 and MTD/D is fixed to 10. When the dose is close to the LED, i.e. less than 3, the number of accepted trials when $\Psi = 100\%$ was lower than when $\Psi = 105\%$. Here, the conservative nature of the acceptance limits clearly distinguishes our method from Schuirmann's and Karalis' methods: The asymmetry of the limits render many trials inconclusive for a true Ψ of 100%, whereas this is not the case for $\Psi = 105\%$. This ensures patients will maintain an efficacious exposure.

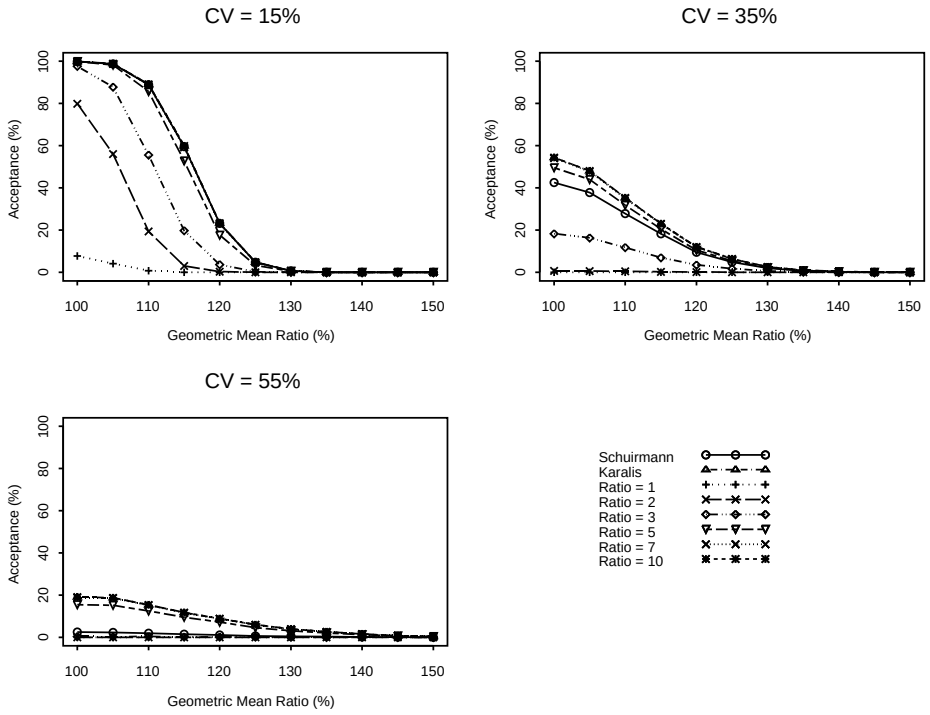


Figure 8.3: Influence of the within-subject variability on the acceptance (%) of bioequivalence trials using Schuirmann's method, Karalis and our new proposal with $MTD/D = D/LED$ from 1 to 10. The sample size is fixed to 24 subjects.

Figure 8.8 shows the most extreme observed value for Ψ as a function of $\%CV$ that leads to the acceptance of bioequivalence. The figure represents different case of $D/LED = MTD/D$. This most extreme value Ψ_e is defined as

$$\Psi_e = U - \sigma_w \sqrt{2/N} t_{0.95,N-2}$$

for the upper range. Ψ_e declines towards 100% for increasing $\%CV$, i.e., if a drug product is more variable, then Ψ has to reduce further to allow for the trial to be bioequivalent. Although it is not easy to see based on the formula, the figure demonstrates that, for a given position of the therapeutic dose within the therapeutic window, the acceptance limits behave, up to a constant, similar to the limits introduced by Karalis et. al. (2005).

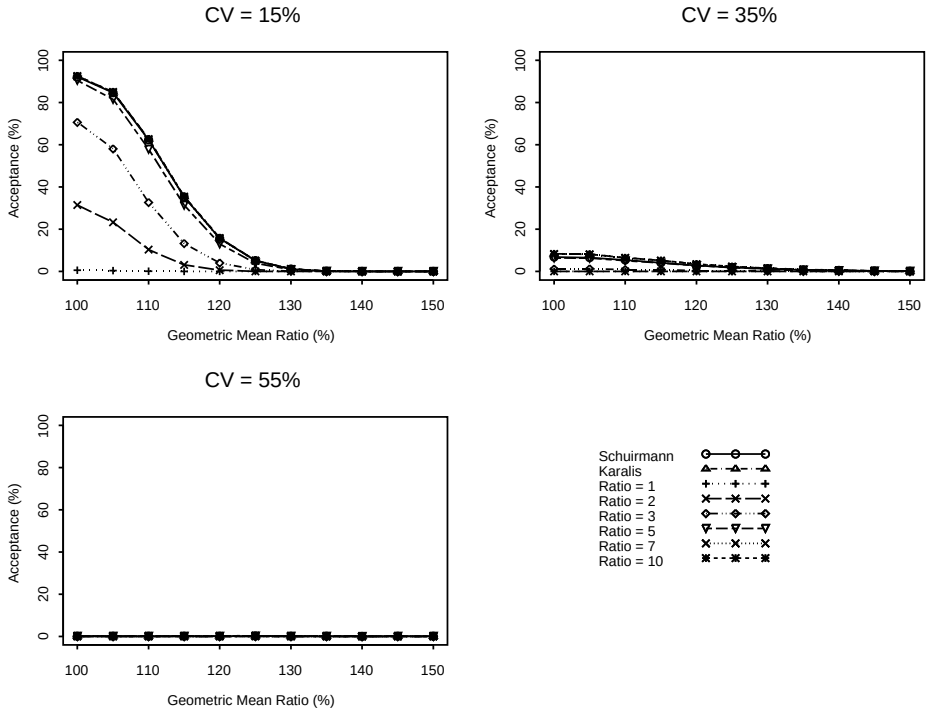


Figure 8.4: Influence of the within-subject variability on the acceptance (%) of bioequivalence trials using Schuirmann's method, Karalis and our new proposal with $MTD/D = D/LED$ from 1 to 10. The sample size is fixed to 12 subjects..

Table 8.1: The proportion of simulated trials for which bioequivalence was concluded erroneously at $\Psi = 125\%$, as a function of %CV and sample size, in the case of $MTD/D = D/LED = \mathcal{R}$, $\theta = 0.3$, $\delta = 0.4$, and $\gamma = 3$.

	Sample size			Sample size			Sample size		
	12			24			36		
%CV	15	35	55	15	35	55	15	35	55
Schuirman	0.0512	0.0190	0.0023	0.0475	0.0485	0.0066	0.0493	0.0529	0.0324
Karalis	0.0515	0.0227	0.0028	0.0478	0.0636	0.0601	0.0497	0.0641	0.0976
$\mathcal{R} = 10$	0.0515	0.0227	0.0028	0.0478	0.0636	0.0601	0.0497	0.0641	0.0976
$\mathcal{R} = 7$	0.0503	0.0225	0.0028	0.0468	0.0624	0.0591	0.0474	0.0637	0.0970
$\mathcal{R} = 5$	0.0394	0.0177	0.0021	0.0327	0.0539	0.0458	0.0308	0.0545	0.0876
$\mathcal{R} = 3$	0.0083	0.0037	0.0001	0.0027	0.0171	0.0020	0.0013	0.0184	0.0324
$\mathcal{R} = 2$	0.0007	0.0002	0	0	0.0005	0	0	0.0022	0
$\mathcal{R} = 1$	0	0	0	0	0	0	0	0	0

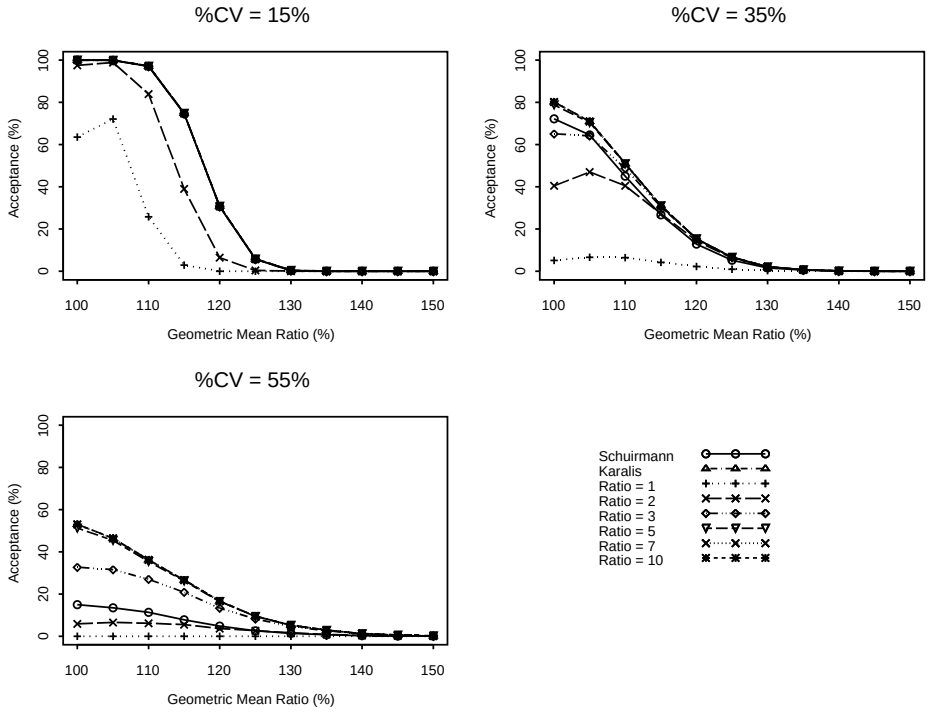


Figure 8.5: Influence of the within-subject variability on the acceptance (%) of bioequivalence trials using Schuirmann's method, Karalis and our new proposal with only D/LED from 1 to 10 and MTD considered large. The sample size is fixed to 36 subjects.

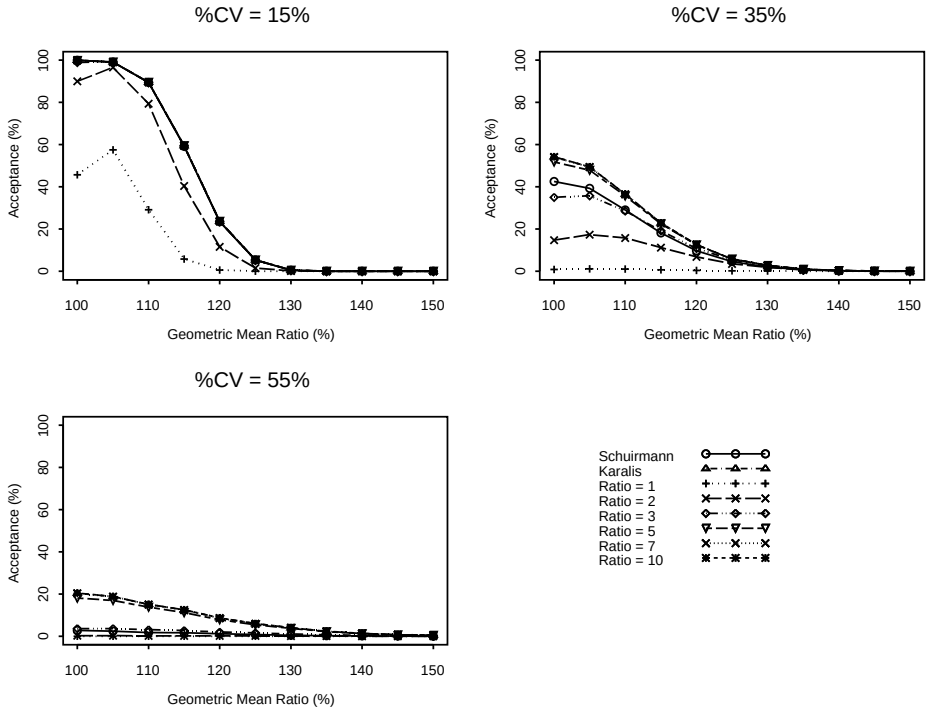


Figure 8.6: Influence of the within-subject variability on the acceptance (%) of bioequivalence trials using Schuirmann's method, Karalis and our new proposal with only D/LED from 1 to 10 and MTD considered large. The sample size is fixed to 24 subjects.

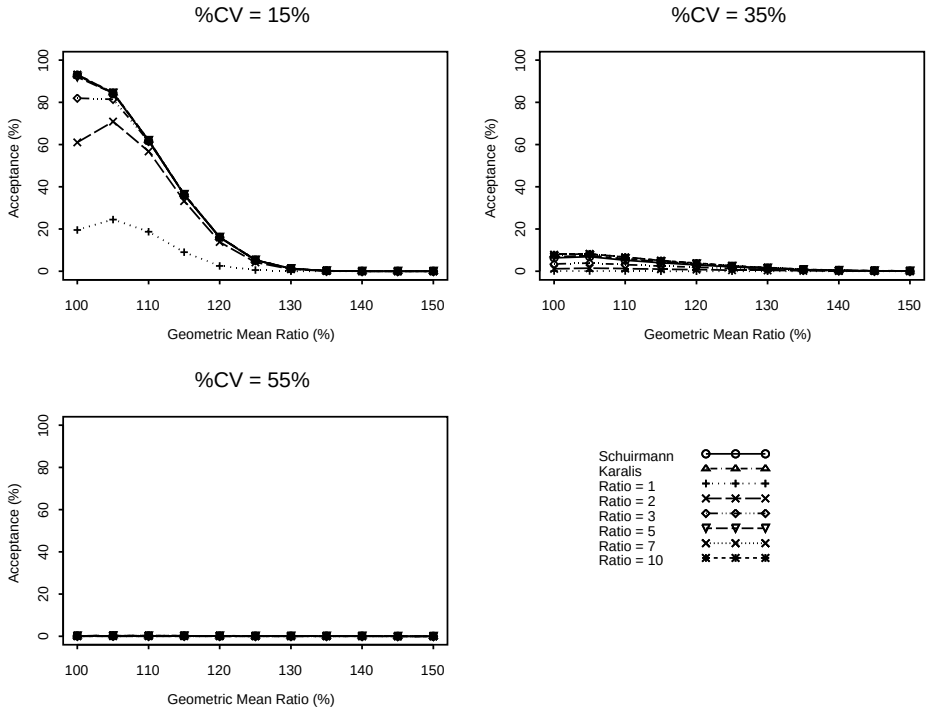


Figure 8.7: Influence of the within-subject variability on the acceptance (%) of bioequivalence trials using Schuirmann's method, Karalis and our new proposal with only D/LED from 1 to 10 and MTD considered large. The sample size is fixed to 12 subjects.

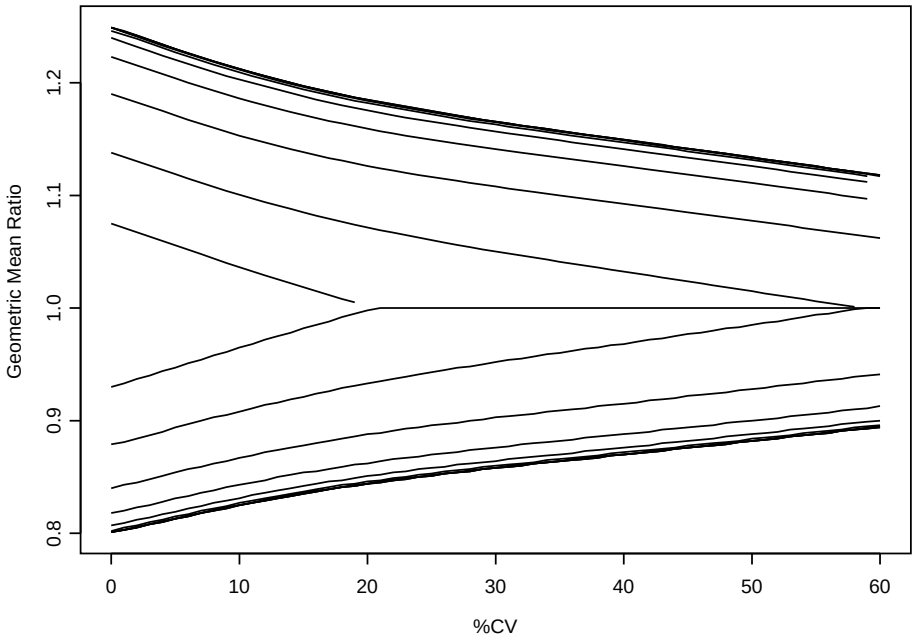


Figure 8.8: The most extreme geometric mean ratio leading to conclusion of bioequivalence as a function of the coefficient of variation. The lines indicate different positions of the therapeutic dose within the therapeutic window. Most narrow therapeutic windows lead to the most stringent limits (closest to 1). The sample size is fixed to 36 subjects.

8.4 Application

The Canadian health authorities recently published a guideline for critical dose drugs (Ministry of Health Canada, 2006). In the appendix of the guideline, a list can be found with a number of drug substances for which a small difference in dose or concentration lead to dose- and concentration-dependent, therapeutic failure and/or serious adverse drug reactions: cyclosporine, digoxin, flecainide, lithium, phenytoin, sirolimus, tracolimus, theophylline, and warfarin. For these drug substances, the more stringent 90–112% acceptance limits for AUC in case of single dose administration are proposed. The three examples below illustrate well the conservative behavior of the new acceptance limits for narrow index drugs. Within the three examples, the parameters $\gamma = 3$, $\theta = 0.3$, and $\delta = 0.4$ were kept fixed.

8.4.1 Theophylline

Theophylline belongs to the list of critical-dose drugs (Ministry of Health Canada, 2006). The data of Mistry et al (1999) is reanalyzed with the different techniques. The study was set up to demonstrate the absence of a drug interaction of indinavir on a single dose of 250 mg theophylline immediate release. Note that the study was not fully powered, i.e., no formal sample size assessment for any of the methods was performed.

The MTD and LED of theophylline can be derived from the literature. Theophylline therapeutic plasma concentrations range from 10 to 20 μg/mL, seizures and cardiac problems can occur at the upper limit (Ministry of Health Canada, 2006). Estimates of the first-order compartmental model (k_e, k_a rate constants, CL clearance) were obtained from Pinheiro and Bates (2000): $\log(k_e) = -2.4327$, $\log(k_a) = -0.45146$, and $\log(CL) = -3.2145$, where dose was denoted in mg/kg. The accumulation factor for multiple dosing is $1/\left[1 - \exp\left(-k_e\tau\right)\right]$, τ corresponding to 8 hours. Solving the equations for a $C_{\max}$ set equal to the above range limits yields an LED of 220 mg and a MTD of 450 mg for a subject of 70 kg. Estimates for the variability are derived from Steinijans et al (1995): $\%CV$ for AUC is 12%, 20% for $C_{\max}$.

The conclusion of Mistry et. al. based on the traditional analysis was an absence of a drug interaction effect: 1.18 (1.13, 1.23) for AUC, and 0.99 (0.92, 1.07) for $C_{\max}$ fall both within the 80–125% Schuirmann acceptance ranges. As the Karalis acceptance limits are always broader or equal to Schuirmann, this method also brings us to concluding that a drug interaction is absent. Our new acceptance limits, taking

the variability as well as the therapeutic window into account, were (0.94, 1.12) for AUC, and (0.88, 1.15) for $C_{\max}$. Therefore, the confidence interval for AUC falls entirely outside the acceptance range and a drug interaction would be concluded for AUC.

8.4.2 Digoxin

Digoxin is another critical dose drug (Ministry of Health Canada, 2006). Martin et al (1997) evaluated the drug interaction of eprosartan on 0.6 mg digoxin. We reevaluate these study results with the new method.

Serum digoxin levels ranging from 0.8 to 2.0 ng/mL are generally considered as therapeutic. Levels greater than 2.0 ng/mL are often associated with toxicity (Ministry of Health Canada, 2006). The IV compartmental model as well as parameter estimates of digoxin for healthy volunteers are found in Wagner (1975). The bioavailability for tablets is 80% (Bochner et al 1977). Estimates for the variability of digoxin are derived from Steinijans et al (1995): $\%CV$ for AUC is 8%, 19% for $C_{\max}$. Using these estimates, the MTD and LED of digoxin can be derived: 0.4 mg as LED, and 1 mg as MTD. These estimated quantities for LED and MTD appear to be rather high from a clinical perspective. The example serves to illustrate the proposed technique rather than making statements about specific drugs. However, this illustrates the need for a thorough understanding and investigation of the therapeutic window of any drug.

The geometric mean ratio in the original analysis was 1.01 (0.81, 1.26) for AUC, and 1.00 (0.86, 1.17) for $C_{\max}$. Our new acceptance limits were (0.90, 1.12) for AUC, and (0.90, 1.13) for $C_{\max}$. Therefore the trial was inconclusive for both parameters.

8.4.3 Phenytoin

In Meyer (2001), three different lots of 100 mg phenytoin sodium capsules were compared. In this study, the observed %CV was low, i.e., 14% and 11% for $C_{\max}$ and AUC respectively. The conclusion based on the traditional analysis was that all 3 lots were bioequivalent.

To apply our method, the MTD and LED of phenytoin were deduced from the literature. Phenytoin exhibits Michaelis-Menten kinetics, which is described by the following equation (Gibaldi and Perrier 1982) for the steady state plasma concentrations c_{ss}:

$$c_{ss} = \frac{FD}{\tau CL_s}, \tag{8.7}$$

where

$$CL_s = \frac{V_m V}{K_m + c_{ss}}, \tag{8.8}$$

and τ represents the dosing interval. CL_s is the clearance parameter, V_m is the theoretical maximum rate of the process, K_m the Michaelis constant, V is the volume of distribution, and F is the bioavailability. Estimates of the Michaelis-Menten constants V_m and K_m for phenytoin are reported as 17.87 mg/h and 4.29 mg/L, respectively (Santos Buelga 2002). In the same study, an average steady state concentration c_{ss} of 12.5 mg/L was observed after multiple dosing of 155 mg. Given the fact that phenytoin is traditionally prescribed as twice daily (b.i.d.), i.e., τ is set to 12 hours, solving (8.7) and (8.8) for the unknown apparent volume of distribution, leads to an estimate of

$$V/F = \frac{D}{\tau}\frac{K_m + c_{ss}}{V_m c_{ss}} = 0.97. \tag{8.9}$$

Phenytoin is associated with severe neurological toxicity from 160 μmol/L onwards, whereas therapeutic plasma concentrations range from 40 to 80 μmol/L (Ministry of Health Canada, 2006). Therefore, a dose associated with 160 μmol/L steady state plasma concentrations will be considered the MTD. Based on the above estimations and equations, one can now calculate the MTD associated with $c_{ss} = 160$ μmol/L, or $c_{ss} = 43.9$ mg/L, given the molecular weight of phenytoin sodium (274.3 g/mol). Solving again (8.7) and (8.8) for D gives an MTD of 190 mg b.i.d., or a total daily dose of 380 mg. Analogue, the lower limit of the therapeutic window is associated with $c_{ss} = 40$ μmol/L, or $c_{ss} = 10.975$ mg/L. This leads to an LED of 150 mg b.i.d. Drug-monitoring is required for phenytoin to ensure patients remain on an optimal exposure. Therefore, the dose corresponding to 60 μmol/L in an average patient will be considered as the therapeutic dose, i.e., 165 mg b.i.d.

Table 8.2 contains the geometric mean ratio, its 90% confidence interval, the equivalence limits using Karalis' equation, and our newly proposed acceptance ranges. The conclusions do not change for the Karalis method and it is inconclusive for all but three cases with novel method because the lower limit of the confidence interval falls below the acceptance limit. This example illustrates that the technique of Karalis only expands the acceptance limits, whereas in our approach, the acceptance limits reduce if the dose is close or outside the edge of the therapeutic window.

8.5 Discussion and Conclusions

Bioequivalence testing is an important topic in drug development. In this kind of trials, the pharmacokinetic parameters AUC and $C_{\max}$ serve as surrogate markers for

Table 8.2: Reconsidering the bioequivalence testing of Phenytoin using the data from Meyer (2001) .

	test vs reference	Ψ	90% confidence interval	Karalis limit	New limit
C_{max}	2 vs 1	0.986	(0.90; 1.04)	(0.781; 1.280)	(0.921; 1.090)
	3 vs 1	0.993	(0.92; 1.05)	(0.781; 1.280)	(0.921; 1.090)
	4 vs 1	0.979	(0.89; 1.02)	(0.780; 1.281)	(0.920; 1.090)
	3 vs 2	0.995	(0.92; 1.06)	(0.782; 1.279)	(0.921; 1.090)
	4 vs 2	0.993	(0.92; 1.05)	(0.781; 1.280)	(0.921; 1.090)
	4 vs 3	0.988	(0.91; 1.04)	(0.781; 1.280)	(0.921; 1.090)
AUC	2 vs 1	0.975	(0.90; 0.99)	(0.787; 1.270)	(0.922; 1.089)
	3 vs 1	0.997	(0.95; 1.04)	(0.788; 1.269)	(0.922; 1.088)
	4 vs 1	0.984	(0.92; 1.01)	(0.788; 1.270)	(0.922; 1.088)
	3 vs 2	0.980	(0.91; 1.00)	(0.787; 1.270)	(0.922; 1.088)
	4 vs 2	0.991	(0.93; 1.03)	(0.788; 1.269)	(0.922; 1.088)
	4 vs 3	0.989	(0.93; 1.02)	(0.788; 1.269)	(0.922; 1.088)

safety and efficacy in the sense that the equivalence of the pharmacokinetic parameters between test and reference implicitly implies that test and reference products have equivalent efficacy and safety. To claim bioequivalence of the parameters, an acceptance range of 80–125% is predefined, which implicitly leads to the conclusion that the observed differences have no efficacy or safety repercussions.

However, the assumption that changes within the 80–125% range have no clinical implications ought to be verified. For narrow-index drugs, even an exposure change of 10% might affect safety and/or efficacy, whereas doubling the exposure for certain other drug products would not affect the safety at all. It is interesting to see that this idea was already partially formulated in Macheras and Rosen (1983) and more prominently in the conclusion of Sheiner (1992):

> "... The main point is that the logical basis for current bioequivalence measurement and regulation is seriously inadequate: only with an appropriate model for dose effect, and a clear delineation of clinical context and values, can one devise, estimate and test bioequivalence measures that

> make clinical and scientific sense. We should judge future contributions to the bioequivalence literature by how well they meet this requirement."

Since then, to our knowledge, no paper has addressed bioequivalence testing in this respect. This paper is an attempt to incorporate therapeutic relevance. Even though the proposed method is more complex than the classical approach, it arguably will have its use in practice and hopefully will trigger discussion and further research.

One might question the regulatory imposed acceptance ranges, since this approach treats all drug products in the same way. One of the concerns is that highly variable drug products, i.e., a within-subject variability of more than 30%, are treated the same way as the rest. This results in studies with unpractical large sample sizes. Boddy (1995) and Karalis (2004, 2005) proposed, respectively, scaled average bioequivalence and bioequivalence with levelling-off properties. Both of these correct the acceptance ranges with respect to the within-subject variability, but do not answer the clinical relevance of the acceptance limits, and rather limits to the logistics and ethics of the method.

Besides the within-subject variability, the newly proposed acceptance ranges take also the therapeutic window into account. More specifically, the proposed approach is highly conservative for doses near the boundaries of the therapeutic window, defined by the ratios MTD/D and D/LED, and more liberal for doses far from the maximum tolerated dose and lowest effective dose.

A simulation study shows that for doses near the MTD, lower acceptance limits are imposed for the upper limit of the 90% confidence interval: this should ensure that patients will not experience toxic exposures for compounds with a narrow therapeutic window. The same recommended for doses close to the least effictive dose: the lower acceptance limit will approach 100% to ensure patients remain on active doses. On the other hand, for doses far from the boundaries of the therapeutic window, the acceptance limits approach the ones of Karalis et al. (2005).

In the extreme case of a therapeutic dose close to the LED and very far from the MTD, the discrepancy with the standard approach is apparent. Formulations with slightly higher exposures (105%) are favored compared to the ones leading to the same exposure (100%). This ought not to be considered a disadvantage. Quite to the contrary, we perceive it as an advantage because it favors formulations that ensure the patients to remain at efficacious and safe exposures. It takes into account the actual position of the therapeutic dose in the therapeutic window.

Based on the simulations, it has been demonstrated that the newly proposed bioequivalence limits differentiate between narrow index drugs and drug products with a

wide therapeutic window. They are very strict when it is of interest for the patient, and more flexible when the therapeutic effect remains unaffected. Traditional methods, on the contrary, apply a uniform method, regardless as to where the marketed dose is positioned in the therapeutic window.

Since the newly proposed bioequivalence limits depend on the MTD and the LED, these quantities need to be determined as accurately as possible in an early stage of drug development. This emphasizes the need for adequate dose finding trials using stochastic methods such as most prominently, the continuous reassessment method (O'Quigley et al. 1990, Patterson et al. 1999). However, also literature can be a good source for estimates of the MTD and the LED as illustrated in the application.

9

A Latent Pharmacokinetic Time Profile to Model Dose-response Survival Data

In the previous chapters, the focus was on plasma concentration-time profiles. The general consensus among pharmaceutical scientists is that this measure of drug exposure is a good surrogate for the clinical effects, i.e., if the plasma concentrations are unchanged, then the clinical effects are expected to remain unaltered. This relation between plasma concentrations and clinical effects was utilized explicitly in Chapter 7. Further, the equivalence of a drug exposure was related to the dose response of the clinical effects in Chapter 8. Whereas changes of the plasma concentration-time profiles are translated into changes of the clinical effects, this relation will be used in the opposite direction in the current chapter to explain the clinical behavior.

The dose response behavior of the time to fall off a rod is measured and modelled. No plasma concentration-time profile is measured. However, as Holford (2006) states, the causal factor in generating a response is not a dose, but rather the resulting drug concentration at the site of action, such as a receptor. Therefore, a sigmoidal model with a latent pharmacokinetic profile is fitted. This is sometimes referred to as a K-PD model (Jacqmin *et al.* 2007). The model is compared to a more traditional

linear model, but also the added value of plasma sampling in the current design is investigated. This chapter is based on Jacobs (2009).

The chapter is organized as follows. The dose-response models are presented in Section 9.2. The case study is discussed in Section 9.3. The set up of the simulation study is described in Section 9.4.

9.1 The rotarod experiment

The development of new medicines requires an enormous amount of research. After potentially promising molecules have been synthesized, they are first tested *in vitro*, followed by tests in animals and finally humans. Each of these steps should enable to select the compounds with a minimum of unacceptable toxicity and a maximum of effect, but without putting animals and patients unnecessary at risk. Therefore, it is important to learn as much as possible with a minimum of *in vivo* testing.

One of such *in vivo* tests with mice is the accelerating rotarod test. It assesses the impact of the new treatment on the motor coordination (Gerald and Gupta, 1977, McIlwain *et al.* 2001). The treatment effect on the mice is measured by evaluating the duration that the mice stay on top of a circular rod that rotates at accelerating speed from 4 to 40 RPM for the first 5 min and then remains at 40 RPM for another 5 min (McIlwain *et al.* 2001). To acquaint the animals to the rod, the study starts with a training phase, followed by a test phase. The animals are placed four times on the rod at 30-minutes intervals during the training phase to allow repeated experience leading to enhanced performance. The second phase corresponds to a dose-response study of the effects on the motor coordination for two compounds over time. These two compounds (phencyclidine (PCP) and d-amphetamine) are expected to disrupt performance on the rod due to their impact on motor coordination. In the experiment considered here, the evaluative occasions are: before administration, at 30, 60, and 90 minutes after administration. Four doses are evaluated for both compounds: 0.0, 2.5, 5.0, and 10.0 mg/kg subcutaneously. A total of 80 mice are included in the study. All animals enter both phases. Ten mice are attributed to each compound-by-dose level combination.

9.2 Methodology

Gerald and Gupta (1977) analyzed their data using medians and Wilcoxon tests. They do not mention how censored data were handled. In this section, we propose a more formal statistical framework to take into account both the censoring and the

dose response.

The response of interest is of the time-to-event type with event falling off the rod. The survival function, $S(t)$, characterizes the distribution of this response and is the probability of having a time-to-event beyond t (Klein and Moeschberger 1997). The time-to-event data are analyzed using an accelerated failure time (AFT) model where the natural logarithm of the time-to-event is modelled as a function of covariates. The first step in developing the AFT model is the choice of the distribution. For the given experiment, the Weibull distribution was chosen, based on the training dataset.

$$S(t) = \exp\left[-(\alpha t)^{\gamma}\right], \tag{9.1}$$

and where $\alpha > 0$ is the scale parameter, $\gamma > 0$ the shape parameter, and t the time to falling off the rod. Instead of modeling α and γ, the following transformation is applied: $\alpha = \exp(-\alpha')$ and $\gamma = \exp(-\gamma')$ to ensure $\alpha > 0$ and $\gamma > 0$. Further, we define T_i as the evaluation time i during the testing phase; D_j correspond to the dose level j. The right-censoring of the data is included in the likelihood function L as described in Klein and Moeschberger (1997).

To calculate the cumulative distribution function from the density $g(t)$, one needs to integrate over time. The parameters α and γ are a function of time and should be included in the integration. However, the repeated time-to-event evaluations take only a limited amount of time in comparison to the entire experiment, or more precisely to the underlying pharmacokinetic profile. Therefore, one could consider that the parameters α and γ are approximately constant during each time-to-event assessment. The integration and likelihood simplify to (9.2).

$$\begin{aligned} G(t) &= \exp\left[-(\alpha t)^{\gamma}\right], \\ g(t) &= \gamma\alpha\left[(\alpha t)^{\gamma-1}\right] G(t), \\ \log\left[L(t)\right] &= 1_{obs}\log\left[g(t)\right] + 1_{cens}\log\left[G(t)\right] \ , \end{aligned} \tag{9.2}$$

where *cens* denotes a right-censored observation, and *obs* indicates that an event was observed. The log-likelihood function was implemented in the procedure NLMIXED using the SAS software.

The dose-response relation is incorporated using two different models: the first model is the linear regression model, where dose is included as a linear effect and time as a quadratic effect. The interaction effect is included as well. Time is expected to be quadratic, because no effect is expected at the time of administration (time 0), whereas a large effect is expected shortly after administration (time 30), fading out at later time points, hence the quadratic term. The interaction is needed because one

cannot automatically expect the same dose effect at each point in time, nor can the same time effect be expected for all doses. Table 9.1 contains the different parameters, which allows for a better understanding of the model.

A more conventional way of modeling a dose-response relationship in pharmacology is the so-called E_{max} *model* (Gabrielson and Weiner 2000). It assumes that the change in time to falling off the rod depends on the underlying drug concentrations of the compounds. These latent drug concentrations at the site of action are included in the model in a sigmoidal way, i.e., an E_{max} model is fitted for both the scale and shape parameters assuming a latent one-compartmental model with first-order absorption (Jacqmin *et al.* 2007):

$$\begin{aligned} \alpha'_{ijk} &= \beta + \frac{e_{max} \quad c_{ij}}{\exp(ec_{50}) + c_{ij}}, \\ \gamma'_{ijk} &= \beta_\gamma + \frac{e_{max_\gamma} \quad c_{ij}}{\exp(ec_{50_\gamma}) + c_{ij}}, \end{aligned} \tag{9.3}$$

where c_{ij} stands for

$$c = \frac{D_j \exp(\phi_2)}{\exp(V)[\exp(\phi_2) - \exp(\phi_1)]} \left\{\exp[-\exp(\phi_1)T_i] - \exp[-\exp(\phi_2)T_i]\right\}. \tag{9.4}$$

Here, e_{max} is the maximal asymptotic effect, and ec_{50} the concentration at which 50% of the maximal effect is attained. α'_{ijk} and γ'_{ijk} correspond to the scale and shape parameter of the Weibull distribution, respectively, where different values for e_{max} and ec_{50} are used in the scale and the shape parameter. ϕ_1 and ϕ_2 are the rate constants determining the latent drug concentration at the site of action model. D_j and T_i still indicate the dose level j and the time point i. c_{ij} stands for the latent drug concentration. This latent drug concentration is considered to correspond to the concentration at the place of action, e.g., the receptors in the brain or in the bloodstream. The parametrization used ensures that positive values are obtained for $\exp(ec_{50})$, $\exp(\phi_1)$, $\exp(\phi_2)$, and $\exp(lvf)$. It also acknowledges the log-normal distribution of the quantities. The number of parameters to be estimated is reduced compared to previous models. Additionally, the model has a pharmacological basis. The dependence of the data, i.e., each mouse is tested four times in the test phase, can be taken into account by the inclusion of a random effect at ec_{50}, ec_{50_γ}, ϕ_1, and/or ϕ_2. Unfortunately, the small sample size prohibited the inclusion of a random effect in the test phase of our case study.

The median time-to-event can be calculated using the following expression, applicable to the case of a Weibull distribution (Klein and Moeschberger 1997):

$$\frac{\Gamma(1 + 1/\gamma)}{\alpha^{1/\gamma}}. \tag{9.5}$$

9.3 Analysis of the Case Study

The case study consists of two parts: the training phase to acquaint the animals to the rod, and the test phase. No treatment is administered during the training phase, but all 80 mice are placed four times on the rod. The mice are subsequently split into two groups and one compound is tested per group of 40 animals. Ten mice are then attributed per dose level (0, 2.5, 5, and 10 mg) and tested before, and 30, 60, and 90 minutes after administration. It is favorable to estimate system parameters such as E_{max} in the K-PD model simultaneously over different compounds to improve the estimation. However, this assumes that the compounds share the same method of action, such as blocking the same receptor. In our case study, the compounds have different methods of action, which leads to different maximal attainable effects. Therefore, the data of the training phase were analyzed as a whole, the test phase was analyzed per compound.

A Weibull distribution fits the training data best in contrast to the log-logistic and log-normal distribution (Akaike Information Criterion (AIC): 3265.7, 3379.0, and 3277.3, respectively). A random intercept is included for the training data to cope with the dependency of the measurements between occasions. The Weibull distribution seems most appropriate and this distribution will also be used for the test phase. In the test phase, the sample size was too small to fit a random effect in the models described in the rest of the paper.

The test data are analyzed using a linear regression model and a pharmacological model for both compounds separately. The first model has both the scale and shape parameter fitted with linear dose and quadratic time incorporated. The interaction terms are included apart from dose-by-time^2 in the shape parameter γ'. Addition of the latter resulted in failure of the model's convergence. For the second model, the scale and shape parameter are fitted using an E_{max} model. An underlying, latent, one compartmental latent drug concentration-time profile at the site of action is estimated (Jacqmin *et al.* 2007).

AIC values indicate that the second model gave a superior fit for both compounds: 1445.8 and 1427.8 for compound A, and 1489.8 and 1475.5 for compound B for the linear and E_{max} models, respectively. The model fit of both models is presented in Figures 9.1 for compound A. It illustrates superiority of the E_{max} model. The dashed line in Figure 9.1 shows the inadequate fit of the linear regression model for the higher doses at 30 seconds, whereas the E_{max} model performs well.

The parameter estimates for the two models can be found in Table 9.1. The parameters of the linear model do not allow a direct physiological interpretation.

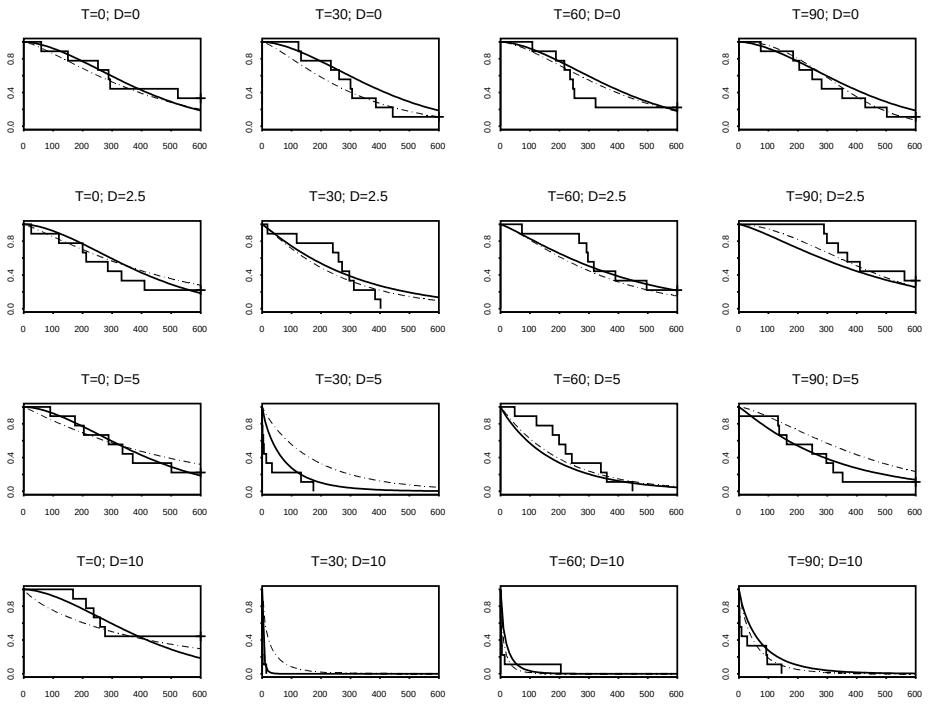


Figure 9.1: Model prediction of the probability to stay on the rod as function over time (seconds) for compound A for the different doses D (mg/kg) and timepoints (minutes) after administration T, where the full line is the E_{max} model, and dashed is used for the linear model.

One can generally say that a reduction of the scale parameter α' is translated into a shorter latency to fall off the rod. However, also the change of the shape parameter γ' has an impact on the median time-to-event. Although the coefficients of T and T^2 are small, it has an important impact on the fit, as time T is expressed in minutes. The quadratic term of T allows for a high time-to-event, both at the start (before administration) and the end of the study at 90 minutes, whereas a short time-to-event is obtained after 30 minutes. The interaction of time with the dose is needed to enforce that there is no change over time for the placebo dose, simultaneously allowing for a significant reduction at 30 minutes. The parameter estimates of the E_{max} model are easier to interpret, as mentioned in Section 9.2. Both the potency (EC_{50}) and the maximal effect E_{max} are directly obtained.

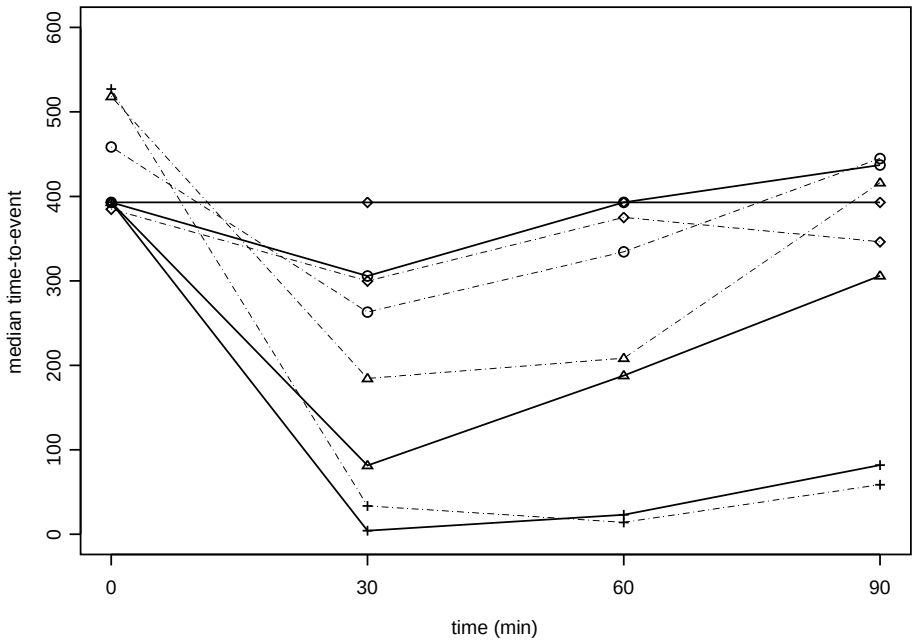


Figure 9.2: Model based median reduction of time to falling off the rod for compound A, where the full line is the E_{max} model, and dashed is used for the linear model. Following symbols are used: $\diamond$ for 0 mg/kg, $\circ$ for 2.5 mg/kg, $\triangle$ for 5 mg/kg, and $+$ for 10 mg/kg.

Figure 9.2 contains the model-predicted median reduction of the time to falling for both models for compound A. The dose-response relationship is readily visible in the figure, with a peak effect at 30 minutes.

9.4 Simulation Study

The observation that a latent pharmacokinetic time profile can be estimated and that it improves the fit compared to a more traditional linear model is worth further scrutiny. Therefore, a simulation study is set up to explore the added value of increasing the number of mice in the study before discussing the results of the experiment. Such additional mice can theoretically be used for plasma concentration sampling or

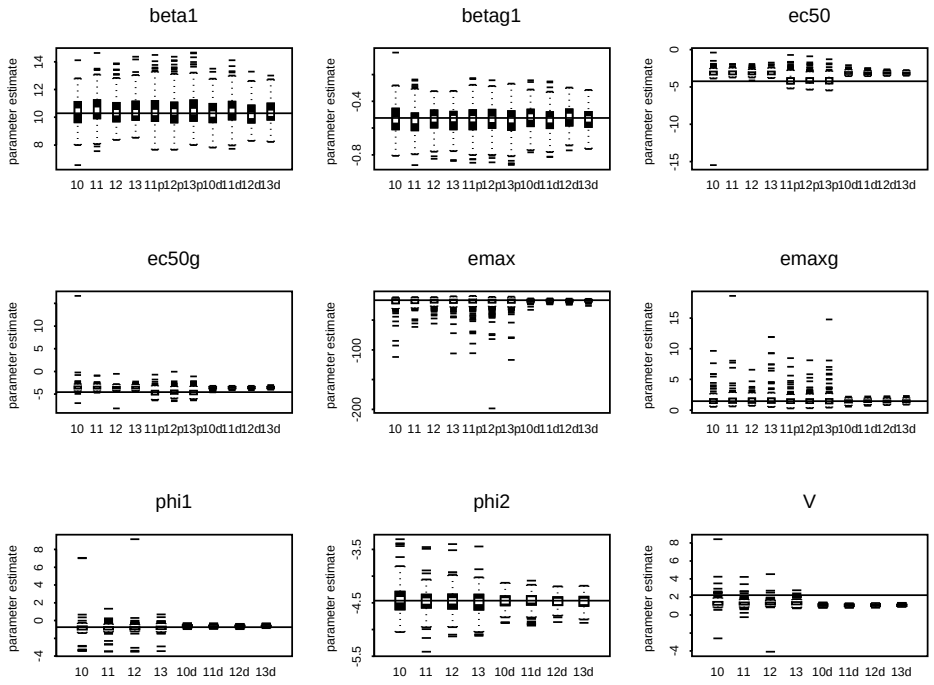


Figure 9.3: Box plots of the parameter estimates (β, β_γ, EC_{50}, EC_{50_γ}, E_{max}, E_{max_γ}, ϕ_1, and ϕ_2) obtained from the study simulation in the case 10, 11, 12, or 13 mice per dose level included in the study for either plasma concentration sampling ('p'), time-to-event sampling, or further dose exploration ('d'). The horizontal line represents the parameter values used for the simulation.

for additional time-to-event data. In the latter case, the additional mice can be attributed to the existing dose levels, or, alternatively, to higher dose levels for a better exploration of the dose-response curve.

Table 9.1: Parameter estimates (standard error) for the different models and compounds. The index γ indicates that the parameter is part of the shape parameter.

Model	Parameter	compound A	compound B
linear	β	7.9784 (1.6784)	8.7640 (1.6043)
	D	-0.3011 (0.1595)	-0.2450 (0.1719)
	T	-0.06011 (0.1207)	-0.1190 (0.09940)
	T^2	0.001342 (0.001545)	0.001502 (0.001201)
	DT	-0.00728 (0.008242)	0.001677 (0.006064)
	DT^2	-7.12E-6 (0.000105)	-0.00006 (0.000065)
	β_γ	-0.2791 (0.2053)	-0.3784 (0.1827)
	D_γ	0.04925 (0.02282)	0.03400 (0.02180)
	T_γ	0.004662 (0.01296)	0.01557 (0.01219)
	T_γ^2	-0.00012 (0.000153)	-0.00019 (0.000145)
	DT_γ	0.000561 (0.000346)	0.000364 (0.000358)
E_{max}	β	10.2922 (1.1088)	10.5209 (1.2918)
	$emax$	-16.7595 (4.3503)	-9.4220 (2.2230)
	$\log(ec50)$	-4.2316 (11.2725)	0.6224 (1.8262)
	β_γ	-0.5252 (0.1070)	-0.5279 (0.1230)
	$emax_\gamma$	1.4692 (0.6529)	1.7611 (0.7990)
	$\log(ec50_\gamma)$	-4.5599 (11.2752)	1.4254 (1.9358)
	ϕ_1	-0.7426 (32.986)	-2.0620 (4.9244)
	ϕ_2	-4.4594 (0.3017)	-4.4693 (1.3075)
	lVf	-2.2085 (22.541)	-2.5522 (3.7463)

Table 9.2: The mean and standard deviation of the parameter estimation in the case of additional mice for time-to-event data, plasma concentration sampling, or an additional dose level.

parameter	mice	time-to-event		plasma concentration		Dose level	
(target value)	per dose	mean	STD	mean	STD	mean	STD
β	10	10.39	1.11	10.39	1.11	10.33	1.00
(10.29)	11	10.56	1.13	10.52	1.16	10.48	1.02
	12	10.46	1.05	10.44	1.10	10.27	0.88
	13	10.49	1.00	10.53	1.14	10.40	0.86
$emax$	10	-21.95	20.52	-21.95	20.52	-17.82	2.21
(-16.76)	11	-19.4	6.66	-20.58	16.56	-17.46	1.52
	12	-20.51	22.03	-19.84	11.27	-17.44	1.59
	13	-20.01	8.96	-19.27	8.08	-17.35	1.56
$\log(ec50)$	10	-3.16	1.00	-3.16	1.00	-3.12	0.22
(-4.23)	11	-3.11	0.33	-4.12	0.66	-3.16	0.18
	12	-3.14	0.29	-4.09	0.58	-3.13	0.17
	13	-3.13	0.32	-4.13	0.58	-3.16	0.16
β_γ	10	-0.53	0.11	-0.53	0.11	-0.52	0.09
(-0.53)	11	-0.55	0.11	-0.54	0.11	-0.54	0.10
	12	-0.54	0.10	-0.53	0.10	-0.52	0.08
	13	-0.54	0.09	-0.54	0.11	-0.53	0.08
$emax_\gamma$	10	3.08	12.88	3.08	12.88	1.41	0.27
(1.47)	11	1.82	2.01	1.63	1.75	1.45	0.24
	12	1.77	2.89	1.65	2.40	1.42	0.24
	13	1.82	1.53	1.61	1.06	1.46	0.26
$\log(ec50_\gamma)$	10	-3.42	1.58	-3.42	1.58	-3.57	0.20
(-4.56)	11	-3.52	0.44	-4.61	0.66	-3.55	0.17
	12	-3.59	0.48	-4.58	0.61	-3.56	0.17
	13	-3.54	0.37	-4.61	0.69	-3.55	0.17
V	10	1.39	0.73			1.09	0.09
(2.21)	11	1.31	0.39			1.08	0.09
	12	1.30	0.47			1.09	0.09
	13	1.32	0.26			1.08	0.08
ϕ_1	10	-0.83	1.00			-0.65	0.11
(-0.74)	11	-0.88	0.44			-0.66	0.09
	12	-0.81	0.76			-0.65	0.10
	13	-0.82	0.34			-0.66	0.09
ϕ_2	10	-4.44	0.29			-4.47	0.13
(-4.46)	11	-4.47	0.24			-4.46	0.14
	12	-4.47	0.23			-4.47	0.11
	13	-4.48	0.22			-4.48	0.12

A Weibull distribution is used to simulate four doses (0, 2.5, 5, and 10) at four time points (before administration, 30, 60, and 90 minutes after administration). Both the shape and scale parameter of the Weibull distribution exhibit a sigmoidal behavior with a latent drug concentration at the site of action. Five hundred studies are simulated per situation. The gain of sampling a number of plasma concentrations in additional mice is assessed by plotting the accuracy of the estimates as a function of the number of additional plasma samples. The simulation is performed under the assumption that only one blood sample can be taken per mouse. Further, it is assumed that blood sampling would influence the time to falling off the rod. Therefore, an additional mouse is required for each plasma concentration sample. The number of plasma concentration samples, and therefore the number of additional mice, is set to vary from 0 up to 12 by steps of 4 due to the number of doses included. The latter would lead to 10 to 13 mice per dose level. The gain of additional mice in the study can as such be explored at three levels: (1) what is the impact on the E_{max} modeling if these additional mice are used for plasma concentrations sampling on the one hand; (2) what is the impact on the modeling if these mice were not used for plasma concentration sampling but for additional rotarod assessments per dose; and (3) what is the consequence of an additional, higher dose level (20 mg)s on the accuracy and precision of the estimates?

Data were simulated according to the proposed latent drug concentration model with and without a measured plasma concentration-time profile. Additionally, a better dose exploration was simulated in a third part of the simulation. Random variation was included in the simulation at two levels; the time-to-event data, and at the latent drug rate or the plasma concentration-time profile. In the simulation, it is assumed that both the latent drug concentration-time profile and the plasma concentration-time profile have a mouse-dependent log-normally distributed elimination. The residual error of the plasma concentration-time profile was assumed to follow a log-normal distribution; there was assumed to be measurement error. Contrary to the plasma concentration-time profile, no measurement error was attributed to the latent drug rate time profile given the latent, i.e., unmeasured, character of the profile.

The assumption of a direct response in this simulation is crucial, i.e., the time of the peak of the plasma concentration-time profile coincides with the time the maximal pharmacodynamic effect. This assumption is required for a fair comparison of the three settings. If the pharmacokinetic and pharmacodynamic peak effect do not coincide, a delay ought to be built in. This is typically achieved with an effects compartment.

Figure 9.3 contains the boxplot with the parameter estimates from the simulation

study. Five hundred studies per situation were simulated. This notwithstanding, model convergence was not attained in a number of studies due to a negative Hessian matrix: only 187, 206, 225, and 226 of the 500 studies had a positive Hessian if 10, 11, 12, and 13 mice per dose group were used. This means that this model estimation fails in more than half of the times in the current design with the number of mice, doses and time points of sampling. If, on the contrary 1, 2, or 3 mice per dose group were additionally recruited for plasma concentration sampling, leading to a total sample size of 11, 12, and 13 mice per dose group, the number of studies with a positive Hessian matrix increased to 354, 450, and 412, respectively. In case of an additional dose level, the number of studies with a positive Hessian matrix was 236, 264, 242, and 239, respectively. The latent drug concentration-time profile turns out to be difficult to estimate, which should not come as a surprise as it is difficult to estimate an unobserved underlying profile.

It is striking that none of the box plots for each of the parameters shows any large differences. This suggests that a small amount of additional mice for either plasma concentration sampling or time-to-event data does not have a large impact on the accuracy of the parameter estimation unless the dose range is explored further. The accuracy and the precision is tabulated as the mean and standard deviation of the parameter estimation in Table 9.2, providing additional detail over the box plots. No major effects are observed for β and β_γ.

The estimates for the simulations with a latent dose concentration and a dose exploration both have precise estimates, but the estimation is slightly biased for EC_{50}, EC_{50_γ}, and V. This is possibly due to confounding. The plasma sampling simulation has slightly bias estimates for E_{max} and E_{max_γ}. The strong decrease of the standard deviation in the case of extra dose exploration suggests that the inflated variability observed in the case study and both other simulations is due to a limited dose range in study. Therefore, the fact that the parameter E_{max} is poorly estimated owes to the poor exploration of the dose-response relationship, rather than the modeling technique. A better dose exploration would solve this issue.

9.5 Discussion

The rotarod test is used to assess the effect in mice on motor coordination of two compounds. The time to fall is measured. Two different models are proposed to model the dose response time-to-event data. Both use a Weibull distribution and are of an accelerated failure type, but the shape and scale parameters include either a linear model or a E_{max} model with latent drug concentration-time profile at the site

of action. The AIC and graphical model fit suggests that the E_{max} model fits the data better.

Unlike the linear model, the parameter estimates of the E_{max} model have a physiological interpretation. Both the potency (EC_{50}) and the maximal effect E_{max} are directly obtained, which facilitates the comparison of different compounds. The E_{max} model being the superior model to the data raises the question whether an alternative design would improve the estimation properties. Therefore, a simulation study was set up to investigate the influence on accuracy and precision if 1, 2, or 3 mice per dose group would be added for either time-to-event data, or plasma concentrations. Does the gain in accuracy and precision justify the 4, 8, or 12 additional mice per compound? Alternatively, the impact of a better dose-response exploration on the estimation properties is considered as well. It turns out that the latent drug concentration-time profile is difficult to estimate in the current design, given the number of non-positive Hessian matrices, which is not surprising; only three time points post administration contain too little information to estimate an underlying, unobserved time profile. The accuracy and precision of the parameters β and β_γ is not altered much, whereas for the other parameters, the standard deviation is inflated if the additional mice are used for time-to-event data compared to when these mice are used for plasma concentration sampling. In this case, the lack of accuracy is also more pronounced. This confirms the results of Jacqmin *et al* (2007). However, the inclusion of the higher dose level improves the precision of the estimation of the parameters tremendously. A small bias is however observed in both latent drug concentration simulations Therefore, a better dose-response exploration would be the best option to improve the current design of the study.

A one-compartmental latent drug concentration-time profile with first-order absorption is used as a latent time profile of the drug at the site of action. What is the impact of miss-specifying the plasma concentration-time profile? In the given design, this is unlikely to have a major impact at the time points the data was collected, because only three assessments are measured after dose administration. Hence, it is unlikely to see, for example, the difference between a one- and a two-compartment model. On the other hand, it would be too demanding for the model to fit such a latent two-compartment drug concentration-time profile given the ten mice per dose group and the three post-administration observations. The impact is considered negligible, because the drug concentration-time profile is only a tool to fit the data with some baseline, followed by a rapid change (increase or decrease) and gradual return to baseline over time.

In conclusion, the study would have benefitted from the addition of a higher dose

level to improve the estimation of the dose-response. The addition of minimally four mice for plasma concentration sampling in the E_{max} model would only be an option if the inclusion of a higher dose level is deemed inappropriate. In that case, historic data might be considered if the data permits. Both would lead to more accurate and precise estimations and as such reduce the risks of bringing the compound into clinical early development. The E_{max} model with latent drug concentration-time profile at the site of action allows for a direct estimation of the potency of the drug candidate. It would lead to an improved comparison of the compounds and bring the study in the more familiar setting of preclinical pharmacology.

10

Concluding Remarks and Further Research

10.1 Concluding Remarks

Many phenomena in our world exhibit a nonlinear behavior, such as physiology, growth curves, etc. One might even consider that all phenomena that are apparently linearly, can only be considered as approximatively linear on a limited, though relevant, interval. The (generalized) linear mixed-effects model has the appealing feature of simplicity, which makes it very suitable for education and research, but the physiological interpretation of the model and the correctness of a simulation is often questionable at the limits of the apparently linear interval. Other phenomena lack completely the linear behavior, such as the case studies in this dissertation. It is clear from these examples that nonlinear mixed-effects modelling is a very useful technique deserving more attention by the statistical community, but the complexity of the methodology might be a hurdle.

IVIVC modelling exists already a number of decades. However, little attention was paid to it. The recent introduction of the convolution-based model can be considered as the first fundamental research beyond the application of the deconvolution-based method on case studies since its introduction. The convolution-based model was

the start to compare it with the existing deconvolution-based method, but also to consider more complicated controlled-release formulations, such as a heterogeneous formulation in Chapter 4. Indeed, to our knowledge, IVIVCs have been limited to the case of homogeneous formulations. The standard convolution-based model is extended by considering the *in-vitro* dissolution profile as a mixture of dissolution profiles, where each profile corresponds to the dissolution of one type of drug product within the formulation. The convolution model is then considered as a mixture of the corresponding plasma concentration-time profiles.

Although IVIVC modelling is a special form of pharmacokinetic modelling, it has some additional difficulties. One obvious complexity is that the input function $F'_{i2k\ell}$ is unobserved, complex, and, in contrast to more standard pharmacokinetic analyses, not estimable without imposing assumptions. Exactly these assumptions contain the strength of the method: it imposes a hypothetical relation between the *in-vitro* dissolution and the *in-vivo* release of the drug product. The applications vary from formulation development, SUPAC-related issues to dissolution specification setting. It ensures that an inexpensive, fast *in-vitro* test guaranties the well-behavior of the *in-vivo* drug exposure. The vast amount of applications is extended in this dissertation to relate the *in-vitro* dissolution specification with a clinical interpretation.

Fitting the IVIVC models requires the use of nonlinear hierarchical models. These models can be considered as the most general type of models. Linear and generalized linear hierarchical models can be considered as special cases of nonlinear models. This has severe drawbacks; the strong asset of nonlinear models is that any chosen shape can be fitted and one is not restricted to (generalized) linear fits. This allows physiological meaningful models. The drawback is, however, that one loses some of the advantages of (generalized) linear models. Fitting a nonlinear mixed effects model requires well-chosen starting values to initiate the likelihood iterations, the iteration algorithm tends to be slower, and the likelihood evaluation requires the integration of the random effects in the likelihood function. For the specific case of IVIVC modelling, the flexibility of nonlinear mixed effects models allows for an additional choice for the statistician; fitting all the data at once, or breaking up to problem by fitting the UIR first and then using these estimates in the second stage of the model. The precision of the two-stage model is, however, less compared to the one-stage.

However, obtaining estimates of the nonlinear mixed effects model is not sufficient. The adequacy of the model fit is crucial for the use in further simulations. The goodness-of-fit of a IVIVC model is traditionally evaluated by a residual analysis and $\%PE$. In this dissertation, it is demonstrated that this is appropriate, but insufficient. Also outlier detection ought to be performed. Local influence is an suitable technique

for this purpose that allows for evaluating the data from a different perspective.

However, one might wonder why plasma concentrations acquire so much attention in drug development. It is very simple: it represents the systematic exposure of a patient to a drug substance and, therefore, it is more relevant than the administered dose. Further, the AUC and C_{max} can be considered as biomarkers for the safety and efficacy of the drug product. The fact that IVIVC models allow for a direct translation of the *in-vitro* dissolution time profile ensures the importance of the technique. However, one should beware of attributing the focus solely on the plasma concentrations. When for example new batches of drug product are tested *in-vivo*, it is assumed that a difference of less than 20% does not have a clinical impact. However, it is know that for certain drug product, small changes in exposure might lead to significant clinical changes. Therefore, standard bioequivalence testing is extended to cope with the size of the therapeutic window. As such, more stringent acceptance limits are imposed for narrow index drugs, whereas the limits are broadened for drug product with a broad therapeutic window.

Are plasma concentrations deemed crucial to understand and model the pharmacodynamic effects? No, it is possible to fit a dose-response model on the pharmacodynamic response without having the plasma concentrations explicitly. Especially in preclinical scenarios with for example mice, the limited amount of blood in the animal can be prohibitive for blood sampling or it have an effect on the pharmacodynamic response. Nevertheless, a dose response can still be fitted to the data assuming a latent pharmacokinetic profile. These models are referred to as K-PD models to stress the absence of plasma concentrations. The use of such a K-PD model allows not only for a physiological model interpretation, it also ensures physiologically possible simulations. This is not necessarily the case for (generalized) linear models. Therefore, the nonlinear mixed effects model has the strong asset to allow for more realistic simulations within pharmacology.

10.2 Further Research

One of the main restrictions of the research during the past four years was the computational time for the IVIVC models, making the technique difficult to implement at this moment. Although other software and algorithms were very briefly investigated, no suitable alternative was found at the time. It would be of interest to reconsider the use of different software and to reduce the computational time as such. The model could be formulated both as a convolution product or as a differential equation in software like monolix, winbugs, or nonmem. Also other methods might be considered

to reduce the computational times, such as the pseudo-likelihood approach, or the implementation of the convolution product as a fourier transformation.

On the other hand, we already illustrated in Section 3.3 that the goodness-of-fit measurement $\%PE$ has some weaknesses. It would be useful to look for alternative approaches to evaluate the model fit. One of the possible alternatives might be to consider the *in-vitro* dissolution as a surrogate marker for the *in-vivo* release. The information theory-based approach of Alonso and Molenberghs (2007) might be considered as a potential alternative to $\%PE$.

A lot of new research can be performed not only in IVIVC modelling, but more generally in cost reduction and optimization of drug development. For example, the simulations in Chapter 9 indicate that the addition of a dose level allows for a better estimation and as such leads to a better selection of the drug candidate. Similar to the IVIVC setting, also other potential relations ought to be found so that the combination of *in-vitro* and *in-vivo* data increases the predictive power. The whole setting of cost reduction via optimization of the drug candidate selection will potentially be one of the big research topics within the pharmaceutical companies. It is up to each statistician and pharmacometrician to prepare for these new challenges and opportunities.

Publications and Reports

Jacobs, T., Straetemans, R., Rossenu, S., Dunne, A., Molenberghs, G., Bijnens, L. (2007) A new convolution-based approach to develop a level A *in-vitro in-vivo* correlation (IVIVC) for an extended release oral dosage form. *Proceedings of the American Statistical Association, Biometrics Section*, Alexandria, 141–146.

Jacobs, T., Rossenu, S., Dunne, A., Molenberghs, G., Straetemans, R., Bijnens, L. (2008) Combined models for data from *in-vitro–in-vivo correlation* experiments. *Journal of Biopharmaceutical Statistics*, **18**, 197–211.

Jacobs, T., De Ridder, F., Rusch, S., Van Peer, A., Molenberghs, G., Bijnens, L. (2008) Including information on the therapeutic window in bioequivalence acceptance limits. *Pharmaceutical Research*, **25**, 2628–2638.

Jacobs, T., Molenberghs, G., Dunne, A., Straetemans, R., Bijnens, L. (2009) Accuracy and Precision of One- versus Two-stage Convolution-based IVIVC Modelling. *Technical Report*.

Jacobs, T., Molenberghs, G., Dunne, A., and Bijnens, L. (2009) Outlying Subject Detection in In-vitro In-vivo Correlation Models. *Technical Report*.

Jacobs, T., De Ridder, F., Dunne, A., Molenberghs, G., Straetemans, R., Mannaert, E., Vermeulen, A., and Bijnens, L. (2009) Can in-vitro dissolution specifications be determined by its clinical relevance? *Technical Report*.

Jacobs, T., Straetemans, Molenberghs, G., R., Bouwknecht, J.A., and Bijnens, L. (2009) A Latent Pharmacokinetic Time Profile to Model Dose-response Survival Data. *Technical Report*.

References

Amidon, G.L., Lennernas, H., Shah, V.P., and Crison, J.R. (1995). A theoretical basis for a biopharmaceutic drug classification: The correlation of in vitro drug product dissolution and in vivo bioavailability. *Pharmaceutical Research*, **12**, 413–419.

Anderson, S., Hauck, W. (1990). Consideration of individual bioequivalence. *Journal of Pharmacokinetics and Biopharmaceutics*, **18**, 259–273.

Balan, G., Timmins, P., Greene, D., Marathe, P. (2001) In vitro-in vivo correlation (IVIVC) models for metformin after administration of modified-release (MR) oral dosage forms to healthy human volunteers. *Journal of Pharmaceutical Sciences*, **90**, 1176–1185.

Beal, S.L., Sheiner, L.B. (1982). Estimating population kinetics. *Critical Reviews in Biomedical Engineering*, **8**, 195–222.

Bochner F., Huffman D., Shen D., Azarnoff D. (1977). Bioavailability of digoxin-hydroquinone complex: A new oral digoxin formulation *Journal of Pharmaceutical Sciences*, **66**, 644–647.

Boddy, A., Snikeris, F., Kringle, R., Wei, G., Opperman, J., Midha, K. (1995) An approach for widening the bioequivalence limits in the case of highly variable drugs *Pharmaceutical Research*, **12**, 1865-1868.

Brendel, K., Comets, E., Laffont, C., Laveille, C., Mentr, F. (2006) Metrics for external model evaluation with an application to the population pharmacokinetics of gliclazide *Pharmaceutical Research*, **23**, 2036-2049.

CDER Guidance for industry (1997). Extended release oral dosage forms: development, evaluation, and application of in vitro/in vivo correlations.

Cheung, R., Kuba, R., Rauth, A., Wu, X. (2004) A new approach to the in vivo and in vitro investigation of drug release from locoregionally delivered microspheres. *Journal of Controlled Release*, **100**, 121-133.

Comets, E., and Mentré, F. (2001). Evaluation of tests based on individual versus population modelling to compare dissolution curves. *Journal of Biopharmaceutical Statistics*, **11**, 107-123.

Committee for Proprietary Medicinal Products (CPMP), the European Agency for the Evaluation of Medicinal Products (EMEA); "Note for guidance on the investigation of bioavailability and bioequivalence"; 2001.

Committee for Proprietary Medicinal Products (CPMP), the European Agency for the Evaluation of Medicinal Products (EMEA); "Concept paper for an addendum to the note for guidance on the investigation of bioavailability and bioequivalence: evaluation of bioequivalence of highly variable drugs and drug products"; 2006.

Cook, D. (1986). Assessment of local influence. *J. R. Statist. Soc. B*, **48**, 133-169.

Dalton, J., Straughn, A., Dickason, D., Grandolfi, G. (2001) Predictive Ability of Level A in Vitroin Vivo Correlation for RingCap Controlled-Release Acetaminophen Tablets. *Pharmaceutical Research*, **18**, 1729–1734.

Davidian, M., and Giltinan, D. (1995). *Nonlinear Models for Repeated Measurement Data* Chapman & Hall/CRC Press.

Dunne, A.J., O'Hara, T., and Devane, J. (1999). A new approach to modelling the relationship between in vitro and in vivo drug dissolution/absorption. *Statistics in Medicine*, **18**, 1865–1876.

Dunne, A., Gaynor, C., and Davis, J. (2005). Deconvolution based approach for level A in vivo-in vitro correlation modelling: statistical considerations. *Clinical Research and Regulatory Affairs*, **22**, 1–14.

Dutta, S., Qiu, Y., Samara, E., Cao, G., Granneman, R. Once-a-day extended-release dosage form of divalproex sodium III: Development and validation of a Level A in vitro-in vivo correlation (IVIVC). *Journal of Pharmaceutical Sciences*, **94**, 1949–1956.

Eddershaw, P., Beresford, A., Bayliss, M. (2000) ADME/PK as part of a rational approach to drug discovery. *Drug Discovery Today*, **5**, 409–414.

Emami, J. (2006). In vitro – In vivo correlation: from theory to applications. *Journal of pharmacy and pharmaceutical sciences*, **9**, 31–51.

Escobar, L., Meeker, W. (1992) Assessing influence in regression analysis with censored data. *Biometrics*, **48**, 507-528.

FIP (1996). Guidelines for dissolution testing of solid oral products. *Drug Information Journal*, **30**, 1071-1084.

U.S. Food and Drug Administration, Center for Drug Evaluation and Research; "Guidance for Industry: Bioavailability and Bioequivalence Studies for Orally Administered Drug Products — General Considerations"; 2003.

Gabrielson, J., Weiner, D. (2000). *Pharmacokinetic and Pharmacodynamic Data Analysis: Concepts and Applications*, Apotekarsocieteten, Stockholm, Sweden.

Gaynor, C., Dunne, A., and Davis, J. (2007). A comparison of the prediction accuracy of two IVIVC modelling techniques. *Journal of Pharmaceutical Sciences*, **00**, 000-000.

Gerald, M.C., Gupta, T.K. (1977). The effects of amphetamine isomers on rotarod performance. *Psychopharmacology*, **55**, 83-86.

Gibaldi, M., Perrier, D. (1982). *Pharmacokinetics*, New York, Marcel Dekker.

Gillespie, W.R., Veng-Pedersen, P. (1985). Gastro-intestinal bioavailability: determination of in vivo release profiles of solid dosage forms by deconvolution. *Biopharmaceutics and Drug Disposition*, **6**, 351355.

Gillespie, W.R. (1997). Convolution-based approaches for in vivo-in vitro correlation modelling. *Advances in Experimental Medicine and Biology*, (423), 53-65.

Hayes, S., Dunne, A., Smart, T., and Davis, J. (2004). Interpretation and optimization of the dissolution specifications for a modified release product with an in vivo–in vitro correlation (IVIVC). *Journal of Pharmaceutical Sciences*, **93**, 571-581.

Holford, N. (2006). Dose response: Pharmacokinetic-pharmacokdynamic approach in Ting, N. . *Dose Finding in Drug Development*, New York: Springer-Verlag.

Hsuan, F. (2000). Some statistical considerations on the FDA draft guidance for individual bioequivalence. *Statistics in Medicine*, **19**, 2879–2884.

Hayes, S., Dunne, A., Smart, T., Davis, J. (2003) Interpretation and optimization of the dissolution specifications for a modified release product with an In VivoIn Vitro Correlation (IVIVC). *Journal of Pharmaceutical Sciences*, **93**, 571-581.

Jacobs, T., Straetemans, R., Rossenu, S., Dunne, A., Molenberghs, G., Bijnens, L. (2007) A new convolution-based approach to develop a level A *in-vitro in-vivo* correlation (IVIVC) for an extended release oral dosage form. *Proceedings of the American Statistical Association, Biometrics Section*, Alexandria, 141–146.

Jacobs, T., Rossenu, S., Dunne, A., Molenberghs, G., Straetemans, R., Bijnens, L. (2008) Combined models for data from *in-vitro–in-vivo correlation* experiments. *Journal of Biopharmaceutical Statistics*, **18**, 197–211.

Jacobs, T., De Ridder, F., Rusch, S., Van Peer, A., Molenberghs, G., Bijnens, L. (2008) Including information on the therapeutic window in bioequivalence acceptance limits. *Pharmaceutical Research*, **25**, 2628–2638.

Jacqmin, P., Snoeck, E., van Schaick, E., Gieschke, R., Pillai, P., Steimer, J.-L., and Girard, P. (2007). Modelling Response Time Profiles in the Absence of Drug Concentrations: Definition and Performance Evaluation of the KPD Model. *Journal of Pharmacokinetics and Pharmacodynamics*, **34**, 57-85.

Jansen, I., Molenberghs, G., Aerts, M., Thijs, H., Van Steen, K. (2003) A Local Influence Approach Applied to Binary Data from a Psychiatric Study. *Biometrics*, **59**, 410 - 419.

Ju, H.L., Liaw, S.-J. (1997) On the assessment of similarity of drug dissolution profiles – a simulation study. *Drug Information Journal*, **31**, 12731289.

Kapur, S., Zipursky, R., Remington, G. (1999) Clinical and theoretical implications of 5-HT_2 and D_2 receptor occupancy of clozapine, risperidone, and olanzapine in schizophrenia. *American Journal of Psychiatry*, **156**, 286–293.

Kapur, S., Zipursky, R., Jones, C., Remington, G., Houle, S. (2000) Relationship between dopamine D_2 occupancy, clinical response, and side effects: a double-blind PET study of first-episode schizophrenia. *American Journal of Psychiatry*, **157**, 514–520.

Karalis, V., Symillides, M., Macheras, P. (2004) Novel Scaled Average Bioequivalence Limits Based on GMR and Variability Considerations. *Pharmaceutical Research*, **21**, 1933–1942.

Karalis, V., Macheras, P., Symillides, M. (2005) Geometric mean ratio-dependent scaled bioequivalence limits with leveling-off properties. *European Journal of Pharmaceutical Science*, **26**, 54–61.

Kenakin, T. (1997). Differences between natural and recombinant G protein-coupled receptor systems with varying receptor/G protein stoichiometry. *Trends in Pharmacological Sciences*, **18**, 456–464.

Klein, J., Moeschberger, M. (1997). *Survival Analysis – Techniques for Censored and Truncated Data*, New York: Springer-Verlag.

Kortejarvi, H., Malkki, J., Marvola, M., Urtti, A., Yliperttula, M., Pajunen, P. (2006) Level A In Vitro-In Vivo Correlation (IVIVC) Model with Bayesian Approach to Formulation Series. *Journal of Pharmaceutical Sciences*, **95**, 1595–1605.

Lee, S.-Y., Tang, N.-S. (2004) Local influence analysis of nonlinear structural equation models. *Psychometrika*, **69**, 573-592.

Lilienfeld, S. (2002). Galantamine–a novel cholinergic drug with a unique dual mode of action for the treatment of patients with Alzheimer's disease. *CNS Drug Reviews*, **8**, 159-76.

Lindsey, J.K. (1997). *Applying generalized linear models.* New York: Springer-Verlag.

Macheras, P., Rosen, A. (1983) The Bioequivalence Factor. *Pharm Acta Helvetiae*, **58**, 233–236.

Mahayni, H., Rekhi, G., Uppoor, R., Marroum, P., Hussain, A., Augsburger, L., Eddington, N. (2000) Evaluation of External Predictability of an In VitroIn Vivo Correlation for an Extended-Release Formulation Containing Metoprolol Tartrate. *Journal of pharmaceutical sciences*, **89**, 1354–1361.

Martin, D., Tompson, D., Boike, S., Tenero, D., Ilson, B., Citerone, D., Jorkasky, D. (1997) Lack of effect of eprosartan on the single dose pharmacokinetics of orally administered digoxin in healthy male volunteers. *British Journal of Clinical Pharmacology*, **43**, 661–664.

McIlwain, K.L., Merriweather, M.Y., Yuva-Paylor, L.A., and Paylor, R. (2001). The use of behavioral test batteries: Effects of training history. *Physiology and Behavior*, **73**, 705-717.

Meyer M.C., Straughn A.B., Mhatre R.M. (2001). Variability in the bioavailability of phenytoin capsules in males and females. *Pharmaceutical Research*, **18**, 394–397.

Ministry of Health, Canada (2006) Bioequivalence requirements: Critical dose drugs.

Mistry, G., Laurent, A., Sterrett, A., Deutsch, P. (1999) Effect of indinavir on the single-dose pharmacokinetics of theophylline in healthy subjects. *Journal of Clinical Pharmacology*, **39**, 636–642.

Modi, N.B., Lam, A., Lindemulder, E., Wang, B., and Gupta, S.K. (2000). Application of in vitro-in vivo correlations (IVIVC) in setting formulation release specifications. *Biopharmaceutics and Drug Disposition*, 21, 321–326.

Molenberghs, G., and Verbeke, G. (2005). *Models for Discrete Longitudinal Data.* New York: Springer-Verlag.

Nyberg, S., Farde, L., Halldin, C., Dahl M-L., Bertilsson, L. (1995). D_2 dopamine receptor occupancy during low-dose treatment with haloperidol decanoate. *American Journal of Psychiatry*, **49**, 538–544.

O'Hara, T., Hayes, S., Davis, J., Devane, J., Smart, T., and Dunne, A. (2001). in vivo–in vitro correlation (IVIVC) modelling incorporating a convolution step. *Journal of Pharmacokinetics and Pharmacodynamics*, **28**, 277–298.

Okumu, A., DiMaso, M., Lbenberg, R. (2008) Dynamic Dissolution Testing To Establish In Vitro/In Vivo Correlations for Montelukast Sodium, a Poorly Soluble Drug. *Pharmaceutical Research*, **25**, 2778–2785.

O'Quigley, J., Pepe, M., Fisher, L. (1990) Continual reassessment method: a practical design for phase I clinical trials in cancer. *Biometrics*, **46**, 33–48.

Patterson, S., Francis, S., Ireson, M., Webber, D., Whitehead, J. (1999) A novel Bayesian decision procedure for early-phase dose finding studies. *Journal of Biopharmaceutical Statistics*, **9**, 583–597.

Pinheiro, J., Bates, D. (2000). *Mixed-Effects Models in S and S-PLUS*, Springer.

Piotrovsky, V., Van Peer, A., Van Osselaer, N., Armstrong, M., and Aerssens, J. (2003). Galantamine population pharmacokinetics in patients with alzheimer's disease: modelling and simulations. *Journal of Clinical Pharmacology*, **43**, 514-523.

Poon, W., Poon, Y. (1999). Conformal normal curvature and assessment of local influence. *J. R. Statist. Soc. B*, **61**, 51-61.

Remington, G., Kapur, S., Zipursky, R. (1998) The relationship between risperidone plasma levels and dopamine D_2 occupancy: a positron emission tomographic study (letter). *Journal of Clinical Psychopharmacology*, **18**, 82–83.

Remington, G., Mamo, D., Labelle, A., Reiss, J., Shammi, C., Mannaert, E., Mann, S, Kapur, S. (2006) A PET study evaluating dopamine D_2 receptor occupancy for long-acting injectable risperidone. *American Journal of Psychiatry*, **163**, 396–401.

Rowland, M., Tucker, G. (1980) *Journal of Pharmacokinetics and Pharmacodynamics*, **8**, 497–507.

Santos Buelga, D., Garcia, M.J., Otero, M.J., Martin Suarez, A., Dominguez-Gil, A., Lukas, J.C. (2002) Phenytoin covariate models for Michaelis-Menten pharmacokinetics in adult epileptic patients *Page meeting*, Paris.

Sheiner, L.B. (1992) Bioequivalence revisited *Statistics in Medicine*, **11**, 1777–1788.

Schuirmann, D. (1987) A comparison of the two one-sided tests procedure and the power approach for assessing the average bioavailability. *Journal of Pharmacokinetics and Biopharmaceutics*, **15**, 657–680.

Shi, L., Ojeda, M. (2004). Local influence in multilevel regression for growth curve. *Journal of Multivariate Analysis*, **91**, 282-304.

Sirisuth, N., Eddington, N. (2000) Influence of Stereoselective Pharmacokinetics in the Development and Predictability of an IVIVC for the Enantiomers of Metoprolol Tartrate. *Pharmaceutical Research*, **17**, 1019–1025.

Sirisuth, N., Augsburger, L., Eddington, N. (2002) Development and validation of a non-linear IVIVC model for a diltiazem extended release formulation. *Biopharmaceutics & Drug Disposition*, **23**, 1–8.

St.-Laurent, R., Cook, D. (1996). Leverage, local influence and curvature in nonlinear regression. *Biometrika*, **80**, 99-106.

Steinijans, V., Sauter, R., Hauschke, D., Diletti, E., Schall, R., Luus, H., Elze, M., Blume, H., Hoffmann, C., Franke, G., Siegmund, W. (1995). Reference tables for the intrasubject coefficient of variation in bioequivalence studies. *International journal of clinical pharmacology and therapeutics*, **33**, 427-430.

Takka, S., Sakr, A., Goldberg, A. (2003) Development and validation of an in vitroin vivo correlation for buspirone hydrochloride extended release tablets. *Journal of Controlled Release*, **88**, 147-157.

Thomas, W., Cook, D. (1990) Assessing influence on predictions in generalized linear models. *Technometrics*, **32**, 59-65.

Vangeneugden, T., Laenen, A., Geys, H., Renard, D., Molenberghs, G. (2004). Applying linear mixed models to estimate reliability in clinical trial data with repeated measurements.*Controlled Clinical Trials*, **25**, 1330.

Veng-Pedersen, P., Gobburu, J.V.S., Meyer, M.C., and Straughn, A.B. (2000). Carbamazepine level-A in vivo-in vitro correlation (IVIVC): a scaled convolution based predictive approach. *Biopharmaceutics and Drug Disposition*, **21**, 1-6.

Verbeke, G., and Molenberghs, G. (2000). *Linear Mixed Models for Longitudinal Data.* New York: Springer-Verlag.

Wagner, J. (1975). *Fundamentals of Clinical Pharmacokinetics.*, Drug Intelligence.

Samenvatting

Vele fenomenen in onze wereld vertonen een niet-lineair gedrag, zoals fysiologie, de groeicurven, enz. Men zou zelfs kunnen stellen dat voor alle fenomenen die op het eerste zicht lineair lijken, dat dit slechts kan worden beschouwd als ongeveer lineair op een beperkt, maar relevant, interval. Het (generalized) linear mixed-effects model heeft het voordeel van de eenvoud, wat het voor onderwijs en onderzoek zeer geschikt maakt, maar de fysiologische interpretatie van het model en de juistheid van een simulatie staat vaak ter discussie aan de grenzen van het interval. Andere fenomenen ontberen dit lineaire gedrag volledig, zoals de case studies in deze verhandeling. Deze voorbeelden maken duidelijk dat niet-linear mixed-effects modelleren een zeer nuttige en krachtige techniek is die meer aandacht verdient door de statistische gemeenschap. De complexiteit van de methodologie beperkt echter de wijde toepassing.

In-vitro In-vivo correlation modelleren bestaat reeds een aantal decennia. Nochtans werd er weinig aandacht aan besteed wat fundamenteel onderzoek betreft. Ondanks de vele toepassingen van de deconvolutie-gebaseerde methode, kan de recente introductie van het convolutie-gebaseerde model als eerste fundamenteel onderzoek worden beschouwd sinds de introductie van het concept. Het gedrag van de twee methoden werd vergeleken in de eerste plaats, maar ook ingewikkeldere vertraagde-werkings formuleringen van medicijnen kunnen worden onderzocht, zoals een heterogene formulering met twee vrijgave mechanismen in Hoofdstuk 4. Zo is bij ons weten de toepassing van IVIVCs tot op heden beperkt geweest tot homogene formuleringen. Het standaard convolution-gebaseerde model is daarom uitgebreid door het *in-vitro* oplossingsprofiel als een combinatie van twee oplossingsprofielen te beschouwen, waarbij elk profiel aan de oplossing van een type van drug product binnen de formulering beantwoordt. Het convolution-gebaseerde model wordt dan beschouwd als een combinatie van de overeenkomstige plasma concentratie-tijd profielen.

Hoewel IVIVC als een speciale vorm van pharmacokinetisch modelleren kan worden beschouwd, heeft het enkele extra complicaties. Een duidelijke complicatie is dat

de *in-vivo* vrijgave $F'_{i2k\ell}$ van het medicijn niet is waargenomen en complexer is, en niet schatbaar is zonder veronderstellingen aan te nemen, in tegenstelling tot meer standaard pharmacokinetische analyzes. Net deze veronderstellingen zijn de sterkte van methode: er wordt een hypothetische relatie tussen de *in-vitro* en de *in-vivo* vijgave van het drugproduct verondersteld en geevalueerd hoe goed het model de gegevens voorspelt. De toepassingen variëren van het ontwikkelen van nieuwe formuleringen voor vertraagde werking van bestaande medicijnen, en de vijgave specificaties in verband met SUPAC gerelateerde kwesties. Als dusdanig laat een IVIVC toe dat aan de hand van een goedkope, snelle *in-vitro* test de patienten van een optimale blootstelling aan het medicijn kunnen genieten. De enorme hoeveelheid toepassingen wordt in deze verhandeling uitgebreid door de *in-vitro* vrijgave specificaties te relateren aan een klinische interpretatie.

Het modelleren van IVIVC modellen vereist het gebruik van niet-lineair hiërarchische modellen. Deze modellen kunnen als meest algemeen soort model worden beschouwd waarvan zowel lineaire als generalized lineaire hirarchische modellen als speciale gevallen van kunnen worden beschouwd. Dit heeft zowel positieve als negatieve consequencies: de flexibiliteit van niet-lineaire modellen laten toe om om het even welke gekozen vorm van model te modelleren en men is niet beperkt tot (generalized) lineaire modellen. Dit laat toe om fysiologisch gebaseerde modellen te fitten. Het nadeel is, echter, dat het de eenvoud van (generalized) lineaire modellen verloren gaat. Het fitten van een niet-lineair mixed effects model vereist welgekozen startwaarden om de likelihood iteraties te initiëren. Verder verloopt de likelihood optimalisatie veel langzamer door het uit integreren van de random effecten in de likelihood functie. In het specifieke geval van IVIVC, is er nog een extra complexiteit voor de statisticus: worden alle gegevens samen gemodelleerd, of worden eerst de gegevens van de unit impuls response (UIR) gemodelleerd om dan in een tweede stap deze schattingen te gebruiken om de gegevens van de vertraagde-werkings formulering te modelleren? De precisie van het model in twee stadia is, echter, minder vergeleken bij het modelleren van alle data in een enkele stap.

Het volstaat echter niet om enkel schattingen voor het niet-lineaire mixed-effects model te verkrijgen. Hoe goed het model de gegevens benaderen is essentieel voor het toepassen van het model in verdere simulaties. De model fit van een IVIVC model wordt traditioneel geevalueerde door een residual analyse en $\%PE$. In deze verhandeling tonen we aan dat dit aangewezen maar ontoereikend is. Ook moet er een outlier analyse worden uitgevoerd. Local influence is een geschikte techniek om de gegevens vanuit een ander perspectief te bekijken. Hiermee wordt nagegaan wat de impact van bepaalde gegevens is op het model.

Men zou zich kunnen afvragen waarom er zo veel aandacht wordt besteed aan plasmaconcentraties tijdens het ontwikkelen van een medicijn. Het is zeer eenvoudig: het stelt de systematische blootstelling van een patiënt aan het product voor en bijgevolg is het relevanter dan de toegediende dosis. In dit opzicht kunnen de totale blootstelling AUC en de piekblootstelling C_{max} als biomarkers voor de veiligheid en de werking van het medicijn worden beschouwd. Het feit dat IVIVC modellen toelaten om het *in-vitro* vrijgave profiel direct te vertalen in de *in-vivo* blootstelling van patienten aan het medicijn maakt deze techniek zo belangrijk. Men mag echter niet alleen zich toespitsen op plasmaconcentraties. Wanneer bijvoorbeeld een nieuwe batch van het medicijn wordt aangemaakt, veronderstelt men dat een verschil van minder dan 20% geen klinische invloed heeft. Nochtans is het geweten dat voor bepaalde medicijnen kleine veranderingen van de blootstelling tot klinisch significante veranderingen kan leiden. Daarom hebben we de standaard methode om bioequivalentie te testen uitgebreid om de grootte van het therapeutische venster in rekening te nemen. Als dusdanig worden de strengere voorwaarden gesteld voor medicijnen met een klein therapeutisch venster, terwijl de grenzen voor een product met een breed therapeutisch venster verbreed worden.

Zijn plasma concentraties strikt genomen noodzakelijk om de farmacologie van een geneesmiddel te begrijpen? Het is mogelijk om een relatie tussen de klinische respons en de dosering van het geneesmiddel te bekomen zonder de specifieke systeem blootstelling te kennen van de patienten. Vooral bij preklinische studies met kleine dieren zoals bijvoorbeeld muizen, kan het onmogelijk zijn om bloedstalen te nemen door de beperkte hoeveelheid bloed in het dier enerzijds, of kan het een effect hebben op de pharmacodynamische uitkomst. Ook in deze situatie kan een relatie tussen de dosering en de reactie worden opgesteld door een latent onderliggend pharmacokinetisch profiel te veronderstellen. Deze modellen worden aangeduid als K/PD modellen om het ontbreken van plasmaconcentraties te beklemtonen. Het gebruik van zo een K/PD model staat niet alleen een fysiologische model interpretatie toe, het laat ook toe om fysiologisch zinvolle simulaties uit te voeren om de impact van veranderingen aan de studie design te bestuderen. Hierin schuilt de kracht van een niet-lineaire model: het laat toe om realistischere simulaties binnen de farmacologie uit te voeren in tegenstelling tot (generalized) lineaire modellen.