CHALMERS

Optimal design of clinical studies using the PFIM software

Master's Thesis in Engineering Mathematics and Computational Science

JACOB LEANDER

Department of Mathematics & Mathematical Statistics
Chalmers University of Technology
Gothenburg, Sweden 2012

AstraZeneca

Abstract

In this thesis PFIM, a software used for evaluation and optimization of clinical studies based on the Fisher information matrix, is investigated. The program is evaluated using three different models, of which two are PK/PD models and the third is a model describing glucose-insulin regulation in the absence of drug.

For an analytical PK/PD model describing therapeutic response the expected standard errors of the parameter estimates given by PFIM correspond to the empirical standard errors obtained from a large simulation in NONMEM. PFIM offers a fast and effective way of calculating the expected standard errors for the parameter estimates using an approximation of the Fisher information matrix. In the glucose-insulin model it is showed that the number of measurements per patient can be decreased while maintaining the same information criterion. The other PK/PD model describes a long-term safety marker with turnover response where the input rate varies over time. The results indicate that PFIM has problem with long running time if a model is specified by an ordinary differential equation system.

In conclusion PFIM can be used to evaluate, reduce and improve designs for clinical studies. For analytical models, the expected standard errors are appropriate. If the model is described by an ordinary differential equation system the expected standard errors obtained from PFIM are slightly lower compared to empirical standard errors. The recommendation is to use PFIM for evaluating and comparing suggested design alternatives. However, the selected design should be evaluated using simulation techniques.

Acknowledgements

This thesis has been done in collaboration with Clinical Pharmacology Sciences and Statistics & Programming at AstraZeneca R&D, Mölndal.

I would like to thank my examiner Ziad Taib and my supervisor Daniel Röshammar. I have a learned a lot from all the fruitful discussions we had during my work at AstraZeneca.

Jacob Leander, Gothenburg June 7, 2012

"All models are wrong, but some are useful" - George E.P Box

Abbreviations and terms

k_a Absorption rate

Cl Clearance

C_{50} Concentration needed to achieve half of the maximal effect

ω^2 Diagonal element of Ω

D Dose

C Drug concentration

k_e Elimination rate

FO First order

FOCE First order conditional estimation

β Fixed parameter vector

ξ Individual design

E_{max} Maximal effect

MBDD Model based drug development

V Volume of distribution

Ω Parameter covariance structure

PD Pharmacodynamics

PK Pharmacokinetics

Ξ Population design

λ Random parameter vector

Σ Residual covariance structure

SAEM Stochastic approximation expectation maximization

Contents

1

Introduction

In medicine and other scientific disciplines there can be problem making inference from experiments based on repeated measurements from a collection of individuals in a given population. There are often few measurements from each individual and the data may be sparse.

In clinical drug development data normally is collected from a limited number of healthy subjects and patients. The data is then used to draw conclusions about the efficacy and safety of newly discovered drugs in the target patient population.

1.1 What is a clinical study?

Generally speaking, a clinical study is an experiment used to evaluate the effectiveness and safety of a treatment. Clinical studies are a necessary part of the drug development process to assure that the treatment is safe and effective before a drug should reach the market.

Pharmaceutical clinical studies are often divided into three phases. Phase I and phase II studies often consist of a small number of healthy subjects and patients with the purpose of exploring the treatment's properties and possible risks. Phase III studies involve large numbers of healthy subjects and patients with the purpose of validating the safety and effect of the treatment in a confirmatory way before entering the market.

1.2 Design of clinical studies

Before conducting a study, there are many factors to consider. An important question is how to design the clinical study. Design aspects can for example be the number of patients in the study, study length, dose range and the number of measurements. The design of a clinical study has great impact on the results and on the statistical inference that can be drawn on the obtained data. For example, if there are too few patients

involved the results obtained are associated with high uncertainty. On the other hand, having an excessive number of patients is associated with a large cost. There can also be a problem of recruiting enough patients to the study.

1.3 Simulation of clinical studies

By using computer aided simulations, a clinical study can be simulated prior to the real life execution. The simulations serve as a tool to evaluate a specific study design and investigate how the data may vary within the considered population. However, this is a time consuming approach. A supplement to the simulation approach is to use a software dedicated to optimal design of clinical trials. One example of such software is PFIM[1]. By using an underlying model describing the data to be collected PFIM can be used to evaluate and optimize clinical studies in order to improve parameter estimation precision when analyzing the generated data from the clinical study.

1.4 Thesis purpose and outline

The purpose of the thesis is to evaluate a software called PFIM used for optimal design of clinical studies. In Chapter 2 the idea of model based drug development is first described with focus on studies evaluating a drug's pharmacokinetic (PK) and pharmacodynamic (PD) properties. In Chapter 3 the mathematical theory needed for the thesis is presented. The mathematical section is brief since the purpose of the thesis is to evaluate the software rather then explore the underlying theory. In Chapter 4 the features and theory of PFIM are presented. In Chapters 5-7 optimal design is applied for three different models to investigate how current study designs perform and can be improved. The first model is a simple PK/PD model which serves the purpose of investigating the appropriateness of the model approximation in PFIM. The second model is a model for glucose-insulin response where it was shown an initial suggested design can be reduced to decrease the number of samples needed in the study. The third model tested is a more complex PK/PD model describing the adverse neutrophil response to drugs as a safety marker in a 6 months long clinical study. In Chapter 8 the results are discussed and conclusions are stated.

2

Model based drug development

In many industries, modeling and simulation are used on a regular basis. By using computer aided simulations the underlying model describing the observed data can better be understood and expensive real life experiments can be avoided or more efficiently conducted. However, in drug development the use of modeling and simulation is limited. In the article by Lalonde *et al.* [2] the authors argue that model based drug development (MBDD) could improve the design of clinical trials and reduce costs.

2.1 Modeling in drug development

Drug development is a highly expensive process. Less than 10% of new compounds that enter clinical trials ultimately make it to the market[2]. By using appropriate models and statistical analysis the clinical trials and conclusions made from them can be improved. This could lead to a better assessment of a drug candidate's probability of success. Probability of success refers to the probability that a drug actually will show the desired effect and the Phase III study is successful. Some of the key components in MBDD are for example pharmacokinetic and pharmacodynamic models, competitor drug comparisons based on literature data, design considerations and quantitative decision criteria.

By having appropriate models and simulation methods, decision making can be improved by combining information from different areas such as multiple response variables, dose levels, patient populations, different drugs, across studies and clinical endpoints. A clinical endpoint refers to the target outcome of the trial, for example the reduction in mortality or the change in blood pressure after treatment. Using these underlying models computer aided simulations can be used for simulating studies to investigate what would be expected in a real life study.

Two major components in model based drug development are the modeling of what happens to the drug in the body (pharmacokinetics) and what effect (therapeutic and adverse) the drug has over time (pharmacodynamics).

2.2 Pharmacokinetics and pharmacodynamics

In many clinical trials one purpose is to estimate the pharmacokinetic and pharmacodynamic properties of a drug. This include the time courses of drug concentration and associated effects and the relationship between them.

Pharmacokinetics

Pharmacokinetics (PK) is commonly referred to as the study of what the body does to the drug. It provides a mathematical basis to describe the propagation of the drug in the body. The different phases after a drug is entering the body consist of absorption, distribution, metabolism and excretion (ADME). Metabolism and excretion are often jointly named elimination.

To model the pharmacokinetics of a drug compartmental models are often used. These models are hypothetical structures that can describe the time course of a drug in the body with an ordinary differential equation system.

The simplest compartmental model is a one-compartment model where the body is modeled as a single, kinetically homogeneous unit. The model has two parameters, volume of distribution V and clearance Cl. The volume of distribution is defined as the theoretical volume in which the total amount of drug would need to be uniformly distributed to produce the desired blood concentration of a drug. The clearance has a unit of $\frac{\text{volume}}{\text{time}}$ and is a measure of the volume cleared of drug every time unit. An illustration of a one-compartment model with volume of distribution V and clearance Cl is seen in Figure 2.1.

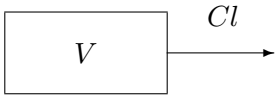

Figure 2.1: One-compartment model with volume of distribution V and clearance Cl.

A common reparametrization of the model is to use the rate $k = \frac{Cl}{V}$, where k has unit $\frac{1}{\text{time}}$. Consider the simplest case where the drug with initial dose D is taken up by the body after an intravenous bolus injection at time $t = 0$. Bolus injection implies that the drug is immediately taken up by the body at the admission time. The expression for the concentration of the drug at time t is

$$C(t) = \frac{D}{V} e^{-kt}.$$

The solution is obtained by setting up the ordinary differential equation for the system. Let A denote the amount of drug in the compartment. Then

$$\frac{dA}{dt} = -kA,$$

with initial condition $A(0) = D$, which gives $A(t) = De^{-kt}$. To obtain the concentration the amount of drug in the compartment is scaled by the volume of distribution, which gives

$$C(t) = \frac{D}{V}e^{-kt} = \frac{D}{V}e^{-\frac{Cl}{V}t}.$$

By extending the one-compartment model to a two-compartment or multi-compartment model more complex phenomena can be modelled. A two-compartment model consists of a central compartment (e.g. liver) and one peripheral compartment (e.g. muscles or skin). It is important to realize that the compartments have no physiological meaning. Instead each of them is a simplified model for the biological processes in the body. In Figure 2.2 a two-compartment model with volumes of distribution V_1 and V_2, inter compartmental rates k_{12}, k_{21} and elimination rate k_e is shown.

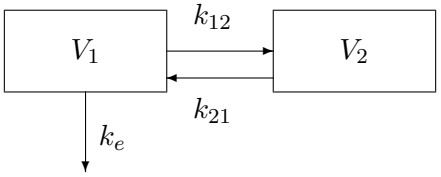

Figure 2.2: A two-compartment model with the two volumes of distribution V_1 and V_2, inter compartmental rates k_{12}, k_{21} and elimination rate k_e.

Let A_1 and A_2 denote the amount of drug in compartment 1 and 2. By using the mass conversion principle the ordinary differential equation system describing the kinetics is

$$\begin{aligned}\frac{dA_1}{dt} &= -k_e A_1 - k_{12}A_1 + k_{21}A_2 \\ \frac{dA_2}{dt} &= k_{12}A_1 - k_{21}A_2.\end{aligned}$$

The concentration in each compartment is obtained by scaling with the volumes of distribution. For the analytic solution to the system, see [3].

By using appropriate compartmental models different drug properties can be described, including different routes of administration and non-linearities in absorption and/or elimination.

Pharmacodynamics

Pharmacodynamics (PD) is often referred to as what the drug does to the body. Here the main question is what effect a certain drug has, for example on blood pressure or tumor size, over time. The effect is dependent of the concentration of the drug. This

relation is often described by a nonlinear, sigmoid-like model. The E_{max} model describes the effect E as a function of dose or concentration C

$$E = \frac{E_{max}C}{C + C_{50}}.$$

Here E_{max} is the maximal effect and C_{50} is the concentration required to achieve half of the maximal effect. An extension of the E_{max} model is the sigmoid E_{max} model where a sigmoid factor γ has been added to the model. That is

$$E = \frac{E_{max}C^{\gamma}}{C^{\gamma} + C_{50}^{\gamma}}.$$

The E_{max} model is applicable to a response variable which increases with drug concentration. However, there are also drugs causing a reduction in the response variable. The model for such behavior is an inhibitory model called the I_{max} model

$$E = 1 - \frac{I_{max}C}{C + C_{50}}.$$

Here I_{max} is the maximal inhibitory effect and C_{50} is the concentration needed to reach half of the maximal effect.

The effect can either act directly or indirectly (time delay between maximal exposure and maximal response) on the response. One example of an indirect response model is the so called turnover response model. In the turnover response model the drug is acting on the production/synthesis rate k_{in} of the response variable or on the elimination/removal rate k_{out} describing the response R (often a biomarker) as illustrated in Figure 2.3.

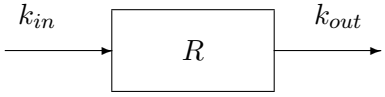

Figure 2.3: A turnover response model with rates k_{in} and k_{out} describing the response R.

The system is described by the ordinary differential equation $\frac{dR}{dt} = k_{in} - k_{out}R$. When no drug is present the solution to the system is $R(t) = \frac{k_{in}}{k_{out}}$. When a drug is present it acts either on k_{in} or k_{out}, leading to a decrease or increase in input or output rate respectively and subsequently a change in the response variable itself. For more information about the indirect response models, see [4].

Joint PK/PD modeling

Having pharmacokinetic and pharmacodynamic models in place it is easy to combine these into a joint PK/PD-model. Hence the PK model describes the drug concentration over time and the PD model the response of the drug. This combined model relates dose and exposure to the response of the drug. This type of modeling is frequently used in

clinical trials where the objective can for example be to select an optimal dose. The goal is to find a dose that is high enough to give a significant desired effect but not higher than needed to avoid negative effects on the patient.

3

Mathematical theory

In this chapter the mathematical theory needed for the thesis is presented. The population modeling approach is described and some important results in optimal design theory is stated.

3.1 Population modeling

In clinical trials, samples are collected from patients to describe markers of drug exposure and response, such as blood pressure or glucose levels. The response is time dependent and often behaves as a nonlinear function of the model parameters. In a population model the parameters may vary between individuals in the population to describe different sources of variability. Examples of parameters that can vary between individuals due covariate patient specific factors (such as body weight, sex, disease status and concomitant medication) are the volume of distribution, clearance and the maximal drug effect. The goal of population modeling is to describe the population parameters and how they vary within the population. Appropriate mathematical models for such modeling are nonlinear mixed effect models.

Nonlinear mixed effect models

The notation of [5] is used. Let y_{ij} denote the jth measurement of the response (drug concentration or drug response) of the ith individual. The measurement is done under condition t_{ij} and possible additional conditions which are denoted $\mathbf{u}_i$. Often t_{ij} denotes time and $\mathbf{u}_i$ is empty. Denote $\mathbf{x}_{ij} = (t_{ij}, \mathbf{u}_{ij})$. For each individual i a vector of characteristics (or covariates) $\mathbf{a}_i$ is defined. It is assumed that the triplet $(\mathbf{y}_{ij},\ \mathbf{u}_i,\ \mathbf{a}_i)$ is independent across i. The model for the response y_{ij} is

$$y_{ij} = f(\mathbf{x}_{ij},\beta_i) + \epsilon_{ij}.$$

Here $f(\mathbf{x}_{ij},\beta_i)$ is a function describing the behavior of the response and $\beta_i \in \mathbb{R}^p$ is a vector of individual parameters for individual i. The individual parameters β_i is related to the so called fixed effects (or population parameters) β by a function

$$\beta_i = g(\mathbf{a}_i,\beta,\mathbf{b}_i)$$

Here $\mathbf{b}_i \sim N(0,\Omega)$ is a vector of so called random effects of individual i. Hence the 'mixed' term in the nonlinear mixed effect model comes from the fact that the model incorporates both fixed (population) and random (individual) parameters. The function $g(\mathbf{a}_i,\beta,\mathbf{b}_i)$ explains how the fixed parameters β varies between individuals $i = 1,\ldots,m$. This is often referred to as inter individual or between-subject variability. Often g is chosen to be of exponential form, $g(\mathbf{a}_i,\beta,\mathbf{b}_i) = \beta e^{\mathbf{b}_i}$, which describes a lognormal distribution of the population parameters. In this thesis only log-normal distributions of parameters are considered. This is a common case in medicine, where the parameters often show a skewed distribution[6].

The model for the errors is $\epsilon_{ij} \sim N(0,\Sigma)$ where Σ can be a general covariance structure, that may depend on values of the response. In this thesis the model is restricted to the form

$$\Sigma = (\sigma_{add} + \sigma_{prop} f(\mathbf{x}_{ij},\beta_i))^2, \tag{3.1}$$

where σ_{add} and σ_{prop} denote the additive and proportional components of the error. When $\sigma_{add} \neq 0$ and $\sigma_{prop} = 0$ the error model is called a homoscedastic model, when $\sigma_{add} = 0$ and $\sigma_{prop} \neq 0$ the model is called heteroscedastic and when $\sigma_{add} \neq 0$, $\sigma_{prop} \neq 0$ the variance model is a combined model. Measurements in medicine often show that the variance varies with the response, and hence a combined error model for the errors is used. For more general variance models, see [7].

3.2 Optimal experimental design

Optimal experimental design deals with constructing experiments in a way that is optimal in some sense. One example of an experiment is a clinical study. To make correct conclusions from the study it is important to be able to estimate the parameters of the mathematical model used for evaluating the data with high precision. Optimal design can be used to decrease the variance of the estimated parameters which can in turn lead to improved decision making. That is, the design of the study should be done in such a way that the study gives as much information as possible. The term 'information' here is general, and will be specified later on.

The theory outlined in this chapter is a very compact description of the theory that is used in this thesis. For the interested reader, see [8].

Design of experiments

Optimal experimental design depends on the model for the data $f(t,\theta)$, the optimality criterion Ψ and the the design region $\mathcal{X}$. A design is a template how to conduct an

experiment. It specifies how many observations that should be collected and at which time points. Assume that the design specifies r_k observations at time point t_k and let N_{tot} be the total number of observations, $N_{tot} = \sum_{k=1}^{n} r_k$. A design ξ is defined to be a set of time points and weights, that is

$$\xi = \begin{Bmatrix} t_1 & t_2 & \dots & t_n \\ w_1 & w_2 & \dots & w_n \end{Bmatrix},$$

where $w_k = r_K/N_{tot}$. In exact design the weights w_k, $k = 1, \dots, N$, are a ratio of integers while in so called continuous design this requirement is dropped. Continuous designs have an advantage since impact of N can be omitted and the methods of calculus can be used to describe the properties of optimal designs. However, any practical design will be exact [9].

Measuring information

In this thesis the underlying model is assumed to be a nonlinear mixed effect model. For an introduction to linear and nonlinear models in design theory, see [8].

Moreover, the observed data is assumed to come from a probability distribution $f_X(x)$. The key to measure information in such models is the so called Fisher information. It is a way of measuring the amount of information that an observable random variable X carries about an unknown parameter $\theta \in \mathbb{R}^p$ upon which the probability of X depends. When the model only contains one parameter ($\theta \in \mathbb{R}^1$) the Fisher information is a scalar. When $\theta \in \mathbb{R}^p$ with $p > 1$, the Fisher information is a matrix defined

$$M_{ij}(\theta) = E\left[\left(\frac{\partial}{\partial\theta_i} \ln f_X(x;\theta)\right)\left(\frac{\partial}{\partial\theta_j} \ln f_X(x;\theta)\right)\right], 1 \leq i,j \leq p$$

Here $\ln f_X(x,\theta)$ denotes the log likelihood for the random variable X. From the definition it is clear that M is a symmetric $p \times p$ matrix. The essence of Fisher information is that it i closely linked to the variance of an unbiased estimator by the Cramér-Rao bound (CRB).

Theorem 1. (Cramér-Rao bound) *Suppose $\theta \in \mathbb{R}^p$ is an unknown deterministic parameter which is to be estimated from measurements x, distributed according to some probability density function $f(x;\theta)$. The covariance of any unbiased estimator $\hat{\theta}$ of θ is then bounded by the inverse of the Fisher information $M(\theta)$. That is, $Cov(\hat{\theta}) \geq M(\theta)^{-1}$.*

The Cramér-Rao bound relates the covariance of the estimated parameter to the Fisher information. Hence minimizing the variance is equivalent to maximizing the Fisher information. An estimator for which equality is fulfilled is called a efficient estimator.

The general equivalence theorem

In continuous designs the goal is to minimize some general function of imprecision $\Psi\{M(\xi)\}$. One example is D-optimality, which will be discussed later, where $\Psi\{M(\xi)\} =$

$\log|M^{-1}(\xi)|$, where $M(\xi)$ is the information matrix for the model. The notation $M(\xi)$ is used to stress the fact that the information matrix depends on the underlying design. See [8] for more information.

The most important theorem in optimal design theory is the so called General Equivalence Theorem, which will be stated. It can be seen as a consequence of derivatives being equal to zero at the minimum of a smooth function over an unconstrained region. Let the measure $\bar{\xi}$ put unit mass at the point x and define the measure ξ' by

$$\xi' = (1-\alpha)\xi + \alpha\bar{\xi}$$

Then

$$M(\xi') = (1-\alpha)M(\xi) + \alpha M(\bar{\xi})$$

The derivative of Ψ in the direction $\bar{\xi}$ is defined [8]

$$\phi(x,\xi) = \lim_{\alpha\to 0}\frac{1}{\alpha}\Big(\Psi\{(1-\alpha)M(\xi) + \alpha M(\bar{\xi})\} - \Psi\{M(\xi)\}\Big) \tag{3.2}$$

A major result in optimal design theory is now stated.

Theorem 2. (The General Equivalence Theorem) *The following three conditions on ξ^* are equivalent.*

- *The design ξ^* minimizes $\Psi\{M(\xi)\}$.*
- *The design ξ^* maximizes the minimum over $\mathcal{X}$ of $\phi(x,\xi)$.*
- *The minimum over $\mathcal{X}$ of $\phi(x,\xi^*) = 0$, this minimum occurring at the points of support of the design.*

As a consequence of the third condition

- *For any non-optimum design ξ the minimum over $\mathcal{X}$ of $\phi(x,\xi) < 0$*

Hence the theorem can provide different criteria for checking if a design is optimal.

D-optimality

In D-optimality the optimality criterion is

$$\Psi\{M(\xi)\} = \log|M^{-1}(\xi)|.$$

D-optimality is the most common objective in optimal design. Some of the most important characteristics of D-optimality are stated.

Let $\lambda_1,\ldots,\lambda_p$ denote the eigenvalues of the information matrix $M(\xi)$. The eigenvalues of $M^{-1}(\xi)$ is then $\frac{1}{\lambda_1},\ldots,\frac{1}{\lambda_p}$. D-optimality corresponds to minimizing the content of the confidence region for the estimated parameters. The confidence region of the estimated parameters is a ellipsoid where the eigenvalues of $M^{-1}(\xi)$ are proportional to

the squares of the lengths of the axes of the confidence ellipsoid. D-optimality minimizes the so called generalized variance of the parameters estimates.[8] That is,

$$\min \prod_{i=1}^{p} \frac{1}{\lambda_i},$$

where λ_i denotes the ith eigenvalue of $M(\xi)$.

By using the definition of the derivative $\phi(x,\xi)$ it can be shown that $\phi(x,\xi) = p - d(x,\xi)$, where $d(x,\xi)$ denotes the standardized variance $d(x,\xi) = f^T(x)M^{-1}(\xi)f(x)$ and $f(x)$ is the vector of partial derivatives of the parameters. Using the third condition in The General Equivalence Theorem it follows that for D-optimality it holds that

$$d(x,\xi) \leq p,$$

where p denotes the number of parameters in the model.

4

The PFIM software

A popular approach for evaluating the appropriateness of a specific study design is by simulations. However, this can be a time consuming approach. An alternative is provided by software for optimal design. PFIM[1] is a software for R[10] dedicated to design evaluation and optimization using multiple response models specifically developed for use in drug development. The program can either be obtained as a package or as an interface tool. In this chapter the models, methods and notation in PFIM are explained.

4.1 PK/PD Models in PFIM

PFIM comes with a rich library of pharmacokinetic and pharmacodynamic models. The pharmacokinetic models support one- two- and three compartment models with linear elimination. The pharmacodynamic part supports the most common models, e.g. E_{max}, I_{max} and turnover response models.[3]

PFIM supports both analytical models or models described by an ordinary differential equation system. Besides the models in the PFIM library, the users can also define their own models either by analytical expressions or ODE systems. The models are assumed to be nonlinear mixed effect models where the individual parameters are either normal or log-normal distributed. There is also an option to include inter-occasion variability, where individual parameters are allowed to vary with time.

The error model is restricted to a combined error model, as described in Equation (3.1), where the user has to specify the additive and proportional component.

4.2 Model definition and evaluation

Having a model for the underlying response, PFIM can be used to evaluate a specific design based on the Fisher information matrix. However, since the model is a nonlinear mixed effect model the Fisher information matrix has to be extended to the so called

population Fisher information matrix (the name PFIM comes from this). The notation of [1] and [11] is used. A single response model is used, but the notation can easily be extended to a multiple response model.

Let ξ_i denote the design for individual i, which is referred to as an elementary design. A population design Ξ is composed of N individuals, where each individual have an individual elementary design ξ_i. Hence the population design is a collection of N elementary designs which is denoted

$$\Xi = \{\xi_1, \ldots, \xi_N\}.$$

If more than one individual has the same elementary design ξ they are said to belong to the same group. Often a population design consists of Q groups, where each group has an elementary design ξ_q, $q = 1, \ldots, Q$.

Assuming y_i is the n_i-vector of observations in the ith individual ($i = 1, \ldots, N$) the model for response is

$$y_i = F(\beta_i,\xi_i) + \epsilon_i,$$

where ξ_i is the design for individual i. Furthermore it is assumed that $\epsilon_i \in N(0,\Sigma_i)$, where Σ_i is defined as a combined error model previously defined with parameters σ_{add} and σ_{prop}. β_i is the vector of individual parameters which is related to the population parameters β by a function $\beta_i = g(\beta,b_i)$. Here b_i is the random effects for individual i with $b_i \in N(0,\Omega)$. Ω is a $p \times p$ diagonal matrix where each diagonal element w_k^2 represents the variance of the kth parameter. Let Ψ denote the vector of all the parameter to be estimated. That is, $\Psi^T = (\beta^T,\omega_1^2, \ldots, \omega_p^2, \sigma_{add}, \sigma_{prop})$. Let λ be the vector of variance terms, $\lambda^T = (\omega_1^2, \ldots, \omega_p^2, \sigma_{add}, \sigma_{slope})$, so that $\Psi^T = (\beta^T,\lambda^T)$.

The population Fisher information matrix is given by

$$\mathbf{M}(\Psi,\Xi) = \mathrm{E}\left(- \frac{\partial^2 \mathbf{L}(\Psi;Y)}{\partial\Psi\partial\Psi^T} \right).$$

Here Ξ denotes the population design and $\mathbf{L}(\Psi;Y)$ the log-likelihood of the vector of all the observations Y for the population parameters Ψ. Assuming independence among the individuals, the population Fisher information matrix is rewritten as a sum of individual Fisher information matrices

$$\mathbf{M}(\Psi,\Xi) = \sum_{i=1}^{N} \mathbf{M}(\Psi,\xi_i). \tag{4.1}$$

Since F is nonlinear in the parameters, no explicit expression for the likelihood or Fisher information matrix can be obtained. To overcome the nonlinearity problem, PFIM uses a linearization of the nonlinear model at the expected value of the random effects b_i. This gives, see for example [12],

$$\mathbf{M}(\Psi,\xi_i) \approx \frac{1}{2} \begin{pmatrix} A(E_i,V_i) & C(E_i,V_i) \\ C^T(E_i,V_i) & B(E_i,V_i) \end{pmatrix},$$

where

$$(A(E_i,V_i))_{ml} = 2\frac{\partial E_i^T}{\partial \beta_m}V_i^{-1}\frac{\partial E_i}{\partial \beta_l} + tr(\frac{\partial V_i}{\partial \beta_l}V_i^{-1}\frac{\partial V_i}{\partial \beta_m}V_i^{-1}),$$
$$(B(E_i,V_i))_{ml} = tr(\frac{\partial V_i}{\partial \lambda_m}V_i^{-1}\frac{\partial V_i}{\partial \lambda_l}V_i^{-1}),$$
$$(C(E_i,V_i))_{ml} = tr(\frac{\partial V_i}{\partial \lambda_l}V_i^{-1}\frac{\partial V_i}{\partial \beta_m}V_i^{-1}),$$

with

$$E_i = F(g(\beta,0),\xi_i),$$
$$V_i = \left(\frac{\partial F(g(\beta,0),\xi_i)}{\partial b_i}\right)\Omega\left(\frac{\partial F(g(\beta,0),\xi_i)}{\partial b_i}\right) + \Sigma_i.$$

Having an explicit expression of the Fisher information matrix the expected standard errors of the estimated parameters are calculated by taking the inverse of the information matrix.

4.3 Design optimization

There are two algorithms that can be used for optimization of designs in PFIM. The first one is the Simplex method which is a continuous optimization method. It only needs specifications of how many measurements should be done and the time interval of allowed times. The second method is the so called Fedorov-Wynn method (FW), which is the method used throughout this thesis. The reader is referred to [13] for a comparison between the two methods.

Fedorov-Wynn algorithm

The FW algorithm is an iterative algorithm that maximizes the determinant of the Fisher information matrix within a finite set of possible designs. Prior to the optimization the user has to specify a set of allowed time points and how many measurements that should be done. The user also has a specify a initial guess Ξ_0. From the set of allowed times a set ζ of so called elementary designs ξ_i is created. Let n_Ψ denote the number of parameters in the model and $\phi(\Xi_k,\xi_i)$ the derivative defined in Equation (3.2). The algorithm relies on the Equivalence Theorem (2) and is as follows

- Start with an initial guess Ξ_0
- at step k, with the current design being Ξ_k, find $\xi^* = \arg\max_{\xi_i \text{ in } \zeta} \phi(\Xi_k,\xi_i)$. If $\phi(\Xi_k,\xi_i) \leq n_\Psi + \epsilon$ terminate, where ϵ is a predetermined tolerance.
- Otherwise, update the design to $\Xi_{k+1} = (1-\alpha^*)\Xi_k + \alpha^*\xi^*$, where $\alpha^* \in [0,1]$ is chosen to maximize the determinant of the current Fisher information matrix.

The Fedorov-Wynn algorithm has many advantages. Since it requires the user to specify which time points that are allowed, clinical constraints can easily be incorporated in the optimization. The user can decide the number of samples per patient and whether the total cost of the study should be measured by the number of patients or the number of samples.

However, it should be noted that the number of elementary designs grow very fast. If the number of allowed time points is n and the design should consist of k points where $k \leq n$ the number of elementary designs is $\binom{n}{k}$. For all these elementary designs the Fisher information matrix has to be calculated prior to the optimization.

4.4 Evaluating the software

In the following three chapters the PFIM software is evaluated using three different models. In Chapter 5 a design for an analytical PK/PD model describing therapeutic drug response is evaluated with the purpose of evaluating the appropriateness of the linearization of the Fisher information matrix. This is done by comparing the relative standard errors for the parameter estimates given by PFIM with those obtained from a large simulation.

In Chapter 6 PFIM is applied to a model describing glucose-insulin response after glucose administration in absence of drug therapy. An initial design is evaluated, optimized and reduced to decrease the number of total samples needed in the study.

The model in Chapter 7 is a more complex PK/PD model describing the adverse neutrophil response during long term therapy with a candidate drug in development.

5

Investigating the linearization of the Fisher information matrix

In this chapter the appropriateness of the linearization of the Fisher information matrix in PFIM is analyzed. Using a drug concentration-response model, the expected standard errors given by PFIM are compared with those given by a large simulation. The model, study design and the parameter values are predefined. No optimization of the design is done here, instead the focus is on the differences in predicted errors for the two methods. The outline in [12] is followed, but with a slightly different underlying model and different parameter values.

5.1 The drug concentration-response model

The pharmacokinetic model is a one compartment model with bolus injection and first-order linear elimination. Assuming the initial dose is equal to 1 mg, the concentration at time t is given by

$$C(t) = \frac{1}{V} e^{-\frac{Cl}{V}t},$$

with clearance Cl (unit Litres/h) and volume of distribution V (unit Litres). The pharmacodynamic model is an E_{max} model given by

$$E(t) = \frac{E_{max} C(t)}{C_{50} + C(t)},$$

with parameters E_{max} and C_{50}.

The random effects are assumed to be of exponential form for both the responses. Moreover, a proportional error model is assumed for both the responses. The vector of parameters Ψ is then

$$\Psi^T = (Cl, V, E_{max}, C_{50}, \omega^2_{Cl}, \omega^2_V, \omega^2_{E_{max}}, \omega^2_{C_{50}}, \sigma_{propPK}, \sigma_{propPD})$$

5.2 The initial design

The parameters used in the model are seen in Table 5.1.

Table 5.1: Parameter values used in the drug concentration-response model.

Cl (L/h)	V (L)	E_{max}	C_{50} (mg/L)	ω^2_{Cl}
4	8	0.5	0.05	0.09
ω^2_V	$\omega^2_{E_{max}}$	$\omega^2_{C_{50}}$	σ_{propPK}	σ_{propPD}
0.09	0.16	0.09	0.1	0.1

The population design consists of 30 patients where all the individuals were sampled at the same times for the two responses. That is $\xi_{PK} = \xi_{PD} = (0.5,1.0,3.0,7.0,10.0,15.0)$, where the values denote hours post dose. The expected PK and PD response together with the sampling times are seen in Figure 5.1.

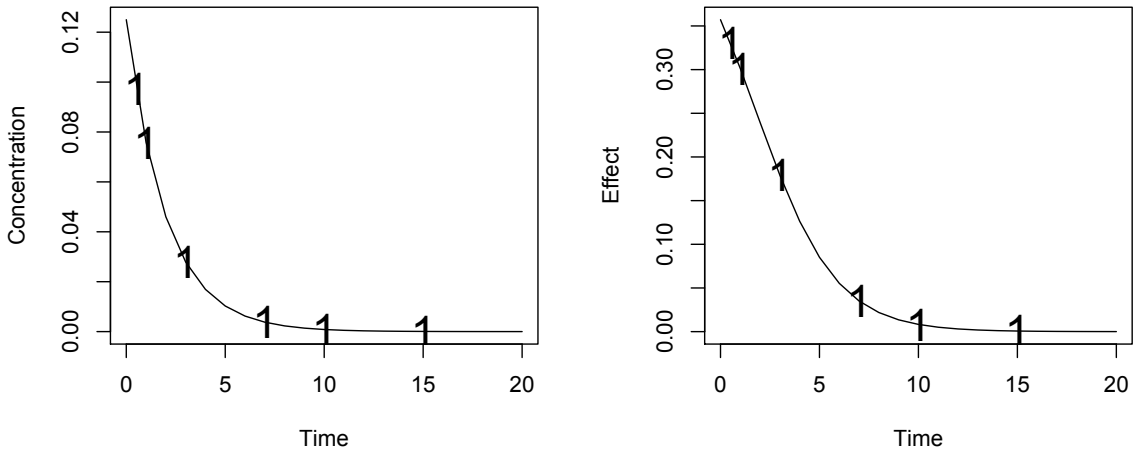

Figure 5.1: Expected responses for the PK (left figure) and PD (right figure) together with the sampling times. The plot is automatically created by PFIM.

5.3 Standard error of parameter estimates

The expected standard errors given by PFIM were calculated using the linearization of the Fisher information matrix as described in Chapter 4. To obtain the empirical standard errors the study was simulated a total of 500 times in the NONMEM[14] software with population parameters estimated in each run. The algorithm used for estimation was the SAEM algorithm, which is a stochastic version of the EM-algorithm. It uses no approximation of the underlying model and can hence be seen as the 'real' standard errors. The relative standard errors (RSE), given in %, from PFIM and SAEM are seen in Table 5.2. The relative standard error is calculated by taking the standard deviation of the parameter estimate and divide by the parameter value.

Table 5.2: Relative standard errors (%) obtained from PFIM and SAEM in NONMEM.

	Parameters									
Method	Cl	V	E_{max}	C_{50}	ω^2_{Cl}	ω^2_V	$\omega^2_{E_{max}}$	$\omega^2_{C_{50}}$	σ_{propPK}	σ_{propPD}
PFIM	3.7	3.7	4.8	4.5	28.0	27.5	29.2	36.6	7.4	7.5
SAEM	5.0	5.2	8.4	6.6	27.8	29.7	29.0	38.8	6.3	6.6

5.4 Discussion

Using the linearization of the Fisher information matrix to calculate the expected standard errors seems like an appropriate method. From Table 5.2 one can see that for most of the parameters the RSE from PFIM was very close to the 'real' RSE given by the SAEM method in NONMEM. The biggest deviation from the true RSE was on the E_{max} parameter, where the difference was 3.6 %.

The calculation of the standard errors in PFIM took only a few seconds while simulating the study in NONMEM to obtain the empirical standard errors took a few hours. The approximation done in PFIM seems to be appropriate for analytical PK/PD models and offers a very efficient approach to evaluate a study design.

6

Insulin-glucose response model

Type 2 diabetes mellitus is a chronic disease which results in a large number of abnormalities in insulin dependent metabolism. There are two factors which play an important role in the glucose disposal. Insulin sensitivity measures the capability of insulin to increase glucose disposal into muscles. Glucose effectiveness measures the ability of glucose to enhance its own disposal at basal insulin level. With any of these factors failing, one may be in the risk zone of impaired glucose tolerance or diabetes. Bergman's minimal model can be used to quantatively measure these factors.[15]

In this chapter this insulin-glucose response model is used to illustrate how PFIM can be used to evaluate and optimize a design to increase the parameter estimation precision. The initial design is also reduced to allow fewer measurement per patient while maintaining the information in the study. Insulin-glucose regulation models have been the topic of two previous master's thesis at AstraZeneca. The interested reader is referred to [16] and [17].

6.1 The model

In Figure 6.1 the underlying compartment model describing the insulin-glucose dynamics is shown.

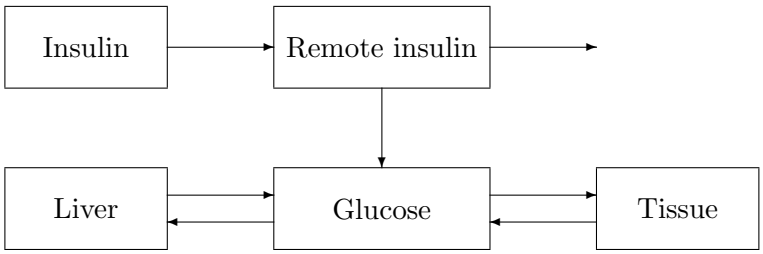

Figure 6.1: Compartment model of Bergman's minimal model describing the dynamics between insulin and glucose.

In an intravenous glucose tolerance test each patient is given an intravenous glucose injection which forces the pancreas β-cells to secrete insulin. This leads to an increase in plasma insulin which increases the uptake of glucose into muscles and tissue. The glucose uptake depends on insulin levels in a second so called remote compartment, describing the time delay in the system. When glucose is taken up by the muscles and tissue the glucose concentration in plasma decreases, which leads to a decrease in β-cell insulin secretion. This implies that there is a feedback effect in the system.[15] The glucose-insulin relationship is described by the system

$$\begin{aligned}
\frac{dG}{dt} &= S_g G_b - (S_g + X(t))G(t), G(0) = G_0 \\
\frac{dX}{dt} &= -p_2(X(t) - S_i(I(t) - I_b)), X(0) = 0 \\
\frac{dI}{dt} &= -n(I(t) - I_b) + \gamma(G(t) - h)^+ t, I(0) = I_0,
\end{aligned}$$

where $G(t)$ is the glucose concentration in plasma, $X(t)$ is remote insulin in action and $I(t)$ is the insulin concentration in plasma. The parameters in the model are described in Table 6.1.

Table 6.1: Parameters in Bergman's minimal model.

S_g:	Glucose effectiveness (min^{-1}).
G_b:	The baseline preinjection level of glucose (mg/dl).
I_b:	The baseline preinjection level of insulin (U/dl).
p_2:	Rate of decrease in tissue glucose uptake ability (min^{-1}).
S_i:	Insulin sensitivity ($min^{-1}(U/ml)^{-1}$).
n:	First order decay rate for insulin in plasma (min^{-1}).
h:	Threshold value of glucose above which the pancreatic cells secrete insulin (mg/dl).
γ:	Rate of pancreatic cells release rate of insulin ($U/ml\ min^{-2}(mg/dl)^{-1}$).
G_0:	The glucose concentration in plasma at time 0 (mg/dl).
I_0:	The glucose concentration in plasma at time 0 (mg/dl).

The parameter values and the variance components are taken from the thesis by Braukovic[17] and are seen in Table 6.2.

Table 6.2: Population parameters used in the model.

S_g	p_2	S_i	n	h	γ	G_0	I_0
0.026	0.025	0.0005	0.27	83	0.001	280	360
G_b	I_b	$\omega^2_{S_g}$	$\omega^2_{S_i}$	ω^2_{η}	ω^2_{γ}	$\omega^2_{G_0}$	$\omega^2_{I_0}$
92	11	0.25	0.25	0.09	0.09	0.01	0.01

It is assumed that G_b and I_b are fixed and are not to be estimated even though there may be some variability also between non-diabetic patients. Furthermore, it is assumed that the residual error is proportional with a coefficient $\sigma_{prop} = 0.02$ for both $G(t)$ and $I(t)$.

6.2 The initial design

To estimate the model parameters data from an intravenous glucose tolerance test was used. The initial design consisted of 30 patients with 13 measurements per patient. The initial design is denoted

$$\Xi_{init} = \{(0, 3, 4, 5, 7, 10, 15, 20, 25, 30, 60, 115, 120), 30\},$$

where the values denote minutes post glucose injection. Samples of G(t) and I(t) are taken to estimate the parameters of the model. It is assumed that the two samples are taken at the same time points. In Figure 6.2 the expected glucose and insulin response are shown together with the initial sampling times.

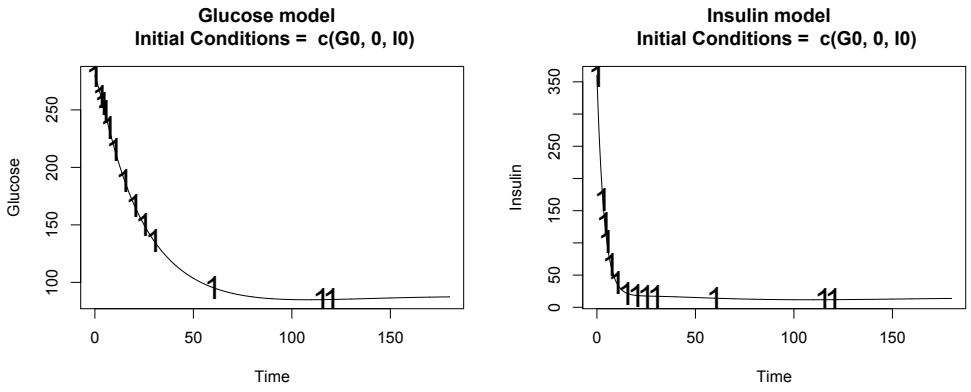

Figure 6.2: The expected glucose and insulin response together with the initial sampling times. The plot is automatically created by PFIM.

One important question in design evaluation is the cost of the design and how to measure the cost. In the glucose tolerance test the cost is the total number of samples in the study (since individual rather than population level parameter estimates are of primary interest). This allows for the possibility to reduce the number of measurements and meanwhile increase the number of patients in the study. Using more patients in the study can lead to better estimations of the random effects parameters in the model. However, it can also be a problem to include a larger number of patients (cost and logistics). Using the number of a samples as a cost the total cost of Ξ_{init} is 780 samples.

Using PFIM the initial design can easily be evaluated. The performance of specific design Ξ is measured by the objective

$$\Phi(\Xi) = (\det M(\Xi))^{1/p},$$

where $M(\Xi)$ is the Fisher information matrix for the design Ξ and p is the number of parameters in the model. As described earlier PFIM estimates the expected standard errors of the parameters for a specific design by approximating the Fisher information matrix. The objective for the initial design was

$$\Phi(\Xi_{init}) = 20848.87,$$

which serves as a reference value when creating new designs.

6.3 Improving the initial design

Here two ways of improving the initial design are considered. First 30 patients are enrolled in the study and every patient is measured 13 times, just as in the initial design, but the sampling time points are then optimized to achieve a better design. In the second proposal fewer measurements than 13 per subject are allowed while maintaining a total of 780 samples in the study, as in the initial design.

Equal number of measurements per patient

By choosing different sampling points the estimation of the model parameters could possibly be improved. It is assumed that there are a total of 20 allowed sampling points. These are the initial 13 points from Ξ_{init} plus an additional of 7 points, which gives the set of allowed sampling points

$$T = \{0, 3, 4, 5, 7, 10, 15, 20, 25, 30, 60, 70, 90, 100, 115, 120, 140, 160, 170, 180\}.$$

From this set exactly 13 time points are chosen for each patient. Hence there are a total of $\binom{20}{13} = 77520$ elementary designs. The optimal design among these time points was obtained using the Fedorov-Wynn algorithm in PFIM and was

$$\Xi_{imp,13} = \left\{ \begin{matrix} (0, 3, 4, 10, 15, 20, 25, 30, 60, 70, 100, 115, 120), 27 \\ (0, 3, 4, 10, 15, 25, 30, 60, 70, 90, 100, 115, 120), 3 \end{matrix} \right\}.$$

The design should be interpreted as 27 subjects having the sampling times (0, 3, 4, 10, 15, 20, 25, 30, 60, 70, 100, 115, 120) and 3 subjects having the sampling times (0, 3, 4, 10, 15, 25, 30, 60, 70, 90, 100, 115, 120). The criterion value for this design was

$$\Phi(\Xi_{imp,13}) = 30637.01,$$

which corresponds to a 47% gain in information compared to the initial 13-point design Ξ_{init}. In Section 6.5 the expected relative standard errors of the design are seen.

Different number of measurements per patient

There is no restriction that each patient should be measured exactly 13 times. Here designs are considered where the total cost is 780 samples and patients are allowed to be measured less then 13 times. This allows the possibility of more subjects entering the study, which could improve the estimation of the model parameters.

Assume that the allowed time points are according to the previously defined set T. To restrict the number of possible elementary designs, a patient was allowed to have between 5 and 8 measurements. This gives a total of $\sum_{i=5}^{8} \binom{20}{i} = 257754$ elementary designs. The optimal design given by the Fedorov-Wynn algorithm was

$$\Xi_{imp,780} = \left\{ \begin{matrix} (0, 4, 15, 30, 60, 70, 100), 33 \\ (0, 15, 30, 60, 100), 16 \\ (0, 10, 30, 60, 70, 100), 12 \\ (0, 15, 30, 60, 70, 100), 1 \end{matrix} \right\},$$

with corresponding criterion value

$$\Phi(\Xi_{imp,780}) = 42385.86.$$

This means that the information has increased with 103 % compared to the initial design Ξ_{init} and still using 780 samples with 62 patients. In Section 6.5 the expected relative standard errors of the design are seen and the parameter precision is improved as imposed by the Cramér-Rao bound.

6.4 Reducing the cost of the study

In this section the goal is to reduce the cost of the study. As before, the two cases with equal and different number of measurements per patient are considered.

Equal number of measurements per patient

Let's assume that the goal is to reduce the number of measurements from 13 to 8 per patient, while keeping 30 subjects in the study. This gives a total of 480 samples and corresponds to a reduced cost of about 38%.

The allowed sampling points are chosen to be the previously defined set T (20 points), which gives $\binom{20}{8} = 125970$ elementary designs. The optimal design $\Xi_{red,8}$ (reduced 8-point design) obtained in PFIM was

$$\Xi_{red,8} = \{0, 4, 15, 30, 60, 70, 100, 115\},$$

with an objective value of

$$\Phi(\Xi_{red,8}) = 24433.26.$$

The optimized 8-sample design shows a higher objective value than the initial design. The gain in information is about 17%. In Section 6.5 the expected relative standard errors of the design are seen and the parameter precision is maintained for the reduced design.

Different number of measurements per patient

Assume that 480 samples are used (as in previous design) but now it is allowed with different number of measurements per subject. Again, the allowed time points is the set T and it is assumed that each patient has a minimum of 5 measurements and a maximum of 8 measurements. This gives a total of $\sum_{i=5}^{8} \binom{20}{i} = 257754$ elementary designs. The optimal design $\Xi_{red,480}$ was

$$\Xi_{red,480} = \left\{ \begin{array}{c} (0,4,15,30,60,70,100),21 \\ (0,15,30,60,100),9 \\ (0,10,30,60,70,100),7 \\ (0,15,30,60,70,100),1 \end{array} \right\}.$$

Hence this study contains 38 subjects with number of measurements ranging from 5 to 7, with a total of 480 samples. The objective value was

$$\Phi(\Xi_{red,480}) = 26083.11,$$

which is very similar in performance as the design having 8 measurements per patient but the design is further reducing the sampling burden for each patient.

6.5 Standard error of parameter estimates

In Table 6.3 the different designs considered in this chapter are summarized for convenient comparison.

Table 6.3: Number of subjects, cost and information criterion for the different designs explored. Ξ_{init} is the initial design, $\Xi_{imp,13}$ is the improved 13-point design, $\Xi_{imp,780}$ is the improved 780-sample design, $\Xi_{red,8}$ is the reduced 8-point design and $\Xi_{red,480}$ is the reduced 480-sample design.

Design name	Subjects	Cost (samples)	Criterion
Ξ_{init}	30	780	20848.87
$\Xi_{imp,13}$	30	780	30637.01
$\Xi_{imp,780}$	62	780	42385.86
$\Xi_{red,8}$	30	480	24433.26
$\Xi_{red,480}$	38	480	26083.62

In Table 6.4 the expected relative standard errors for the designs given my PFIM are seen.

Table 6.4: Expected relative standard errors (%) of the parameters for the initial 13-point design (Ξ_{init}), the improved designs and the reduced designs given by PFIM.

Design	S_g	p_2	S_i	n	h	γ	G_0	I_0
Ξ_{init}	3.9	8.3	3.0	0.8	0.2	1.5	1.7	1.7
$\Xi_{imp,13}$	3.5	7.0	1.6	0.9	0.2	1.3	1.6	1.7
$\Xi_{imp,780}$	3.3	6.5	1.4	1.0	0.2	1.3	1.2	1.2
$\Xi_{red,8}$	4.1	8.2	1.8	1.1	0.2	1.6	1.7	1.8
$\Xi_{red,480}$	4.1	8.3	1.7	1.3	0.2	1.6	1.6	1.6
Design	$\omega^2_{S_g}$	$\omega^2_{S_i}$	ω^2_n	ω^2_γ	$\omega^2_{G_0}$	$\omega^2_{I_0}$	$\sigma_{prop,G}$	$\sigma_{prop,I}$
Ξ_{init}	27.6	26.5	26.1	26.8	26.3	26.8	4.1	4.4
$\Xi_{imp,13}$	27.2	26.2	26.2	26.8	26.4	26.8	4.0	4.4
$\Xi_{imp,780}$	19.5	18.5	18.6	19.3	18.6	18.8	4.5	6.0
$\Xi_{red,8}$	27.7	26.4	26.3	27.3	26.5	27.0	5.4	6.5
$\Xi_{red,480}$	24.9	23.6	23.6	24.7	23.7	24.0	5.7	7.7

PFIM gives the expected standard errors by calculating the inverse of the approximated Fisher information matrix. To obtain realistic errors simulations are used. The empirical standard errors were calculated by simulating the study 50 times and then re-estimate the parameters with the FOCE algorithm, which is implemented in the package **nlmeODE**[18] and **nlme**[19] in R. For information about the algorithms, see [7].

In each simulation the estimated parameters were obtained. The standard deviations are then calculated from these 50 estimates. This was done for the initial design and the reduced 8-point design which has similar information criterions. In Figure 6.3 the expected relative standard errors given by PFIM are shown together with the empirical standard errors for the initial design.

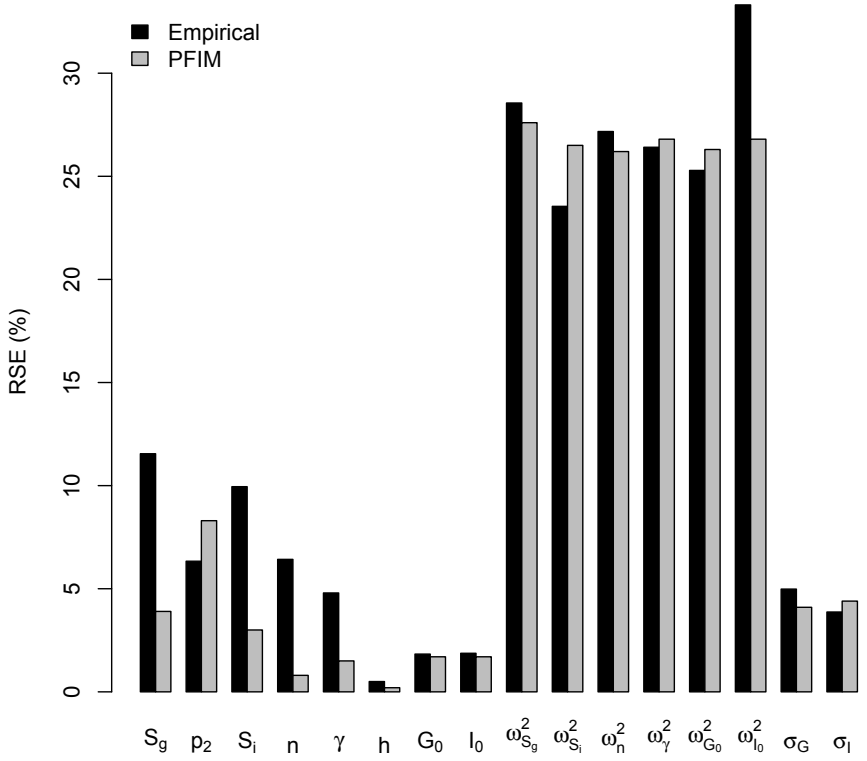

Figure 6.3: Empirical relative standard errors obtained by simulation (black bars) and relative standard errors given by the PFIM software (grey bars) for the initial design Ξ_{init}.

Finally, the empirical standard errors for two different designs are compared. The designs considered are the initial design Ξ_{init} and the reduced design $\Xi_{red,8}$. In Figure 6.4 the relative standard error (in %) for the two designs are shown.

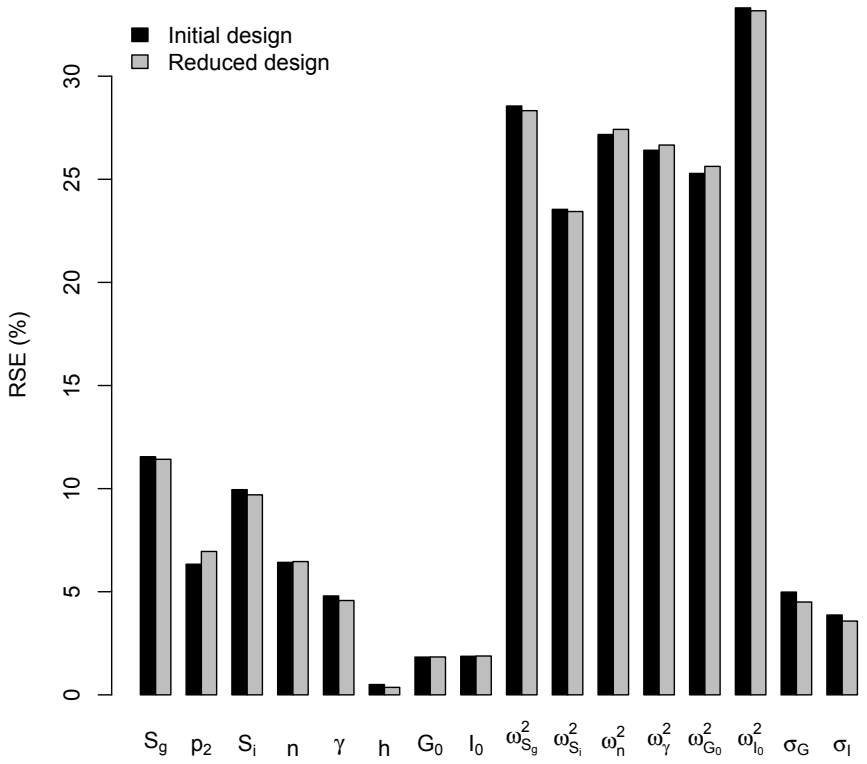

Figure 6.4: Empirical RSE for the initial 13-point design (black bars) and the reduced 8-point design (grey bars).

It should be noted that for the reduced design there was problem with convergence in 7 cases out of 50. The cases where the estimated parameters were obviously wrong were omitted from the calculations of RSE. Also, the NLME package has no combined error model. Instead, a exponential error model was used, which is very similar to a proportional error model. For more information about the exponential error model, see [7].

6.6 Discussion

From the results above the conclusion is that PFIM can be used both to find a better design with the same cost and also reduce the cost of study while maintaining high information. Using the Fedorov-Wynn algorithm all possible designs can be compared given the allowed sampling times and number of measurements.

In Table 6.4 the expected relative standard errors (RSE) for the designs explored are shown. From the results one can conclude that the improved designs have a decrease in RSE for almost all the parameters. When different number of measurements per patient is allowed while including more subjects the RSE for the variance components are decreased. For the reduced 8-point design the estimated parameter precision is maintained while reducing the cost by 38%.

PFIM gives the expected standard error by calculating the inverse of the Fisher information matrix, as imposed by the Cramér-Rao bound. When comparing the standard errors with the empirical standard errors given by the FOCE algorithm the conclusion is that the empirical standard errors are generally higher then the standard errors given by PFIM. The increase in RSE compared to PFIM may be due to the high non-linearity of the model or problems with ODE systems in PFIM.

When the empirical standard errors for two different designs with similar information criterion were compared the result was that the RSEs were very similar as well. Hence the standard errors of the parameter estimates seem to be reflected by the Fisher information as expected. This suggests that one may use PFIM to improve and reduce designs and then evaluate the final design by simulations.

7

PK/PD modeling of blood neutrophil data

In this chapter PFIM is used to evaluate and improve an initially suggested design in a 6 months long study with a new drug in development. As a side effect, the drug is known to decrease the production of neutrophils, which are the most common type of white blood cells in mammals. Neutrophils forms an essential part of the immune system. In this study the neutrophil level is a safety parameter which has to be monitored with high precision.

First a PK/PD model describing the time course of neutrophils and the initial study design are introduced. The initial design and an improved design are evaluated in PFIM. In addition the initial design is optimized with respect to the pharmacokinetic sampling schedule.

7.1 The underlying model

A two-compartment pharmacokinetic model with linear elimination was assumed. In Figure 7.1 the model parameterized with volumes of distribution V_1 and V_2, clearance Cl and inter compartmental clearance Q is shown.

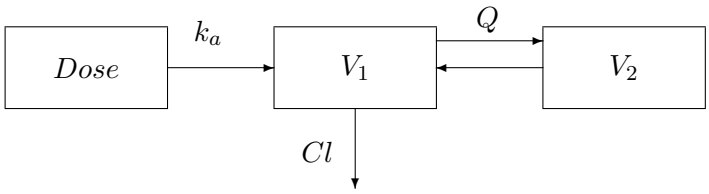

Figure 7.1: The pharmacokinetic model with parameters k_a, V_1, V_2, Cl and Q.

The drug absorption is modeled as a dosing compartment from which the drug enters the system after administration. The absorption rate is the rate by which the drug enters the central compartment, which is an appropriate model when the drug is administrated orally. It is assumed that the patient takes the drug with dose D every 12th hour. The model is reparametrized using the rates $k_{12} = Q/V_1$, $k_{21} = Q/V_2$ and $k_e = Cl/V_1$. The amount of drug in the three compartments are denoted A_0 (dose), A_1 (V_1) and A_2 (V_2).

The pharmacodynamic model is a turnover response model as illustrated in Figure 7.2.

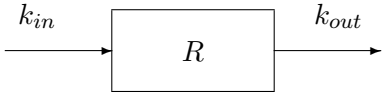

Figure 7.2: A turnover response model with production rate k_{in} and elimination rate k_{out}.

Since the underlying drug is known to decrease the production of neutrophils an I_{max} model where the drug is causing a reduction of the production rate k_{in} is used. The differential equation describing the response $R(t)$ is

$$\frac{dR}{dt} = k_{in}\left(1 - \frac{I_{max}C}{C + C_{50}}\right) - k_{out}R,$$

where C is the drug concentration in the central compartment, I_{max} is the maximal drug effect and C_{50} is the concentration needed to reach half of the maximal effect.

To model the circadian rhythm where the rate k_{in} varies over time a cosine function describing the fluctuations in neutrophil levels over the day is added. Rewriting $k_{in} = R_{base}k_{out}$, the differential equation describing the response is

$$\frac{dR}{dt} = Acos(2\pi(t-s)/f) + k_{out}\left(R_{base}\left(1 - \frac{I_{max}C}{C + C_{50}}\right) - R\right),$$

where A is the amplitude, s the shift and f the frequency. This leads to the ordinary differential equation system

$$
\begin{aligned}
\frac{dA_0}{dt} &= -k_a A_1 \\
\frac{dA_1}{dt} &= k_a A_1 - k_e A_2 - k_{12} A_2 + k_{21} A_3 \\
\frac{dA_2}{dt} &= k_{12} A_2 - k_{21} A_3, \\
\frac{dR}{dt} &= Acos(2\pi(t-s)/f) + k_{out}\bigg(R_{base}\big(1 - \frac{I_{max} A_1/V_1}{A_1/V_1 + C_{50}}\big) - R\bigg),
\end{aligned}
$$

with initial conditions

$$
\begin{aligned}
A_0(0) &= D \\
A_1(0) &= 0 \\
A_2(0) &= 0 \\
R(0) &= R_{base}.
\end{aligned}
$$

Moreover, the parameters k_a, Cl, R_{base}, C_{50}, k_{out} and s vary in the population according to a lognormal distribution. A proportional error model with $\sigma_{prop} = 0.1$ for both the PK and PD response was assumed. Hence there are a total of 20 parameters in the model. The parameter values used in the model are seen in Table 7.1.

Table 7.1: Parameters values with units used in the model.

k_a (1/h)	Cl (L/h)	V_1 (L)	V_2 (L)	Q (L/h)	R_{base}	I_{max}
0.54	8	24	12	0.64	4.8	0.5
C_{50} (mg)	k_{out} (1/h)	A	s (h)	f (h)	$\omega^2_{k_a}$	ω^2_{Cl}
0.1	0.21	0.11	12	24	2.25	0.16
$\omega^2_{R_{base}}$	$\omega^2_{C_{50}}$	$\omega^2_{k_{out}}$	ω^2_{shift}	σ_{propPK}	σ_{propPD}	
0.0576	1.44	0.16	0.0676	0.1	0.1	

7.2 The initial study design

The suggested clinical study is a 6 month study with 450 patients to be enrolled. The patients are divided into three groups with 150 patients in each. Each group is given a different dose.

The patients visit the clinic a total of 10 times. On each visit the PD response is measured just before the dose is administered. At four visits the PK response is measured

pre-dose and at one visit three post-dose PK samples are taken. This gives a total of 10 PD samples and 7 PK samples. The pre-dose samples are taken immediately before the drug administration and the 3 post dose samples are taken somewhere between 1-2h, 3-5h and 7-10h after drug administration. Hence there is no strict sampling schedule but intervals.

7.3 Evaluating the initial design in PFIM

When using the full ordinary differential system with a turnover response model it was not feasible to use the analytic option in PFIM. Instead, the user-defined model option which allows specification of the ODE system was used. Unfortunately a model described by an ODE system together with a multiple dose parallell group study has not yet been implemented in PFIM. Instead the ODE system must be tweaked a little, making the model even more complicated.

The evaluation of a specific design using the full PK/PD model took a long time in PFIM. Since the study is a 6 month study it would been preferred to evaluate the the study and optimize the sampling times at every time the patient visit the clinic. However, this was not possible due to the complicated model and the need of simulating the ODE system for a long time.

Due to the limitations of PFIM, only the initial PK/PD design and one new design were evaluated to obtain the expected standard errors of the parameter estimates. Instead of simulating a study length of 6 months the samples was allocated in steady state with 24 hours in between. A new design with the sampling points at different post-dose times was also considered. The initial sampling times for PK and PD was

$$\xi_{PK}^{init} = (72,96,97,98,103,120,144)$$
$$\xi_{PD}^{init} = (0,48,72,96,120,144,168,192,216,240),$$

where the values denote hours after the first dose. We denote the initial population design

$$\Xi_{init} = \{(\xi_{PK}^{init}, \xi_{PD}^{init}), 450\}$$

In Figure 7.3 the expected responses for the PK and PD profiles together with the sampling points are shown. Notice that the system reaches steady-state after 3 doses.

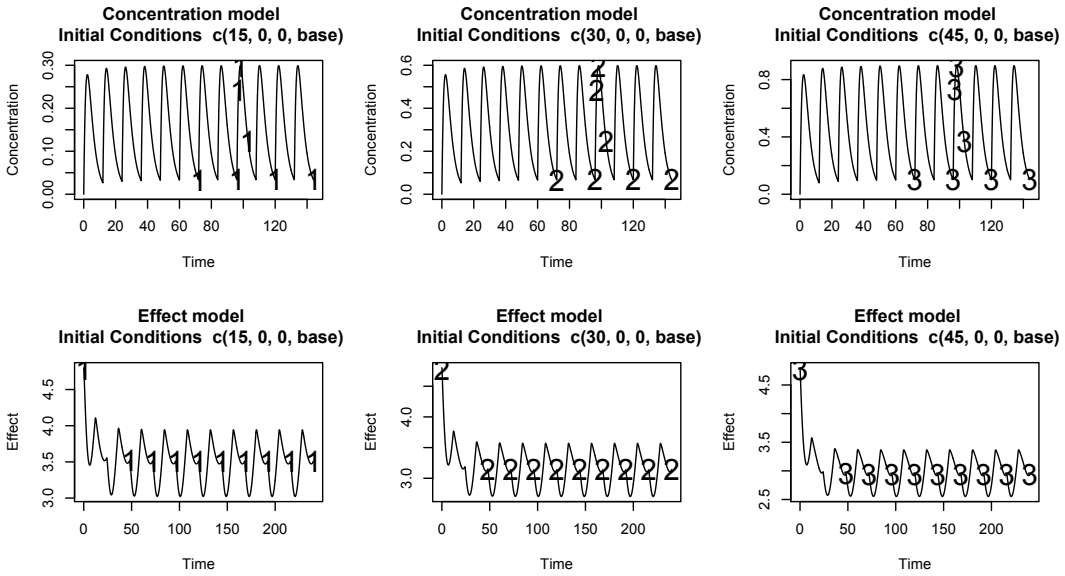

Figure 7.3: The expected PK (upper) and PD (lower) response together with the initial sampling times for the three dose groups. The plot is automatically created by PFIM.

The second design Ξ_{new} has sampling times

$$\xi_{PK}^{new} = (72,97,122,147,172,197,222)$$
$$\xi_{PD}^{new} = (0,48,72,97,122,147,172,197,222,247),$$

where the values denote hours from the first dose. This design implies that the patients are measured at different post-dose times at every visit. The information criterions for the two designs was

$$\Phi(\Xi_{init}) = 808.45$$
$$\Phi(\Xi_{new}) = 1535.31,$$

which implies that the new design is much more informative. In Table 7.2 the expected relative standard errors given by PFIM for the two designs are shown, indicating an informative initial design.

Table 7.2: Expected relative standard error (%) for the initial design Ξ_{init} and the new design Ξ_{new} given by PFIM.

Design	k_a	Cl	V_1	V_2	Q	R_{base}	I_{max}	C_{50}	k_{out}	A
Ξ_{init}	1.1	1.5	2.4	13.6	7.6	1.2	4.8	12.8	12.7	14.8
Ξ_{new}	0.8	1.4	2.2	29.0	7.5	1.2	2.8	8.0	4.8	12.8
Design	s	f	$\omega^2_{k_a}$	ω^2_{Cl}	$\omega^2_{R_{base}}$	$\omega^2_{C_{50}}$	$\omega^2_{k_{out}}$	ω^2_s	σ_{propPK}	σ_{propPD}
Ξ_{init}	12.7	0.1	7.6	7.5	8.1	21.0	227.5	173.1	2.5	1.9
Ξ_{new}	3.8	0.3	7.5	7.4	7.9	12.2	125.6	19.9	2.1	2.0

7.4 Improving PK sampling

Due to the complicated ODE model and the limitations of PFIM when it comes to running time only the sampling for the PK samples in steady-state is considered.

For the two compartment model there is a closed form expression for the concentration at time t after repeated doses with dose interval τ. This model is implemented in PFIM library and it is used to find better sampling times. Note that in the real study samples are taken at different times, for example after 1 month and after 2 months. The steady-state model here serves as a tool to find how long after the dose administration samples should be collected. The expression for the concentration at time t after the drug's administration is [3]

$$C(t) = D\left(\frac{Ae^{-\alpha t}}{1-e^{-\alpha\tau}} + \frac{Be^{-\beta t}}{1-e^{-\beta\tau}} - \frac{(A+B)e^{-k_a t}}{1-e^{-k_a\tau}}\right),$$

where

$$\begin{aligned}
\beta &= \frac{1}{2}\big(k_{12}+k_{21}+k_e - \sqrt{(k_{12}+k_{21}+k_e)^2 - 4k_{21}k_e}\\
\alpha &= \frac{k_{21}k_e}{\beta}\\
A &= \frac{k_a}{V_1}\frac{k_{21}-\alpha}{(k_a-\alpha)(\beta-\alpha)}\\
B &= \frac{k_a}{V_1}\frac{k_{21}-\beta}{(k_a-\beta)(\alpha-\beta)}.
\end{aligned}$$

The pharmacokinetic response is measured a total of seven times in steady-state. Four of the samples are taken at the same time as the administration of the drug, that is at $t = 0$. The three post-dose samples are taken somewhere between 1-2h, 3-5h and 7-10h after the drugs administration. Hence there are no strict sampling schedule since the times are given as intervals. Using PFIM library of pharmacokinetic models the

steady state behavior of the PK response easily can be implemented. The initial design was set to

$$\Xi_{init} = \Big\{(0,0,0,0,1,3,7), 450\Big\},$$

with corresponding information criterion

$$\Phi(\Xi_{init}) = 385.0.$$

A design using PFIM with the Fedrov-Wynn algorithm was optimized. In the first step it was assumed that 7 samples were allowed to be taken from the set of time points $\{0,1,\ldots,10\}$. The optimized design was

$$\Xi_1 = \begin{Bmatrix} (0,1,4,5,6,8,9), 254 \\ (0,1,4,5,6,7,10), 196 \end{Bmatrix},$$

with an information criterion of

$$\Phi(\Xi_1) = 536.3.$$

If all the patients are allocated to the first group or second group rep, we get designs Ξ_2 and Ξ_3 respectively which has information criterions of

$$\Phi(\Xi_2) = 535.6.$$
$$\Phi(\Xi_3) = 535.1.$$

By defining a larger set of allowed time points the design can be improved further. The set of allowed time points is now set to $\{0, 0.5, 1, \ldots, 10\}$. The optimized design was

$$\Xi_4 = \begin{Bmatrix} (0,0.5,4.5,5,5.5,8.5,9), 46 \\ (0,0.5,4.5,5,5.5,6,9.5), 161 \\ (0,0.5,4,4.5,5,8.5,9), 243 \end{Bmatrix},$$

with an information criterion of

$$\Phi(\Xi_4) = 571.4,$$

If the patients are allocated to the same groups the design Ξ_5, Ξ_6 and Ξ_7 are obtained with criterions

$$\Phi(\Xi_5) = 570.5$$
$$\Phi(\Xi_6) = 567.5$$
$$\Phi(\Xi_7) = 570.1$$

.

In Table 7.3 the expected standard errors for the different designs given by PFIM are shown.

Table 7.3: Expected relative standard errors (%) for the considered designs given by PFIM.

Design	k_a	Cl	V_1	Q	V_2	ω^2_{ka}	ω^2_{Cl}	σ_{prop}
Ξ_{init}	1.1	1.7	3.1	14.3	103.0	8.0	7.6	2.6
Ξ_1	0.9	1.6	2.6	10.4	58.5	7.7	7.5	2.2
Ξ_2	0.9	1.6	2.6	10.4	57.0	7.7	7.4	2.2
Ξ_3	0.9	1.6	2.6	10.4	61.1	7.7	7.5	2.2
Ξ_4	0.8	1.6	2.5	10.0	54.6	7.6	7.4	2.2
Ξ_5	0.8	1.6	2.5	9.9	53.2	7.6	7.4	2.2
Ξ_6	0.8	1.6	2.6	10.2	60.4	7.7	7.5	2.2
Ξ_7	0.8	1.5	2.5	9.9	52.2	7.6	7.4	2.2

7.5 Discussion

The PK/PD model considered in this chapter combined with a long study reveals a major drawback of PFIM. When a need for a long study arises, the evaluation in PFIM takes a very long time. No optimization seems feasible for such models unless using a high performance computational cluster.

In the first section the initial design and a new design was evaluated in PFIM. The initial design showed a very high RSE on the parameters $\omega^2_{k_{out}}$ and ω^2_s although the design was informative when looking at the other parameters. When changing the sampling points to all be taken at different post-dose times the RSE was reduced but a high RSE on $\omega^2_{k_{out}}$ was still noticed. The RSE on V_2 was increased for the new design.

When optimizing the PK sampling in steady state some interesting results were noticed. The results in Table 7.3 show that that the different designs exhibit a large difference in RSE for the peripheral volume of distribution. The initial design had a RSE of 103 % while the improved designs had a RSE ranging from 52.2-61.1% for that specific parameter. For the other parameters the RSE was low, ranging from 1.1-14.3 %. From the results the conclusion is that an improvement of the PK sampling in the study could be done by taking samples at 7 different post-dose times. There is still high uncertainty for the peripheral volume of distribution.

Note that when modeling PK and PD response simultaneously the parameter precision for the peripheral volume was decreased due to the relationship between the two responses.

8

Final discussion and conclusions

For scientific and ethical reasons it is most important clinical experiments are sufficiently designed to answer the pre-specified questions of interest. There should not be too few individuals recruited to reach inconclusive results and a study should not be over dimensioned and expose an excess of individuals at risk. In this thesis the PFIM software used for optimal design of clinical studies has been evaluated. The purpose of PFIM is to improve the information from clinical studies by maximizing the Fisher information which in turn can improve the precision of estimated parameters.

PFIM has been evaluated using three different nonlinear mixed effect models. The first model was a simple analytical PK/PD model where the goal was to compare the expected standard errors given by PFIM with those from a large simulation in NONMEM. The results showed that the RSE were very similar and hence PFIM can be an appropriate alternative when it comes to the evaluation of a specific design. It offers a fast way of calculating the expected standard errors (a few seconds) compared to the simulation method (a few hours).

The second model was Bergman's model for glucose-insulin response, which was described by an ordinary differential equation system. Here an initial 13-point design was reduced to a 8-point design while maintaining the information in the study. Simulation results showed that the standard errors given by PFIM were slightly underestimated compared to the empirical standard errors. When comparing the initial and the reduced design, with similar information criteria, the empirical standard errors were very similar. This is an interesting result which can have great impact on the cost of future studies and the wellbeing of the patients included in them. Instead of having to be sampled at 13 times the patient can be sampled 8 times.

The final example included a PK/PD model describing a safety endpoint in a 6 month long study. When applying this model in PFIM some major drawbacks were noticed. When the model was described by an ordinary differential equation system the evaluation of a specific design took a very long time in PFIM, so no optimization of the

suggested design study was feasible. Two different designs were evaluated using PFIM and the results showed that the RSE could be reduced by taking samples at different post-dose times every visit. The results also showed that the suggested design was sufficiently informative giving confidence in the suggested study design. When optimizing the PK sampling at steady state, a major reduction in RSE for the peripheral volume of distribution was obtained.

During the work with PFIM some limitations of the software have been noticed. It only offers the optimality criterion to be D-optimality. When there is a need for estimating some parameters better than others one would like to be able to choose a subset of parameters to estimate, so called D_s-optimality. For example, there is often a need for more precise estimation of the parameters Cl and E_{max} than for k_a and k_{out} where the former does not impact overall drug exposure and the latter is a disease rather than a drug specific system parameter which may be known from the literature. The model also has to be pre-specified and all the parameters must be known prior to the optimization. There is no option where one could specify uncertainty in the prior estimates, as implemented in so called Bayesian designs where a distribution can be specified for the parameters.

There is a number of other programs used for optimal design of clinical studies. The software PopED[20] has been evaluated in the thesis by Dosne[16] and offers a wider range of optimality criteria, including D_s-optimality. The user can also specify an uncertainty in the prior parameter estimates (ED-optimality). PopED also offers a wide range of approximation types, including the First Order, First Order Conditional and First Order Conditional Interaction. Comparing this to the methods in PFIM, the conclusion is that PFIM has a very limited range of options, since it only incorporates the First Order approximation and D-optimality.

For future work, it would be interesting to investigate how improved parameter estimation affects the decision making in clinical trials. Ultimately a clinical trial translates to making the correct decision, for example choosing the right dose. By taking this work one step further one would like to investigate the impact of study designs in decision making. Another type of optimality criteria which could be suitable for such purposes is V-optimality, which seeks to minimize the average prediction variance.

Finally, the conclusion is that PFIM can be used to evaluate, optimize or reduce clinical study designs. For a model specified by an ordinary differential equation system, great care has to be applied to the results given by PFIM. The suggestion is that one always should evaluate a final design by simulations, for example in the NONMEM software.

Bibliography

[1] C. Bazzoli, S. Retout, F. Mentré, Design evaluation and optimization in multiple response nonlinear mixed effect models: Pfim 3.0, Computer Methods and Programs in Biomedicine 98 (2010) 55–65.

[2] R. Lalonde, K. Kowalski, M. Hutmacher, Model-based drug development, Clinical pharmacology and therapeutics 82 (2007) 21–32.

[3] A. Dubois, J. Bertrand, M. F, Mathematical expression of the pharmacokinetic and pharmacodynamic models implemented in the PFIM software (2011).

[4] E. Ette, P. Williams, Pharmacometrics, The Science of Quantitative Pharamcology, 2007.

[5] M. Davidian, D. Giltinan, Nonlinear models for repeated measurement data: An overview and update, JABES 8 (2003) 387–419.

[6] E. Limpert, W. Stahel, M. Abbt, Log-normal distributions across the sciences: Keys and clues, BioScience 5 (2001) 341–352.

[7] J. Pinheiro, D. Bates, Nonlinear mixed effect models with S-plus, 1995.

[8] A. Atkinson, A. Donev, R. Tobias, Optimum Experimental Design, with SAS, 2007.

[9] A. Atkinson, Examples of the use of an equivalence theorem in constructing optimum experimental designs for random-effects nonlinear regression models (2007).

[10] R Development Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0 (2008).
URL `http://www.R-project.org`

[11] K. Ogungbenro, G. Graham, I. Gueorguiva, L. Aarons, The use of fedorov exchange algorithm to optimize sampling times for a population pharmacokinetic experiments, Computer Methods and Programs in Biomedicine 80 (2005) 115–125.

[12] C. Bazzoli, S. Retout, F. Mentré, Fisher information matrix for nonlinear mixed effects multiple response models: Evaluation of the appropriateness of the first order linearization using a pharamcokinetic/pharamcodynamic model, Statistics in medicine 28 (2009) 1940–1956.

[13] S. Retout, E. Comets, A. Samson, F. Mentré, Design in nonlinear mixed effect models: optimization using the fedorov-wynn algorithm and power of the wald test for binary covariates., Statistics in Medicine 26 (2007) 5162–79.

[14] S. Beal, L. Sheiner, A. Boeckmann, R. Bauer, NONMEM User's Guides, Icon Development Solutions, Ellicott City, MD, USA (2009).

[15] K. Andersen, H, A Bayesian approach to Bergman's minimal model.

[16] A. Dosne, Optimal design for a diabetes dose finding study, Master's thesis (2011).

[17] M. Brauckovic, Modelling insulin sensitivity using nonlinear mixed effect models, Master's thesis (2010).

[18] C. Tornoe, nlmeODE: Non-linear mixed effects modeling in nlme using differential equations, r package version 1.0 (2010).

[19] J. Pinheiro, D. Bates, S. DebRoy, D. Sarkar, R Development Core Team, nlme: Linear and Nonlinear Mixed Effects Models, r package version 3.1-103 (2012).

[20] M. Foracchia, A. Hooker, P. Vicini, A. Ruggeri, Poped, a software for optimal experiment design in population kinetics, Comput Methods Programs Biomed 74 (2004) 29–46.