Stockholms universitet

Response Adaptive Optimal Design in Clinical Trials - A Simulation Study Motivated by a Real Data Example

Carolina Blomqvist

Kandidatuppsats i matematisk statistik
Bachelor Thesis in Mathematical Statistics

Kandidatuppsats 2010:7
Matematisk statistik
Juni 2010

www.math.su.se

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm

Stockholm University

Mathematical Statistics
Stockholm University
Bachelor Thesis **2010:7**
http://www.math.su.se

Response Adaptive Optimal Design in Clinical Trials - A Simulation Study Motivated by a Real Data Example

Carolina Blomqvist*

Juni 2010

Abstract

In this paper we summarize the theory of clinical trials, in particular their design, optimal design and adaptive design. We emphasize the two-stage optimal response-adaptive design and implement the theory of optimization on the nonlinear, multiparameter *Emax*-sigmoid model using two different optimization criteria, two optimization methods, three different parameter vectors and three different sample sizes. The aim of the paper is to study the distributional properties of an optimal design through simulations. We also propose an optimization method for computing the optimal design. The results reveal that a large variation in parameter estimates of the dose-response curve yields a large variation of the estimated optimal design. To explore this further we consider four different dose-response curve models. One of these models proves to give a reduction in variation and hence a reliable optimal design. However, our conclusion is that this model might not be realistic and that further investigation of the asymptotics of the estimated parameters should be carried out.

*Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Sweden. E-mail:carolina_blomqvist@hotmail.com. Supervisor: Ola Hössjer.

Response Adaptive Optimal Design in Clinical Trials - A Simulation Study Motivated by a Real Data Example

By Carolina Blomqvist

2010-05-20

Summary

In this paper we summarize the theory of clinical trials, in particular their design, optimal design and adaptive design. We emphasize the two-stage optimal response-adaptive design and implement the theory of optimization on the nonlinear, multiparameter E_{max}-sigmoid model using two different optimization criteria, two optimization methods, three different parameter vectors and three different sample sizes. The aim of the paper is to study the distributional properties of an optimal design through simulations. We also propose an optimization method for computing the optimal design. The results reveal that a large variation in parameter estimates of the dose-response curve yields a large variation of the estimated optimal design. To explore this further we consider four different dose-response curve models. One of these models proves to give a reduction in variation and hence a reliable optimal design. However, our conclusion is that this model might not be realistic and that further investigation of the asymptotics of the estimated parameters should be carried out.

Preface

This paper constitutes a thesis of 15 ECTS and leads to a Bachelors' degree in Mathematical Statistics at the Department of Mathematics at Stockholm University. The paper has been carried out in collaboration with the Department of Clinical Information Science, AstraZeneca, Sweden.

I would like to express my sincere thanks to my mentors Ola Höjsser at the Department of Mathematics at Stockholm University and Frank Miller and Olivier Guilbald at AstraZeneca for all the guidance and help during the development of this paper.

Carolina Blomqvist,
Stockholm 2010-05-12

Contents

1 Introduction

The process of developing new drugs is time-consuming and expensive and it is hence important for researchers to outline the clinical trial in a way that gives as much information as possible about effective doses and safety of the drug. Knowledge about the dose-response relationship is crucial in order to decide which doses should be investigated in the rest of the clinical development process, and this knowledge is oftenly obtained in a *dose-finding trial*, also called a Phase IIB trial. It is based on this trial that a decision is made of whether or not it is sufficient to continue to investigate the new drug, and if it is, the dose-finding trial determines the appropriate doses to be investigated in the following Phase III trials.

In dose-finding trials, a traditional design is oftenly used, allocating the patients into different dose groups and one placebo-group equally. The problem with this type of design is that it may throw away possibilities to investigate the dose-response relationship more efficiently. What we will look at in this paper is an alternative to the traditional design for dose-finding trials called an adaptive optimal design, where we after some preliminary investigation reallocate the patients to the dose groups with respect to some optimality criterion. We will not focus on the comparison between the efficiency in an adaptive optimal design versus a traditional design, but rather look at the process of developing the adaptive optimal design and the potential difficulties it may cause.

1.1 Background material

Before a more detailed description of the problem is stated it is a good idea to look more into what a clinical trial is, what parts it consists of, what an optimal design of a clinical trial is and why the design is such an important part of the clinical trial. We will also give an explanation of what adaptive designs are and in particularly focus of the explanation the type of design that we will use in this paper, namely the response-adaptive design.

What is a clinical trial?

A definition of a clinical trial and its' main purpose is formulated as:

"A clinical trial is a clinical investigation in which treatments are administered, dispensed or used involving one or more human subjects for evaluation of the treatment"[1]

[1] Chow et. al. (2004). *Design and Analysis of clinical trials*, Wiley & Sons, second edition

A drug development process, which consists of several clinical studies, can be divided in drug development phases, as follows.

- **Pre-clinical:** this is the phase where pharmaceutical companies decide whether a new drug should be further investigated or not, based on early evidence of toxicity, efficacy and pharmacokinetic responses (see the word list in Appendix D for explanation of the term pharmacokinetics). The tests can be either *in vivo* (on animals or cells) or *in vitro* (in test tubes) and one oftenly tests a wide range of doses.

- **Phase 0:** this phase is also called human microdosing (see Appendix D) study since doses given to a small group of human patients (10-15) by definition are too low to give any real therapeutic effect. The main goal with phase 0 is to see whether the investigational drug behaves as expected (from the pre-clinical phase) in the human body and to collect preliminary data of how the body processes the drug and how the drug works in the body. The procedure with a phase 0 is relatively new in the area of clinical trials, it was first recommended in the United States Food and Drug Administration's (FDA) 2006 *Guidance on Exploratory Investigational New Drug (IND) Studies* and it has been discussed whether or not this phase is useful or redundant.

- **Phase I:** if no phase 0 trials has been carried out, then this is the first introduction of the new drug to human subjects. The primary objectives are to determine how the drug is absorbed in the human body, possible early side effects at different dose levels, the maximal tolerable dose MTD (see Appendix D) and above all; the safety of the drug. To determine the appropriate dose levels for the following phases one often uses dose escalation; different groups of patients are given different doses. Phase I trials are usually relatively small, around 20-100 subjects are enrolled.

- **Phase II:** this phase usually includes larger trials with approximately 100-300 patients. In this phase one continues to monitor safety and the key issue is to determine the dosing ranges and doses for the phase III trials. These dosing ranges and dose levels are oftenly determined by dose-finding trials and it is based on these trials that we make a decision for proceeding or not proceeding to phase III. Is it safe to continue the testing of the new drug or should the development of the new drug be stopped now? It is not unlikely that the new drug fails in this development phase when it is discovered that the drug doesn't work as expected or it has unexpected, potentially dangerous, side affects

- **Phase III:** this phase involves trials with a few hundred or sometimes a thousand or more patients. The trials often has one or several control

groups and the patients are usually randomized into dose groups. It is also common to use multicenter trials as phase III trials, which is several clinical trials performed at different clinics or institutions. This can be a benefit when the trials involves a lot of patients, but it can also impose variation between the clinics (not giving a reliable result as a consequence). The primary goal of the studies is to give additional information regarding the safety and effectiveness of the drug. Phase III trials are only conducted if the preliminary evidence of the safety and efficacy of the drug has been demonstrated in the previous phase II trials.

- **Phase IV:** this final phase is usually carried out first when the drug has been approved for marketing. It is an opportunity to learn about interaction with other substances/drugs and about possible rare side affects.

Planning and executing a clinical trial is a very large project and it contains a sufficient number of details and procedures that are beyond the scope of this paper, hence we refer to the literature in the reference list for more reading.

As mentioned earlier, a good understanding of the dose-response relationship is crucial, not only for cost-reduction reasons, but also because any compound is potentially toxic if consumed in high enough doses. Water may for example seem harmless but states a good example of a compound that is toxic in large amounts since it is potentially lethal for humans if we drink more than 10 liters at once. The emphasis of this paper will therefore be on the dose-finding phase IIB trial usually performed at the end of the phase II trial, and we will look at a particular type of design of this trial called a response-adaptive design.

Design of clinical trials

Before we can explain the concept of optimal design we need to establish what a design of a clinical trial is. In short and simple terms the design of a clinical trial is the specification of which treatments, such as which doses of the new drug or placebo, will be used in the trial and how many patients will be treated with each dose. Further, designs of clinical trials can be divided into two main groups as follows.

- **Non-adaptive designs**: the design is fully pre-specified, i.e. the allocation of patients to each dose group is pre-specified and may not be altered during the course of the trial.

- **Adaptive designs**: allocations to dose groups can be altered during the trial based on interim analyses of this ongoing trial. Adaptive

designs are sometimes called sequential designs.

Within these two main groups there are several types of designs. The description of these designs goes beyond the scope of this paper, but for the interested reader we refer to the literature in the reference list for more reading.

The design we will use in this paper belongs to the group of adaptive designs and we will describe these in more detail in a section below.

Optimal design

The design that we will use in this paper belongs to a group called *optimal designs*. An optimal design is a design that depend on a statistical model and is optimal with respect to some statistical criterion. The advantages of optimal designs are that they reduce cost by using statistical models which enable estimation of parameters with fewer replicates than standard non-optimal designs and they are applicable to designs where the design space is constrained (which sometimes isn't even possible with non-optimal designs). The disadvantage is that they are model dependent due to the fact that most optimal design criteria are based on some function of the information matrix, hence an optimal design that is best for one model might be inefficient for another model. Depending on model and optimization criterion, these designs can also be computationally more challenging than non-optimal designs. It is also worth noting that optimal design is a general statistical concept which has been not only applied to clinical trials, but has many areas of usage.

There are a wide variety of optimality criteria for optimal designs. We will not go through any other criteria then the ones used in the paper, but it can be mentioned that the majority of optimality criteria in use have in common that they are functions of the eigenvalues of the information matrix (which equals the inverse of the variance-matrix) and the goal is to minimize the variance of some component of the estimated parameter vector or the estimated dose-response curve. In the single- parameter case this is a simpler task since we don't have to deal with matrices. However, in this paper we will look at a multiparameter model and use two different optimality criteria for comparison. Hence we will have to deal with matrices and the problems that this may cause (ill-conditioning is one example).

For more reading about optimal designs, we refer to Silvey S.D. (1980). *Optimal Design*, which gives a short but comprehensive presentation of this area.

Adaptive designs

An adaptive design is, as already implied, a design that allows for modifications of the design during the ongoing clinical trial and it offers a possibility to improve the conventional methods used in the area of clinical trials. For example, in a typical clinical trial comparing a standard drug to a new drug, subjects are assigned to the different treatments with half of the subjects to one treatment and the other half to the other treatment. In a response-adaptive design, patients responses can be evaluated as they become available during the trial and used to reallocate the rest of the patients in the study such that the responses from different treatments can be determined as efficiently as possible .

As already mentioned, development of new drugs is expensive and time-consuming and the idea of adaptive designs are that they can help to "extract" as much information as possible and more efficiently than standard designs. The question is then; why don't we always use adaptive designs? Unfortunately, there are some drawbacks when it comes to these types of designs. One drawback is that adaptive designs usually are more complicated to design and analyze, and sometimes an adaptive design is not even applicable on the study in question. Another drawback is that unblinding (see Appendix D for explanation of the term "blinding") during the ongoing clinical trial can potentially cause bias introduced by the change of design after the analysis of the interim data has become available, which can damage the credibility of the study. To reduce this damage, results of the unblinded analysis should not be available to anyone directly operating the study.

There are a number of different adaptive designs. In our case, we will look at a two-stage response-adaptive design where we first use a smaller part of the patients enrolled in the study and assign them equally to each dose group and then adapt the allocation of patients to the dose groups after the interim data from the first stage has been analyzed.

1.2 Purpose

Since clinical trials are expensive and time-consuming, improvements of the already existing designs are always desirable. Adaptive designs have been studied for years and can in some cases be an improvement of the efficiency of the study. Unfortunately, adaptive procedures are more complicated to design and to analyze, and in some cases more difficult to apply. As mentioned in the section about optimal designs of clinical trials, optimal designs can be more complicated and computationally more challenging than non-optimal designs. We will combine these two and look at an adaptive optimal

design. The purpose with this paper is to implement the theory for a given model and to show some of the risks and complications of working with optimal designs.

2 Description of the problem

The problem formulation is based on the one used in "Optimal Designs for estimating the interesting part of a dose-effect curve" by Miller et al. (2007), the difference is that they use a Bayesian approach on the adaptive design, whereas we use a frequentistic approach. In short, we want to implement a two-stage response-adaptive optimal design for a given model using two different optimization criteria in order to obtain the optimal allocation of the subjects to the different dose groups for the second stage of the clinical trial. Since the problem formulation is based on the article by Miller et al, the investigation preceeding this paper is the planning of the same AstraZenca phase IIB dose-finding trial as in their article and for confidentiality reasons, details regarding this investigation will be left out. However, the information given in the paper should be enough for the understanding of the problem and the methods of how to solve it.

3 Method of adaptive optimization - a step-by-step procedure

The procedure of optimizing the allocation of patients to dose groups in the second stage of the trial can be compressed into a few steps. We begin by going through these steps briefly to get an outline of the method, and then work through these steps more thoroughly. Last but not least, we will implement the method on a specific model.

In short steps the procedure can be arranged as follows:

1. Choose a model for the dose-response relationship, i.e. a function $f(x_i, \theta)$ measuring the effect of the treatment where x_i is the dose group for patient $i = (1, \ldots, N)$ and θ is an unknown parameter vector.

2. Choose the sample size N.

3. Start out with $n/k = N/(3k)$ patients in each dose group, where k is the number of dose groups. This corresponds to a uniform design

$$\mathbf{w}_{unif} = (1/k, \ldots, 1/k).$$

   More generally, a design $\mathbf{w} = (w_0, \ldots, w_{k-1})$ refers to the proportion w_j of patients allocated to dose group number j. Hence the weights should satisfy $w_j \geq 0$ and $\sum_{j=0}^{k-1} w_j = 1$. When we have response data from the $n = N/3$ patients, compute an estimate $\hat{\theta}$ of the parameters $\theta = (E_0, E_{max}, \alpha, ED_{50})$. The estimation of the parameters is carried out with nonlinear least squares estimation (NLS estimation) since the model we use is nonlinear.

4. Given an optimality criterion $\Phi(\mathbf{w}, \theta)$, find the design $\mathbf{w}_{opt} = \mathbf{w}_{opt}(\hat{\theta})$ that maximizes the optimality criterion with estimated θ, $\Phi(\mathbf{w}, \hat{\theta})$. The challenge here is that this optimization is constrained, due to properties of $\mathbf{w}$. We go through these constraints in more detail below.

5. Reallocate the rest of the $N - n$ patients according to the optimal design from the previous step.

6. Collect data from the remaining $N - n$ patients, allocated to doses according to Step 5. Compute a final estimate $\hat{\theta}_{final}$ of θ using NLS estimation and all N patients.

In this paper we will focus on finding the estimated optimal design $\mathbf{w}_{opt}(\hat{\theta})$ and thus confine ourselves to Steps 1-5. We will also use simulated data, drawn from the model, in Step 3. Thus we are able to repeat Steps 1-5 a large number of times, and may so infer the distribution of the estimated optimal design.

3.1 The model

As mentioned in the section about optimal designs, these types of designs require a model to be specified. The model that we use to describe the dose-response curve was chosen by pharmacometric modellers from AstraZeneca based on results from pre-clinical trials, phase I clinical trials and on data from a drug with similar characteristics as the investigational drug. The model that they chose is called an E_{max}-sigmoid model and the model function is:

$$Y_i = f(x_i, \theta) + \epsilon_i = E_0 + \frac{E_{max} x_i^{\alpha}}{ED_{50}^{\alpha} + x_i^{\alpha}} + \epsilon_i, \qquad i = 1, \ldots, N \tag{1}$$

where

- N is the total number of patients enrolled in the trial, Y_i is the response for patient i, the set of parameters $\theta = (E_0, E_{max}, \alpha, ED_{50})$ and normally distributed, independent, random errors ϵ_i with mean 0 and standard deviation σ.

- the response is a continuous variable measured in some unit where higher values indicate a positive effect.

- dose (measured in mg) $x = 0$ is the placebo dose and $x = 100$ is the highest dose (any dose over 100 mg was considered to have a too high safety risk) determined by pharmacometric modellers based on the same information that the choice of model is based on. The reason why only doses $x \in \{0, 20, 40, 60, 80, 100\}$ are possible is that

the available manufactured tablets only are available in multiples of 20 mg.

- E_0 is the effect of the placebo dose.
- E_{max} is the maximum effect for a very high dose compared to a placebo dose.
- α is a parameter that alternates the shape of the curve.
- ED_{50} is the dose with half of the maximum effect.

3.2 Choice of sample size

In clinical trials the sample size is usually calculated based on statistical power analysis. This is used to calculate the minimum sample size N required to obtain a test with a pre-specified significance level and at least some given power for a certain assumed effect. This is a very important part of a clinical trial since a too small sample size might not give reliable results and a too large sample size might be too expensive (or impossible). It is well known that larger sample size increase statistical power, but in the area of clinical trials large sample sizes are more exceptional.

However, power analysis goes beyond the scope of this paper and we consider three different sample sizes, $N = 360$, $N = 720$ and $N = 1440$, as given. Looking at three different sample sizes gives us an idea of how the optimal design changes as we increase the sample size. Naturally, we would get a more reliable asymptotic result if we increased the sample size even more, but in most practical cases this wouldn't be feasible.

3.3 Initial estimation of parameters

Now that we have our sample size N we can estimate the parameters $\theta = (E_0, E_{max}, \alpha, ED_{50})$. Since the model that we use is nonlinear we must use a nonlinear estimation method and the one we will use is the common nonlinear least squares estimation method (NLS estimation). The estimated parameters in this adaptive design are only made for a third of the entire sample size in this first stage, as mentioned in the quick step-by-step preview. Before we move on, let's have a look at the theory of NLS estimation.

3.3.1 Nonlinear Least Squares estimation (NLS estimation)

Assume that we have n observations $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$ from a nonlinear model where y_i is an observation from

$$Y_i = f(x_i, \theta) + \epsilon_i, \qquad i = 1, ..., n,$$

$E[\epsilon_i] = 0$, x_i is a predictor and θ is a m-dimensional vector of parameters where the true θ belongs to $\Theta \subseteq \Re^m$ and $n \geq m$. Define the residuals

$$r_i(\theta) = y_i - f(x_i, \theta).$$

The least squares estimate of θ, denoted by $\hat{\theta}$, minimizes the residual sum of squares

$$S(\theta) = \sum_{i=1}^{n} r_i(\theta)^2 \qquad \theta \subseteq \Theta. \tag{2}$$

Assuming ϵ_i to be independent and identically distributed with variance σ^2 it can be shown that, in addition with some other regularity conditions on $f(\cdot\,, \theta)$, $\hat{\theta}$ is a consistent estimate of θ and that $\hat{\theta}$ is asymptotically normally distributed when $n \to \infty$. If we also assume that ϵ_i are normally distributed variables, then $\hat{\theta}$ is the maximum likelihood estimate (ML estimate) of θ. For a proof of consistency and asymptotic normality, we refer to Seber and Wild (1989).

Differentiating $S(\theta)$ with respect to the jth component θ_j of θ

$$\tfrac{\partial S(\theta)}{\partial \theta_j} = 2\sum_{i=1}^{n} r_i(\theta)\tfrac{\partial r_i}{\partial \theta_j}, \quad j = 1, \ldots, m,$$

gives us the gradient functions of the model. Setting the gradients equal to zero

$$\tfrac{\partial S(\theta)}{\partial \theta_j} = \mathbf{0}, \quad j = 1, \ldots, m,$$

gives us the *normal equations* for the parameters, but solving these equations is not always easy. Due to the fact that these equations in general are nonlinear, they oftenly do not have an explicit solution and thus we need iterative methods. Most statistical packages have these functions built in, ready to use, but a tricky part is that they generally demand initial guesses (or starting values) of the values of the parameters. Different methods are more or less sensitive to starting far away from the values that minimize the function, but guessing values as close as possible is always an advantage since it requires less iterations (and hence less computational power).

Estimation of the parameters in our case was initially made with nlin in Matlab, which uses the Levenberg-Marquardt method, but due to problems that we will bring up later we changed method to a derivative-free one that uses a version of an algorithm called the Nelder-Mead algortihm (also called the simplex search algorithm), namely fminsearch in Matlab. For information on how this algorithm works, see Appendix A.

3.4 Finding the optimal design

Once the estimators of the parameters are available, we can use them to search for an optimal design for the second part of the dose-finding trial. In optimal design theory, a design is optimal with respect to some optimality criterion, as mentioned in the section about optimal designs. In this paper we look at two different criteria, both of these depending on a function $d(x, \mathbf{w}, \theta)$. This function in turn is proportional to the asymptotic variance of the nonlinear least squares estimate $f(x, \hat{\theta}) - f(0, \hat{\theta})$ of the dose-response function increment $f(x, \theta) - f(0, \theta)$, given a design $\mathbf{w}$ and parameter vector θ, as $n \to \infty$. These estimates are asymptotically unbiased (see Seber and Wild (1989) for a proof), hence the value of the function d says something about the precision of the estimate and it can be shown that the function depends sensitively on the design $\mathbf{w}$. Consequently, to obtain the optimal design we minimize this estimated variance. To learn the structure of this function d, see Appendix B.

Now, let's look at the two different criteria that we will use.

Criterion 1: This criterion is an alternated version of Criterion 1 in Miller et al. (2007). The criterion determines the design $\mathbf{w}$ such that the function

$$\Phi_1(\mathbf{w}, \theta) = \left(\sum_x d(x, \mathbf{w}, \theta) \right)^{-1} \tag{3}$$

is maximized. This optimality criterion is used to produce a design that best can answer questions regarding the estimated dose-response compared to placebo dose $x = 0$ for doses

$$x \in \{\frac{x_{max}}{k-1}, \frac{2x_{max}}{k-1}, \ldots, x_{max}\}$$

up to the maximal tolerable dose (in our case 100 mg).

Criterion 2: This criterion is the same as Criterion 2 of Miller et al. (2007),

$$\Phi_2(\mathbf{w}, \theta) = d(x_{max}, \mathbf{w}, \theta)^{-1} \tag{4}$$

In the same way as in Criterion 1, this function gives the optimal design when it is maximized. It is a somewhat simpler criterion since it is used to produce a design that answers questions regarding the estimated dose-response at the maximal tolerable dose compared to placebo, hence it only looks at a single dose compared to placebo.

The inverse of the criteria, $1/\Phi(\mathbf{w}, \theta)$, can be minimized using any minimization method without particular difficulties. Problems however arise when we have a constrained design space (i.e. when the weights have boundaries), since constrained optimization is quite complicated. In our case, the weights have a lower and a upper boundary, complicating the optimization. More about this under "Implementation and results" .

3.5 Reallocating patients using the optimized design

Reallocating the patients is actually quite straightforward, it is just simple calculus:

1. Calculate the number of patients in each dose group with the new optimized weights, for the entire sample size N. Let the optimized weights be $\mathbf{w}_{opt} = \mathbf{w}_{opt}(\hat{\theta})$, then

   $$\mathbf{n} = (n_0, \ldots, n_{k-1}) = N \cdot \mathbf{w}_{opt}$$

   where n_j is the total number of patients assigned to the jth dose.

2. Round off so that the sum of the elements in $\mathbf{n}$ is equal to N. Denote the rounded off number of patients in each dose group by $\tilde{\mathbf{n}}$.

3. Subtract the $N/3k$ patients that were allocated equally to each dose group in the first stage. We then get

   $$\mathbf{n}_{opt} = \tilde{\mathbf{n}} - \tfrac{N(1,\ldots,1)}{3k}$$

   as our optimized number of patients in each dose group for the second part of the study.

4 Implementation and results

4.1 Implementation

Now that we have an outline of the method we can apply it on the E_{max}-sigmoid model (1) with $\sigma = 10$. As mentioned earlier, we consider sample size as given and we study three different sample sizes; $N = 360$, $N = 720$ and $N = 1440$.

We begin by estimating the parameters for $n = N/3$ patients equally allocated to the $k = 6$ dose groups. This consequently gives us $n/6$ patients in each dose group and an initial design

$$\mathbf{w}_{unif} = \begin{pmatrix} 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{pmatrix} \tag{5}$$

In order to estimate the parameters in the model with the NLS method, we need initial guesses of the values of the parameters. The seven different combinations (here called scenarios) of parameter values in the table below were suggested by experts from AstraZeneca. Observe that we hereafter will call the parameter values θ from the scenarios the "true" θ, despite that we actually do not know the true values. However, the scenarios are suggested by experts and they are the best guesses we have, hence we will call them the true parameter values since we simulate data from this model.

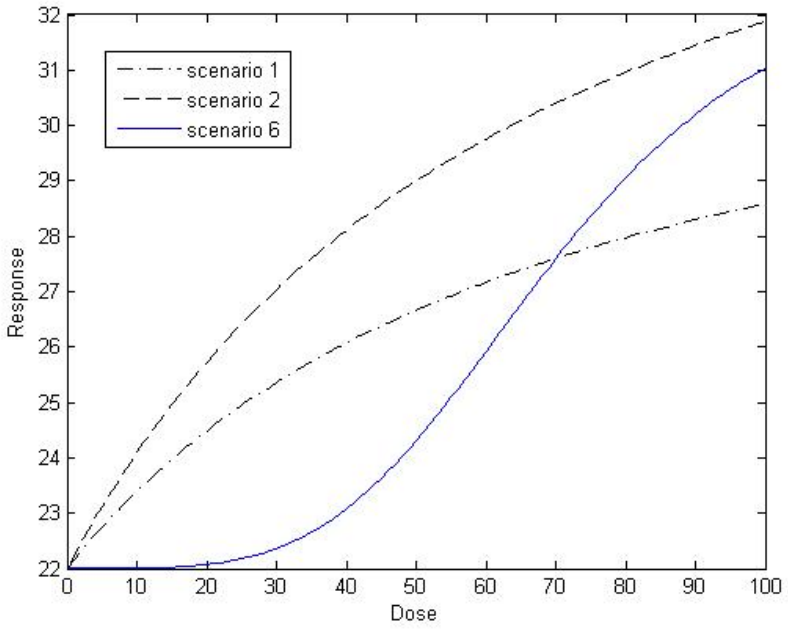
Table 1: Scenarios for values of the parameters

Scenario	E_0	E_{max}	α	ED_{50}
1. Prior guess	22	11.2	1	70
2. High E_{max}	22	16.8	1	70
3. Low ED_{50}	22	11.2	1	35
4. High ED_{50}	22	11.2	1	200
5. Intermed α	22	11.2	2	70
6. High α	22	11.2	4	70
7. Low E_{max}, ED_{50}	22	7.0	1	35

For more information about the reasoning regarding why these particular values were chosen, we refer to Miller et al. (2007). Although all of these scenarios are interesting we will not study all of them (due to time limitations), but settle for the most interesting scenarios 1, 2 and 6. Before

proceeding with parameter estimation it is rewarding to have a look at how the dose-response curves look like for these three scenarios.

Figure 1: Dose-response curve for three different scenarios



From this it is easy to see how the parameter values affect the shape of the curve in these specific scenarios; a higher E_{max} pushes the curve upwards and a higher alpha makes the curve more s-shaped. The latter should be observable in the optimized design; it should put a larger part of patients in the higher dose groups.

The estimation of the parameters was initially made with nlin in Matlab, which uses the Levenberg-Marquardt algorithm (which in turn uses the Jacobian matrix for estimation). However, it quickly became obvious that we had a problem. There was a very large variation in estimated parameter values (mainly in $\widehat{E_{max}}, \hat{\alpha}$ and $\widehat{ED_{50}}$) between different runs and in some runs the algorithm stopped due to an ill-conditioned Jacobian, we therefore tried the more robust and derivative-free Nelder-Mead algorithm. This algorithm doesn't use gradient functions but instead searches directly for the minimum value of the residual sum of squares $S(\theta)$ in (2). This naturally removed the problem with an ill-conditioned Jacobian and the algorithm runned smoothly. However, the estimated values of the parameters still had a large variation between different runs. It was also discovered that one of the estimated parameters, $\widehat{ED_{50}}$, in some runs took on negative values,

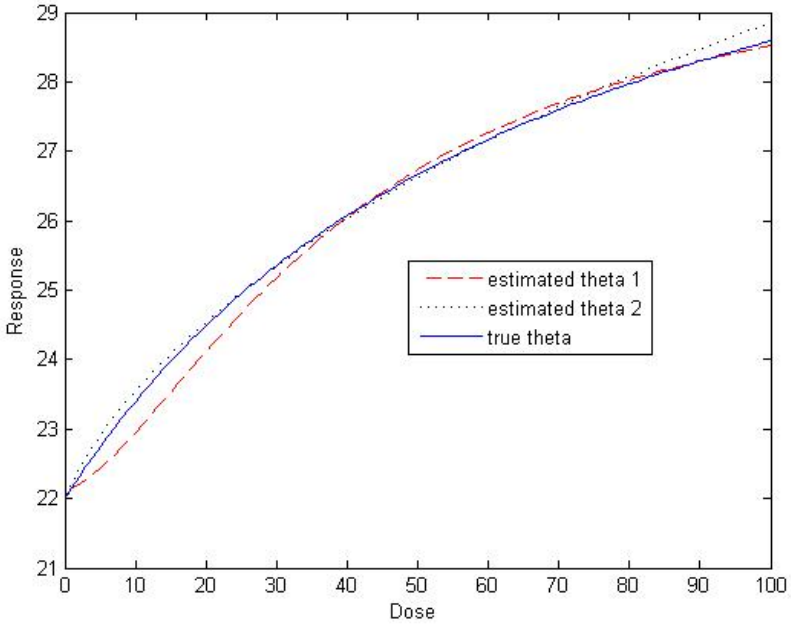
causing the estimated response to be imaginary valued. An interesting feature was that, although different sets of estimated parameters were quite different, they gave almost the same estimated dose-response curve. The main reason for this is that the estimated parameters that had the largest variation were $\widehat{E_{max}}$ and $\widehat{ED_{50}}$, and looking at the model function one can see that if these two increase or decrease simultaneously, it will not affect the curve very much. Below is an example with two different data sets, both simulated from the sigmoid model (1) with $n = 360/3 = 120$ patients, true parameter vector θ corresponding to Scenario 1 and normally distributed errors ϵ_i with $\sigma = 10$. The corresponding parameter estimates $\hat{\theta}_1$ and $\hat{\theta}_2$ were both obtained with the NLS method using the true θ as starting point for the iterations. The two runs gave quite different estimated parameter values and despite that, they gave similar estimated dose-response curves.

Table 2: Estimated parameter values for two different runs and the true θ

	E_0	E_{max}	α	ED_{50}
True θ	22	11.2	1	70
$\hat{\theta}_1$	22.1368	7.9986	1.5447	41.0948
$\hat{\theta}_2$	21.9264	21.3534	0.7650	261.4315

Below is a figure that illustrates that the estimated dose-response curves aren't all that different for these sets of parameters.

Figure 2: Estimated dose-response curve for three different sets of parameters



The problem with large variation in estimated parameter values wouldn't have been a big issue if we had been interested in the behavior of the dose-response curve. However, we need the estimated parameter values to calculate the optimal design (via the gradients, see appendix B for a refreshment), and if the estimated values of the parameters have a large variation it will cause a large variation in optimal designs. Therefore, we had to find the reason or reasons for these large variations. Digging a bit deeper into what the problem was, the situation was diagnosed to depend on either

- overparametrization of the model, or
- a too large standard deviation of the random errors

A standard deviation of $\sigma = 10$ can cause negative simulated responses, which is impossible in reality. The only way to find out which ones of these (or perhaps both) were the problem, was to repeat the simulations with one parameter fixed and/or a reduced σ. It turned out that both of these facts caused large variation in estimated optimal design. Reducing σ to 1 removed the problem with negative simulated responses, but we still had a large variation in parameter estimates between different runs. From this we were also able to draw the conclusion that increasing the sample size N would have almost the same effect as decreasing σ , since the precision of the estimates is *proportional* to $\sigma/\sqrt{N}$ (see Seber and Wild, 1989). Letting one parameter be fixed, α in our case, and keeping $\sigma = 10$ reduced the variation but the problem with negative simulated responses naturally remained.

For the sake of comparison, we have looked at the following combinations (from now on called models) for each of the two criteria and for the sample sizes $N = 360$, $N = 720$ and $N = 1440$.

Table 3: Four possible models

Model	Number of unknown parameters	Standard deviation σ
1	4	10
2	4	1
3 (α fixed)	3	10
4 (α fixed)	3	1

In Models 3 and 4, the value of α is fixed to 1 if we look at Scenario 1 or 2, and fixed to 4 if we look at Scenario 6.

Once we had our estimated parameters we used these to calculate the optimal design $\mathbf{w}_{opt} = \mathbf{w}_{opt}(\hat{\theta})$ for the second part of the trial, using our two criteria and optimization theory. There are a wide variety of ready-built functions for solving optimization problems without constraints in statistical packages, but as soon as we have a constrained parameter space it complicates things since most of these functions cannot handle constraints. As mentioned earlier, in our case we had constraints on the design space, namely that the proportion of patients in each dose group w_j had to lie between

$$\frac{1}{18} \leq w_j \leq \frac{13}{18} \qquad \text{and} \qquad \sum_{j=1}^{6} w_j = 1 \tag{6}$$

where $j = 1, 2, 3, 4, 5, 6$ correspond to doses $0, 20, 40, 60, 80, 100$. The lower boundary is the minimum proportion of patients for a dose group in the second part of the trial and it origins from the fact that we have already assigned a third of the patients to each dose group in the first part of the trial, according to the initial design $\mathbf{w}_{unif}$ (5). Rewriting the number of patients that are allocated to various doses in the first step as

$$n\mathbf{w}_{unif} = \begin{pmatrix} N/18 & N/18 & N/18 & N/18 & N/18 & N/18 \end{pmatrix},$$

one can more easily see that the smallest value w_j can take is $1/18$. The upper boundary is calculated as

$$1 - (5/18) = 13/18$$

where 5/18 is the sum of the other five weights when they take the smallest value they can take (i.e. the lower boundary). Last but not least, the sum of the weights naturally have to equal one.

The optimization was carried through with the following two methods.

Method 1: This method is to use the ready-built function fminsearch in Matlab with the function to be minimized inside an additional function, where the additional function controls that the constraints are fulfilled. In our case the function to be minimized is the criterion $1/\Phi$, and this is called upon in a function that adds a large "punishment" to the value of $1/\Phi$ whenever

$$\min(\mathbf{w}) < \tfrac{1}{18} \qquad \text{and/or} \qquad \max(\mathbf{w}) > \tfrac{13}{18}$$

This additional "check-function" also controls that the sum of the weights equalled one. Hence this function forces the algorithm to search within the chosen boundaries at the same time as it minimizes $1/\Phi$.

Method 2: The second algorithm we built ourselves. The main idea is to change one coordinate of $\mathbf{w}$ at a time in a cyclic manner, with a geometrically decreasing step size. If $\mathbf{e}_j$ is a unit vector with a one in position j and zeros elsewhere we replace $\mathbf{w}$ by

$$\mathbf{w}' = (\mathbf{w} + \Delta w_j \mathbf{e}_j)/(1 + \Delta w_j) \tag{7}$$

as the current best design if the criterion function $1/\Phi$ decreases, i.e. $1/\Phi(\mathbf{w}', \hat{\theta}) < 1/\Phi(\mathbf{w}, \hat{\theta})$. In addition, $\mathbf{w}'$ must be a valid two-stage design, i.e. satisfy the boundary conditions imposed by (6). We run the algorithm for a pre-specified number of steps m_{max} where one "sweep" or step of the algorithm consists of iterating (7) for $j = 1, \ldots, 6$. The step size is

$$\Delta w_j = \Delta w_0 \times \rho^m$$

if (7) belongs to sweep number $m = 0, 1, \ldots, m_{max}$ and where ρ^m is some change factor that geometrically decreases the step size for each sweep number m. In other words; the step size is constant *within* a sweep but decreases with each sweep. The algorithm stops when Δw_j reaches a pre-specified stop value.

This algorithm is a modified first order exchange algorithm and for more

reading about the original algorithm we refer to Atkinson A.C. and Donev A.N. (1992), *Optimum Experimental Designs* in the reference list.

We let Method 2 use the four different combinations of starting step size Δw_0, stopping step size $\Delta w_{m_{max}}$ and the number of steps according to the table below. Then we let the program chose the combination that produced an optimal design with the smallest value of the criterion.

Table 4: Four combinations of Δw_0, $\Delta w_{m_{max}}$ and m_{max} for Method 2

Starting step size Δw_0	Stopping step size $\Delta w_{m_{max}}$	Number of steps m_{max}
0.1	10^{-5}	100
0.05	10^{-5}	100
0.07	10^{-5}	100
0.09	10^{-5}	100

In the same way, we let the program choose between Method 1 and Method 2 based on the value of the criterion. In approximately 90 percent of the times the program chose the optimal design produced by Method 2. This is most likely due to the fact that Method 2 is "tailor-made" for this particular problem and hence gives a better result. Although we made the constrained problem unconstrained by the additional function in Method 1, fminsearch is built for unconstrained optimization and the additional function might not remove as much of the problem as we would like it too.

Next step was then to reallocate the rest of the patients using our optimal design. The reallocation of patients to obtain the optimal number of patients $\mathbf{n}_{opt}$ for the second part of the trial was made according to the steps in section 4.5. Apart from one small challenge, this was just straightforward calculus. The minor challenge here was to control the rounding off in step 2, since for some optimal designs the sum of the elements in $\mathbf{n}$ equaled $N + 1$ or $N - 1$ instead of N. Whenever this occurred, we corrected it with a function that added or subtracted 1 to the element in $\mathbf{n}$ that was closest to be rounded off in the opposite direction.

The final part was then to repeat the program a large number of times. Due to time limitations and the fact that the program took some time to run (optimization is usually quite expensive in computational power which affects the speed of the program), we settled with 500 iterations.

4.2 Results

4.2.1 Fitted response curves

Finally, let's have a look at the most interesting part; the results. The presentation of the results will be made using boxplots with explanatory text. The first boxplots illustrates the behavior of the estimated dose response curves $f(\cdot\ ;\hat{\theta})$ that we observed in figure 2, but for more iterations than two. The boxplots below show the estimated responses given by (1) with 500 simulated sets of $\hat{\theta}$ and the dose-response curve for the true θ, for all the four models in Table 3, Scenario 1 and with total sample size $N = 360$, hence $n = 120$ data points are used for computing $\hat{\theta}$. We choose not to show these plots for a larger N since the effect of increasing sample size was almost the same as the effect of reducing σ. Also, we choose not to show plots for scenario 2 and scenario 6 since they basically show the same thing as the plots for scenario 1, but for the interested reader these plots can be studied in Appendix C. Observe that which criterion we choose here has no significance, since these have no impact on the initial estimate $\hat{\theta}$ of the parameters.

Figure 3: Boxplot of estimated mean responses, 500 simulations, Model 1, Scenario 1, $N = 360$

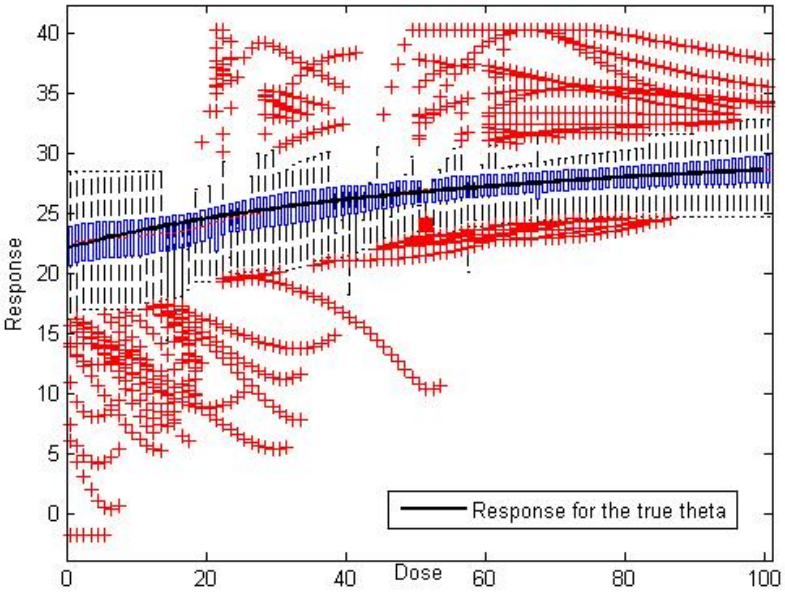


Figure 4: Boxplot of estimated mean responses, 500 simulations, Model 2, Scenario 1, $N = 360$

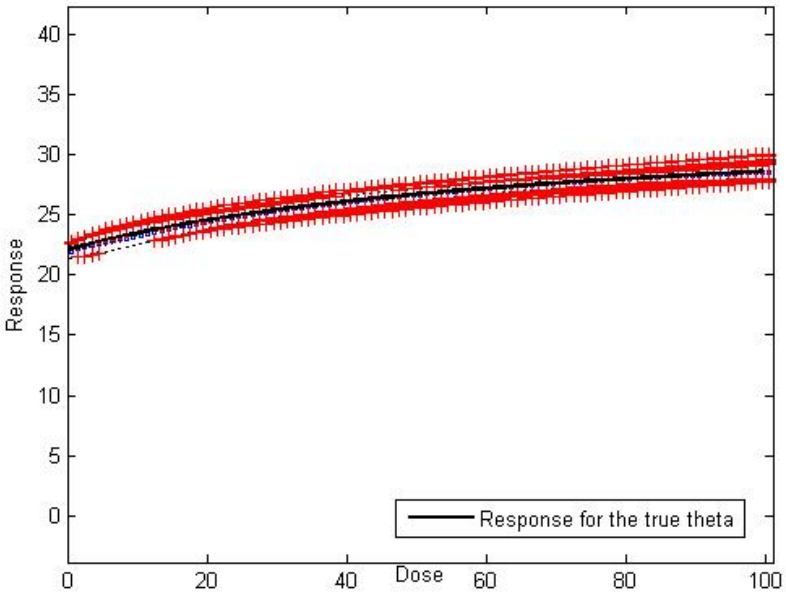


Figure 5: Boxplot of estimated mean responses, 500 simulations, Model 3, Scenario 1, $N = 360$

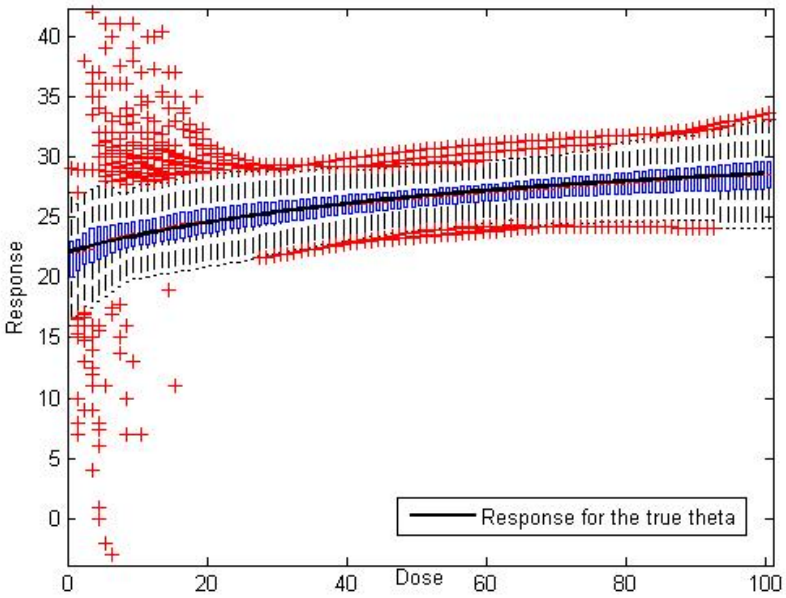
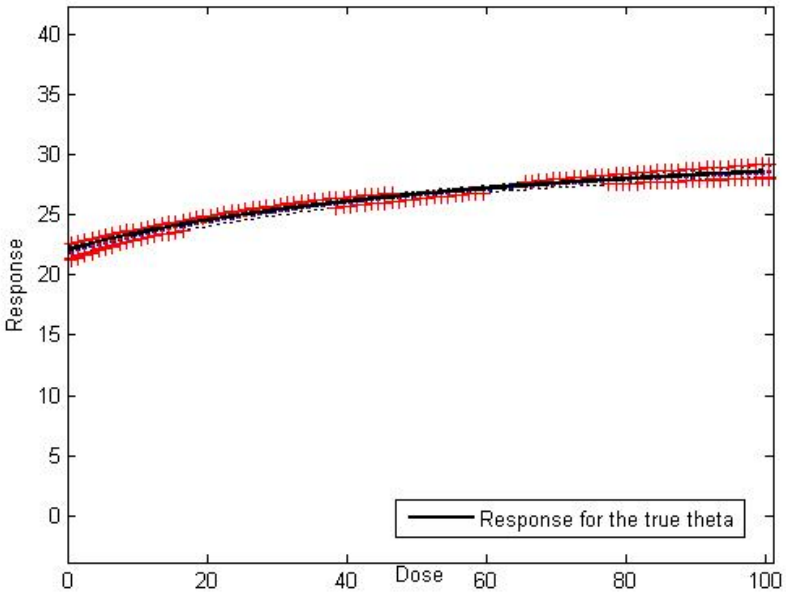


Figure 6: Boxplot of estimated mean responses, 500 simulations, Model 4, Scenario 1, $N = 360$



The reason to the odd shape of the boxplots in Figure 3 is that the estimated parameter $\widehat{ED_{50}}$ has taken on negative values, causing estimated responses to be imaginary valued. However, this will not interfere with the result of which we are interested in showing since we are mainly interested in showing that there *is* a large variation, rather than the pattern of the variation. The important thing to observe in these plots is that although the variation of the estimated responses in some cases is large (which depends on a large variation in $\hat{\theta}$), the boxplots are still narrow around the true response curve, which was what we showed for two iterations in figure 2. The smallest variation of the estimated responses is for Model 4 and we can expect that the smallest variation in optimal designs will be found for this model too.

4.2.2 Estimated optimal designs

Now, let us consider the true optimal designs $\mathbf{w}_{opt}(\theta)$ and the estimated optimal designs $\mathbf{w}_{opt}(\hat{\theta})$ produced by the same set of simulated $\hat{\theta}$ that was used to calculate the estimated responses above.

Below are boxplots for 500 simulated optimal weights for Models 1 and 4 (the two "extreme" ones), both of the criteria, Scenario 1 and $N = 360$, $N = 720$ and $N = 1440$. The reason that we will not show plots for the other scenarios is the same as above, and again these plots can be found in Appendix C. Worth to observe in the plots below is that both the criteria produces optimal designs with a large proportion of patients in the placebo dose group, which is important since we are interested in the differences between the placebo dose and all other doses.

Figure 7: Boxplot of estimated optimal weights, 500 simulations, Criterion 1, Model 1, Scenario 1, $N = 360$

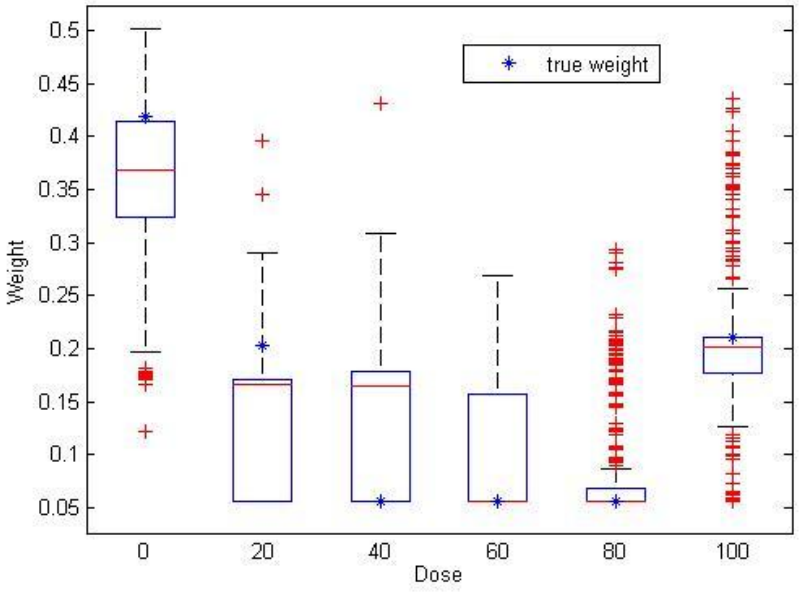


Figure 8: Boxplot of estimated optimal weights, 500 simulations, Criterion 1, Model 1, Scenario 1, $N = 720$

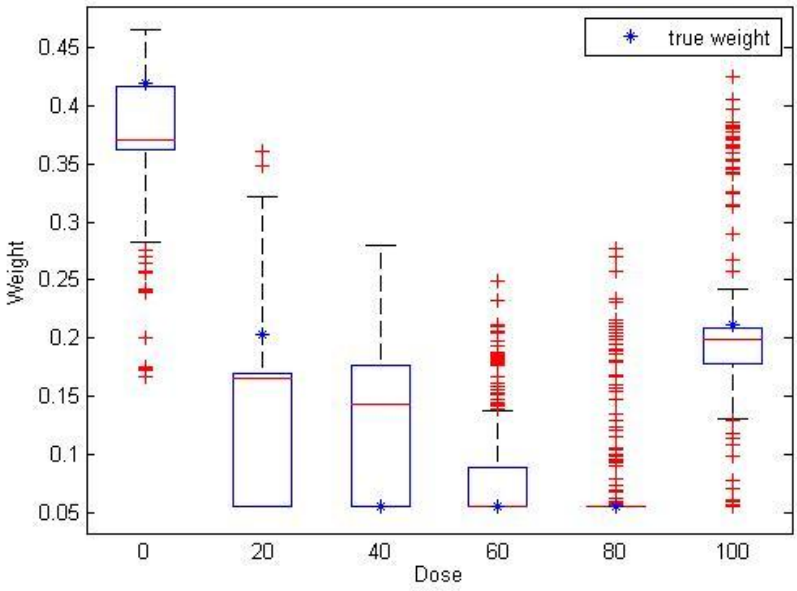


Figure 9: Boxplot of estimated optimal weights, 500 simulations, Criterion 1, Model 1, Scenario 1, $N = 1440$

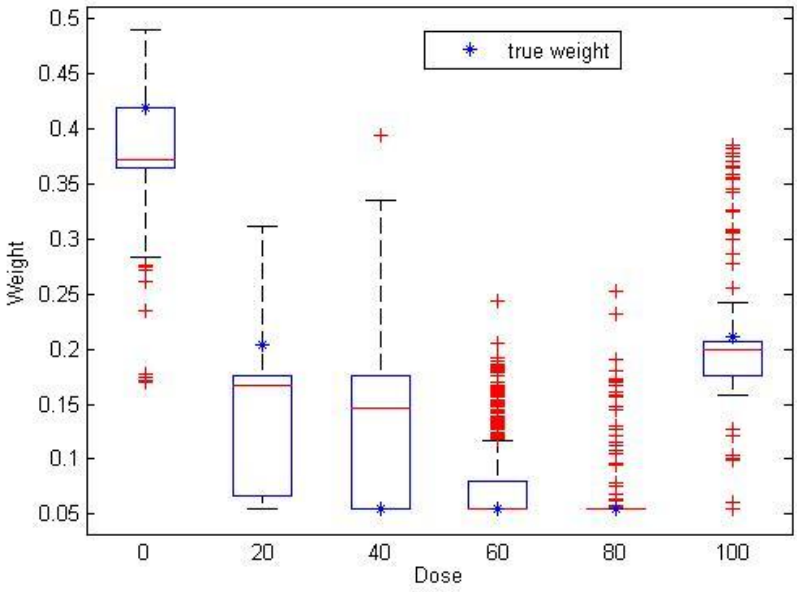


Figure 10: Boxplot of estimated optimal weights, 500 simulations, Criterion 1, Model 4, Scenario 1, $N = 360$

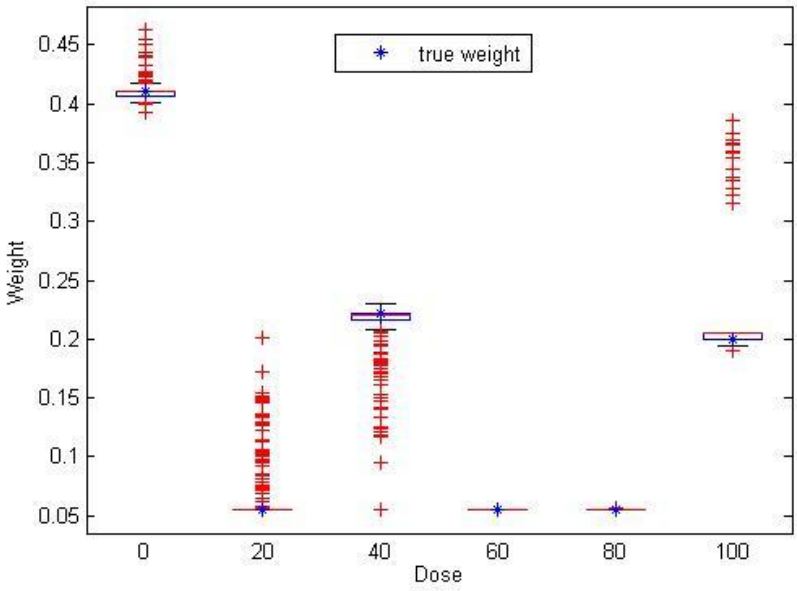


Figure 11: Boxplot of estimated optimal weights, 500 simulations, Criterion 1, Model 4, Scenario 1, $N = 720$

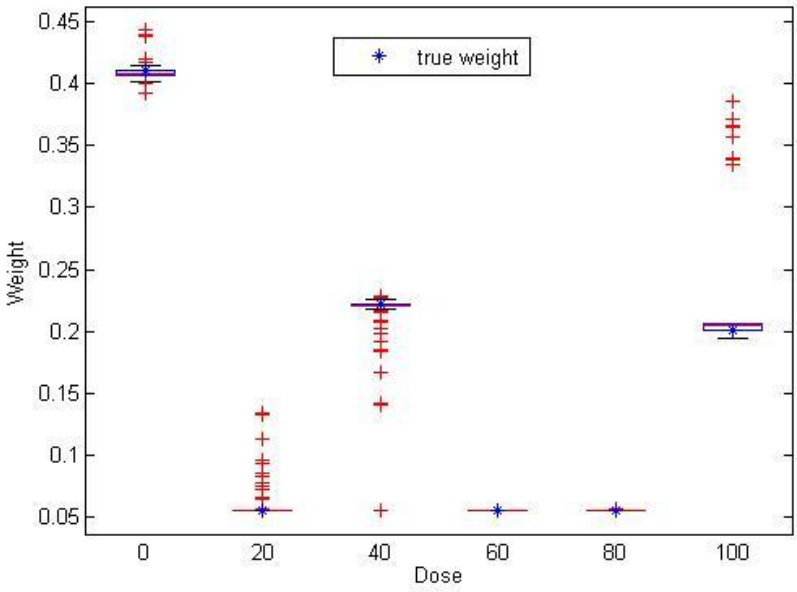


Figure 12: Boxplot of estimated optimal weights, 500 simulations, Criterion 1, Model 4, Scenario 1, $N = 1440$

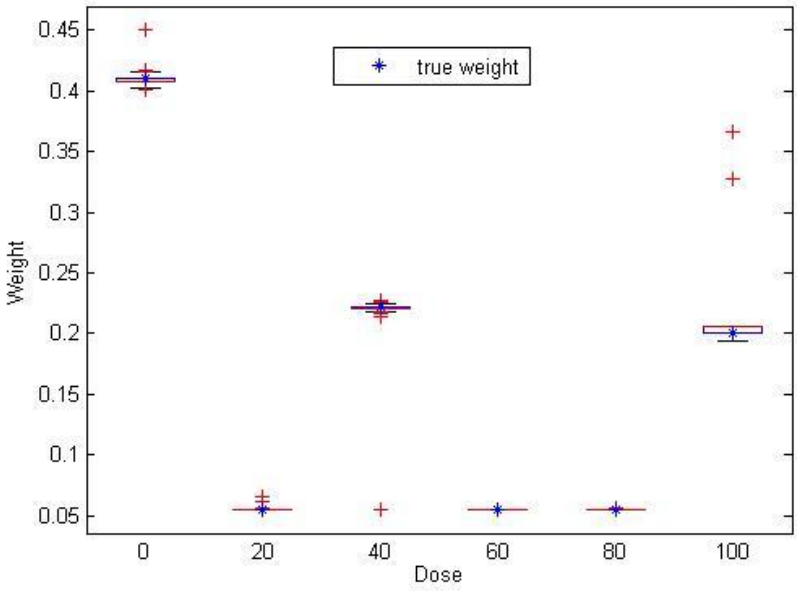


Figure 13: Boxplot of estimated optimal weights, 500 simulations, Criterion 2, Model 1, Scenario 1, $N = 360$

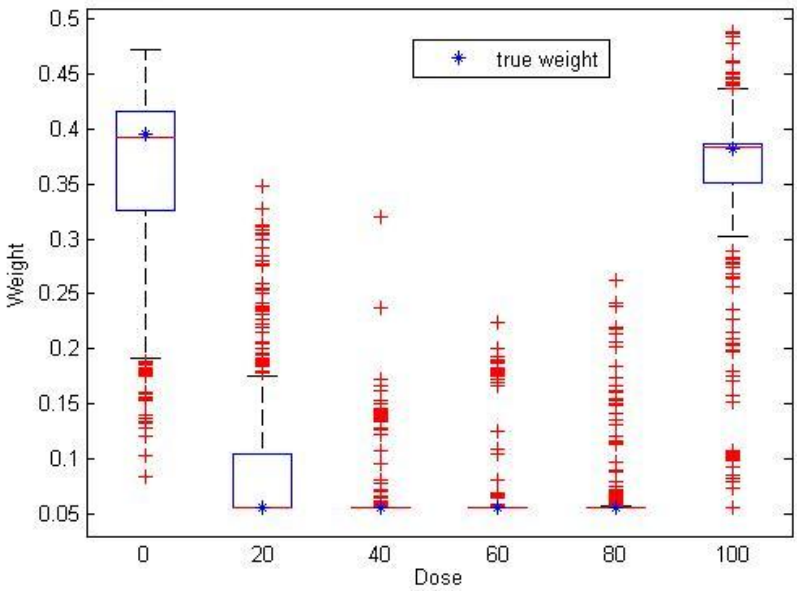


Figure 14: Boxplot of estimated optimal weights, 500 simulations, Criterion 2, Model 1, Scenario 1, $N = 720$

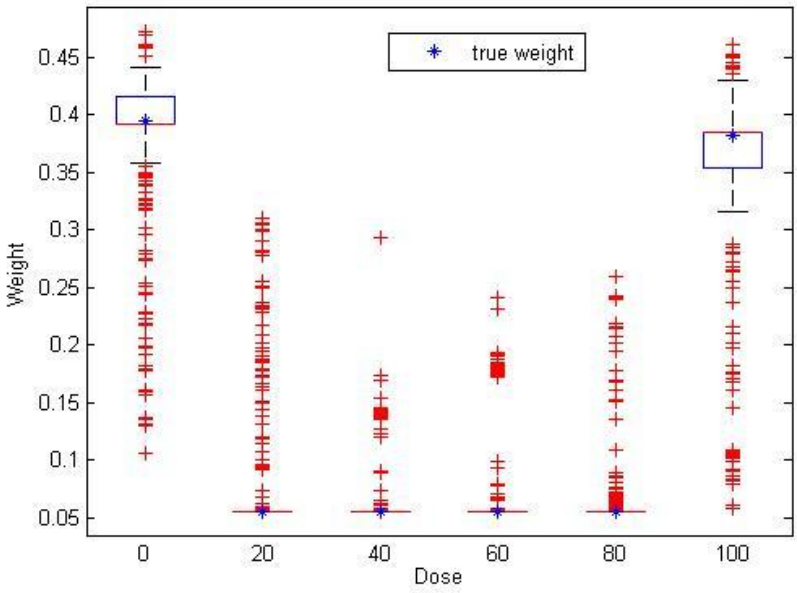


Figure 15: Boxplot of estimated optimal weights, 500 simulations, Criterion 2, Model 1, Scenario 1, $N = 1440$

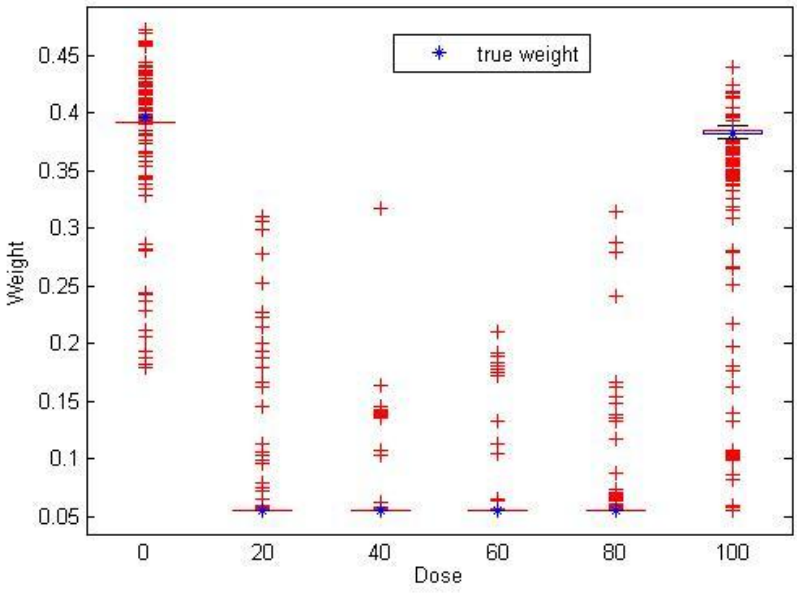


Figure 16: Boxplot of estimated optimal weights, 500 simulations, Criterion 2, Model 4, Scenario 1, $N = 360$

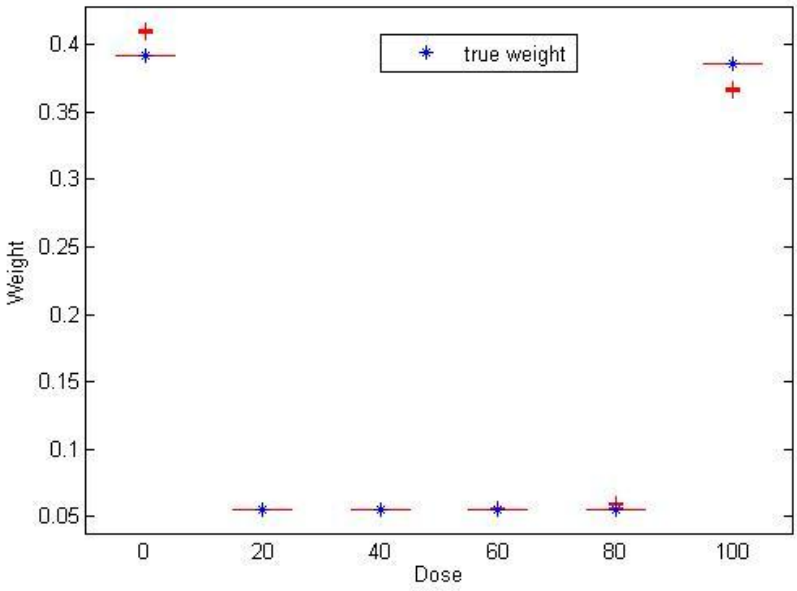


Figure 17: Boxplot of estimated optimal weights, 500 simulations, Criterion 2, Model 4, Scenario 1, $N = 720$

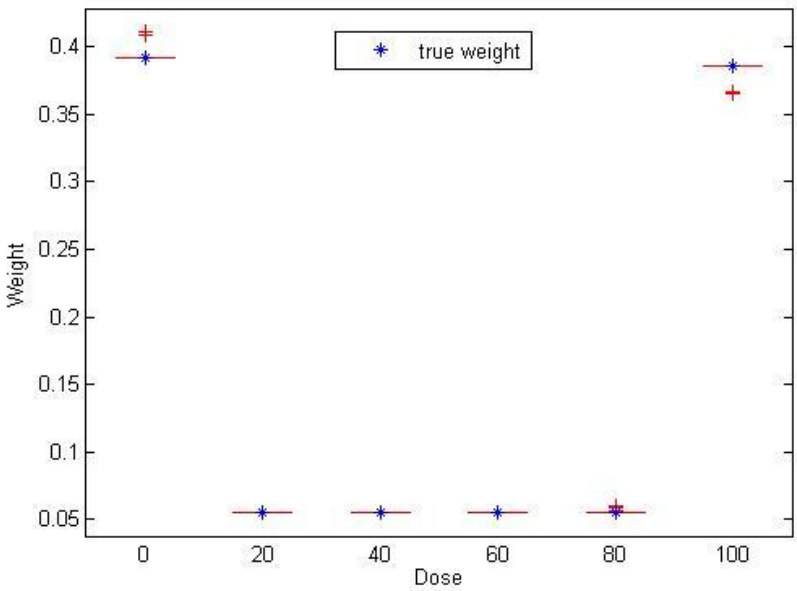
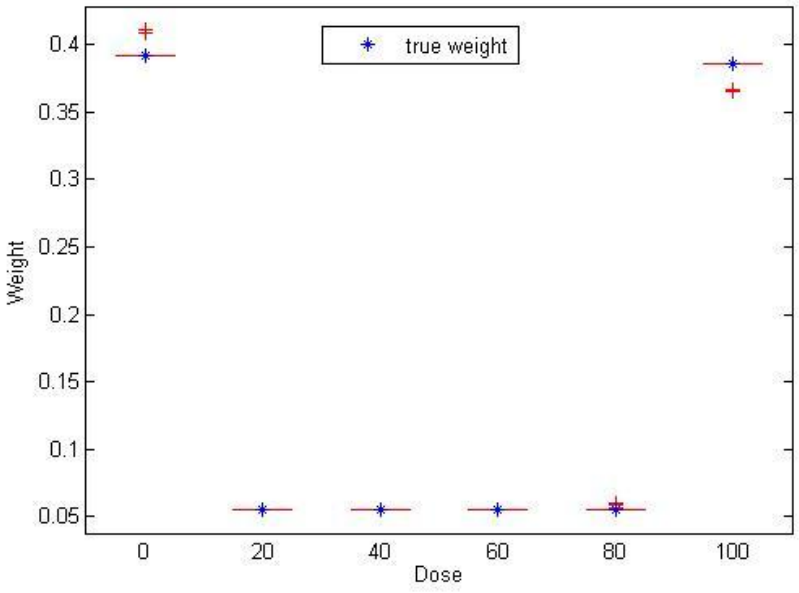


Figure 18: Boxplot of estimated optimal weights, 500 simulations, Criterion 2, Model 4, Scenario 1, $N = 1440$



Part of the reason to why we only show the two extreme models is that the difference between them gives a quite clear result; large variation for estimated parameters causes a large variation in estimated optimal designs. As we saw in the Figure 3 and 6, the variation for estimated responses, and thus the variation in parameter estimates, was much smaller in Model 4 than in Model 1. Since our optimality criteria Φ_1 and Φ_2 both reflect the efficiency of estimating a dose-response curve increment rather than the parameters themselves, it is natural that the estimated optimal design varies in a similar way as the estimated response curves. Hence there is a larger variation in estimated optimal design in Figures 7-9 compared to Figures 10-12, and in Figures 13-15 compared to Figures 16-18.

Worth noting in these plots is the difference in the estimated optimal designs between the criteria, which is easiest to observe in Figures 10-12 and Figures 16-18 (i.e. for Model 4). Criterion 2 produces estimated optimal designs with a larger proportion of the patients in the placebo dose group and the group given the maximal dose. This is not so surprising, since this criterion is used to produce a design that answers questions regarding the estimated dose-response at the maximal tolerable dose compared to placebo. Criterion 1, however, produces estimated optimal designs that has a larger proportion of patients in three dose groups; $x = 00$, $x = 40$ and $x = 100$. Looking at the description of the criterion in section 3.4, this is not so sur-

prising either since the criterion is designed to produce a design that best can answer questions regarding the estimated dose-response compared to placebo dose $x = 0$ for doses all doses up to the maximal tolerable dose.

4.2.3 Estimated criterion values

Last but not least, let's have a look at the estimated criterion values for both criteria, Model 1 and Model 4, $N = 360$, $N = 720$, $N = 1440$ and all three scenarios. In the boxplots below, the asterisk represents the criterion value based on the true θ, $1/\Phi(\mathbf{w}_{opt}, \theta) = 1/\Phi(\mathbf{w}_{opt}(\theta), \theta)$, and the boxplots shows the estimated criterion values based on $\hat{\theta}$, $1/\Phi(\mathbf{w}_{opt}, \hat{\theta}) = 1/\Phi(\mathbf{w}_{opt}(\hat{\theta}), \hat{\theta})$, for 500 simulations.

Figure 19: Boxplot of estimated values of $1/\Phi_1$, 500 simulations, Model 1

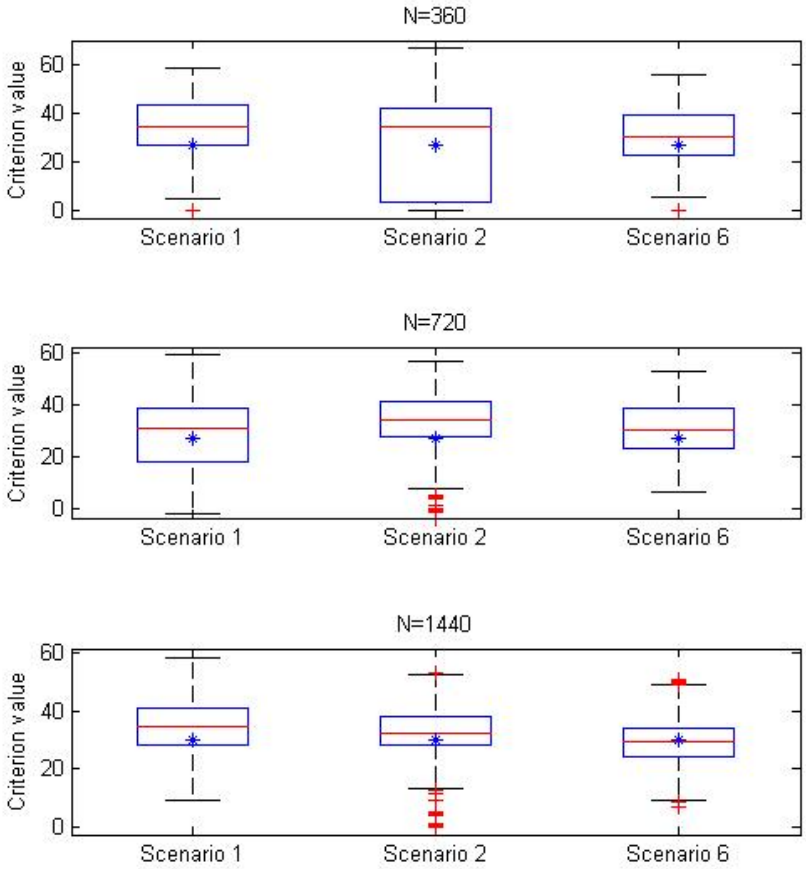


Figure 20: Boxplot of estimated values of $1/\Phi_1$, 500 simulations, Model 4

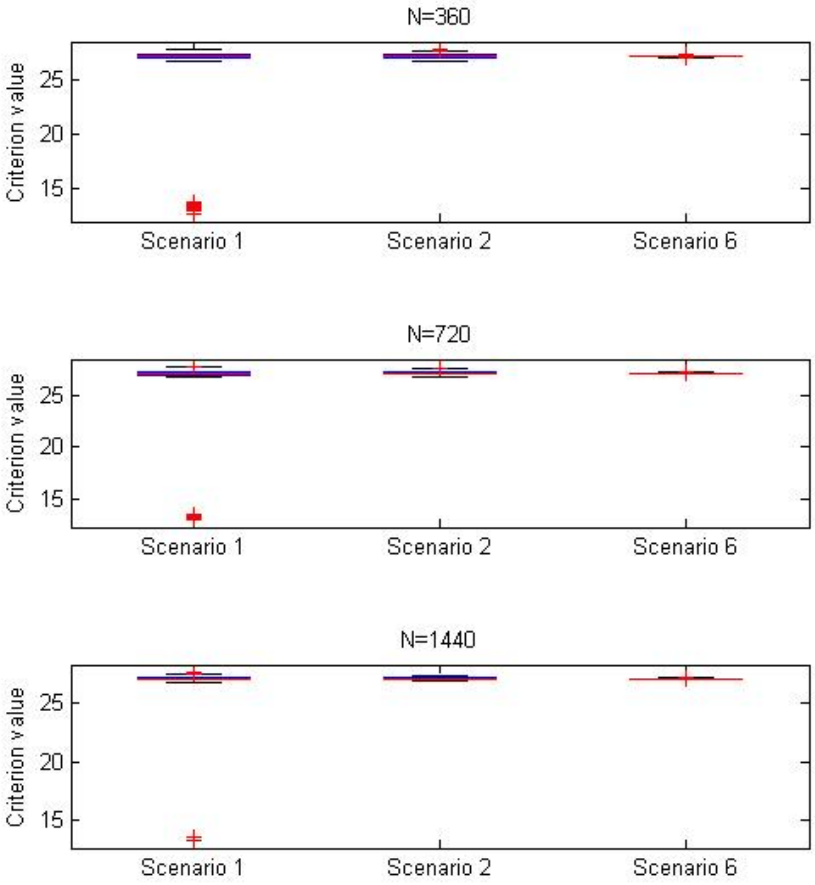


Figure 21: Boxplot of estimated values of $1/\Phi_2$, 500 simulations, Model 1

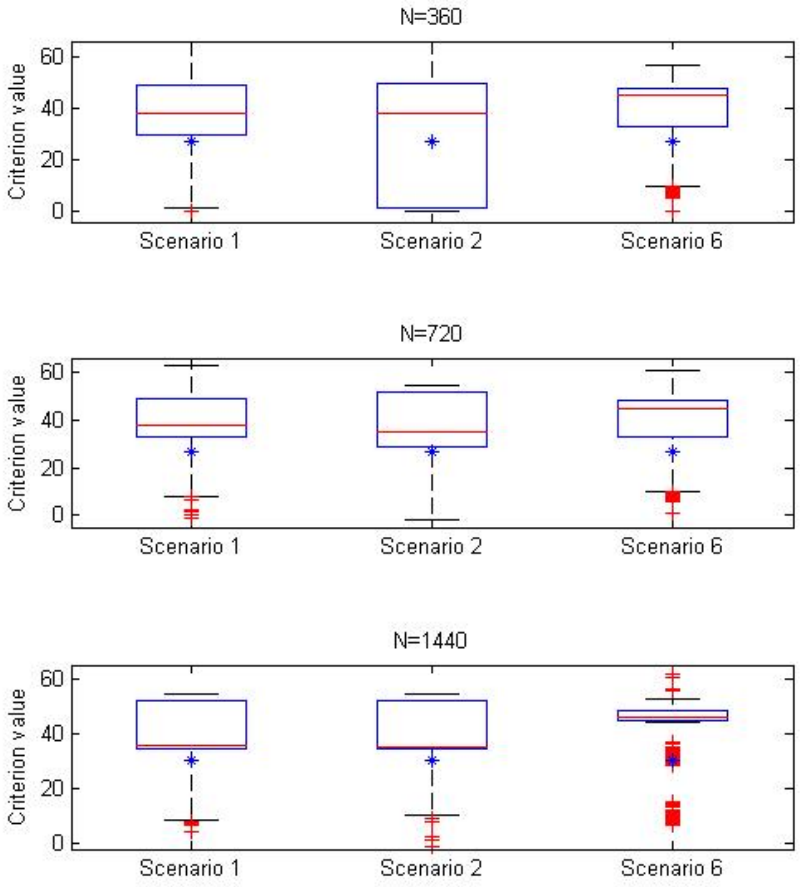
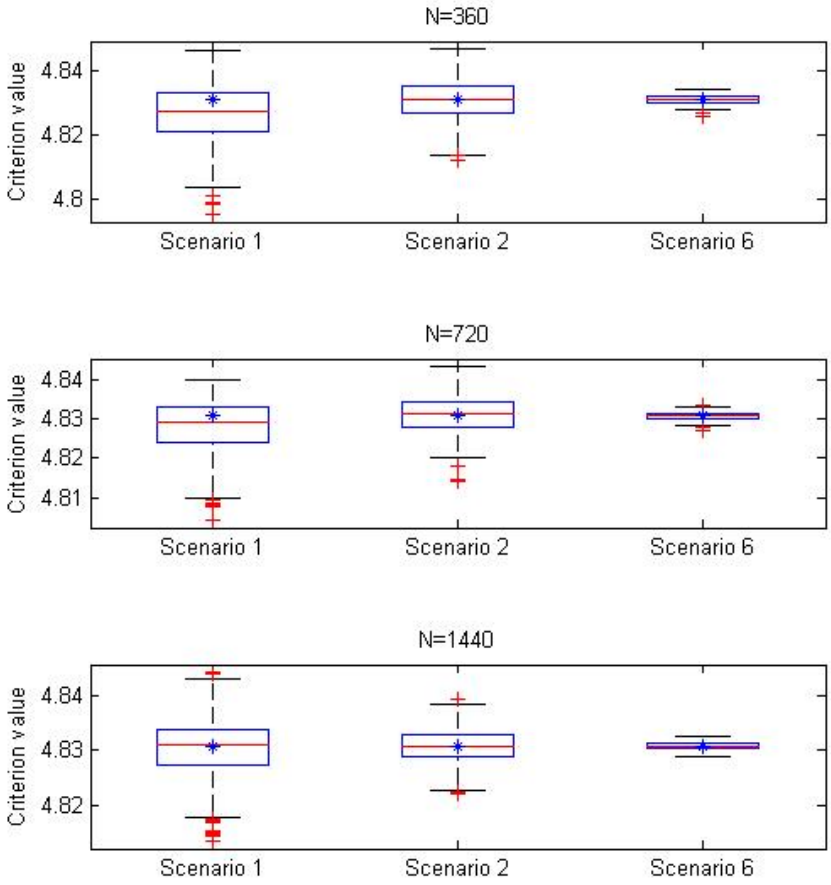


Figure 22: Boxplot of estimated values of $1/\Phi_2$, 500 simulations, Model 4



Again, since our optimality criteria depends on $\mathbf{w}_{opt}$, which in turn depends on $\hat{\theta}$, it is natural that the estimated values of the criteria are afflicted with a large variation if we have a large variation in parameter estimates. These plots are just another way of visualizing what we have already observed in the plots above.

5 Conclusions and discussion

When we started this project, we didn't expect the difficulties that the optimization of the design for this model brought. However, the contents of this paper states a clear message; when dealing with optimal designs, one has to be extremely careful. Our results shows us that, to be able to obtain a reliable optimal design, it is fundamental that we have a good estimation of the parameter values to base our optimization on.

From the plots above we can draw the conclusion that for this particular problem, Model 4 is the model that we should use if we want a reliable optimal design. But is this model realistic? When working with the development process of a new drug, we work with human subjects who are different in metabolism and in other aspects. A standard deviation of 1 unit (in combination with a clinical relevant effect of 5 units) is usually unrealistically low in this context. Also, we might not have a good foundation for fixing the value of one parameter, due to several reasons (no similar compounds to compare to, pre-clinical studies weren't possible to conduct, etc.).

What do we do then? When we work with an adaptive design we need to obtain a reliable optimal design, which is impossible with the large variation in parameter estimates as Model 1 gives. Increasing the total sample size N is one solution, but as mentioned earlier, in most cases this isn't possible.

An interesting follow-up to this paper would be to continue with step 6 from section 3; collect data from the remaining $N - n$ patients, allocated to doses according to Step 5 and then compute a final estimate $\hat{\theta}_{final}$ of θ using NLS and all N patients. What would be interesting here is to look at the first- and second order asymptotics of $\hat{\theta}_{final}$.

A The Nelder-Mead simplex search method

The Nelder-Mead simplex search method is a well-known technique that was suggested by John Nelder and Roger Mead (1965). It is a derivative-free method for solving unconstrained optimization problems of minimizing nonlinear functions. The method is called a direct search-method since it, instead of using gradients to search for the minimum, uses the function values of the model.

A *simplex* is a generalization of the concept of a triangle or tetrahedron to arbitrary dimension. For example, a 2-simplex is a triangle, a 3-simplex is tetrahedon, and an n-simplex is an n-dimensional *polytope* (a geometric object with flat sides, existing in any number n of dimensions) with $n + 1$ specific kind of points that mark the corners of the polytope.

The Nelder-Mead simplex method starts at an initial set of points that create the starting simplex S, say $\theta_0, \theta_1, ..., \theta_{n+1}$, with corresponding function values $f(\theta_0), f(\theta_1), ..., f(\theta_{n+1})$. There is only one restriction on the starting simplex S; the initial set of points can not lie in the same hyperplane. The method is then to move the simplex through space toward the minimum by shifting the corner with the highest function value. Once the process hits a higher function value than the present or when the function value has reached some pre-specified value, it stops. The whole idea is hence to change the location of the simplex at each step to find the smallest value of the function.

There are naturally advantages and disadvantages of this method, as with all methods of optimization. The largest advantage is that it is a fairly easy technique to understand and implement. A disadvantage is that it has a relatively slow convergence if the initial guesses of the starting points are far from the ones that minimize the function, which in turn demands a lot of computational power (it is computationally expensive, one can say).

B The variance function $d(x, w, \theta)$

Let

- x = dose $\in \{0, 20, 40, 60, 80, 100\}$
- $\mathbf{w}$ = the weights for the different doses (the *design* of the study)
- $\theta = (E_0, E_{max}, \alpha, ED_{50})$ = vector of parameters
- $f(x, \theta)$ = the true response from dose x
- σ^2 = the variance of the observations

From the theory of nonlinear least squares estimation we know that if the errors of the model are *iid* normally distributed, then the estimation $(\hat{\theta}, \hat{\sigma^2})$ of (θ, σ^2) has an asymptotic bivariate normal distribution:

$$\sqrt{N}\left(\begin{pmatrix} \hat{\theta} \\ \hat{\sigma} \end{pmatrix} - \begin{pmatrix} \theta \\ \sigma \end{pmatrix}\right) \to N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2 M^{-1}(\mathbf{w}, \theta) & 0 \\ 0 & 2\sigma^4 \end{pmatrix}\right)$$

where the convergence is in distribution and N denotes the total sample size of the clinical trial. The information matrix M is calculated as

$$M(\mathbf{w}, \theta) = \sum_{j=1}^{k} w_j g(x_j, \theta) g'(x_j, \theta)$$

where $j = 1$ indicates dose 0 (placebo) and $j = k$ indicates the max dose, and

$$g(x_j, \theta) = \begin{pmatrix} \frac{\partial f(x_j,\theta)}{\partial E_0} & \frac{\partial f(x_j,\theta)}{\partial E_{max}} & \frac{\partial f(x_j,\theta)}{\partial \alpha} & \frac{\partial f(x_j,\theta)}{\partial ED_{50}} \end{pmatrix}' =$$

$$\begin{pmatrix} 1 & \frac{x_j^\alpha}{ED_{50}^\alpha + x_j^\alpha} & \frac{E_{max} ED_{50}^\alpha x_j^\alpha (log x_j - log ED_{50})}{(ED_{50}^\alpha + x_j^\alpha)^2} & \frac{-E_{max} \alpha ED_{50}^{\alpha-1} x_j^\alpha}{(ED_{50}^\alpha + x_j^\alpha)^2} \end{pmatrix}'$$

are the gradient functions. The variance of the estimated effect of dose x_j is then approximately

$$g'(x_j, \theta) \sigma^2 M(\mathbf{w}, \theta)^{-1} g(x_j, \theta)$$

This finally gives us the variance of the estimated difference between any dose compared to a placebo dose as

$$d(x_j, \mathbf{w}, \theta) = (g(x_j, \theta) - g(0, \theta))' \sigma^2 M(\mathbf{w}, \theta)^{-1} (g(x_j, \theta) - g(0, \theta))$$

where $j = 1, 2, 3, 4, 5$ correspond to doses $20, 40, 60, 80, 100$.

C Figures

C.1 Fitted response curves

Figure 23: Boxplot of estimated mean responses, 500 simulations, Model 1, Scenario 2, $N = 360$

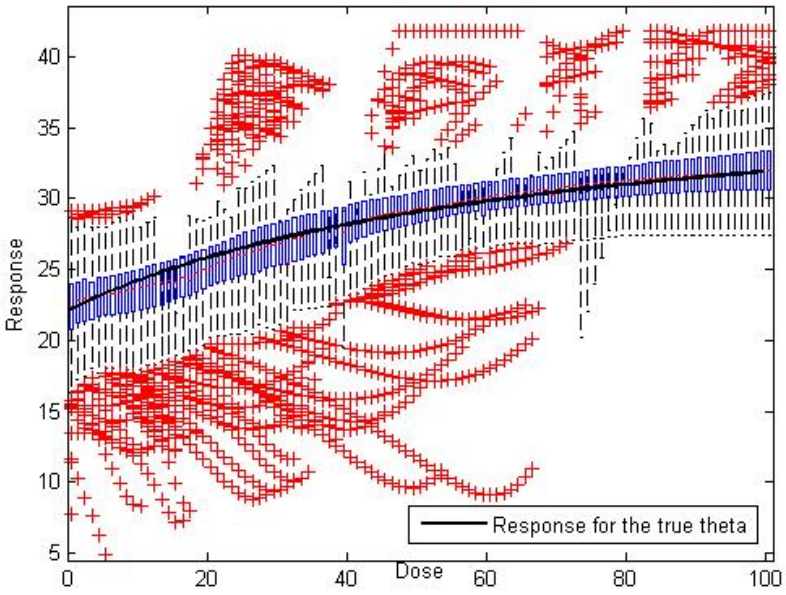


Figure 24: Boxplot of estimated mean responses, 500 simulations, Model 2, Scenario 2, $N = 360$

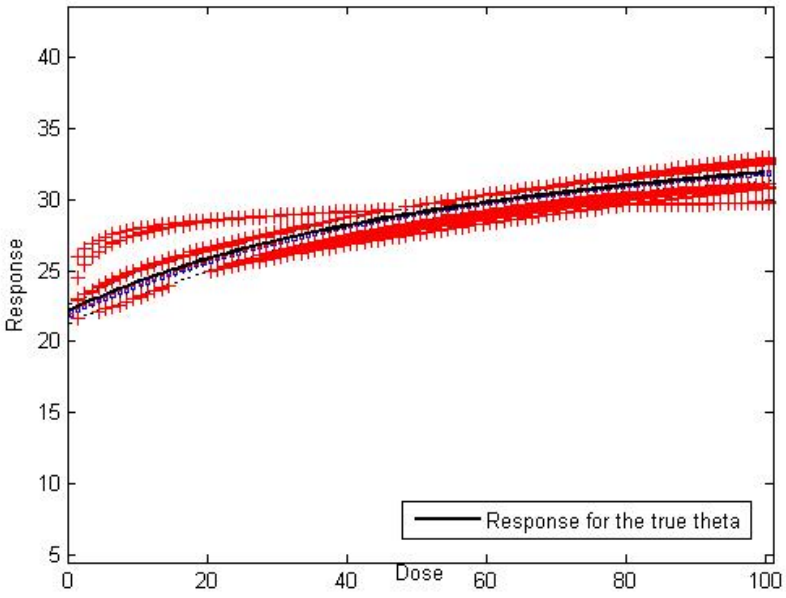


Figure 25: Boxplot of estimated mean responses, 500 simulations, Model 3, Scenario 2, $N = 360$

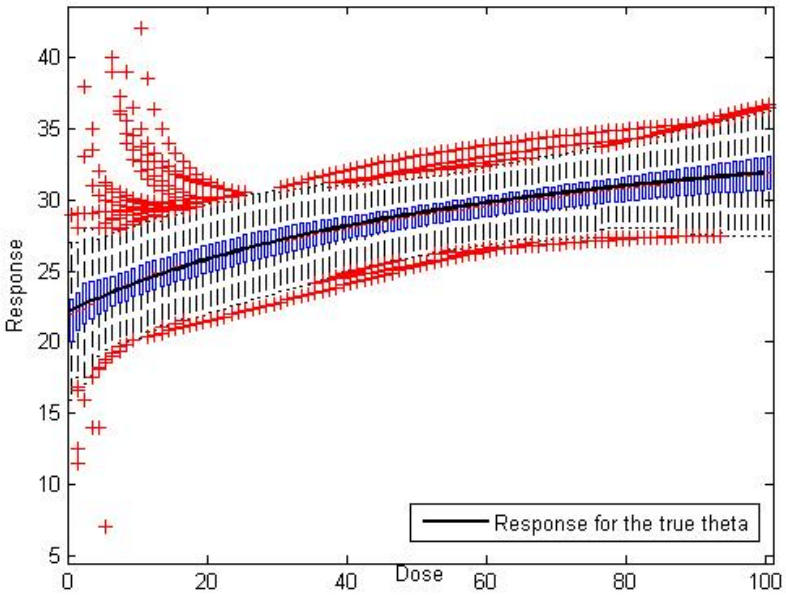


Figure 26: Boxplot of estimated mean responses, 500 simulations, Model 4, Scenario 2, $N = 360$

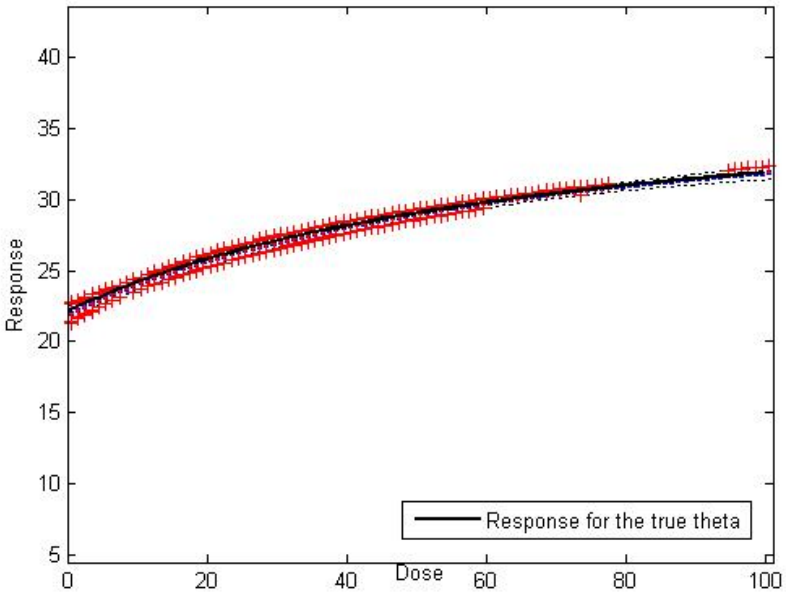


Figure 27: Boxplot of estimated mean responses, 500 simulations, Model 1, Scenario 6, $N = 360$

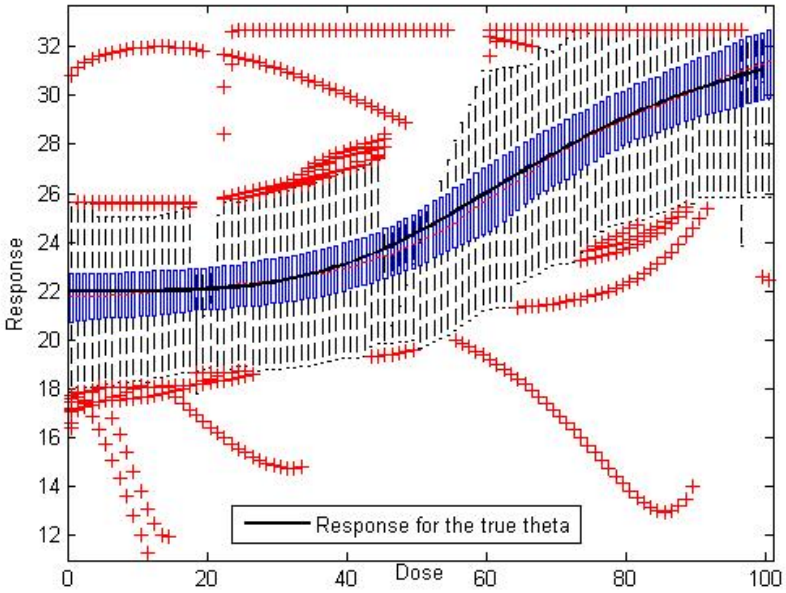


Figure 28: Boxplot of estimated mean responses, 500 simulations, Model 2, Scenario 6, $N = 360$

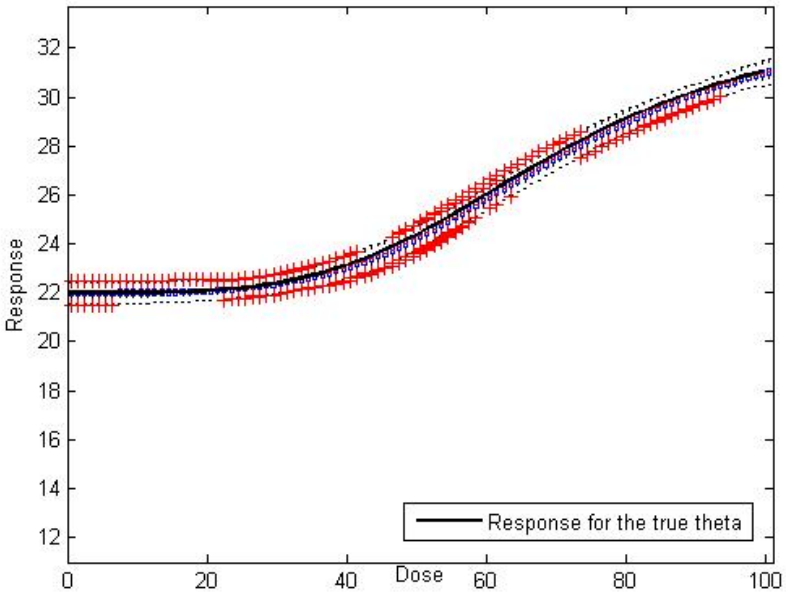


Figure 29: Boxplot of estimated mean responses, 500 simulations, Model 3, Scenario 6, $N = 360$

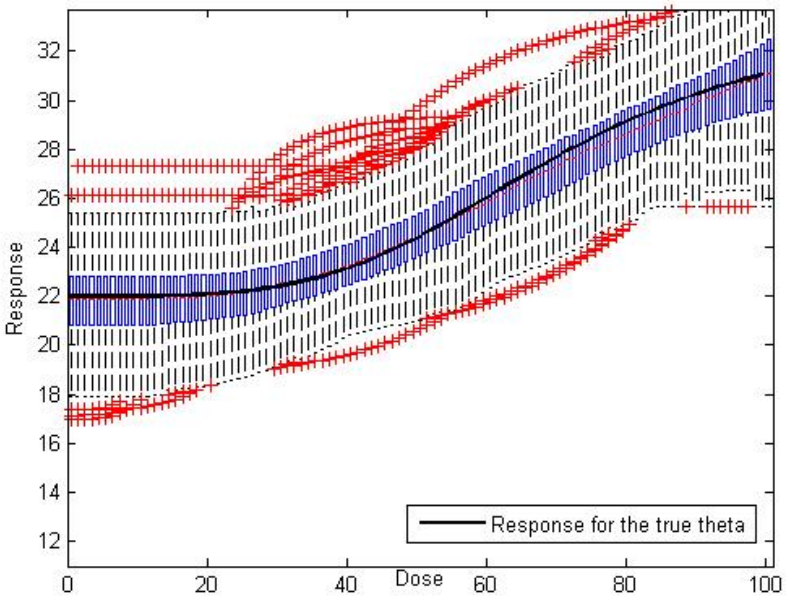
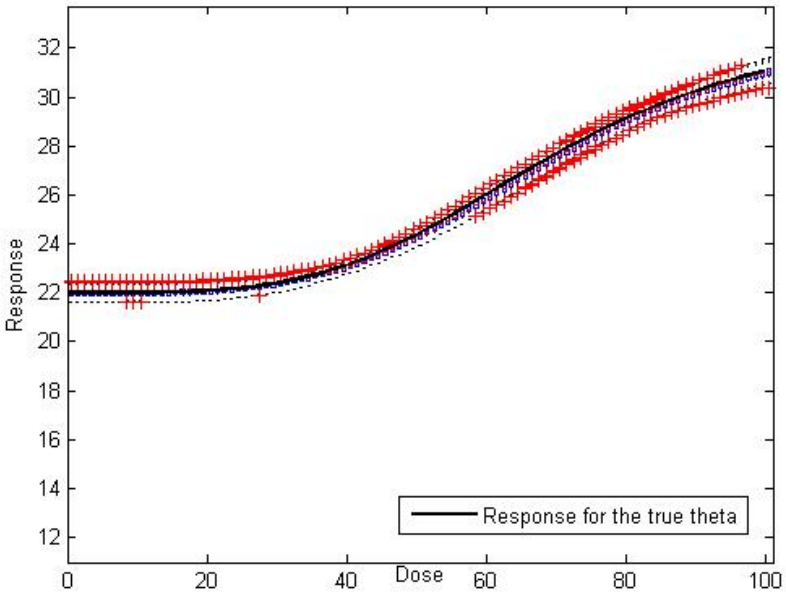


Figure 30: Boxplot of estimated mean responses, 500 simulations, Model 4, Scenario 6, $N = 360$



C.2 Estimated optimal designs

Figure 31: Boxplot of estimated optimal weights, 500 simulations, Criterion 1, Model 1, Scenario 2, $N = 360$

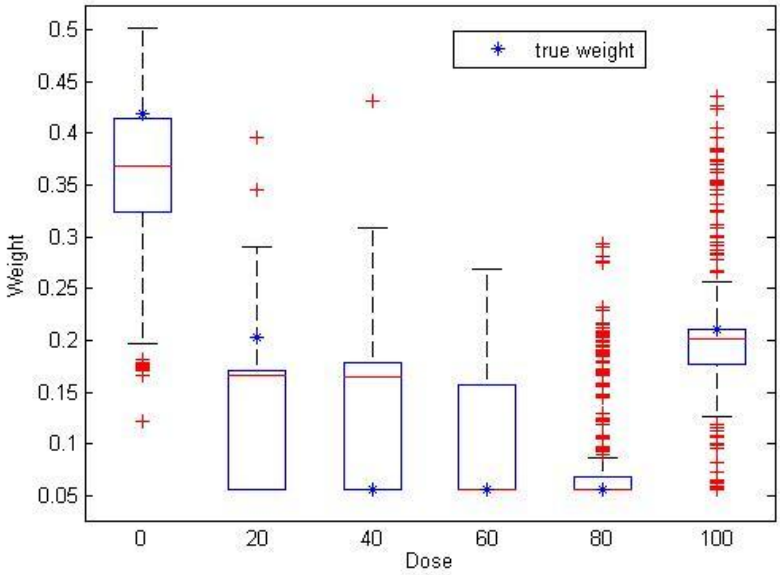


Figure 32: Boxplot of estimated optimal weights, 500 simulations, Criterion 1, Model 1, Scenario 2, $N = 720$

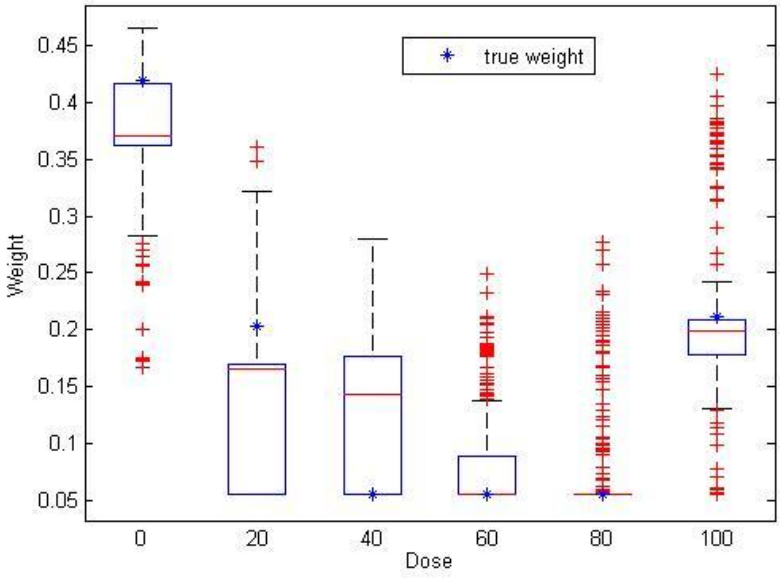


Figure 33: Boxplot of estimated optimal weights, 500 simulations, Criterion 1, Model 1, Scenario 2, $N = 1440$

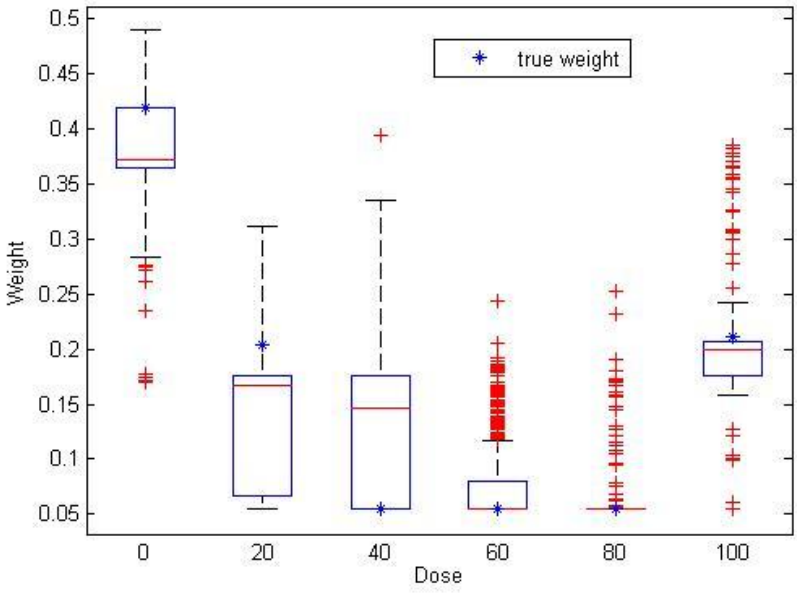


Figure 34: Boxplot of estimated optimal weights, 500 simulations, Criterion 1, Model 4, Scenario 2, $N = 360$

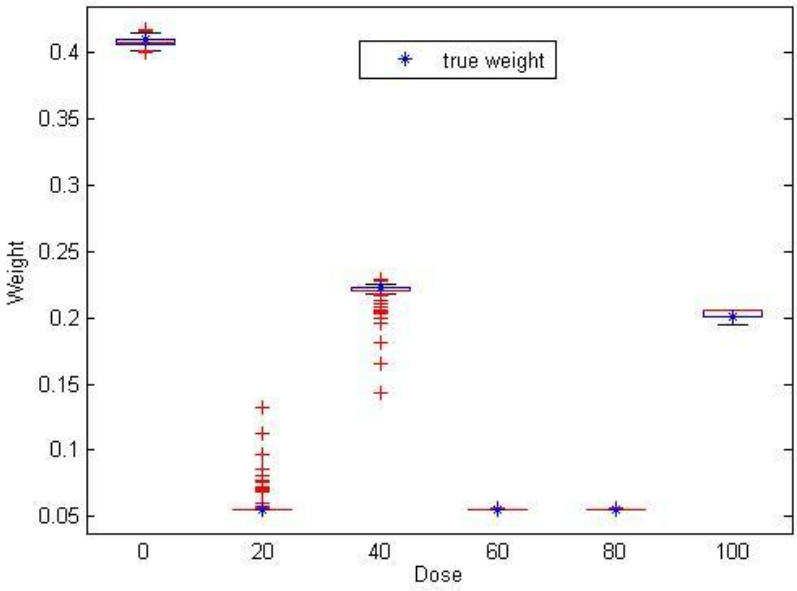


Figure 35: Boxplot of estimated optimal weights, 500 simulations, Criterion 1, Model 4, Scenario 2, $N = 720$

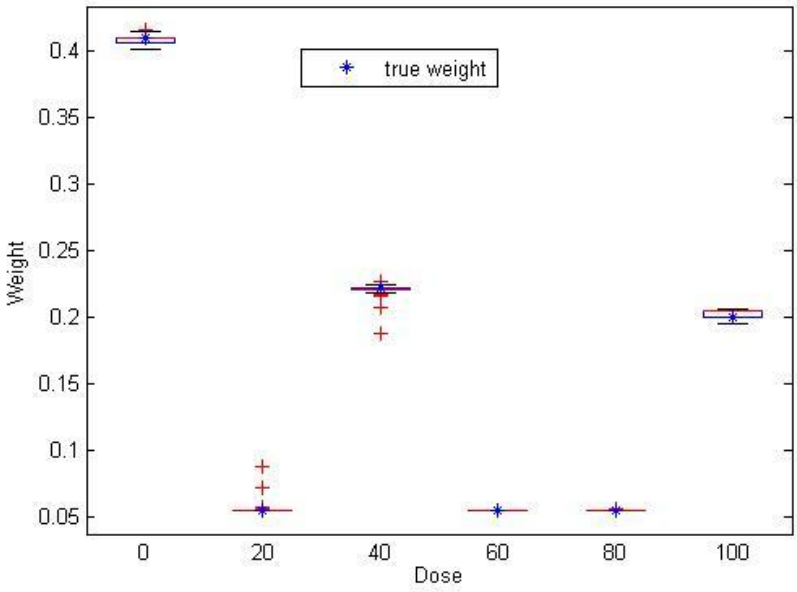


Figure 36: Boxplot of estimated optimal weights, 500 simulations, Criterion 1, Model 4, Scenario 2, $N = 1440$

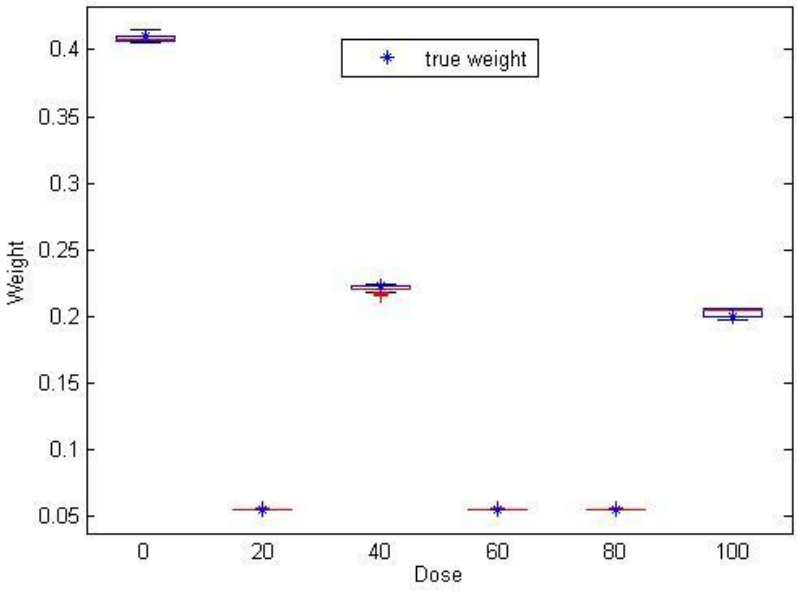


Figure 37: Boxplot of estimated optimal weights, 500 simulations, Criterion 2, Model 1, Scenario 2, $N = 360$

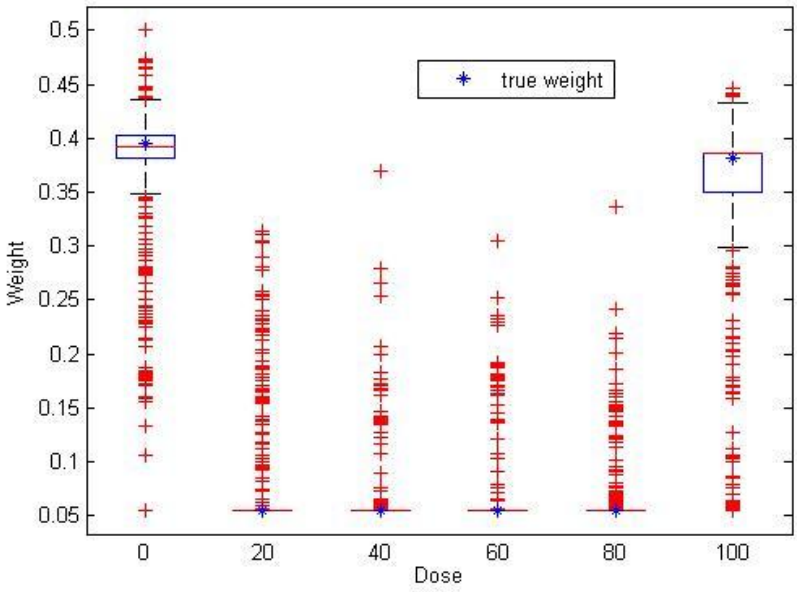


Figure 38: Boxplot of estimated optimal weights, 500 simulations, Criterion 2, Model 1, Scenario 2, $N = 720$

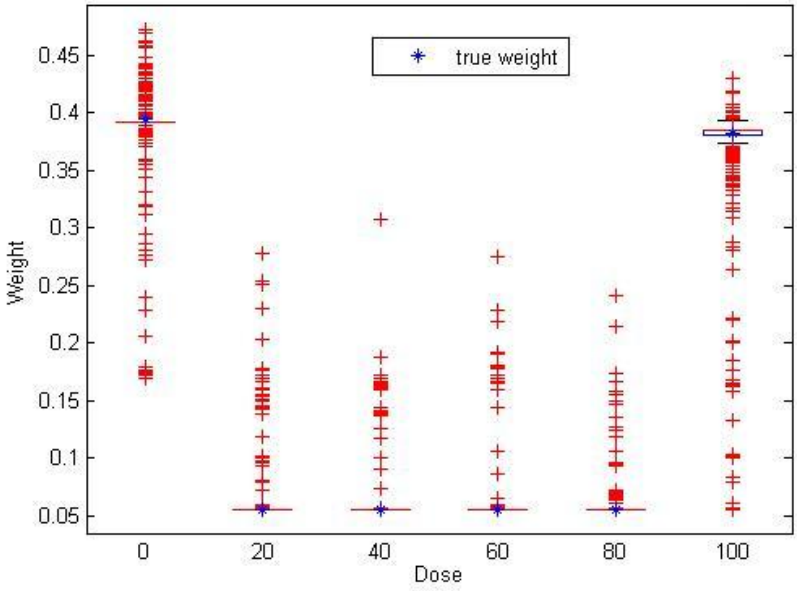


Figure 39: Boxplot of estimated optimal weights, 500 simulations, Criterion 2, Model 1, Scenario 2, $N = 1440$

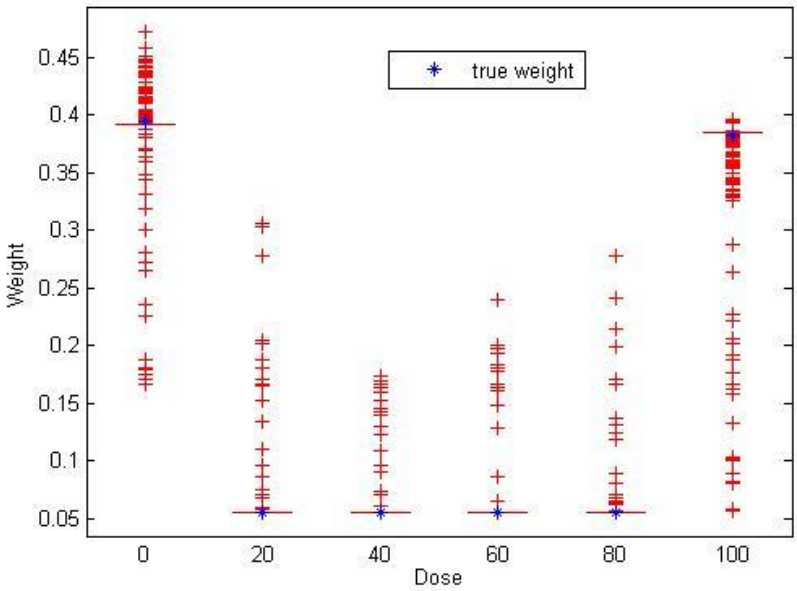


Figure 40: Boxplot of estimated optimal weights, 500 simulations, Criterion 2, Model 4, Scenario 2, $N = 360$

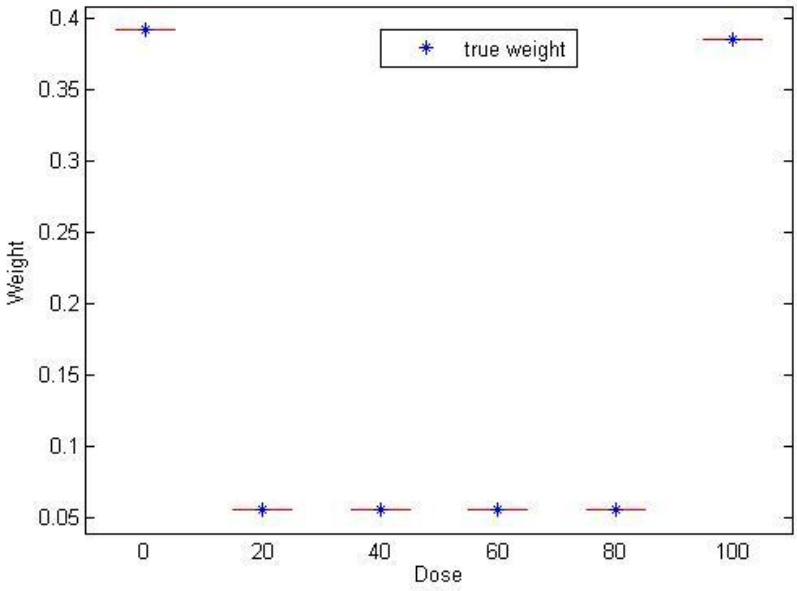


Figure 41: Boxplot of estimated optimal weights, 500 simulations, Criterion 2, Model 4, Scenario 2, $N = 720$

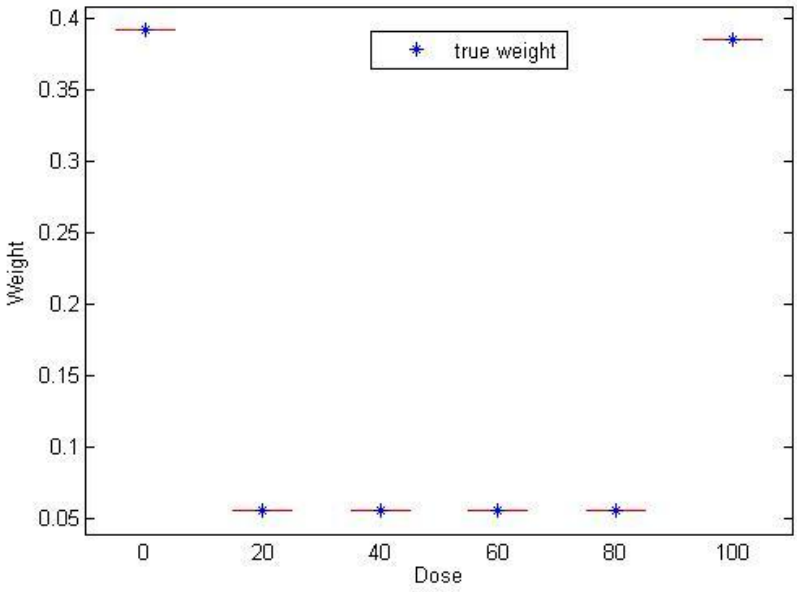


Figure 42: Boxplot of estimated optimal weights, 500 simulations, Criterion 2, Model 4, Scenario 2, $N = 1440$

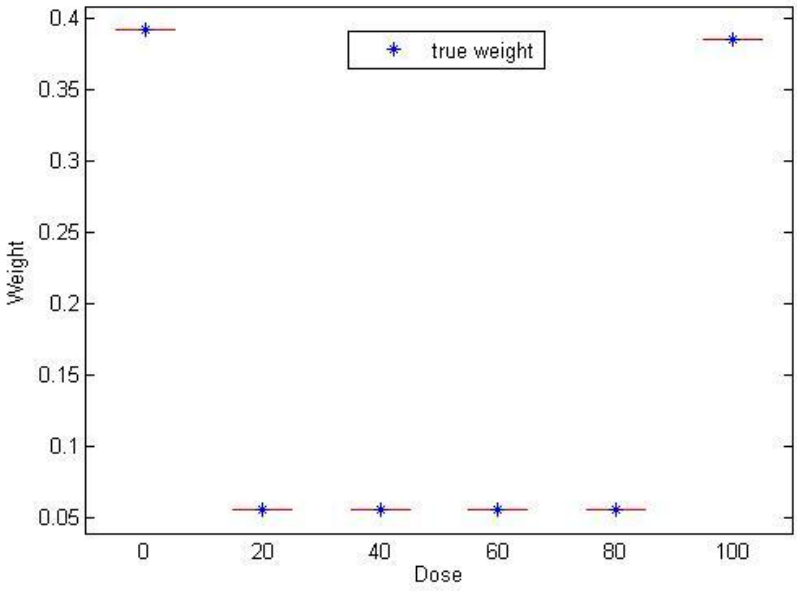


Figure 43: Boxplot of estimated optimal weights, 500 simulations, Criterion 1, Model 1, Scenario 6, $N = 360$

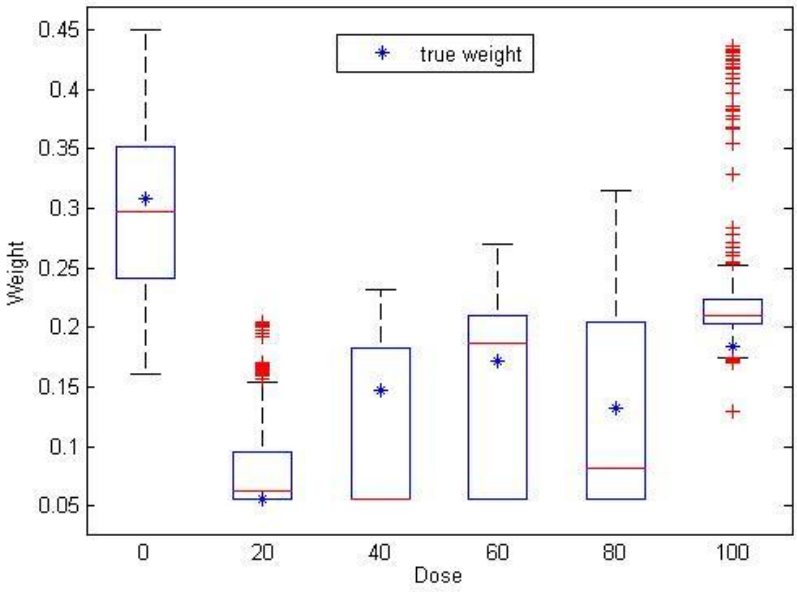


Figure 44: Boxplot of estimated optimal weights, 500 simulations, Criterion 1, Model 1, Scenario 6, $N = 720$

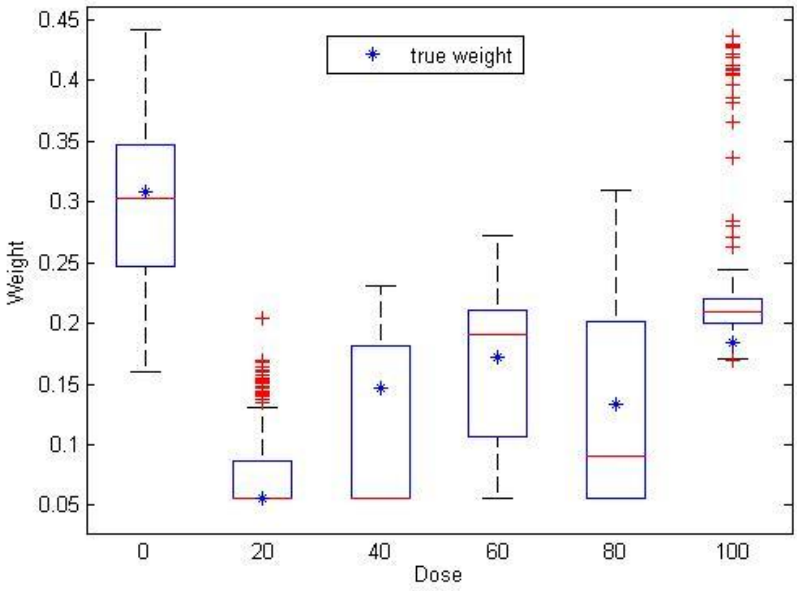


Figure 45: Boxplot of estimated optimal weights, 500 simulations, Criterion 1, Model 1, Scenario 6, $N = 1440$

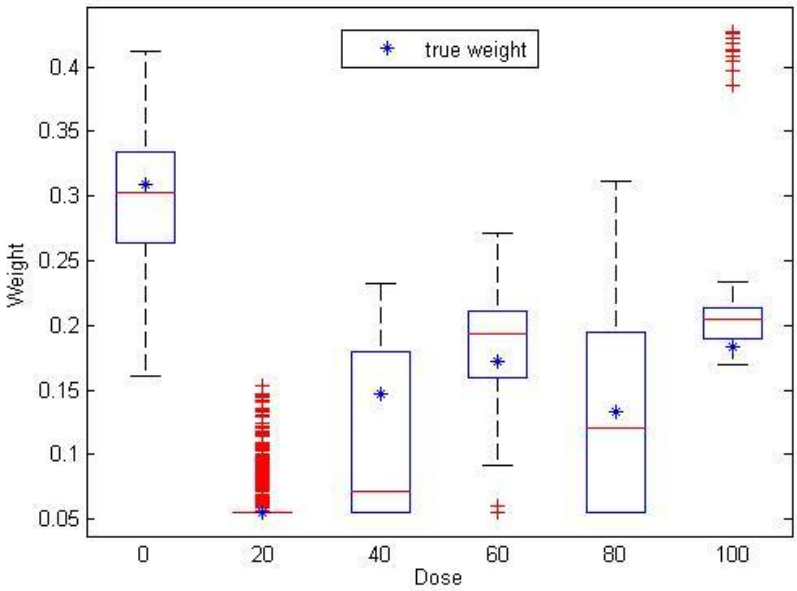


Figure 46: Boxplot of estimated optimal weights, 500 simulations, Criterion 1, Model 4, Scenario 6, $N = 360$

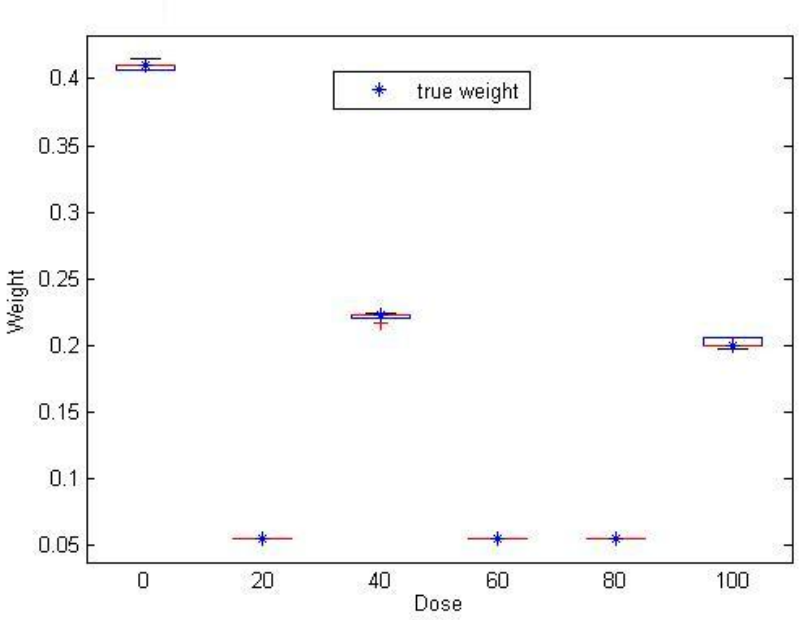


Figure 47: Boxplot of estimated optimal weights, 500 simulations, Criterion 1, Model 4, Scenario 6, $N = 720$

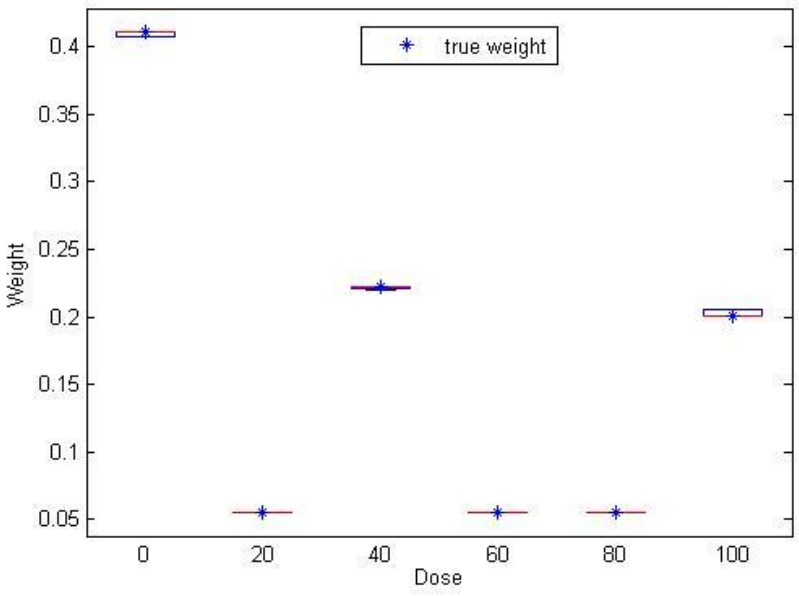


Figure 48: Boxplot of estimated optimal weights, 500 simulations, Criterion 1, Model 4, Scenario 6, $N = 1440$

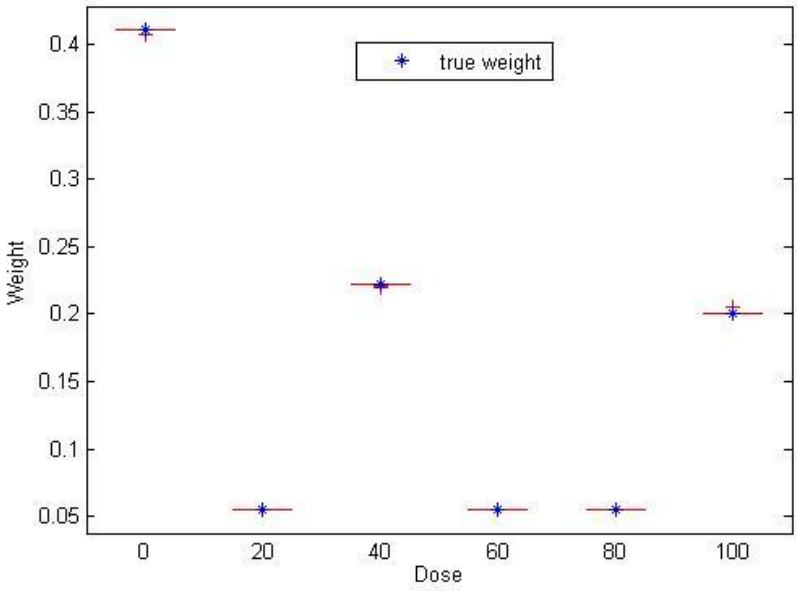


Figure 49: Boxplot of estimated optimal weights, 500 simulations, Criterion 2, Model 1, Scenario 6, $N = 360$

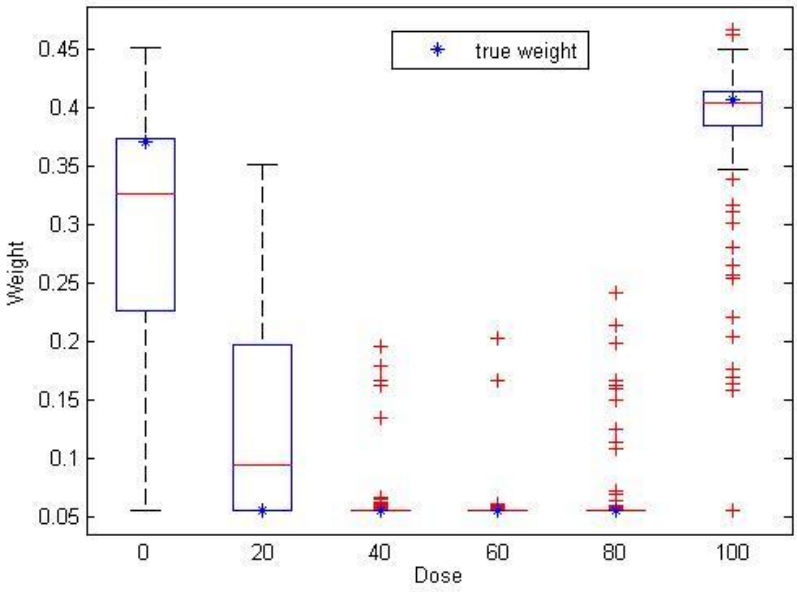


Figure 50: Boxplot of estimated optimal weights, 500 simulations, Criterion 2, Model 1, Scenario 6, $N = 720$

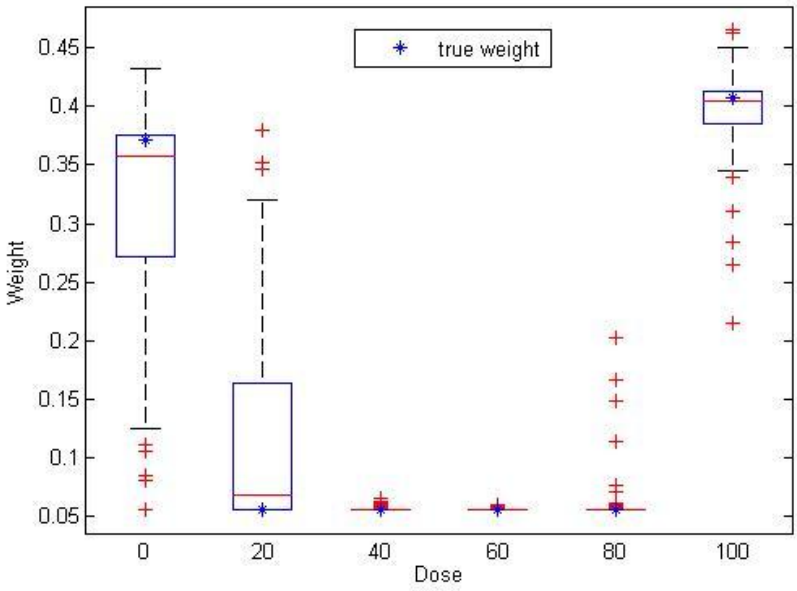


Figure 51: Boxplot of estimated optimal weights, 500 simulations, Criterion 2, Model 1, Scenario 6, $N = 1440$

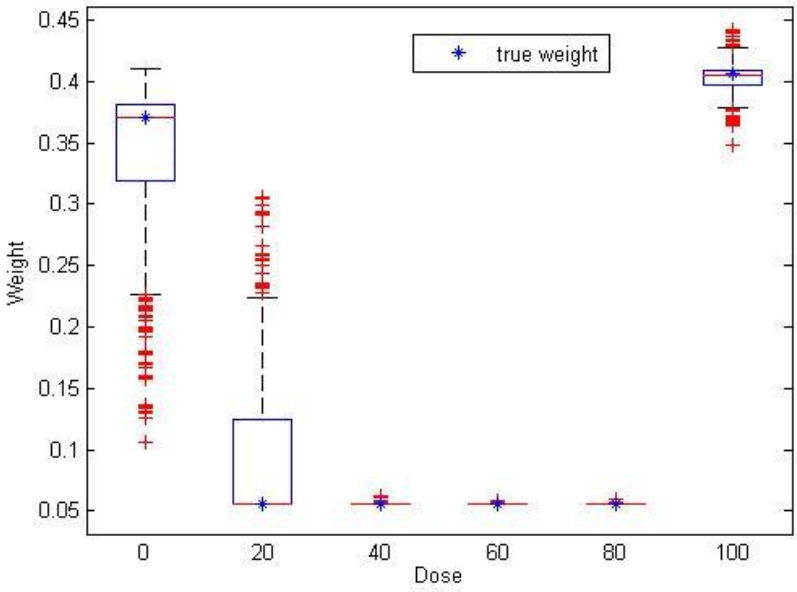


Figure 52: Boxplot of estimated optimal weights, 500 simulations, Criterion 2, Model 4, Scenario 6, $N = 360$

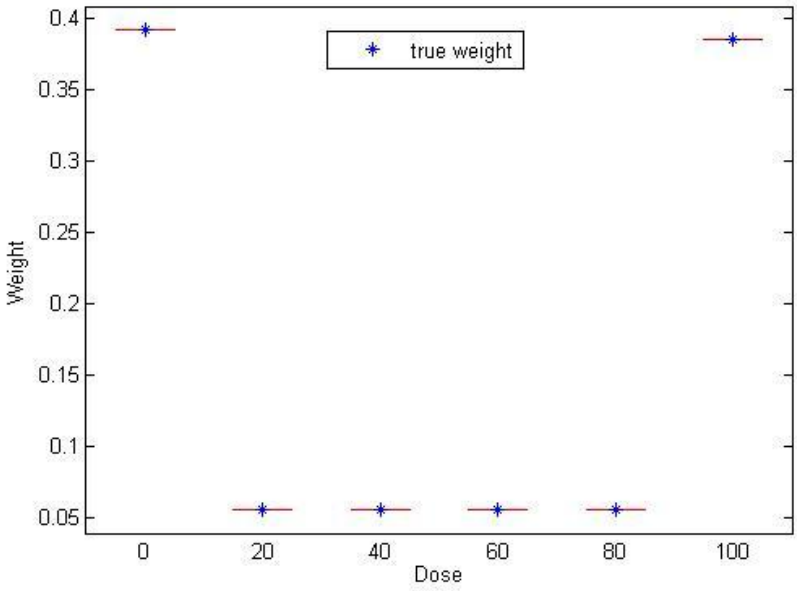


Figure 53: Boxplot of estimated optimal weights, 500 simulations, Criterion 2, Model 4, Scenario 6, $N = 720$

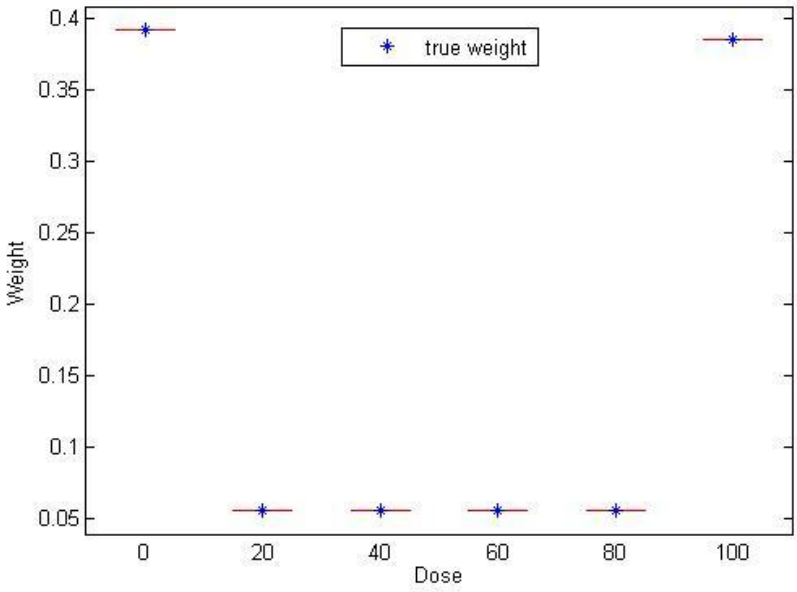
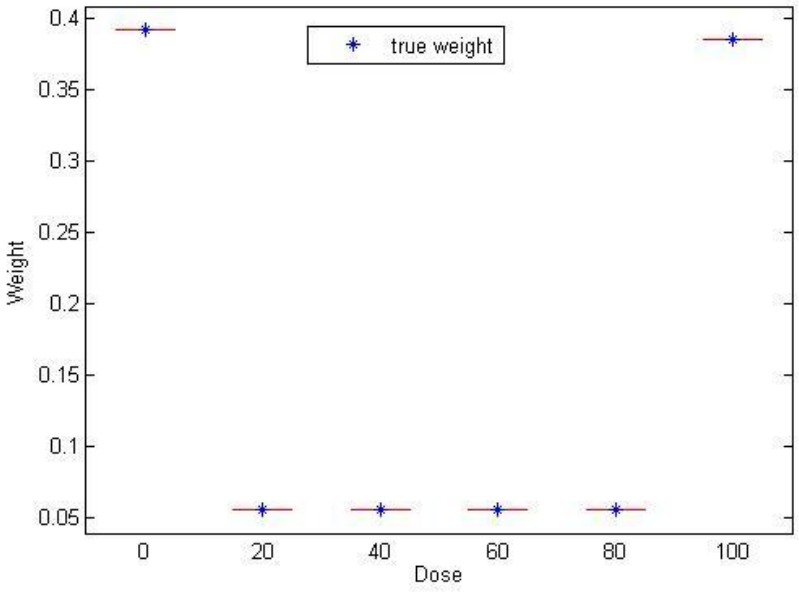


Figure 54: Boxplot of estimated optimal weights, 500 simulations, Criterion 2, Model 4, Scenario 6, $N = 1440$



D Word list

- **Pharmacokinetics**=what impact the body has on the drug, for example how the body absorbes it, how it is distributed in the body and the duration of the effect of the drug.
- **Microdosing**="a technique for studying the behaviour of drugs in humans through the administration of doses so low they are unlikely to produce whole-body effects, but high enough to allow the cellular response to be studied"[1]
- **ICH-E4** = the primary source of regulatory guideline in the area of clinical trials
- **Maximal tolerable dose (MTD)** ="the maximal tolerable dose beyond which no further beneficial effect is seen"(ICH-E4)
- **Blinding**=a blinded experiment is a experiment where some of the persons involved in the experiment are kept unknowing of certain information that might otherwise lead to conscious or unconscious bias, causing damage on the credibility of the results

[1] Wikipedia

References

Literature and articles

Atkinson A.C. and Donev A.N. (1992). *Optimum Experimental Designs.* Oxford: Claredon Press (Oxford statistical science series).

Bretz F., Hsu J., Pinheiro J. and Liu Yi (2008). Dose Finding - A challenge in Statistics. *Biometrical Journal*, **50(4)**: 480-504.

Chow S.C. and Liu J.P. (2004). *Design and Analysis of clinical trials.* John Wiley & Sons, second edition.

Dette H. and Kwiecien R. (2004). A comparison of sequential and non-sequential designs for discrimination between nested regression models. *Biometrika*, **91(1)**: 165-176.

Miller F., Guilbaud O. and Dette H. (2007). Optimal designs for estimating the interesting part of a dose-effect curve. *Journal of Biopharmaceutical Statistics*, **17**: 1097-1115.

Ruberg S.J. (1995). Dose response studies I. Some design considerations. *Journal of Biopharmaceutical Statistics*, **5(1)**: 1-14.

Ruberg S.J. (1995). Dose response studies II. Analysis and interpretation. *Journal of Biopharmaceutical Statistics*, **5(1)**: 15-42.

Seber G.A.F. and Wild C.J. (1989). *Nonlinear Regression.* John Wiley & Sons.

Silvey S.D. (1980). *Optimal Design.* Chapman and Hall.

Internet

www.wikipedia.org (Non-linear least squares, Nelder-Mead method, Levenberg-Marquardt algorithm and Optimal design)

http://kurser.math.su.se/course/view.php?id=347 (lecture notes from the course in Computer intensive statistical methods, 7.5 hp, Stockholm University)