Personalized health care: switching to a subpopulation in Phase III

FELIX LEOPOLDO RIOS

Master’s Thesis at Astra Zeneca
Supervisors: Carl-Fredrik Burman and Niclas Sjögren
Examiner: Timo Koski

TRITA xxx 2012-03
Acknowledgments

I would like to express my sincere gratitude to my supervisors at Astra Zeneca, Niclas Sjögren and Carl-Fredrik Burman for introducing me to the drug development industry and for input to this thesis. I would also like to thank my supervisor at the Department of Mathematical Statistics at the Royal Institute of Technology, Professor Timo Koski for his feedback and support during the process.

Felix Leopoldo Rios

Stockholm March 2012
Abstract

Since different patients may have different causes of getting a disease, treating every patient having a certain disease in the same manner is not always be the best way to go. A treatment having effect in one type of patients may not have the same effect in a different type of patients. This makes it possible to partition a patient population into subpopulations in which a drug has distinct expected response. In this thesis the patient population is partitioned into two subpopulations where we have prior knowledge that one of them has a higher expected response to a drug than the other. Based on responses to a drug in Phase II, it has been analyzed in which of the populations Phase III should continue. The results show that the decision is highly dependent on the utility function on which the analysis is based. One interesting case is when the vast majority of the patient population belongs to the subpopulation with the higher expected response and a utility function that takes into account the prevalence of the populations. In that case the simulations show that when the difference in expected response between the subpopulations is large, it is a safer choice in continuing in Phase III in the subpopulation having the higher expected response than in the full population even though the expected utility will be less. This is an expected result which indicates that the approach used to model the situation studied in this report is reasonable.
Referat

Personalized health care: Byta till en delpopulation i fas III

Contents

1 Introduction ............................................. 1
   1.1 Clinical trials in drug development ................. 1
   1.2 The grand problem .................................. 2
   1.3 Methods for making clinical trials more efficient .... 3
       1.3.1 Adaptive design ................................. 3
       1.3.2 Optimal Phase II/III design ..................... 3
       1.3.3 Seamless Phase II/III design ..................... 3

2 Problem description ..................................... 5
   2.1 Personalized health care ............................. 5
   2.2 Switching to a subpopulation in Phase III ............ 5
   2.3 Problem ........................................... 7

3 Method .................................................. 9
   3.1 Assumptions ......................................... 9
       3.1.1 Classifier ....................................... 9
       3.1.2 Measuring effect .................................. 10
       3.1.3 Dose response curve - $E_{\text{max}}$ model ....... 10
       3.1.4 Utility .......................................... 13
       3.1.5 Relation the between subpopulations ............... 13
       3.1.6 Safety function .................................. 14
       3.1.7 Trial design ..................................... 14
   3.2 Mathematical preliminaries and definitions .......... 15
       3.2.1 Test statistic ................................... 15
       3.2.2 Probability of success (PoS) ..................... 16
       3.2.3 Expected PoS .................................... 17
       3.2.4 Computing expected utility ....................... 17
       3.2.5 Optimal design in Phase III ..................... 18
       3.2.6 Expected utility from optimal design in Phase III . 18
   3.3 Method ............................................. 19
       3.3.1 Overview of the algorithm ....................... 19
       3.3.2 Detailed explanation of the algorithm ............. 19
   3.4 Mathematical justification ........................... 21
3.4.1 Computing and sampling from the posterior distribution . . . 21
3.4.2 Rejection sampling . . . . . . . . . . . . . . . . . . . . . 25

4 Results 31
4.1 Simulation study . . . . . . . . . . . . . . . . . . . . . . . . . . . 31
4.1.1 Parameters for the simulation studies . . . . . . . . . . . . . . . 31
4.1.2 Interpretation of plots . . . . . . . . . . . . . . . . . . . . . . 35
4.1.3 Simulation study with utility function $U_{P,A}$ . . . . . . . . . 35
4.1.4 Simulation study with utility function $U_{P,B}$ . . . . . . . . . 43
4.2 Discussion and possible future extensions . . . . . . . . . . . . . . 53
4.2.1 Positive outcome . . . . . . . . . . . . . . . . . . . . . . . . . 53
4.2.2 The utility functions . . . . . . . . . . . . . . . . . . . . . . . 53
4.2.3 Sampling posterior . . . . . . . . . . . . . . . . . . . . . . . . 53
4.3 Conclusion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 55

Bibliography 57
Chapter 1

Introduction

1.1 Clinical trials in drug development

Clinical trials are conducted in drug development to show that a drug has the desired benefit-risk profile in human. The trials are made on a restricted number of subjects and mathematical statistical models are used to make inference about the trial and conclusions for future outcomes. The outcome of a trial refers to the treatment’s safety and the efficacy properties. When measuring safety, the possible adverse effect a drug may have to a patient is measured. Efficacy refers to the positive effect the treatment has to the patient.

There are a lot of parameters that define a clinical trial. Some of the parameters that have to be taken into account are, for example, the sample size, which doses should be tested, and who should take part of the trial. A clinical study is usually divided into four Phases, I, II, III and IV which are themselves individual clinical trials, with new recruitment of subjects into each trial.

In order to be able to advertise a drug at the market it has to be approved by the regulatory authorities (RA). The RA’s require that a drug has to successfully pass Phase I - III in order to get a label. Here is a brief explanation of the four phases.

Phase I is the first time the drug is tested in human. (Before Phase I preclinical studies have been made in animals). The primary objective in Phase I is typically to evaluate safety properties and the aim is to find the maximal dose that does not have any serious side effects. The number of subjects taking part in a Phase I study is usually about 20 - 80, who typically are healthy volunteers usually from the drug developing company, see [10]. If it can be found in Phase I that there exists doses that do not have any serious side-effects, the study continues to Phase II.

In Phase II the main interest is to find out if the drug has a positive effect

---

1Since not every phase always includes sick people, patients are often called subjects.
2Examples of regulatory authorities are Food and Drug Administration (FDA) and European Medicine Agency (EMEA).
3In some therapeutic such as oncology, Phase I is done in real patients due to ethical reasons.
and which dose is to be tested in Phase III. It is common that Phase II is divided into two separate trials Phase IIa and Phase IIb, where the latter is aimed to find the most suitable dose to be tested in Phase III and the former is aimed at showing proof of concept. Phase II traditionally involves about 100 - 200 patients [10]. The last phase that has to be successfully passed before the drug is to be advertised at the market is Phase III, the confirmatory phase.

In Phase III usually one dose is tested against a control group (in which the subjects usually get placebo or an already existing drug at the market). Which dose to be tested has been evaluated at Phase II. Phase III is the most extensive study and thus usually the most expensive and can involve about 1000 patients [10].

Phase IV takes place when the drug is already available at the market and the drug company wishes to study e.g. long term adverse effects such as mortality and morbidity [10].

The vast majority of the drugs ever tested never reaches the market [11]. This is of course costly for the companies developing the treatments, hence much of the research going on is spent on evaluating how to make more drugs to be approved in Phase III and how to stop non effective drugs in an as early stage as possible. Since Phase II and Phase III are the two most comprehensive trials, most of the research in the area of clinical trial is concentrated at optimizing these phases.

In the last decade it has been increasingly apparent that a treatments effect between different genomic patient subpopulations can be dissimilar [15]. This has raised the interest of developing stratified medicine for patients in specific subpopulations [14]. In this report we assume that the patient population can be partitioned into two distinct subpopulations showing different response to a drug in Phase II. The question that is studied is, should the trial continue in Phase III in a subpopulation or in the full population. The approach of developing stratified medicine for targeted populations falls into the category of Personalized health care and will be elucidated in more detail in Section 2.1.

The method used in this thesis is an extension of the method described by Professor C. Jennison in [7] for analyzing jointly optimal design for Phase II/III clinical trials.

1.2 The grand problem

The grand problem in the pharmaceutical industry is the problem of finding an optimal design for a whole clinical study including all the phases. To fully solve the grand problem is hard since there is a great number of parameters that has impact on a study and that not all studies are comparable. However, one can solve parts of it by making suitable assumptions. In Section 1.3 some approaches to make clinical trials more efficient are briefly explained.

---

When talking about a trial design we talk about the parameters that that define the trial.
1.3 Methods for making clinical trials more efficient

1.3.1 Adaptive design

Sometimes it can be of interest to do an interim analysis of an ongoing study. Since it is not possible to just look at data in an ongoing study without further considerations of bias that will be introduced, there is a special category in the research of clinical trials named Adaptive design. Adaptive design in clinical studies is a design protocol which takes into account information gathered during an ongoing study (e.g. Phase III) when deciding how to continue the study [3].

For example when testing a new drug there could be a need for making an interim analysis during a study due to safety reasons. If a drug shows to be toxic at the interim analysis, the study can be stopped and patients can be saved from adverse effects. Such adaptive design falls into the category of Early stopping [3]. There are other adaptive designs too that can be read about in e.g. [2, 8, 3, 13].

One problem that all adaptive designs have to take into account is the Type I errors\(^5\) that the interim analysis will infer to the final analysis if one does not correct for that.

1.3.2 Optimal Phase II/III design

Since Phase II and Phase III are the most extensive studies there is an interest in finding an optimal design of Phase II and Phase III. Since Phase III relies on the outcome of Phase II it is meaningful to study the interaction of these phases to find a jointly optimal Phase II/III trial design. This is studied in [4, 7].

1.3.3 Seamless Phase II/III design

In this approach Phase II and Phase III is seen as interwoven into one seamless trial, viewing the gap between Phase II and Phase III as an interim analysis. This means that the subjects are the same in Phase II and Phase III with probably some further recruitment of subjects into Phase III. In this way it is possible to use the information gathered in Phase II together with the information gathered in Phase III in the decision whether to approve a drug or not. More on this subject can be read about in [9, 17].

The problem of handling Type I errors arises with this approach as with the adaptive designs. Fortunately this problem has been studied extensively and the results from the research can be applied here.

---

\(^5\)In statistical analysis Type I errors are also called False positive. That is rejecting a null hypothesis when it actually is true.
Chapter 2

Problem description

2.1 Personalized health care

All patients may have different causes of getting a disease. Therefore treating every patient showing symptoms of a certain disease with the same treatment might not always be the best way to go [14]. A treatment having effect in one type patients may not have the same effect in a different type of patients.

Personalized health care involves manufacturing stratified medicine for patients that could be grouped together in some way. The groups could e.g. be patients having a similar benefit-risk profile that depends on a certain biomarker (BM) [14].

The part of the whole patient population consisting of subjects having a certain BM will in the reminder of this text be denoted by $P_{BM^+}$. The rest of the patient population will be denoted by $P_{BM^-}$ and the whole population will be referred to as $P_F$. If it is possible to partition $P_F$ into two subpopulations with respect to the benefit-risk profile, the response curve for a certain dose of a drug in $P_F$ could have the shape of the black curve in Figure 2.1.

One problem that arises when trying to identify patients with a better benefit-risk profile is that the classification is not always 100% reliable, meaning that a patient having a certain BM may not always show positive result on a test for having that BM. Moreover, a patient not having a BM can sometimes show positive response in a test for having that BM. Therefore the partition into subpopulations is not always correct.

2.2 Switching to a subpopulation in Phase III

Traditionally the main strategy has been to show efficacy in $P_F$ since it leads to being able to market the drug for a wider range of patients. Consequently conclusions drawn from a clinical trial are usually a result averaged over the whole sample. Even though a drug may be efficient in $P_{BM^+}$, making conclusions based on the average over the whole population might lead to incorrectly conclude that a drug is inefficient due to dilution. One way to tackle this problem is to evaluate the
CHAPTER 2. PROBLEM DESCRIPTION

Figure 2.1. The black curve visualizes how $P_F$ could respond to a treatment. The right hump could e.g. correspond to responses from a biomarker positive subpopulation, $P_{BM^+}$ and the left hump responses from a biomarker negative subpopulation, $P_{BM^-}$.

outcome in Phase II in order to guide us in the decision of either stopping the development of the candidate drug, continue in Phase III with only patients from $P_{BM^+}$ or $P_F$ and in this way increase the probability of a successful outcome in Phase III. Since reaching a broader patient population normally leads to higher commercial revenues, a risk is taken when deciding to continue the study in a targeted subpopulation only. Nevertheless, a risk is also taken when deciding to continue in $P_F$ since the trial might fail due to dilution.
2.3. Problem

Let us say that $\mathcal{P}$ can be partitioned into two distinct subpopulations, $\mathcal{P}_{BM^+}$ and $\mathcal{P}_{BM^-}$ having prevalence’s $\beta$ and $1 - \beta$ respectively. The distribution of the response to a drug in $\mathcal{P}$ could look then like a mixture of distributions as in Figure 2.1. Given that we have a prior knowledge that a drug has greater effect in $\mathcal{P}_{BM^+}$ than in $\mathcal{P}_{BM^-}$ and simulated responses from Phase II, we focus on analyzing the expected utility in each of the following two scenarios.

- Continue in Phase III recruiting patients only from $\mathcal{P}_{BM^+}$.
- Continue in Phase III recruiting patients from $\mathcal{P}$.

More precise explanation of these alternatives and some further assumptions are given in Section 3.1.

Note

If the difference in response between the subpopulations is small it might be hard to decide on which of the populations to choose. In this report, a number of Phase II studies is simulated with different $\beta$’s and different expected responses in the biomarker negative subpopulation.
Chapter 3

Method

In Section 3.1 we will go through the assumptions made about the circumstances under which the method will operate. In Section 3.2 some mathematical preliminaries for the algorithm given in Section 3.3.2 are described. In Section 3.4 a mathematical explanation of parts of Section 3.3.2 is presented.

3.1 Assumptions

3.1.1 Classifier

As mentioned in Section 2.1 the classification of subpopulations can sometimes involve errors. Assume the classifier classifies a member of \( P_{BM^+} \) as a member of \( P_{BM^+} \) with probability \( C_{BM^+ \in BM^+} \) and as member of \( P_{BM^-} \) with probability \( C_{BM^+ \in BM^-} \). In the same manner we assign the probability of a member of \( P_{BM^-} \) to be classified as a member of \( C_{BM^- \in BM^-} \) and to be incorrectly classified as a member of \( P_{BM^+} \) to \( C_{BM^- \in BM^+} \). Table 3.1 might aid to understanding the situation. A perfect classifier would give a table with probabilities \( C_{BM^+ \in BM^+} = 1 \) and \( C_{BM^+ \in BM^-} = C_{BM^- \in BM^+} = 0 \).

In this thesis it is assumed that a perfect classifier is used.

<table>
<thead>
<tr>
<th>( P_{BM^+} )</th>
<th>( P_{BM^-} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P_{BM^+} )</td>
<td>( C_{BM^+ \in BM^+} )</td>
</tr>
<tr>
<td>( P_{BM^-} )</td>
<td>( C_{BM^- \in BM^+} )</td>
</tr>
</tbody>
</table>

Table 3.1. Properties of the classification of subpopulations. The probability of a patient that belongs to the subpopulation in row \( i \) to be classified as a member of the subpopulation in column \( j \) is given in cell \( i,j \).
CHAPTER 3. METHOD

3.1.2 Measuring effect

Measuring effects of a drug could be done in different manners and can regard different endpoints. Which measurement that should be used depends on the disease and the drug to be tested.

One endpoint commonly used in oncology is measuring the survival time. Such endpoints are called time-to-event endpoints.

When treating Alzheimer’s disease there is a test measuring cognitive functionality called ADAS-Cog (Alzheimer’s Disease Assessment Scale-Cognitive Sub scale). The outcome of the test is summarized in a score. Hence an endpoint when testing treatments for Alzheimer’s could be measuring results in the ADAS-Cog test.

We will assume here that it is possible to summarize the effect of a treatment with a numerical value. The numerical value will be referred to as the outcome or the response of the treatment.

3.1.3 Dose response curve - E\text{max} model

It is assumed that J active doses and one placebo \( (d_0 = 0, d_1, d_2, \ldots, d_J) \) is tested in Phase II and that the relation between the doses is linear on the log-scale so it is reasonable to assume that \( d_j = j \). The responses \( Y_{P,d_j} \) to a certain dose \( d_j \) in a subpopulation \( P \) in both Phase II and Phase III are distributed as

\[
Y_{P,d_j} \sim N(\mu_{P,d_j}, \sigma^2),
\]

where \( N(\mu_{P,d_j}, \sigma^2) \) is the normal distribution with mean \( \mu_{P,d_j} \) and variance \( \sigma^2 \). The variance \( \sigma^2 \) is assumed to be known and the mean \( \mu_{P,d_j} \) is defined in (3.2) below, i.e.,

\[
\mu_{P,d_j} = E_0 + E_m \cdot g_{d_j}(ED_{50}, \gamma)
\]

where

\[
g_{d_j}(ED_{50}, \gamma) = \frac{d_j^\gamma}{d_j^\gamma + ED_{50}^\gamma}.
\]

Equation (3.2) is known as the E\text{max} model\textsuperscript{1} [17] where the parameters are described below

- \( E_0 \) corresponds to the placebo effect (dose \( d_0 \))
- \( E_0 + E_m = 'maximal response of the drug'\)
- \( \gamma \) is a shape parameter for the model
- \( ED_{50} \) is the dose that gives half of the response \( E_m \).

Figure 3.1 shows three different E\text{max} curves with parameters having the following values: \( E_0 = 5, ED_{50} = 4, \gamma = 4, J = 8 \) and different values for \( E_m \). The green curve is placebo and has \( E_m = 0 \). The blue curve is a population with \( E_m = 5 \) and the red curve is from a population with \( E_m = 15 \). With this setting, the response

\textsuperscript{1}The notation might be a bit confusing here, E\text{max} is the name of the model, \( E_m \) is a parameter in that model.
3.1. ASSUMPTIONS

![Plot of Emax dose response curves with parameters: E0 = 5, ED50 = 4, γ = 4 and dγ = 8. The red curve has Em = 15. The blue curve has Em = 5. The green curve has Em = 0.](image)

**Figure 3.1.** Plot of E_{max} dose response curves with parameters: E_0 = 5, E_{D50} = 4, \( \gamma = 4 \) and \( d_\gamma = 8 \). The red curve has \( E_m = 15 \). The blue curve has \( E_m = 5 \). The green curve has \( E_m = 0 \).  

distributions for dose \( d_3 \) for the three different \( E_{max} \) models have the shapes as in Figure 3.2.
Figure 3.2. Three plots of response (denoted by Y) distributions for dose $d_3$ from three different $E_{max}$ curves. The corresponding $E_{max}$ curves are the same as in 3.1. That is, the red curve has $E_m = 15$. The blue curve has $E_m = 5$ the green curve has $E_m = 0$. The rest of the parameters are: $E_0 = 5, ED_{50} = 4, \gamma = 4, d_j = 3$ and standard deviation $\sigma = 2$. 

12
3.1. ASSUMPTIONS

3.1.4 Utility

The utility of a Phase II/III trial is based on the financial gain for a successful Phase III trial. Further aspects considered are the costs for making the trial and the risks involved by taking a higher dose. To model the utility of a clinical trial is hard and we will present two different models in this report defined in (3.3) and in (3.4). \( U_{P,A} \) is suggested in [7] and is given in (3.3). \( U_{P,B} \) is based on \( U_{P,A} \) but takes the prevalence of targeted population \( P \) into account. More specifically

\[
U_{P,A}(d_j, n_3) = (1 - \Gamma(d_j))g - 2n_3c_3 - n_2c_2 \tag{3.3}
\]

and

\[
U_{P,B}(d_j, n_3) = \text{prev}_P \cdot ((1 - \Gamma(d_j))g - 2n_3c_3) - n_2c_2 \tag{3.4}
\]

where

- \( d_j \) is one of the doses tested in Phase II
- \( n_2 \) and \( n_3 \) are the sample sizes in Phase II and Phase III respectively
- \( \Gamma(d_j) \) is the probability that a trial with dose \( d_j \) fails in Phase III due to safety reasons
- \( g \) is the financial gain for an approved drug in Phase III
- \( c_2 \) and \( c_3 \) are the costs of one subject in Phase II and Phase III respectively
- \( \text{prev}_P \) is the prevalence of the population \( P \), i.e. \( \text{prev}_{P_{BM+}} = \beta, \text{prev}_{P_{BM-}} = (1 - \beta) \) and \( \text{prev}_{P_{F}} = 1. \)

The function \( \Gamma \) is a loss function hence it is to be minimized in order to maximize the utility. \( \Gamma \) is described in detail in Section 3.1.6.

3.1.5 Relation the between subpopulations

Furthermore it is assumed that the subpopulation’s \( E_{\text{max}} \) response curves only differ in the \( E_m \) parameter. The relation between \( E_m \) in \( P_{BM-} \) denoted by \( E_{m,BM-} \) and \( E_m \) in \( P_{BM+} \) denoted by \( E_{m,BM+} \) is \( E_{m,BM-} = r \cdot E_{m,BM+} \) where \( 0 \leq r \leq 1 \). Therefore we can write the parameter \( \mu_{P,d_j} \) in the \( E_{\text{max}} \) model given in (3.2) for the subpopulations \( P_{BM+} \) and \( P_{BM-} \) as

\[
\mu_{BM+,d_j} = E_0 + E_{m,BM+} \cdot g_{d_j}(ED_{50}, \gamma)
\]

and

\[
\mu_{BM-,d_j} = E_0 + r \cdot E_{m,BM+} \cdot g_{d_j}(ED_{50}, \gamma)
\]

where \( g_{d_j}(ED_{50}, \gamma) = \frac{d_j^{\gamma}}{d_j^{\gamma} + ED_{50}^{\gamma}} \).
As an example, in Figure 3.1 the ratio $r = \frac{E_{m,BM^-}}{E_{m,BM^+}}$ between the blue curve and the red curve is $\frac{5}{15}$.

In $P_F$, since we consider a mixture of two normal distributions, the variance in this distribution will differ from $\sigma^2$. Expressions for the mean $\mu_{F,d_j}$ and the variance $\sigma^2_{F,d_j}$ are given in (3.5) and (3.6).

$$\mu_{F,d_j} = \beta \cdot \mu_{BM^+,d_j} + (1 - \beta) \cdot \mu_{BM^-,d_j}$$  \hspace{1cm} (3.5)

$$\sigma^2_{F,d_j} = \beta \cdot (\mu^2_{BM^+,d_j} + \sigma^2_{BM^+,d_j}) + (1 - \beta) \cdot (\mu^2_{BM^-,d_j} + \sigma^2_{BM^-,d_j}) - \mu^2_{F,d_j}$$  \hspace{1cm} (3.6)

### 3.1.6 Safety function

We let the probability of a Phase III trial to fail due to safety reasons with dose $d_j$ be the quadratic function defined in (3.7).

$$\Gamma(d_j) = m \cdot \left(\frac{d_j}{d_f}\right)^2$$  \hspace{1cm} (3.7)

Figure 3.3 shows a smoothed plot of the safety function $\Gamma(d_j)$ with $d_0 = 0$, $d_f = 8$ and a maximum, $m = 0.2$. The model of the safety function is of course highly dependent on the drug to be tested.

### 3.1.7 Trial design

We test $J$ active doses plus placebo in Phase II. The proportion of the total number of subjects in Phase II assigned to each dose is denote by $b_{d_j}$. It is assumed that even in the subpopulations, the proportion of subjects that is assigned to each dose is the same as in the full population.

In Phase III, one active dose is tested against placebo with $n_3$ subjects in each dose arm.
3.2. Mathematical preliminaries and definitions

3.2.1 Test statistic

As assumed in Section 3.1.3, responses $Y_{P,d_j}$ in subpopulation $P$ exposed by dose $d_j$ in Phase III are normally distributed as

$$Y_{P,d_j} \sim N(\mu_{P,d_j}, \sigma^2),$$

where $\mu_{P,d_j}$ is the expected response level defined by the E$_{max}$ model and the variance $\sigma^2$ is assumed to be known.

We test the null hypothesis

$$H_{0,P} : \mu_{P,d_j} - \mu_{P,d_0} \leq 0$$

i.e., the effect of the treatment is no better than placebo against the alternative hypothesis

$$H_{1,P} : \mu_{P,d_j} - \mu_{P,d_0} > 0$$

meaning that the effect is better than placebo.
CHAPTER 3. METHOD

We estimate \( \mu_{P,d_j} - \mu_{P,d_0} \) by \( (\mu_{P,d_j} - \mu_{P,d_0})^* \) defined in (3.8).

\[
(\mu_{P,d_j} - \mu_{P,d_0})^* = \mu_{P,d_j}^* - \mu_{P,d_0}^* = \frac{1}{n_3} \sum_{i=1}^{n_3} Y_{i,P,d_j} - \frac{1}{n_3} \sum_{i=1}^{n_3} Y_{i,P,d_0} = \\
Y_{P,d_j} - Y_{P,d_0}.
\]

Since \( Y_{P,d_j} \sim N(\mu_{P,d_j}, \frac{\sigma^2}{n_3}) \) and \( Y_{P,d_0} \sim N(\mu_{P,d_0}, \frac{\sigma^2}{n_3}) \) we have that

\[
\frac{Y_{P,d_j} - Y_{P,d_0}}{\sqrt{\frac{2\sigma^2}{n_3}}} \sim N\left(\frac{\mu_{P,d_j} - \mu_{P,d_0}}{\sqrt{\frac{2\sigma^2}{n_3}}}, 1\right).
\]

Now we define a Z statistic in (3.9).

\[
Z = \frac{Y_{P,d_j} - Y_{P,d_0}}{\sqrt{\frac{2\sigma^2}{n_3}}}
\]

3.2.2 Probability of success (PoS)

As given in (3.9) we have the test statistic \( Z = \frac{Y_{P,d_j} - Y_{P,d_0}}{\sqrt{\frac{2\sigma^2}{n_3}}} \). Assume that we have the rejection level at \( 1 - \alpha \). Under the null hypothesis we have that \( (\mu_{P,d_j} - \mu_{P,d_0}) \leq 0 \), hence a successful outcome is equal to an observation of \( Z > z_{1-\alpha} \) where \( z_{1-\alpha} \) is the inverse cumulative standardized normal distribution evaluated at \( 1 - \alpha \) usually denoted by \( \Phi^{-1}(1 - \alpha) \).

The probability of a successful outcome for \( \mu_{P,d_j} \) and \( n_3 \) patients in Phase III, \( \text{PoS}(\mu_{P,d_j}, n_3) \) can be written as

\[
\text{PoS}(\mu_{P,d_j}, n_3) = P(Z > z_{1-\alpha}) = P\left(Z - \frac{\mu_{P,d_j} - \mu_{P,d_0}}{\sqrt{\frac{2\sigma^2}{n_3}}} > z_{1-\alpha} - \frac{\mu_{P,d_j} - \mu_{P,d_0}}{\sqrt{\frac{2\sigma^2}{n_3}}} \right) = \\
P\left(X > z_{1-\alpha} - \frac{\mu_{P,d_j} - \mu_{P,d_0}}{\sqrt{\frac{2\sigma^2}{n_3}}} \right) = 1 - \Phi\left(z_{1-\alpha} - \frac{\mu_{P,d_j} - \mu_{P,d_0}}{\sqrt{\frac{2\sigma^2}{n_3}}} \right) = \\
\Phi\left(\frac{\mu_{P,d_j} - \mu_{P,d_0}}{\sqrt{\frac{2\sigma^2}{n_3}}} - z_{1-\alpha} \right),
\]

(3.10)

where \( X \sim N(0,1) \).

In this context, the mean, \( \mu_{P,d_j} \) will be a sample from the posterior distribution that will be explained in Section 3.3.1.
3.2. MATHEMATICAL PRELIMINARIES AND DEFINITIONS

3.2.3 Expected PoS

The expected PoS for a Phase III trial in a subpopulation \( \mathcal{P} \) with \( n_3 \) patients exposed by dose \( d_j \) is denoted by \( \mathbb{E}[\text{PoS}]_{\mathcal{P},d_j,n_3} \) and is defined in (3.11). This gives

\[
\mathbb{E}[\text{PoS}]_{\mathcal{P},d_j,n_3} = \int \text{PoS}(\mu_{\mathcal{P},d_j}, n_3) dF(\mu_{\mathcal{P},d_j})
\]

where \( F \) is the cumulative distribution function for \( \mu_{\mathcal{P},d_j} \). This is in some texts referred to as the expected power [11].

In the full population \( \mathbb{E}[\text{PoS}]_{\mathcal{F},d_j,n_3} \) has to be adjusted by the prevalence's of the subpopulations. Hence we have

\[
\mathbb{E}[\text{PoS}]_{\mathcal{F},d_j,n_3} = \int I_{BM^+}(\mu_{\mathcal{P},d_j}) \beta \text{PoS}(\mu_{BM^+}, d_j, n_3) + I_{BM^-}(\mu_{\mathcal{P},d_j})(1 - \beta) \text{PoS}(\mu_{BM^-}, d_j, n_3) dF(\mu_{\mathcal{P},d_j})
\]

(3.12)

where

\[
I_{BM^+}(\mu_{\mathcal{P},d_j}) = \begin{cases} 
1 & \text{if } \mu_{\mathcal{P},d_j} \text{ is a sample from } BM^+ \\
0 & \text{if } \mu_{\mathcal{P},d_j} \text{ is a sample from } BM^- 
\end{cases}
\]

and

\[
I_{BM^-}(\mu_{\mathcal{P},d_j}) = \begin{cases} 
1 & \text{if } \mu_{\mathcal{P},d_j} \text{ is a sample from } BM^- \\
0 & \text{if } \mu_{\mathcal{P},d_j} \text{ is a sample from } BM^+ 
\end{cases}
\]

Numerically, with \( S \) samples \( [\mu^i_{\mathcal{P},d_j}, \ldots, \mu^S_{\mathcal{P},d_j}] \) from the posterior we calculate (3.12) as

\[
\frac{1}{S} \sum_{i=1}^{S} (I_{BM^+}(\mu^i_{\mathcal{P},d_j}) \beta \text{PoS}(\mu^i_{BM^+}, d_j, n_3) + I_{BM^-}(\mu^i_{\mathcal{P},d_j})(1 - \beta) \text{PoS}(\mu^i_{BM^-}, d_j, n_3))
\]

(3.13)

and (3.11) as

\[
\frac{1}{S} \sum_{i=1}^{S} \text{PoS}(\mu^i_{\mathcal{P},d_j}, n_3).
\]

(3.14)

3.2.4 Computing expected utility

We define the expected utility, \( \mathbb{E}[U_{\mathcal{P},X}(d_j, n_3)] \) for a Phase III trial in a population \( \mathcal{P} \) with \( n_3 \) patients exposed by dose \( d_j \) and one of the utility functions \( U_{\mathcal{P},X}(d_j, n_3) \) defined in Section 3.1.4 as

\[
\mathbb{E}[U_{\mathcal{P},X}(d_j, n_3)] = \int \text{PoS}(\mu_{\mathcal{P},d_j}, n_3) U_{\mathcal{P},X}(d_j, n_3) dF(\mu_{\mathcal{P},d_j}) = U_{\mathcal{P},X}(d_j, n_3) \int \text{PoS}(\mu_{\mathcal{P},d_j}, n_3) dF(\mu_{\mathcal{P},d_j}) = \mathbb{E}[\text{PoS}]_{\mathcal{P},d_j,n_3} \cdot U_{\mathcal{P},X}(d_j, n_3),
\]

(3.15)

where \( X \) is one of \( A \) or \( B \) and \( \mathbb{E}[\text{PoS}]_{\mathcal{P},d_j,n_3} \) is the expected PoS in Phase III.
3.2.5 Optimal design in Phase III

For \( n_2 \) patients in Phase II, the optimal design in Phase III for a population \( \mathcal{P} \) is the dose \( d_j \) and number of patients in Phase III, \( n_3 \) that maximizes the product \( U_{\mathcal{P},X}(d_j, n_3) \cdot \mathbb{E}[\text{PoS}|\mathcal{P}, d_j, n_3] \), i.e.,

\[
\arg \max_{d_j, n_3} (U_{\mathcal{P},X}(d_j, n_3) \cdot \mathbb{E}[\text{PoS}|\mathcal{P}, d_j, n_3]).
\] (3.16)

3.2.6 Expected utility from optimal design in Phase III

Let \( d_j \) and \( n_3 \) define the optimal Phase III design in the sense described above. The expected utility for the optimal design is then given in (3.17) below.

\[
\mathbb{E}[U_{\mathcal{P},X}(d_j, n_3)]^{opt} = U_{\mathcal{P},X}(d_j, n_3) \cdot \mathbb{E}[\text{PoS}|\mathcal{P}, d_j, n_3].
\] (3.17)
3.3 Method

The idea is to see how \( E[U_{P,X}(d_j, n_3)]^{opt} \) varies for different values of the ratio \( r = \frac{E_{m,BM} - E_{m,BM+}}{E_{m,BM+}} \) of the maximal expected response in the subpopulations, and based on that to analyze the expected utility of the options described in Section 2.3.

We approach the problem in a natural way by simulating a number of mean responses from different Phase II studies for a certain ratio, \( r \) and use this data to calculate the expected utility for the optimal Phase III design in each study.

3.3.1 Overview of the algorithm

We do the procedure described above by using the Bayesian framework. As our data comes from the \( E_{max} \) model described in (3.2) we are interested in getting a posterior for the unknown parameters, \( E_0, E_m, ED_{50} \) and \( \gamma \) given the response \( y \) from Phase II. Having good knowledge about these parameters we can calculate the expected utility for a trial design.

3.3.2 Detailed explanation of the algorithm

For syntactical ease the four dimensional parameter to the mean in the \( E_{max} \) model \((E_0, E_m, ED_{50}, \gamma)\) will from here on be referred to as \((\theta_1, \theta_2, \phi_1, \phi_2)\).

As will be seen in Section 3.4.1 it turns out to be a good idea to divide the 4-dimensional parameter \((\theta_1, \theta_2, \phi_1, \phi_2)\) into two 2-dimensional parameters \((\theta_1, \theta_2)\) and \((\phi_1, \phi_2)\) since this partition enables sampling from the posterior \(f(\theta_1, \theta_2, \phi_1, \phi_2 | y)\) by first sample from the posterior \(f(\phi_1, \phi_2 | y)\) and then sample from the conditional posterior \(f(\theta_1, \theta_2 | \phi_1, \phi_2, y)\).

However, since the posterior \(f(\phi_1, \phi_2 | y)\) is still hard to sample from, rejection sampling is used in order to sample from this distribution. The generic concept of rejection sampling is explained in Section 3.4.2.

Next follows the algorithm for computing the expected utility for an optimal designed Phase III study with the assumptions made in Section 3.1.

Generate parameters for the \( E_{max} \) model in Phase II

1. Simulate the parameters \( \theta_1^*, \theta_2^*, \phi_1^* \) and \( \phi_2^* \) for the \( E_{max} \) model from their prior distributions.
CHAPTER 3. METHOD

Generate mean response from Phase II

We simulate mean response from one Phase II study with \( J \) active doses plus placebo and \( n_2 \) subjects.

1. For each dose \( d_j \), simulate random normal distributed errors \( \epsilon_{d_j} \sim N(0, \sigma^2) \).

2. Generate \( J + 1 \) mean responses by \( y_{P,d_j} = \theta_1^* + \theta_2^* \frac{d_j \phi_2^*}{d_j \phi_2^* + \theta_1^*} + \frac{\epsilon_{d_j}}{\sqrt{n_2 d_j}} \). We put the responses in a vector \( y = y_P = (y_{P,d_0}, y_{P,d_1}, \ldots, y_{P,d_J}) \).

Define \( g \) in rejection sampling

The aim is to define a new function \( g(\phi_1, \phi_2) \) that is greater than \( f(\phi_1, \phi_2|y) \) at each point and easier to sample from.

Let \( \text{max}(a, b) \) be the maximum of \( a \) and \( b \) where \( a, b \in \mathbb{R} \).

1. Define a 2-dimensional grid \( G_1 \) on which \( \log f(\phi_1, \phi_2|y) \) is be calculated.

2. Calculate the posterior \( \log f(\phi_1, \phi_2|y) \) at the grid \( G_1 \).

3. Calculate the sample mean and variance of \( \phi_1 \) and \( \phi_2 \) and define a new more suitable grid, \( G_2 \) based on these values.

4. Calculate the posterior \( \log f(\phi_1, \phi_2|y) \) at the grid \( G_2 \).

5. Let \( (i, j), (i+1, j), (i, j+1) \) and \( (i+1, j+1) \) be four points at the grid \( G_2 \). Now, \( \log g(x_1, x_2) = \text{max}(\log f(i, j|y), \log f(i+1, j|y), \log f(i, j+1|y), \log f(i+1, j+1|y)) \) where \( x_1 \in [i, i+1], x_2 \in [j, j+1] \).

Sample from posterior

1. Let \( C \) be a suitable constant such that \( C \geq 1 \). For each sample \( (\phi_1^*, \phi_2^*) \) from \( \log C \cdot g(\phi_1, \phi_2) \) we calculate \( \log f(\phi_1^*, \phi_2^*|y) \).

2. We keep the sample with probability \( \log f(\phi_1^*, \phi_2^*|y) \cdot g(\phi_1^*, \phi_2^*) \).

3. If the sample is kept, we draw a sample \( (\theta_1^*, \theta_2^*) \) from \( f(\theta_1, \theta_2|\phi_1^*, \phi_2^*, y) \).

4. The sample \( (\phi_1^*, \phi_2^*) \) and \( (\theta_1^*, \theta_2^*) \) will serve as a sample from \( f(\theta_1, \theta_2, \phi_1, \phi_2|y) \).

Compute expected utility of optimal Phase III design

1. For each sample \( (\theta_1^*, \theta_2^*, \phi_1^*, \phi_2^*) \) we form the corresponding mean for each dose \( d_j \) by \( \mu_{P,d_j}^* = \theta_1^* + \theta_2^* \frac{d_j \phi_2^*}{d_j \phi_2^* + \theta_1^*} \) according to the \( E_{max} \) model.

2. For each \( \mu_{P,d_j}^* \) and \( n_3 \) compute PoS(\( \mu_{P,d_j}^*, n_3 \)) = \Phi \left( \frac{\mu_{d_j}^* - \mu_{d_0}^*}{\sqrt{n_3}} - z_{1-\alpha} \right) .

20
3.4. MATHEMATICAL JUSTIFICATION

3. Using the samples from the posterior, compute the $E[\text{PoS}]_{\mathcal{P},d_j,n_3}$ for each $d_j$ and $n_3$.

4. Calculate the expected utility for the optimal design $E[U_{\mathcal{P},X}(d_j,n_3)]^{opt}$.

Note

In the full population, $\mathcal{P}_F$ we do not sample from the posterior explicitly. Instead we compute $E[\text{PoS}]_{\mathcal{P},d_j,n_3}$ based on samples from the posterior for $\mathcal{P}_{BM+}$ and $\mathcal{P}_{BM-}$.

3.4 Mathematical justification

We put the parameters for the $E_{\text{max}}$ model in two column vectors as

$$
\begin{bmatrix}
\phi_1 \\
\phi_2
\end{bmatrix} = \phi
$$

and

$$
\begin{bmatrix}
\theta_1 \\
\theta_2
\end{bmatrix} = \theta.
$$

3.4.1 Computing and sampling from the posterior distribution

The outcome $Y_{\mathcal{P},d_j}$ of a certain dose $d_j$ and subpopulation $\mathcal{P}$ in Phase II is distributed as

$$
Y_{\mathcal{P},d_j} \sim N(\mu_{\mathcal{P},d_j}, \sigma^2),
$$

where $\sigma^2$ is known and $\mu_{\mathcal{P},d_j}$ is the mean in the $E_{\text{max}}$ model described in (3.2) and is given below in the new notation for quick reference

$$
\mu_{\mathcal{P},d_j} = \theta_1 + \theta_2 \cdot g_{d_j}(\phi) = \begin{bmatrix} 1 & g_{d_j}(\phi) \end{bmatrix} \begin{bmatrix} \theta_1 \\
\theta_2 \end{bmatrix},
$$

where

$$
g_{d_j}(\phi) = \frac{d_j \phi_2}{\phi_2^2 + \phi_1^2}.
$$

Distribution of response vector

Since the estimator in Phase II is the sample mean for each dose, response $y_{\mathcal{P},d_j}$ to a dose $d_j$ in Phase II has the distribution

$$
y_{\mathcal{P},d_j} \sim N\left( \begin{bmatrix} 1 & g_{d_j}(\phi) \end{bmatrix} \begin{bmatrix} \theta_1 \\
\theta_2 \end{bmatrix}, \frac{\sigma^2}{n_2 b_{d_j}} \right).
$$

Where $b_{d_j}$ is the proportion of the subjects in Phase II assigned to dose $d_j$. We put all the responses in a vector $y = (y_{d_0}, y_{d_1}, \ldots, y_{d_J})^T$. Since each $y_{d_i}$ in $y =$
(y_{d_0}, y_{d_1}, \ldots, y_{d_J})^T$ is independent of the others, the covariance matrix $\Lambda_y = \text{Cov}(y)$ is a $(J+1) \times (J+1)$ diagonal matrix hence,

$$
\Lambda_y = \text{Cov}(y) = \begin{bmatrix}
\sigma^2 / n^{2b_{d_0}} & 0 & \ldots & 0 \\
0 & \sigma^2 / n^{2b_{d_1}} & 0 & \vdots \\
\vdots & 0 & \ddots & 0 \\
0 & \ldots & 0 & \sigma^2 / n^{2b_{d_J}}
\end{bmatrix}.
$$

(3.18)

This yields the distribution of $y$ as

$$
y \sim N(X_\phi \theta, \Lambda_y),
$$

where $X_\phi$ has the form as in (3.19).

$$
X_\phi = \begin{bmatrix}
1 & g_{d_0}(\phi) \\
1 & g_{d_1}(\phi) \\
\vdots & \vdots \\
1 & g_{d_J}(\phi)
\end{bmatrix}
$$

(3.19)

**Posterior**

We wish to compute the posterior distribution $f(\theta, \phi|y)$ for the parameter $(\theta, \phi)$. Since $f(\phi, \theta|y) = f(\phi|y)f(\theta|\phi, y)$ we can a sample from $f(\phi, \theta|y)$ in the two steps as suggested in [6] and given below.

1. Draw a sample $\phi^*$ from the posterior $f(\phi|y)$.
2. Draw a sample $\theta^*$ from the conditional posterior $f(\theta|\phi^*, y)$.

Recall that the sample from $f(\phi|y)$ was obtained by rejection sampling. To sample from $f(\theta|\phi, y)$ turns out to be easy in this case since it has a normal distribution.

For $\phi$ it turns out to be nice to work with the logarithm of the distribution. Next follows a detailed derivation of the posterior with proper distributions.
3.4. MATHEMATICAL JUSTIFICATION

Calculating $f(\theta|\phi^*, y)$

$$f(\theta|\phi^*, y) = L(y|\theta, \phi^*)\pi(\theta|\phi^*)$$

We have a 2-dimensional normal distribution as prior for $\theta|\phi^*$ defined by

$$\theta|\phi^* \sim N(\mu_b, \Sigma_b),$$

where

$$\Sigma_b = \begin{bmatrix} \sigma_{b_1}^2 & 0 \\ 0 & \sigma_{b_2}^2 \end{bmatrix}$$

so

$$\pi(\theta|\phi^*) = \frac{1}{2\pi \sqrt{\det(\Sigma_b)}} e^{-\frac{1}{2}(\theta - \mu_b)^T \Sigma_b^{-1} (\theta - \mu_b)}.$$ 

The likelihood $L(y|\theta, \phi^*)$ is defined by

$$L(y|\theta, \phi^*) = \frac{1}{(2\pi)^{J+1/2} \sqrt{|\Lambda_y|}} e^{-\frac{1}{2}(y - X_\phi \theta)^T \Lambda_y^{-1} (y - X_\phi \theta)},$$

i.e. we have a $J + 1$ dimensional normal distribution

$$y|\theta, \phi^* \sim N(X_\phi \theta, \Lambda_y).$$

Combining $\pi(\theta|\phi^*)$ and $L(y|\theta, \phi^*)$ we get the posterior as

$$\pi(\theta|\phi^*)L(y|\theta, \phi^*)$$

$$= \frac{1}{2\pi \sqrt{\det(\Sigma_b)}} e^{-\frac{1}{2}(\theta - \mu_b)^T \Sigma_b^{-1} (\theta - \mu_b)} \cdot \frac{1}{(2\pi)^{J+1/2} \sqrt{|\Lambda_y|}} e^{-\frac{1}{2}(y - X_\phi \theta)^T \Lambda_y^{-1} (y - X_\phi \theta)}$$

$$\propto \exp(-\frac{1}{2}((\theta - \mu_b)^T \Sigma_b^{-1} (\theta - \mu_b) + (y - X_\phi \theta)^T \Lambda_y^{-1} (y - X_\phi \theta)))$$

$$\propto \exp\left(-\frac{1}{2} \left( \sum_{i=1}^{2} \frac{(\theta_i - \mu_{bi})^2}{\sigma_{bi}^2} + \sum_{j=0}^{J} \frac{(y_j - (\theta_1 - \theta_2 g_{1j}(\phi^*))^2)}{d_{b_1} n_2 \sigma^2} \right) \right)$$

$$\propto \exp\left(-\frac{1}{2} \left( \sum_{i=1}^{2} \frac{(\theta_i - \mu_{bi})^2}{\sigma_{bi}^2} + \frac{n_2}{\sigma^2} \sum_{j=0}^{J} b_{d_j}(y_j - (\theta_1 - \theta_2 g_{1j}(\phi^*))^2) \right) \right)$$

$$\propto \exp\left(-\frac{1}{2} \left( \sum_{i=1}^{2} \frac{(\theta_i - \mu_{bi})^2}{\sigma_{bi}^2} + \frac{n_2}{\sigma^2} \sum_{j=0}^{J} b_{d_j}(y_j^2 + (\theta_1 - \theta_2 g_{1j}(\phi^*))^2 - 2y_j(\theta_1 - \theta_2 g_{1j}(\phi^*)))^2 \right) \right)$$

From Bayesian regression analysis [12] this is

$$\propto \exp\left(-\frac{1}{2}(\theta - \mu_c)^T \Lambda^{-1}(\theta - \mu_c)\right),$$

(3.20)
where
\[ \Lambda^{-1} = \Sigma_b^{-1} + X^T \phi^* \Lambda_y^{-1} X_{\phi^*} \] (3.21)

and
\[ \mu_c = (\Sigma_b^{-1} + X^T \phi^* \Lambda_y^{-1} X_{\phi^*})^{-1} (\Sigma_b^{-1} \mu_b + X^T \Lambda_y^{-1} y). \] (3.22)

**Calculating** \( f(\phi|y) \)

By Bayes rule we have that
\[ f(\phi|y) \propto L(y|\phi)\pi(\phi). \]

We have a 2-dim prior for \( \phi \) defined by
\[ \phi \sim N(\mu_a, \Sigma_a), \]

where
\[ \Sigma_a = \begin{bmatrix} \sigma_{a1}^2 & 0 \\ 0 & \sigma_{a2}^2 \end{bmatrix}, \]

that is
\[ \pi(\phi) = \frac{1}{2\pi \sqrt{|\Sigma_a|}} e^{-\frac{1}{2}(\phi - \mu_a)^T \Sigma_a^{-1} (\phi - \mu_a)} \]

and
\[ \log \pi(\phi) \propto -\frac{1}{2}(\phi - \mu_a)^T \Sigma_a^{-1} (\phi - \mu_a). \] (3.23)

The marginal likelihood of observing data \( y \) under \( \phi \) is defined by
\[ L(y|\phi) = \int L(y|\phi, \theta)\pi(\theta) d\theta. \]

From [7] we have that
\[ \log L(y|\phi) \propto \frac{1}{2} (\mu_c^T \Lambda^{-1} \mu_c - \log |\Lambda^{-1}|), \] (3.24)

where \( \Lambda^{-1} \) is given in (3.21) and \( \mu_c \) is given in (3.22). Combining (3.23) and (3.24) yields the logarithm for the posterior of \( \phi \) as
\[ \log f(\phi|y) \propto \log L(y|\phi) + \log \pi(\phi). \]

We get the posterior by
\[ f(\phi|y) \propto e^{\log f(\phi|y)}. \]
3.4. MATHEMATICAL JUSTIFICATION

3.4.2 Rejection sampling

Rejection sampling is used as a method for sampling from a complicated distribution with a density function, say $f$, defined on $\mathbb{R}^n$. $f$ is too complicated to sample from directly so a new distribution $g$ is considered from which sampling is easily done. The function $g$ should be such that $g(x) \geq C \cdot f(x)$ for every $x$ where $C$ is a predefined constant. If we draw a sample $x$ from $g$, then $x$ will serve as a sample from $f$ with probability $\frac{f(x)}{Cg(x)}$.

To make the method accept as many proposals, $x$ as possible the definition of $g$ is crucial. Next we present the methodology used in this study to construct $g$. Then we show how to sample from $f$ in practice.

Construct $g$

To keep as many samples as possible, $C \cdot g$ has to be close to $f$. To achieve this we first calculate $f(x)$ on a predefined grid as in Figure 3.4. Based on these values we define values for $g(x)$ that is greater than or equal to $f(x)$ for every $x$ in the grid. This will correspond to the blue points in Figure 3.5. Then in Figure 3.6, an example of a density $g$ that has been defined based on $f(x)$ for points $x$ in the grid. Figure 3.7 shows a plot of $g$ multiplied by the constant $C$ where $C = 1.2$.

Sample from $f$

Now we can sample from $f$ by first draw a sample $x$ from $C \cdot g$ and then draw a sample $u$ from the uniform distribution, $U(0,1)$. If $u < \frac{f(x)}{Cg(x)}$ we keep $x$ as a sample from $f$. 

25
Why does this work?

We can denote the probability of getting an accepted sample \( x \) from the rejection sampling method as \( P(x|\text{accepted } x) \). The aim is to prove that

\[
P(x|\text{accepted } x) = f(x).
\]

Let

\[
P(x) = g(x).
\]

We can write

\[
P(x|\text{accepted } x) = \frac{P(\text{accepted } x \cap x)}{P(\text{accepted } x)} = \frac{P(\text{accepted } x| x)P(x)}{P(\text{accepted } x)}
\]

and by the criterion for accepting a sample

\[
P(\text{accepted } x| x) = P\left( u \leq \frac{f(x)}{C \cdot g(x)} \right) = \frac{f(x)}{C \cdot g(x)}
\]

where \( u \sim U(0, 1) \). By the law of total probability

\[
P(\text{accepted } x) = \int P(\text{accepted } t| t) g(t) dt = \int P(\text{accepted } x) g(t) dt = \int \frac{f(t)}{C} g(t) dt = \frac{1}{C} \int f(t) dt = \frac{1}{C}.
\]

Thus \( P(x|\text{accepted } x) \) can be rewritten as

\[
P(x|\text{accepted } x) = \frac{f(x) g(x)}{\frac{1}{C}} = f(x).
\]
3.4. MATHEMATICAL JUSTIFICATION

Figure 3.4. $f(x)$ is calculated at a grid.
Figure 3.5. Points for the density $g$ are selected to be greater than or equal than $f$ at every point of the grid.
3.4. MATHEMATICAL JUSTIFICATION

Figure 3.6. A continuous probability density $g$ that is easy to sample from. $g$ is defined based on samples from $f$. 
Figure 3.7. A continuous probability density $C \cdot g$ that is easy to sample from. $C \cdot g$ is defined so that it should be greater than $f$ in each point. Here $C = 1.2$. 
Chapter 4

Results

4.1 Simulation study

Two simulations studies have been made in order to analyze the expected utility for the optimal Phase III design, \( E[U_{P,X}(d_j, n_3)]^{opt} \) with the dose \( d_j \) and \( n_3 \) subjects in each of the dose arms in Phase III. The first simulation study is using the utility function \( U_{P,A} \) and the second is using the utility function \( U_{P,B} \). In order to analyze the impact of the prevalence \( \beta \) of the biomarker positive subpopulation, different values of \( \beta \) has been considered. Furthermore, the impact of the ratio \( r \) has been evaluated for each studied value of \( \beta \).

In Section 4.1.1 the parameter values that define the simulations are given together with a short explanation. In Section 4.1.2 the results of the simulations visualized in the figures 4.1-4.16 are described. In Section 4.1.3 and 4.1.4 the results from the simulations are discussed and interpreted.

4.1.1 Parameters for the simulation studies

Costs and utilities

We define the financial gain \( g \) for a new drug that can be advertised at the market to be 12000 times as high as the costs \( c_2, c_3 \) for one patient in Phase II and Phase III respectively.

\[
g = 12000 \\
c_2 = 1 \\
c_3 = 1
\]
Number of doses in Phase II

The number of active doses in Phase II is 3. So we have

\[ J = 3 \]

and the doses, here formatted as a vector are

\[ [d_0, d_1, d_2, d_3] = [0, 1, 2, 3], \]

where 0 is placebo. This is the number of doses that traditionally is set in Phase II but it can vary between therapeutic areas and between companies.

Number of Phase II subjects

The number of subjects in Phase II is 350. So

\[ n_2 = 350. \]

This number is proposed in [7] and since these simulation are based on the same procedure as in [7], it seems reasonable to use this value here.

The number of subjects assigned to dose \( d_j \) is \( b_{d_j} \cdot n_2 \). We set

\[ b_{d_0} = 0.3 \]

and for \( j > 0 \) we set

\[ b_{d_j} = \frac{1 - 0.3}{3} = \frac{0.7}{3}. \]

Number of Phase III subjects

The number of subjects in each dose arm in Phase III is 1000. So

\[ n_3 = 1000. \]

This number is also proposed in [7]. Note that, by using one single value here, the optimal design in Phase III will only be dependent on the dose.

Priors for the parameters in the \( E_{\text{max}} \) model

We assume normally distributed priors for the individual parameters in the \( E_{\text{max}} \) model.

Since \( E_0 \) is the effect of placebo, we assume that we can be quite certain about this parameter, measuring this parameter we can take information from different studies of the same disease. We can also measure this parameter without having even developed any drug. Hence we give this parameter a relatively small variance.

\[ E_0 \sim N(5, 2) \]
4.1. SIMULATION STUDY

Information about $E_m$ is probably also something that we can be quite sure about since the maximal effect of our treatment is not likely to be dramatically larger or smaller than other drugs studied. Recall that it is not only the positive effect that is regarded when testing a drug but also the possible adverse effects.

$$E_{m,BM^+} \sim N(15, 3)$$

The $ED_{50}$ parameter is certainly more dependent on the drug to be tested, thereby we put a relatively high variance in the prior for it.

$$ED_{50} \sim N\left(\frac{d_J - d_0}{2}, (d_J - d_0)^2\right) = N(\frac{3}{2}, 3^2)$$

The shape parameter $\gamma$ is also highly dependent on the tested drug so we set this variance relatively high.

$$\gamma \sim N\left(\frac{d_J - d_0}{J}, 1\right) = N(1, 1)$$

Variance for outcomes in Phase II

In the model we have that response $Y_{P,d_j}$ for a subpopulation $P$ and dose $d_j$ is normally distributed with mean $\mu_{P,d_j}$ and variance $\sigma^2$. That is $Y_{P,d_j} \sim N(\mu_{P,d_j}, \sigma^2)$. Here the standard deviation $\sigma$ is selected to be 9. So

$$\sigma = 9.$$  

Prevalence’s

Different values of the prevalence $\beta$ of the biomarker positive subpopulation $P_{BM^+}$ has been used in the simulations in order to analyze its influence of the results. Below is a list of the values of $\beta$ that has been used.

$$0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9.$$  

$E_m$ ratios

The different ratios between the $E_m$ parameter in $P_{BM^-}$ and $P_{BM^+}$ denoted by $r$ that are compared are

$$0, 0.2, 0.4, 0.6, 0.8, 1.$$  

Rejection level in Phase III

I practice the RA almost always requires that a drug passes two Phase III trials with significance $\alpha = 0.025$ in order to approve it. This study is focused on one single Phase III study. However, if we do two independent studies, under the null hypothesis the probability having significant result result in both studies is $\alpha \cdot \alpha = 0.025^2 = 0.000625$. I these simulations, a slightly more conservative rejection level $\alpha$ has been used.

$$\alpha = 0.0005$$
Safety function

We set the maximum value of the safety function, $\Gamma$ to be 0.2 so

$$m = 0.2$$

which yields

$$\Gamma(d_j) = m \cdot \left( \frac{d_j}{d_J} \right)^2 = 0.2 \cdot \left( \frac{d_j}{3} \right)^2.$$ 

Number of simulated samples from posterior

The number of simulated samples from the posterior for the mean in the $E_{max}$ model is 300 for each subpopulation. So when analyzing the full population, there will be 600 samples, since the samples are concatenated. This is chosen to be as large as possible and the choice is a consequence of computer limitation.

Number of simulated Phase II studies

Number of simulated Phase II studies is 100. This number is also a consequence of computer limitations.

Classifier

It is assumed that the diagnostic test always makes the correct classification in $P_{BM+}$ and $P_{BM-}$, that is, using the notation in Section 3.1.1, 

$$C_{BM+ \in BM+} = C_{BM- \in BM-} = 1$$

and

$$C_{BM+ \in BM-} = C_{BM- \in BM+} = 0.$$
4.1. SIMULATION STUDY

4.1.2 Interpretation of plots

Figures 4.1-4.16 show boxplots from the simulation studies. In each figure, the left plot shows how the expected utility for the optimal design, $E[U_{P,A}(d_j, n_3)]^{opt}$ varies between the simulated Phase II studies for each ratio $r$. The right plot shows how the expected PoS when choosing the optimal design varies between the simulated Phase II studies. The red plots are results from $P_{BM+}$ and the green plots are results from $P_F$. For the sake of reference, the results from $P_{BM-}$ also given, these are plotted in blue.

4.1.3 Simulation study with utility function $U_{P,A}$

Figures 4.1-4.7 correspond to the utility function $U_{P,A}$ that does not take the prevalence into account and is defined in (3.3).

Since there is no drawback built in to $U_{P,A}$ for going in $P_{BM+}$, the simulations tell us that there will always be a best choice to go in $P_{BM+}$ in Phase III.

However, we can see that where the effect is similar in $P_{BM+}$ and $P_{BM-}$, the utility is obviously similar in the three populations. This can be seen for values of $r$ that are close to 1 and for high values on the $P_{BM+}$ prevalence $\beta$. 

35
Figure 4.1. The prevalence of $P_{BM^+}$ is $\beta=0.1$ and the utility function is $U_{P,A}$. The ratio between the $E_m$ parameter in the subpopulations is $r$. More specifically, $r = \frac{E_m,BM^-}{E_m,BM^+}$. 
Figure 4.2. The prevalence of $P_{BM^+}$ is $\beta=0.2$ and the utility function is $U_{p,A}$. The ratio between the $E_m$ parameter in the subpopulations is $r$. More specifically $r = \frac{E_{m,BM^-}}{E_{m,BM^+}}$. 

4.1. SIMULATION STUDY
Figure 4.3. The prevalence of $P_{BM^+}$ is $\beta=0.4$ and the utility function is $U_{P,A}$. The ratio between the $E_m$ parameter in the subpopulations is $r$. More specifically $r = \frac{E_{m,BM^-}}{E_{m,BM^+}}$. 
4.1. SIMULATION STUDY

Figure 4.4. The prevalence of $P_{BM^+}$ is $\beta=0.5$ and the utility function is $U_{P,A}$. The ratio between the $E_m$ parameter in the subpopulations is $r$. More specifically $r = \frac{E_m, BM^-}{E_m, BM^+}$. 

39
The prevalence of $P_{BM+}$ is $\beta=0.6$ and the utility function is $U_{P,A}$. The ratio between the $E_m$ parameter in the subpopulations is $r$. More specifically
\[ r = \frac{E_{m,BM-}}{E_{m,BM+}}. \]
4.1. SIMULATION STUDY

Figure 4.6. The prevalence of $p_{BM^+}$ is $\beta=0.8$ and the utility function is $U_{\text{BM},A}$. The ratio between the $E_m$ parameter in the subpopulations is $r$. More specifically $r = \frac{E_m, BM^-}{E_m, BM^+}$. 

41


**Figure 4.7.** The prevalence of $\mathcal{P}_{BM^+}$ is $\beta=0.9$ and the utility function is $U_{\mathcal{P},A}$. The ratio between the $E_m$ parameter in the subpopulations is $r$. More specifically $r = \frac{E_{m,BM^-}}{E_{m,BM^+}}$. 

---

**CHAPTER 4. RESULTS**
4.1. SIMULATION STUDY

4.1.4 Simulation study with utility function $U_{P,B}$

Figures 4.8-4.16 correspond to the utility function $U_{P,B}$ that takes the prevalence into account and is defined in (3.4).

Looking at the results of $P_{BM}$ in Figure 4.8 where the prevalence of $P_{BM}$ is 10% ($\beta = 0.1$). It is not surprising that the expected utility is below zero since the maximal telly is $12000 \cdot 0.1 = 1200$, hence the costs exceed the utility. The reason why the expected PoS is low is that the procedure of choosing the optimal Phase III design guides us to take lower doses in order to minimize the risk of failing due to safety reasons. Lower doses reduces the expected PoS.

The trend these simulations points to that for low values of $\beta$ it is better to choose to continue in $P_F$ in Phase III since it leads to higher expected utilities.

However, when the prevalence of $P_{BM}$ is larger than the prevalence of $P_{BM}$ (i.e $\beta > 0.5$) the choice is not obvious. From the simulations we can see the importance of analyzing the expected utility and the corresponding PoS in conjunction. If we analyze the expected utility of a Phase III trial design in isolation there is a risk that the decisions we make about in which population to go involves unnecessary high risks.

Consider $\beta = 0.9$ (see Figure 4.16), the expected utility is larger in $P_F$ for all values of the ratio $r$. When analyzing the corresponding expected PoS for each ratio, the expected PoS is higher in $P_{BM}$ than in $P_F$. Such information could lead a developer to continue in Phase III in $P_{BM}$ rather than in $P_F$ even though $P_F$ could possibly bring more utility.

Another example is when $\beta = 0.6$ and $r = 0.2$, the median of the expected utility in $P_{BM}$ is 1900 and in $P_F$ it is 3050. If we look at the related PoS in conjunction, we see that the median is 0.58 in $P_{BM}$ and 0.44 in $P_F$ which could lead us to choose to continue in $P_{BM}$ in these situations, see Figure 4.13. The same pattern is seen in Figure 4.14 and Figure 4.15.

**Note**

Since the cost of a Phase II/III trial is $2n_3c_3 + n_2c_2 = 2\cdot 1000 \cdot 1 + 350 \cdot 1 = 2350$ and the gain for a successful Phase III trial, $g = 12000$ we know that the total gain cannot be larger than $12000 - 2350 = 12000 - 2350 = 9650$. 

Figure 4.8. The prevalence of $\mathcal{P}_{BM^+}$ is $\beta=0.1$ and the utility function is $U_{\mathcal{P}_B}$. The ratio between the $E_m$ parameter in the subpopulations is $r$. More specifically $r = \frac{E_{m,BM^-}}{E_{m,BM^+}}$. 
4.1. SIMULATION STUDY

Figure 4.9. The prevalence of $P_{BM^+}$ is $\beta=0.2$ and the utility function is $U_{P,R}$. The ratio between the $E_m$ parameter in the subpopulations is $r$. More specifically $r = \frac{E_{m,BM^-}}{E_{m,BM^+}}$. 

45
Figure 4.10. The prevalence of $\mathcal{P}_{BM^+}$ is $\beta=0.3$ and the utility function is $U_{\mathcal{P},B}$. The ratio between the $E_m$ parameter in the subpopulations is $r$. More specifically $r = \frac{E_{m,BM^-}}{E_{m,BM^+}}$. 
4.1. SIMULATION STUDY

Figure 4.11. The prevalence of $P_{BM^+}$ is $\beta=0.4$ and the utility function is $U_{P,B}$. The ratio between the $E_m$ parameter in the subpopulations is $r$. More specifically $r = \frac{E_{m,BM^-}}{E_{m,BM^+}}$. 
Figure 4.12. The prevalence of $P_{BM+}$ is $\beta=0.5$ and the utility function is $U_{P,B}$. The ratio between the $E_m$ parameter in the subpopulations is $r$. More specifically $r = \frac{E_{m,BM-}}{E_{m,BM+}}$. 
Figure 4.13. The prevalence of $\mathcal{P}_{BM^+}$ is $\beta = 0.6$ and the utility function is $U_{\mathcal{P}, B}$. The ratio between the $E_m$ parameter in the subpopulations is $r$. More specifically $r = \frac{E_{m, BM^-}}{E_{m, BM^+}}$. 
Figure 4.14. The prevalence of $P_{BM^+}$ is $\beta = 0.7$ and the utility function is $U_{P,B}$. The ratio between the $E_m$ parameter in the subpopulations is $r$. More specifically $r = \frac{E_{m,BM^+}}{E_{m,BM^-}}$. 
4.1. SIMULATION STUDY

**Figure 4.15.** The prevalence of $\mathcal{P}_{BM^+}$ is $\beta=0.8$ and the utility function is $U_{\mathcal{P}, B}$. The ratio between the $E_m$ parameter in the subpopulations is $r$. More specifically $r = \frac{E_{m, BM^-}}{E_{m, BM^+}}$. 

The prevalence of $\mathcal{P}_{BM^+}$ is $\beta=0.8$ and the utility function is $U_{\mathcal{P}, B}$. The ratio between the $E_m$ parameter in the subpopulations is $r$. More specifically $r = \frac{E_{m, BM^-}}{E_{m, BM^+}}$. 

51
Figure 4.16. The prevalence of $\mathcal{P}_{BM^+}$ is $\beta=0.9$ and the utility function is $U_{P,B}$. The ratio between the $E_m$ parameter in the subpopulations is $r$. More specifically $r = \frac{E_{m,BM^+}}{E_{m,BM^-}}$. 
4.2 Discussion and possible future extensions

4.2.1 Positive outcome

In this study, what we regard as a successful outcome is a statistical significant effect of a dose in Phase III. In a subpopulation this seems to be fair to claim, but for $\mathcal{F}$ the RA’s might be of another opinion. So, what is enough for a positive outcome in $\mathcal{F}$? The quick answer might be, a significant outcome in Phase III. However, if the RV’s have the prior knowledge that the drug has been tested in both $\mathcal{P}_B^{M^+}$ and $\mathcal{P}_B^{M^-}$, they will probably need to be convinced that the drug is effective in both $\mathcal{P}_B^{M^+}$ and $\mathcal{P}_B^{M^-}$ to approve the drug in $\mathcal{F}$. A possible alternative method of computing $\text{E}[\text{PoS}]_{d,j,n_3}$ which might meet this requirement better could be as in (4.1) below.

$$\text{E}[\text{PoS}]_{d,j,n_3} = \min(\text{E}[\text{PoS}]_{\mathcal{P}_B^{M^+},d,j,n_3}, \text{E}[\text{PoS}]_{\mathcal{P}_B^{M^-},d,j,n_3}),$$

where $\min(a, b)$ is the minimum value of $a$ and $b$ for $a, b \in \mathbb{R}$.

4.2.2 The utility functions

As mentioned above, the choice of utility function is a complex problem. It depends on the disease, the type of molecule, the competition of other drugs etc. The two utility functions given in this report should be seen as examples. When using the method described in this study to find a trial design, one should tailor the utility function for that specific situation.

It is reasonable to believe that at least when $\beta = 0.5$ and $r = 0.9$ it should be better to choose $\mathcal{F}$ which would support for $U_{P,B}$. On the other hand, for $\beta = 0.1$ and $r = 0.1$, it could at least in some cases be reasonable to believe that continuing in $\mathcal{P}_B^{M^+}$ would at least bring some utility as it does with $U_{P,A}$. With the assumptions made in this study, the first argument would guide us to use $U_{P,B}$ and the latter argument would guide us to $U_{P,A}$.

One example of a parameter of importance that has not been taken into account in the models used in this study is the recruitment of patient process, in particular in Phase III. If the prevalence of the biomarker positive subpopulation is low (low values of $\beta$) and we choose to continue to Phase III with that population only, it is presumably harder and thereby also more costly to recruit subjects.

Another parameter is that the financial gain for developing a drug is highly dependent on the patent life time still remaining on the product upon market entry. Such impact of the utility could also be included in the utility model.

4.2.3 Sampling posterior

Other method for sampling from the posterior than the one used in this study could also be used. The sampling procedure could for example be done with MCMC$^1$

---

$^1$Markov Chain Monte Carlo
methods, e.g. GIBBS sampler.

Since the method used here can be viewed as an extension of [7], to use the sampling procedure that was already implemented in [7], seemed like the approach to be preferred instead of implementing some alternative method. Moreover, the burn-in phase for a MCMC simulation could potentially take long time and in that case be a limitation for testing different $E_{\text{max}}$ models.

**Optimal Phase II/III design**

In [7], the topic of finding a jointly optimal Phase II/III design with respect to the number of patients in Phase II ($n_2$), the number of patients in Phase III ($n_3$), and which dose ($d_j$) to use in Phase III is studied. One interesting and possible extension of this work would be to analyze the impact of the values of $n_2$, $n_3$ and $d_j$ on the expected utility, in addition to the subpopulations impact studied in this report.

**Noisy classifier**

Another interesting way to extend this work is adding support for classifiers that is not perfect. This would e.g. alter the way how we recruit patients in Phase III since we cannot be sure if the recruited patients actually belongs to the targeted subpopulation. It would also have influence on the posterior, since some patients treated as belonging to $P_{BM+}$ might actually belong to $P_{BM-}$ and vice versa which will have impact of the responses in Phase II.

**Test strategy**

In this report we decided on the two test strategies in Phase III defined in Section 2.3. Other strategies that could be evaluated are e.g.

- To first run a Phase III trial in $P_{BM+}$ only and then in $P_F$.
- Do two separate Phase III trials in parallel, one in $P_{BM+}$ and one in $P_F$.
- Do an interaction test meaning that we test if there is effect in $P_{BM+}$ or $P_F$. Such test will have impact on the statistical analysis in Phase III.
4.3 Conclusion

With the assumptions made in this study, to use the utility function that takes the prevalence of the populations into account seems to be preferred. The utility function that does not take this factor into account exclusively guides us to choose to continue in the biomarker positive subpopulation which seems rather unrealistic.

When using the utility function that takes the prevalence of the populations into account, it is not obvious which population that should continue to Phase III. It depends on the risk the designer of the study is willing to take. In particular when the prevalence of the biomarker positive subpopulation is larger than the prevalence of the biomarker negative, the risks of failing when going in the full population could guide us to continue in Phase III in the biomarker positive subpopulation even though it might bring less utility.

In practice when using this method, the parameters that define the study should of course be tailored for that specific therapeutic area and circumstances.
Bibliography


