Adaptive Designs for Complex Dose Finding Studies

by

Pavel Mozgunov

BA (Hons), Higher School of Economics, 2014

Submitted for the degree of

Doctor of Philosophy

at Lancaster University

April 2018

Adaptive Designs for Complex Dose Finding Studies

by Pavel Mozgunov, BA (Hons).

Submitted for the degree of Doctor of Philosophy at Lancaster University,

April 2018

Abstract

The goal of an early phase clinical trial is to find the regimen (dose, combination, schedule, etc.) satisfying particular toxicity or (and) efficacy characteristics. Designs for trials studying doses of a single cytotoxic drug are based on the fundamental assumption "the more the better", that is, the toxicity and efficacy increase with the dose. This monotonicity assumption can be violated for novel therapies and for more advanced trials studying drug combinations or schedules. It also becomes common to consider a more complex endpoint rather than a binary one as they can carry more information about the drug. Both the violation of the monotonicity assumption and the complex outcomes give rise to important statistical challenges in designing novel clinical trials which require an extensive attention.

In the first part of this thesis, we consider a specific class of combination trials which involve novel therapies and can benefit from the monotonicity assumption. We also propose a general tool evaluating the performance of novel designs in the context of complex clinical trials. Further, we consider a problem of Bayesian inference on restricted parameter spaces. We propose novel loss functions for parameters defined on the positive real line and on the interval demonstrating their performances in standard statistical problems. Based on the obtained results, we propose a novel allocation criterion for model-based designs that results in a more ethical allocation of patients.

In the second part of this thesis, we consider a more general setting of early phase trials in which an investigator has no (or limited) information about the monotonic

orderings of regimens' responses. Using an information-theoretic approach we derive novel regimen selection criteria which allow the avoidance of any parametric or monotonicity assumptions. We propose novel designs based on these criteria and show their consistency. We apply the proposed designs to Phase I, Phase II and Phase I/II clinical trials and compare their performances to currently used model-based methodologies.

Dedicated to my grandparents, Vladimir Chumakov and Iraida Chumakova

Acknowledgements

I wish to thank two people without whom this work would not be possible: my supervisor, Prof Thomas Jaki, for providing a tremendous help and advice during my period of research and for broadening my mind; and my wife, Mrs Kristina Evtifeva, for the continuous support and inspiration. I would like to thank both of them for the endless patience in reading my manuscripts and in correcting my grammar.

I would also like to thank Dr Xavier Paoletti and Prof Mauro Gasparini for many useful discussions and for fruitful collaborations which resulted in several works upon which Chapter 2 and Chapter 3 are based, respectively.

I would like to thank Prof Mark Kelbert for introducing me to Information Theory and to its emerging area of weighted information measures which motivated the novel information gain criterion in Chapter 4.

The research described in this thesis was carried on while I held a Marie Curie Research Fellowship as part of the IDEAS training network which I would like to thank for the great opportunity to work with brightest minds around Europe.

Declaration

I declare that the work in this thesis has been done by myself and has not been submitted elsewhere for the award of any other degree.
The word count for this thesis is 49712 words.

Pavel Mozgunov

List of contributed papers

Section 2.2 has been submitted for publication as Mozgunov, P., Jaki, T. and Paoletti, X. (2018) Randomized dose escalation designs for drug combination cancer trials with immunotherapy.

Section 2.3 has been submitted for publication as Mozgunov, P., Jaki, T. and Paoletti, X. (2018) A Benchmark for Dose Finding Studies with Continuous Outcomes.
The first part of Chapter 3 has been submitted for publication as Mozgunov, P., Jaki, T. and Gasparini, M. (2018) Loss Functions in Restricted Parameter Spaces and Their Bayesian Applications.

Section 3.5 has been submitted for publication as Mozgunov, P. and Jaki, T. (2018) Improving a safety of the Continual Reassessment Method via a modified allocation rule.

The first part of Chapter 4 has been submitted for publication as Mozgunov, P. and Jaki, T. (2018) An information-theoretic approach for selecting arms in clinical trials.

Section 4.5 has been published as Mozgunov, P. and Jaki, T. (2018) An information-theoretic Phase I/II design for molecularly targeted agents that does not require an assumption of monotonicity. *J. R. Stat. Soc. C. doi:10.1111/rssc.12293.*

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Background

Clinical trials are an integral part of drug development. A clinical trial is defined as a prospective study comparing effects of interventions in a human being. In the specific context of drug trials, four different phases of testing are conventionally identified (Friedman et al., 2015):

- Phase I, first in human clinical trials, primarily evaluating the basic safety of a drug
- Phase II, evaluating whether a drug has the desirable biological activity/activities
- Phase III, confirmatory trials, assessing the effectiveness of a drug and comparing it to a standard of care or a placebo
- Phase IV, a post-marketing phase, which identifies and evaluates long-term effects of a drug

Phase III and Phase IV clinical trials are usually conducted using a large number of patients and have been extensively studied in the literature (see e.g. Chow and Liu, 1999, and references therein). Commonly, Phase III clinical trials investigate

one or two dosing regimens that were found to be promising in terms of toxicity and efficacy during Phase I and Phase II clinical trials (Friedman et al., 2015). Therefore, successes in later phases depend, to a great extent, on the knowledge about the novel compound obtained in early phase clinical trials, which are argued to be one of the most important steps in the drug development. A well-designed early phase study is therefore essential to provide valuable information (Chevret, 2006). This underlines the need for effective statistical methods for the accurate selection of optimal dosing regimens. Early phase clinical trials are, however, conducted using a small number of patients or healthy volunteers and face particular ethical restrictions due to the limited knowledge about adverse events associated with a new drug. This raises many statistical challenges which need to be addressed appropriately. Due to its great importance, the methodology for early phase clinical trials is a growing field and is attracting a lot of attention both from statisticians and clinicians (Chevret, 2006). In this thesis, we propose novel methods for early phase clinical trials prompted by emerging practical needs. The proposals in this thesis are mainly motivated by statistical challenges arising in the context of oncology studies. However, the majority of novel methodologies presented are generic and can be applied beyond cancer trials and beyond early phase designs.

Early phase clinical trials are commonly called *dose finding* clinical trials as they aim to identify the optimal (in some sense) dosing regimen of a drug. In recent years, the term "doses" has expanded to not only be understood as doses of a single drug but as different treatment levels which can be, for example, the combinations of several drugs or schedules of administration (Bornkamp, 2017). For the sake of terminology, we will refer to a "dose" as a dose of a single agent, and we use the term "regimen" as a generic name of the objects of study in an early phase clinical trial. Depending on the context, the regimen can be dose, combination, dose-schedule, combination-schedule. Goals of early phase *regimen finding* oncology clinical trials are described below.

Phase I clinical trials are the first studies in humans evaluating the toxicity of regimens. There is limited e affects a human body, and the initial step here is to find "safe" regimens. Given a discrete set of regimens, the goal of Phase I clinical trial is to identify the maximum tolerated regimen (MTR) defined as the regimen having the maximum acceptable level of toxicity γ_t (Friedman et al., 2015) where t stands for toxicity and γ_t might be a percentage, a score or a medical characteristic depending on the trial's endpoint. Phase I clinical trials are usually referred to as *regimen escalation trials*. It is a common practice to summarise the data about the regimen's toxic severity in a single binary outcome, a limiting toxicity, conventionally called dose-limiting toxicity, DLT, (Le Tourneau et al., 2009). In this case, γ_t is a maximum acceptable proportion of patients suffering toxicity. Then, the MTR defines the set of regimens that have an acceptable probability of DLT and can be studied in subsequent phases. We will use terms maximum tolerated dose (MTD) and maximum tolerated combination (MTC) in contexts of single agent and combination trials, respectively.

Once the set of safe regimens is defined, the second step is to ensure that at least one of them has a desirable therapeutic effect (Yin, 2013). To answer this question, the efficacy of regimens is evaluated in Phase II clinical trials. Given a discrete set of regimens, the goal of Phase II clinical trial is to find the regimen corresponding to a pre-specified efficacy level γ_e (where e stands for efficacy), which, again, can be a percentage, a score or a medical characteristic. Phase II clinical trials are usually referred as *regimen ranging trials*. Upon the completion of Phase II trials, the regimens (usually one or two) for a Phase III confirmatory clinical trial are selected (Friedman et al., 2015).

For many years Phase I and Phase II clinical trials were conducted separately. However, there was a recent shift to integrate these phases into a single study. This integration allows for the acceleration of development of the trial and a reduction in costs (Yin, 2013). The integrated trial also provides more observations

for both toxicity and efficacy endpoints and can result in a more reliable recommendation for Phase III study, particularly, for novel agents (Wages and Conaway, 2014). The goal of a Phase I/II clinical trial evaluating both toxicity and efficacy simultaneously is to find the optimal biologic regimen (OBR) corresponding to pre-specified levels of toxicity and efficacy.

Despite different formulations of research problems, the goals of all these trials are similar from the statistical perspective. The common goal of an early phase study is to find the regimen having characteristics "as close as possible" to the target $\boldsymbol{\gamma}$ which can be either a scalar (e.g. γ_t or γ_e) or a vector (e.g. $(\gamma_t, \gamma_e)^{\mathrm{T}}$). We would use the term *target regimen* (TR) to emphasise that a particular design can be applied to different types of trials.

1.2 Motivation

The primary motivation behind methods proposed in this thesis is the growing complexity of early phase clinical trials. It is becoming more common to consider more complex dosing regimens rather than doses of a single agent only. For instance, therapies using a combination of drugs have become the mainstream approach to diseases such as cancer (Khalil et al., 2016) and tuberculosis (Kerantzas and Jacobs, 2017). It was found that administering several agents simultaneously can noticeably improve their safety profiles and therapeutic effects. However, potential synergistic (or antagonistic) effects give rise to additional challenges in designing early phase combination trials compared to single agent ones. A lot of *dose* finding trials are based on the fundamental assumption "the more the better", that is, the toxicity of the drug increases with the dose. This is also known as the *monotonicity assumption* (Clertant and O'Quigley, 2017). This means that clinicians can naturally order doses according to the monotonically increasing toxicity. This, however, might not hold for combinations. Consider two dual agents combinations: T_1 and T_2. Assume that, relative to T_2, the combination T_1 has a

higher dose of the first compound, but a smaller dose of the second compound. In the vast majority of trials, clinicians cannot define which of these combinations is more toxic prior to the trial (Wages et al., 2011). The same argument holds when considering efficacy levels of combinations.

The schedule of administration can add an additional complexity as well (Kodama et al., 2009). As an illustration, consider a six days course of treatment and two dose-schedules: (i) 10 mg a day every day and (ii) 20 mg a day every two days. The total amount of drug received by a patient is the same by the end of six days, but it is unclear which of these dose-schedules is more toxic. A smaller dose is expected to reduce toxicity, but a more regular administration is expected to increase it. Again, clinicians cannot define which of these effects is greater in many clinical trials.

The problem of unknown ordering of toxicities and efficacies can also appear in the single agent context. While the paradigm "the more the better" holds for cytotoxic agents, it can be violated for molecularly targeted agents (MTA) which include hormone therapies, signal transduction inhibitors, gene expression modulators, apoptosis inducers, angiogenesis inhibitors, immunotherapies, and toxin delivery molecules, among others. For MTAs either dose-efficacy or dose-toxicity relationships can have a plateau (Morgan et al., 2003; Postel-Vinay et al., 2011; Robert et al., 2014; Paoletti et al., 2014) or a dose-efficacy relationship can exhibit an umbrella shape (Conolly and Lutz, 2004; Lagarde et al., 2015).

The problems in all of these advanced clinical trials are similar in their nature as they suffer from the unknown ordering of regimens and cannot fully benefit from the monotonicity assumption. Additionally, more complex outcomes rather than binary ones are becoming of growing interest in clinical trials as they can carry more information about the profile of a drug (see e.g. Thall and Cook, 2004; Lee et al., 2010, 2017). To conduct more complex clinical trials, specific methods for

the adequate identification of the TR are essential. In this work, we propose novel designs that are able to address either some or all of the stated challenges and extend several well-established early phase designs in order to guarantee a more accurate and ethical regimen selection.

1.3 Outline of Thesis

In Chapter 2 we consider several types of combination trials which can benefit from the monotonicity assumption. We show, however, that the straightforward application of standard Phase I dose escalation model-based designs may fail to address the goals specific to the combination context and propose an alternative. Further, we generalise the non-parametric optimal benchmark O'Quigley et al. (2002), a tool to evaluate the performance of regimen finding designs, to the setting of complex trials with multiple endpoints having discrete or continuous distributions.

In Chapter 3 we consider a problem of estimation in restricted parameter spaces that arise in many areas of regimen finding. We stress that standard criteria for the choice of estimators and for the allocation of patients might not be a proper choice in the context of early phase clinical trials. We demonstrate the performance of the proposed estimator in some classic statistical problems and construct a novel allocation criterion for a Phase I model-based design which improves its ethics.

In Chapter 4 the general setting of regimen finding clinical trials with the monotonicity assumption violation and multinomial outcomes is considered. We propose a class of novel designs which relax the monotonicity assumption and can be applied in complex Phase I, Phase II and Phase I/II trials. We demonstrate their applications motivated by actual clinical trials.

A conclusion, limitations and directions for further research are provided in Chapter 5.

Chapter 2

Model-based Dose Escalation Designs and Optimal Benchmarks

2.1 Background

Historically, Phase I oncology clinical trials investigated increasing doses of a single cytotoxic compound, e.g. chemotherapy or radiotherapy. Based on their mechanism of action, it is known that a higher dose has higher toxicity and the monotonicity assumption is satisfied. The algorithm-based "3+3" dose escalation design (Carter, 1973; Storer, 1989) was one of the first methods used to conduct dose escalation studies of cytotoxic cancer treatments in humans. Despite an enormous number of papers showing that the 3+3 design is inferior to model based alternatives (see e.g. O'Quigley et al., 1990; Ratain et al., 1993; Kang and Ahn, 2001, 2002), it is still the most common design used in a clinical practice (Rogatko et al., 2007; Le Tourneau et al., 2009; Chiuzan et al., 2017) due to its

simple escalation/de-escalation rules which do not require any support from a statistician. The 3+3 design uses the information from the most recently enrolled patients only and ignores earlier data obtained in the trial. This results in a systematic underestimation of the MTD, an unethical allocation of patients (patients are assigned to low doses far from the target toxicity) and a high risk of erroneous conclusions (Reiner et al., 1999). Instead, more statistical approaches that use all available information have been shown to lead to noticeable improvements in the probability of correct MTD selection (Wheeler, 2017). Below we recall the core of model-based dose escalation designs that formed the basis of many of regimen escalation methodologies.

2.1.1 Continual Reassessment Method Type Designs

Consider a single agent Phase I clinical trial with binary toxicity outcome, dose-limiting toxicity (DLT) or no DLT, N patients and m dose levels $d_1, \ldots, d_m$. Let Y_{ij} be a binary (Bernoulli) random variable taking value $y_{ij} = 0$ if patient i has experienced no DLT given dose d_j and $y_{ij} = 1$ otherwise. This random variable is characterised by toxicity probability p_j such that $p_j = \mathbb{P}(Y_{ij} = 1)$, $i = 1, \ldots, N$. It is assumed that the toxicity probability is an increasing function of dose, $p_1 < \ldots < p_m$. The goal of the trial is to find the MTD, the dose corresponding to the target level of toxicity, γ_t.

Original Continual Reassessment Method One of the first model-based dose escalation designs that borrows information across doses was the Bayesian Continual Reassessment Method (CRM) by O'Quigley et al. (1990). Relying on the monotonicity assumption, the idea of the method is that the identification of the MTD can be efficiently achieved by means of a working model without characterizing the true dose-toxicity relationship.

Assume that toxicity probability p_j has the functional form

$$p_j = \psi(d_j, \theta) \tag{2.1.1}$$

where $\theta \in \mathbb{R}^d$ is a d-dimensional vector of parameters and d_j is a scalar unit-less dose level (also refereed as standardised levels). Denote the prior distribution of θ by $f_0(\cdot)$. Assume that n patients have been already assigned to doses $d(1), \ldots, d(n)$ and binary responses $y_1, \ldots, y_n$ were observed, respectively. The CRM updates the posterior distribution of θ using Bayes's Theorem

$$f_n(\theta) = \frac{f_{n-1}(\theta)\phi(d(n), y_n, \theta)}{\int_{\mathbb{R}^d} f_{n-1}(u)\phi(d(n), y_n, u)\mathrm{d}u} = \frac{f_0(\theta)\prod_{i=1}^n \phi(d(i), y_i, \theta)}{\int_{\mathbb{R}^d} f_0(u)\prod_{i=1}^n \phi(d(i), y_i, u)\mathrm{d}u} \tag{2.1.2}$$

where

$$\phi(d(i), y_i, \theta) = \psi(d(i), \theta)^{y_i}(1 - \psi(d(i), \theta))^{1-y_i}.$$

The posterior mean of the toxicity probability given dose d_j after n patients is equal to

$$\hat{p}_j^{(n)} = \mathbb{E}(\psi(d_j, \theta)|y_1, \ldots, y_n) = \int_{\mathbb{R}^d} \psi(d_j, u) f_n(u)\mathrm{d}u. \tag{2.1.3}$$

Patients in the dose escalation trial are assigned cohort-by-cohort, where a cohort is a small group of typically 1 to 4 patients. Then, the dose d_j that minimises

$$|\hat{p}_j^{(n)} - \gamma_t|, \tag{2.1.4}$$

or, equivalently, the squared distance $\left(\hat{p}_j^{(n)} - \gamma_t\right)^2$, among all $d_1, \ldots, d_m$ is recommended for the next group of patients. The procedure is repeated until the maximum number of patients, N, has been treated. Note that the original CRM implementation plugs the mean value of θ in the model $\psi(d_j, \hat{\theta})$ instead of using the posterior mean (2.1.3). While no noticeable differences in these approaches are found for a 1-parameter model (Iasonos et al., 2016), a difference can, however, be expected using models with more parameters. Therefore, we will use the

estimator (2.1.3).

There have been many developments building on the CRM since its original proposal. O'Quigley and Shen (1996) have considered a likelihood approach to CRM and Shen and O'Quigley (1996) have formulated the conditions of its consistency (the probability that the design would select the true MTD tends to 1 as $N \to \infty$). Korn et al. (1994); Zohar and Chevret (2001); O'Quigley (2002) have considered early stopping rules to accommodate practical challenges of implementations in actual clinical trials. Cheung and Chappell (2002); Lee and Cheung (2009); O'Quigley and Zohar (2010); Cheung (2011) have investigated choices of standardised levels $d_1, \ldots, d_m$ and parameters of the prior distribution f_0. Cheung and Chappell (2000) have proposed a modification called the time-to-event CRM which accommodates late-onset toxicities and leads to conducting a trial in a timely manner. Cheung (2005) has shown that the CRM does not lead to counter-intuitive escalation/de-escalation decisions that would not be accepted by clinicians. The CRM was subsequently extended to contexts of more complex trials, for example, to Phase I combination trials (Wages et al., 2011) and to Phase I/II dose finding trials (Braun, 2002).

While it is generally agreed that the Bayesian CRM design leads to an accurate MTD selection, the choice of $\psi(d_j, \theta)$ is debated (Iasonos et al., 2016). The common parametric forms of the model function are given below.

Choice of the Working Model in Model-based Designs The common choice for the 1-parameter working dose-toxicity model is

$$\psi(d_j, \theta) = d_j^{\exp(\theta)} \tag{2.1.5}$$

or $\psi(d_j, \theta) = d_j^{\theta}$ where θ is a scalar parameter (O'Quigley and Shen, 1996). The use of 1-parameter models is motivated by challenging estimation given small sample sizes.

On the other hand, it is argued that the 1-parameter model (2.1.5) does not reflect the anticipated dose-toxicity relation (Neuenschwander et al., 2008). Instead, the logistic function was suggested as an appropriate alternative with a clinically relevant interpretation. Whitehead and Williamson (1998) proposed to use the 2-parameter logistic model

$$\psi(d_j, \theta_1, \theta_2) = \frac{\exp(\log(\theta_1) + \theta_2 d_j)}{1 + \exp(\log(\theta_1) + \theta_2 d_j)} \tag{2.1.6}$$

where θ_1, θ_2 are scalar parameters. There are variants of the logistic model: with parameter θ_1 being known and fixed and with both θ_1 and θ_2 being unknown. The same parametric model was also used by Babb et al. (1998). This model is seen to be appropriate in the context of actual clinical trials and is extensively employed in practice (Neuenschwander et al., 2008) despite potential inconsistency problems raised by Iasonos et al. (2016). Paoletti and Kramar (2009) have compared several model choices of $\psi(\cdot)$ in a comprehensive simulation study. Routinely, models with more than two parameters are not routinely considered in the setting of dose escalation trials.

Regardless of the choice of working model, the dose escalation designs above are based on the "strict" assumption of monotonicity (the toxicity strictly increases with the dose). In Section 2.2 we consider particular types of combination trials for which the monotonicity assumption holds in a "non-strict" sense - the toxicity does not decrease with the dose. While such trials can potentially benefit from application of single-agent dose escalation designs, we show that they do not address all goals of combination trials and might lead to misleading conclusions. Consequently, we propose modifications to model-based designs to accommodate the combination specific setting.

2.1.2 An Optimal Benchmark for Studies with a Binary Endpoint

A variety of dose finding methods for Phase I clinical trials aiming to find the MTD were developed in the literature since the proposal of the CRM. A conventional way to assess the performance of a design is to conduct an extensive simulation study. One of the key characteristics of any regimen finding method is its accuracy which is usually computed as the proportion of times the true target regimen (e.g. the MTD) is selected (also referred as the proportion of correct selections - PCS). The majority of novel proposals are studied in scenarios chosen by investigators themselves. This, clearly, adds subjectivity to the assessment of a method's operating characteristics as one can always find scenarios in which the TR selection is easier than in others.

To solve this problem, O'Quigley et al. (2002) proposed the *non-parametric optimal benchmark* that provides an upper limit of accuracy (in terms of the PCS) for dose finding methods based on a binary toxicity endpoint. The benchmark uses the *complete information* concept which assumes that outcomes of each patient can be observed at all dose levels (in contrast to an actual trial in which a patient can be assigned to one dose only). The benchmark shows how "difficult" the TR identification is in the chosen scenario and provides the objective context for the performance evaluation of the design under investigation. Since its proposal, the benchmark has proven its great usefulness to assess newly proposed designs comprehensively (see e.g. Paoletti and Kramar, 2009; Yin and Yuan, 2009). Additionally, based on the benchmark, Cheung (2013) derived sample size formulae for the CRM.

In the setting of a Phase I single agent trial with a binary endpoint aiming to identify the MTD, the benchmark was proposed as follows. For a given patient the complete information consists of the vector of outcomes at all dose levels assuming

that $p_1, \ldots, p_m$ are known. In other words, for a given patient one knows the maximum toxicity probability that this patient can tolerate. Formally, the information about the response of patient i at all dose levels is summarised in a single value $u_i \in (0, 1)$, profile of the patient, which is drawn from a uniform distribution, $\mathcal{U}(0, 1)$. The value u_i was also defined as a *tolerance* of patient i by Finney (1952) who, historically, was one of the originators of the idea of having an increasing series of probabilities generated by shareholding a continuous distribution which formed a basis for the complete information concept. For instance, $u_i = 0.3$ means that patient i can tolerate doses d_j with $p_j \leq 0.3$, but would experience a DLT if given dose $d_{j\prime}$ with $p_{j\prime} > 0.3$. It follows that u_i is transformed to $y_{ij} = 0$ for doses with $p_j < 0.3$ and to $y_{ij} = 1$ otherwise. The procedure is repeated for N patients which results in the vector of responses corresponding to each dose level $\boldsymbol{y}_j = (y_{1j}, \ldots, y_{Nj})$, $j = 1, \ldots, m$. Let $T(\boldsymbol{y}_j, \gamma_t)$ be a summary statistic for the dose level d_j upon which the decision about the MTD selection is based. Conventionally, $T(\boldsymbol{y}_j, \gamma_t)$ is chosen such that its minimum (or maximum) value corresponds to the estimated MTD. Therefore, d_j for which $T(\boldsymbol{y}_j, \gamma_t)$ is minimised (maximised) is selected as the MTD in a single trial. The procedure is repeated for S simulated trials and then proportions of each dose selections are computed.

In a context of Phase I clinical trial with a binary endpoint

$$T(\boldsymbol{y}_j, \gamma_t) = \left| \frac{\sum_{i=1}^{N} y_{ij}}{N} - \gamma_t \right| \tag{2.1.7}$$

is a conventional choice for the MTD selection criterion. Wages and Varhegyi (2017) proposed a Web application for the benchmark evaluation using this criterion.

Motivated by more complex studies, for instance, Phase I/II clinical trials evaluating binary toxicity and binary efficacy endpoints simultaneously (Thall and Russell, 1998) or Phase I trials with multiple grades of toxicities (Lee et al., 2011),

Cheung (2014) generalized the benchmark to both of these cases. This has broadened the benchmark application significantly. However, there are a growing number of Phase I and Phase I/II clinical trials involving continuous endpoints, but no corresponding benchmark exists. For example, Bekele and Thall (2004); Yuan et al. (2007); Ivanova and Kim (2009); Bekele et al. (2010); Ezzalfani et al. (2013); Wang and Ivanova (2015) considered a continuous toxicity endpoint while, for example, Bekele and Shen (2005); Hirakawa (2012); Yeung et al. (2015, 2017) studied Phase I/II trials with binary toxicities and continuous efficacies.

In Section 2.3, we propose a new benchmark which can be applied to dose finding studies with continuous outcomes and shares the same concept of the complete information as the original approach. We demonstrate the application of the benchmark in contexts of recently proposed Phase I and Phase I/II dose finding designs.

2.2 Designs for Combination Trials Involving an Immunotherapy

2.2.1 Problem Formulation

The monotonicity assumption upon which the CRM type designs are based can be violated for MTAs, e.g. an immunotherapy. For example, the majority of Phase I trials of immune-checkpoint protein blockers, such as anti-programmed-death-receptor-1 (PD1), have never actually reached the MTD (see e.g. Robert et al., 2014; Patnaik et al., 2015; Yamamoto et al., 2017). At the same time, immunotherapies have been shown to have a strong anti-tumour activity and a better safety profile than cytotoxic cancer therapies (Disis, 2010; Kyi and Postow, 2014). It was also found that the therapeutic effect of an immunotherapy can be noticeably improved by combining it with a backbone chemotherapy or other targeted agents (Sharma and Allison, 2015). Therefore, there are many Phase I

combination trials investigating either (i) the added value of an immune checkpoint blocker to a backbone therapy or (ii) the added value of a new agent to an immune checkpoint blocker (Pardoll, 2012). In both cases, one agent, called a standard of care, is administered at a fixed dose and another agent is dose-escalated. While such setting allows for the straightforward adaptation of single agent designs, it also hides potential difficulties.

The conventional goal of Phase I drug combination trial is to find the maximum tolerated combination (MTC). However, many studies of immunotherapies have never actually reached the maximum tolerated level as stated above. An immunotherapy can also have a complex mechanism of its interactions with other compounds (Sharma and Allison, 2015). Consequently, when considering combination trials involving an immunotherapy, the information beyond the conventional (and only) objective of a Phase I study can be also important for a more accurate choice of the combination(s) to study in subsequent phases.

We focus on the setting that covers two important types of Phase I clinical trials with corresponding research questions added to the MTC selection:

1. The standard of care is a toxic agent (e.g. chemotherapy) and is given at the full (single agent) dose and an immune-checkpoint blocker is dose-escalated.

   A clinician would tolerate only a slight increase τ in the toxicity of combination compared to the toxicity of the standard of care (Paller et al., 2014) and the question whether "the increase in toxicities is acceptable?" should be tested.

2. The standard of care corresponds to a low toxicity level (e.g. an immune-checkpoint blocker) and is given at the full (single agent) dose and either a toxic agent or an MTA is dose-escalated.

   (a) An interest is to determine a plateau region between the estimated MTC

and the standard of care which would help to define the therapeutic index (a ratio of the dose that produces highest acceptable toxicity to the dose needed to produce the desired therapeutic response) more accurately. A plateau is defined by not exceeding the difference in associated toxicities by more than τ.

(b) A clinician has an expectation of an additional toxicity τ over the standard of care under the assumption of compounds independence. An interest lies in checking for an interaction effect defined as an additional toxicity over τ.

The objectives of these trials are (i) to identify the MTC, (ii) to quantify the expected difference between single and combination treatments and (iii) to determine the shape of the dose-toxicity relationship. These clinical trials share a common interest in the comparison of toxicity levels associated with the estimated MTC and the standard of care alone but they differ in their motivation. To unify notation, we study the extra toxicity (ET) beyond the expected difference τ. We show that standard single agent dose-escalation methods currently used for such trials may fail to address secondary questions of the trial.

To achieve all objectives, we suggest to adapt two modifications to the Bayesian model-based design. We propose to include the standard of care given alone as a control arm and, to randomise each patient to the control arm or to the combination selected by the Bayesian model-based design. We demonstrate that such randomisation procedure leads to a reliable statistical evaluation of the ET (added over the standard of care) and of the general interaction effects between compounds. We also show that the Emax model provides a well-established tool to detect and to evaluate different patterns in the dose-toxicity relationship and its parameters match the information needed to address stated objectives.

It is important to mention that the question of using the control group in Phase I

trials is widely debated in the literature (e.g. Saad et al., 2017). The main argument against is the monotonicity assumption that makes comparing toxicity levels unnecessary and results in considering the control group as not contributing to the goal of the trial (the MTC selection). However, this assumption has been found to be inappropriate for many immunotherapies. Importantly, patients allocated to the control arm receives the standard of care which makes the randomisation an ethically viable option. It will be also demonstrated that modelling the data observed for the control group simultaneously with the escalated arm contributes to toxicity estimates for all dose levels and can serve the MTC selection objective as well.

2.2.2 Motivating Trials

The proposed design is motivated by two recent combination trials that could potentially benefit from its implementation.

Gemcitabine is a standard chemotherapy for an advanced pancreatic cancer (Aglietta et al., 2014) which has a narrow therapeutic index (Crane et al., 2002). Tremelimumab is a fully humanized monoclonal antibody against CTLA-4 that may allow effective immune responses against tumour cells. In several clinical studies, anti-CTLA4 agents have been shown to induce durable tumour responses through modulation of the immune system in patients with metastatic melanoma (Buchbinder and Desai, 2016). The hypothesis was that the combination of these two agents "might provide a synergistic anti-tumour activity without increasing toxicity" (Aglietta et al., 2014). Gemcitabine 1000 mg/m^2 was administered in all patients while escalating doses of tremelimumab (6, 10 and 15 mg/kg) were sequentially tested to identify the MTC. This trial was conducted using the 3+3 design, leading to a high risk of erroneous conclusions due to the assumption of the increasing dose-toxicity relationship. Furthermore, as the expected probability of the DLT for Gemcitabine is more than 20%, any (if the dose-response curve is flat) or no combination (if the dose-response curve is steep) can be the MTC.

Finally, the highest dose was recommended for further investigations, but the question whether the toxicity is increased over a clinically meaningful difference τ was never formally tested.

Sorafenib is a treatment for advanced cellular cell carcinoma. However, its efficacy remains limited as the time to progression is around six months. Despite this agent is being prescribed at the MTD (800mg/kg), its therapeutic index makes it possible to reduce the dose in case of adverse reactions (Wilhelm et al., 2006). SPLASH is a dose-escalation study of Avelumab in combination with sorafenib in patients with advanced cellular cell carcinoma that is about to be initiated at the Hospital Gustave Roussy. While Avelumab will be given at a fixed dose, Sorafenib will be escalated from 200 mg/kg up to 800mg/kg. The MTD is defined as the highest dose having the probability of DLT during cycle 1 closest to 25%. The expected DLT probability for Avelumab alone is 8%.

2.2.3 Randomisation and Emax Model

Framework Consider a clinical trial in which combinations of agents A and B are studied. The drug A is a standard of care given at the fixed dose, a, and B is dose-escalated. We specify increasing doses of B: $0 = b_0 < b_1 < b_2 < \ldots < b_m$. Then, $\tilde{d}_0 = \{a, b_0\}, \tilde{d}_1 = \{a, b_1\}, \ldots, \tilde{d}_m = \{a, b_m\}$ are $m+1$ combinations available in the trial, where $\tilde{d}_0$ corresponds to the agent A given as a single agent and is subsequently referred to as the *control arm*, and the combinations $\tilde{d}_1, \ldots, \tilde{d}_m$ are referred to as the *investigational arms*. A clinician observes binary outcomes, DLT and no DLT. Let p_j be the probability for a patient to experience a DLT given the combination $\tilde{d}_j$. It is assumed that toxicity is a non-decreasing function of combination, $p_0 \leq p_1 \leq \ldots \leq p_m$ and reliable prior information for the toxicity probability of the control arm, p_0, is available.

Let γ_t be the maximum acceptable toxicity probability. The primary goal of

the dose-escalation clinical trial is to find the MTC $\tilde{d}_{j^\star}$ such that

$$j^\star = \underset{j=0,\ldots,m}{\operatorname{argmin}} |p_j - \gamma_t|$$

using estimated toxicity probabilities $\hat{p}_0, \ldots, \hat{p}_m$. The secondary goal, specific to combination trials, is to test if the difference of toxicity probabilities associated with the estimated MTC and the standard of care is as expected. The ET is defined in terms of the expected difference, τ. One would conclude the ET when the probability that the difference in the two estimated risks of toxicity exceeds τ is larger than some credible level α. Formally, if the ET is present one would like the probability

$$\mathcal{P} \equiv \mathbb{P}\left(\mathbb{P}\left(\hat{p}_{j^\star} - \hat{p}_0 \geq \tau\right) > \alpha\right) \tag{2.2.1}$$

to be equal to 1. In contrast, if there is no ET, the probability (2.2.1) is desired to be equal to 0.

Note that as only one drug is varied, the problem can be considered as a unidimensional MTC search and the CRM (Section 2.1.1) can be applied. It, however, is designed for the MTC selection only, therefore, we introduce two design features into it.

Randomisation between control and investigational arms Patients in dose-escalation trials are assigned cohort-by-cohort. According to the original allocation rule, the CRM design tends to assign the majority of patients in the neighbourhood of the MTC. This leads to a sparse allocation of patients on other combinations and on the control arm. This will make it difficult to test for differences in toxicity risk associated with the estimated MTC and the control arm. We introduce the following randomisation procedure between the investigational (combinations) and control (standard of care) arms.

Denote the cohort size by $c = c_1 + c_2$ and let c_1 be the number of patients in cohort

assigned to the estimated MTC by CRM according to the criterion (2.1.4), and c_2 be the number of patients in cohort assigned to the control arm, $\tilde{d}_0$. This results in at least $N_2 = \frac{c_2}{c_1+c_2}N$ patients on the control arm and at most $N_1 = \frac{c_1}{c_1+c_2}N$ on the investigational arm by the end of the trial. For instance, taking $c_1 = 3$ and $c_2 = 1$ (denoted by 3:1), one will end up with at least 25% of the total sample size being assigned to the control. Note that the model-based design is allowed to select the control arm as the estimated MTC if the associated toxicity is the closest to the target γ_t. This facilitates avoiding exposing patients to high toxicity if the first combination has an unacceptable DLT probability. Therefore, the total number of patients on the control arm can be more than $\frac{c_2}{c_1+c_2}N$.

Importantly, in the proposed design the values of c_1 and c_2 are fixed before the trial. While the choice of these values is investigated in Section 2.2.6, their choice can be also guided by a prior information that a clinician has about the standard of care. For example, if a clinician is certain about the toxicity risk associated with the standard of care, one can allocate more patients to the investigational arms and use lower value of c_2.

The modified allocation rule (the randomisation of patients between control and investigational arm) raises an important question of the suitable choice of the parametric model $\psi(d_j, \theta)$ which is discussed below.

4-parameter Emax model In the context of the single agent trial, the 1-parameter power model is argued to be an appropriate choice (Iasonos et al., 2016). There are, however, at least two reasons why 1-parameter models are not suitable for the type of trials considered here.

Firstly, the parametric form (2.1.5) implies the "strict" monotonicity assumption - the toxicity increases as dose increases. It does not allow modelling a plateau in the combination-toxicity relations. Therefore, one can expect that such a model would tend to finding the ET regardless the scenario which is undesirable in the

considered combination trials.

Secondly, one of the main arguments behind using a 1-parameter model (and, consequently, the main critique of models with more parameters) is that the CRM design tends to collect observations in the neighbourhood of the MTC only. This means that the approximation of the dose-toxicity relation is of interest in the neighbourhood of one point only. Clearly, 1-parameter models are rich enough to achieve it. Note, however, that given the randomisation step proposed above, the majority of patients will be now assigned in the neighbourhood of two points: the control arm and the estimated MTC. Therefore, more flexible models are essential to consider.

Motivated by the ability to model a plateau in a combination-toxicity relation, we propose to use the 4-parameter Emax model

$$\psi(d_j, E_0, E_{max}, \lambda, ED_{50}) = E_0 + \frac{d_j^{\lambda} E_{max}}{d_j^{\lambda} + ED_{50}^{\lambda}},$$

where E_0 is the toxicity probability associated with the control arm, $E_{max} + E_0$ is the maximum toxicity probability attributable to the combination, ED_{50} is the combination which produces $E_0 + \frac{E_{max}}{2}$ toxicity and λ is the slope factor. Following the assumption of the non-decreasing toxicity probability, we specify $\lambda \geq 0$. To adjust the single agent model to the combination setting, the unit-less variable d_j, corresponding to the same toxicity probability as the combination $\tilde{d}_j$, is used. To construct d_j one needs to represent them in terms of prior estimates of toxicity probabilities $\hat{p}_j^{(0)}$ associated with combinations $\tilde{d}_j$ $j = 0, \ldots, m$

$$d_j = \hat{ED}_{50}^{(0)} \times \left(\frac{\hat{p}_j^{(0)} - \hat{E}_0^{(0)}}{\hat{E}_{max}^{(0)} + \hat{E}_0^{(0)} - \hat{p}_j^{(0)}} \right)^{\frac{1}{\hat{\lambda}^{(0)}}}$$

where $\hat{\lambda}^{(0)}$, $\hat{E}_0^{(0)}$, $\hat{E}_{max}^{(0)}$ and $\hat{ED}_{50}^{(0)}$ are point prior estimates of model parameters. The prior point estimate of the toxicity probability on the control arm is $\hat{p}_0^{(0)}$.

Therefore, by definition, $\hat{p}_0(0) \equiv \hat{E}_0^{(0)}$ that leads to $d_0 = 0$. Modelling E_0 directly guarantees that the sequential update of other parameters does not contribute to the toxicity probability estimation on the control arm. In this case, the model takes the trivial form

$$\psi(d_0, \cdot) = E_0.$$

Intuitively, it reflects that the toxicity probability of the standard therapy does not depend on the mechanism of its interaction with compound B.

The parameters of the Emax model allows to model the toxicity on the control arm and the plateau. Modelling a plateauing dose-toxicity relationship (e.g. immunotherapies) is not possible with other working models, e.g. two-parameter logistic form. We investigate the choice of the parametric model in the numerical study below.

2.2.4 Simulation Setting

We explore the performance of the Bayesian model-based dose-escalation method incorporating randomisation to the control arm into the Emax model by simulations in different scenarios. Motivated by a recent combination trials review by Riviere et al. (2015) we consider a setting with $N = 48$ patients and $m = 7$ combinations. We set the target toxicity probability $\gamma_t = 0.25$, the expected difference $\tau = 0.05$ and the confidence level $\alpha = 0.90$.

Three main characteristics, (i) the proportion of correct selections (PCS), (ii) the proportion of times the ET is concluded and (iii) a goodness-of-fit measure, are considered. A goodness-of-fit measure is used to capture the overall shape of the dose-toxicity relationship. We use the scenario-normalized mean squared error (NMSE) defined as

$$NMSE = \frac{1}{S} \sum_{s=1}^{S} \sqrt{\frac{\sum_{j=0}^{m} (p_j - \hat{p}_j^{(s)})^2}{\sum_{j=0}^{m} (p_j - \hat{p}_j^{opt})^2}} \tag{2.2.2}$$

where S is the number of replications, $\hat{p}_j^{(s)}$ is the toxicity probability estimate for combination $\tilde{d}_j$ obtained by the design in s^{th} simulation and $\hat{p}_j^{opt}$ is the toxicity probability estimate for combination $\tilde{d}_j$ obtained by the non-parametric optimal benchmark approach described in Section 2.1.2. The normalisation facilitates comparisons betwen different scenarios as the optimal benchmark incorporates their specificities. The value of the NMSE being equal to 1 corresponds to the curve estimated as precise as by the optimal benchmark.

Prior Specification The standardised levels, d_j, are constructed using the skeleton $\hat{\mathbf{p}}^{(0)}$.

$$\hat{\mathbf{p}}^{(0)} = [0.08, 0.25, 0.35, 0.45, 0.55, 0.65, 0.70, 0.75]^{\mathrm{T}}.$$

The first value, 0.08, corresponds to the mean prior toxicity probability for the control arm. The prior MTC being the first combination ensures that the trial is started at the lowest combination. Other values are obtained by taking an adequate spacing between prior values. It has been shown by O'Quigley and Zohar (2010) that the CRM design is robust and efficient in this case. In contrast to the skeleton that is the same regardless which parameter model is used, the prior distributions of the model parameters needs to be calibrated such that all competing designs carry the same amount of the prior information.

To ensure that the proposed approach and competitive designs are evaluated under comparable set-ups the credible intervals for the prior toxicity probabilities associated with the control arm and the prior MTC are used. Taking into account that one usually has a reliable information about the standard therapy, but limited information about combinations, we specify prior distributions to satisfy the following conditions:

1. The control arm $\tilde{d}_0$: the expected toxicity probability is $\hat{p}_0^{(0)} = 0.08$ and the upper bound of the 95% credibility interval is 0.25.

2. The prior MTC $\tilde{d}_1$: the expected toxicity probability is $\hat{p}_1^{(0)} = 0.25$ and the upper bound of the 95% credibility interval is 0.80.

Following these conditions the prior distributions of the Emax model parameters are specify as

$$E_0 \sim \mathbb{B}(0.8, 10 - 0.8),\ E_{max}|E_0 \sim \mathbb{U}[0, 1 - E_0],\ ED_{50} \sim \Gamma(0.4, 0.4),\ \lambda \sim \Gamma(1, 1) \tag{2.2.3}$$

where $\mathbb{B}(u, v)$ denotes the Beta distribution with parameters u, v and $\Gamma(u, v)$ denotes the Gamma distribution with the mean $\frac{u}{v}$ and the variance $\frac{u}{v^2}$. As the choice of prior distributions may have quite a significant impact on the estimates of parameters, we first rely on the elicitation given in Equation (2.2.3), but also investigate the robustness to different prior distributions in Section 2.2.6. We refer to the design using the Emax model with proposed randomisation as "EmaxR".

Cohort Size for the Control Arm and Combination Skipping The cohort size is fixed to be $c_1 = 3$ for the investigational arm and $c_2 = 1$ for the control arm. Therefore, at least 12 patients (25% of the total sample size) will be allocated to the standard therapy and at most 36 patients are allocated to the investigational arm. We allow combinations to be skipped. Other randomisation ratios, 2 : 1 and 4 : 1, and an impact of no skipping constraint are studied in Section 2.2.6.

Scenarios Nine qualitatively different scenarios, depicted in Figure 2.1, are used to study the performance of the proposed approach. Scenarios 1-5 are consistent with the prior information about the control arm. Scenario 1 corresponds to a commonly used logistic dose-toxicity shape while scenario 2 considers a flat dose-toxicity curve with the MTC being $\tilde{d}_7$ and no ET beyond τ. In scenario 2, the model monotonicity assumption implies that the toxicity probability on the control has the largest difference with combination $\tilde{d}_7$ and the model is expected to have a tendency to claim the ET. Then, scenario 2 is used to test if the ran-

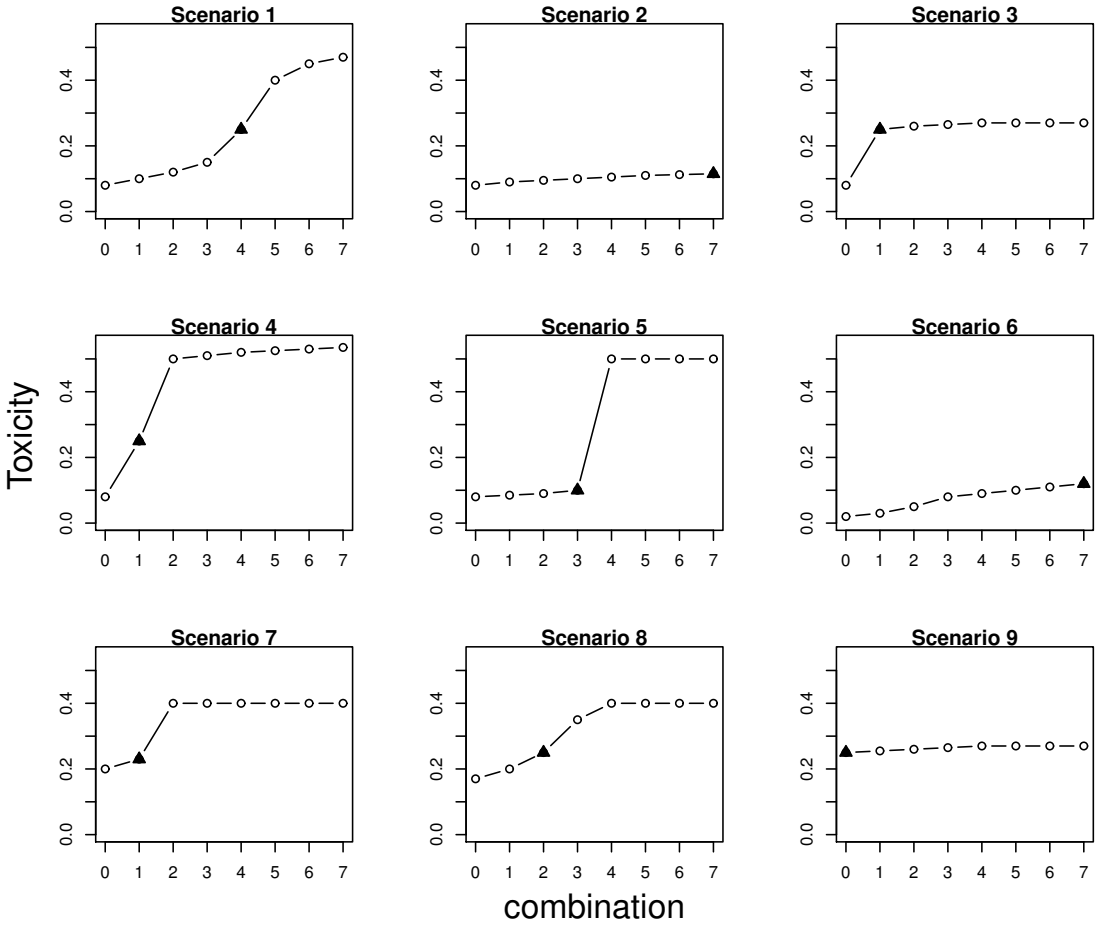

Figure 2.1: Considered combination-toxicity scenarios. The MTC is marked by a triangle.

domisation to the control arm together with the Emax model can prevent false conclusions regarding the ET. Scenarios 3-4 correspond to steep dose-toxicity relationships that have a plateau around and outside the target toxicity probability, respectively. These scenarios are used to investigate under what conditions it is easier to find the ET. Scenario 5 also corresponds to a sharp increase in toxicities, but for medium combinations having unacceptably high toxicity such that there is no ET between the MTC and the control arm.

Scenario 6 reflects the case of a misspecified prior toxicity probability for the control arm - the true toxicity probability for $\tilde{d}_0$ is below the prior value. The prior information dictates that there is no ET between the MTC and the control ($0.11 - 0.08 < \tau = 0.05$), although it is present. In scenarios 7-9 the toxicity

probability for the control is underestimated by the prior information and the control arm has a toxicity probability close to the maximum acceptable one. Both cases with (scenario 8) and without (scenarios 7, 9) ET are considered.

Competing Models The performance of the proposed approach is compared to designs that are currently used for the considered types of trials: the 1-parameter power model

$$p_j(d_j, \theta) = d_j^{\theta}$$

denoted by P1 and the 2-parameter logistic model given in Equation (2.1.6) without randomisation denoted by L2. We also explore the possibility of randomisation to the control arm using the 2-parameter logistic model denoted by L2R. As there are only at most 36 patients on the investigational arm in the randomised setting against 48 in the non-randomised one, we also consider the 2-parameter logistic model without randomization and $N = 36$. This would allow spotting the influence of the randomisation on the operating characteristics. Prior distributions for the parameters of P1 and L2, L2R models $\theta \sim \Gamma(5.5, 3)$ and

$$(\log(\theta_1), \log(\theta_2))^{\mathrm{T}} \sim \mathcal{N}\left(\left(\log\left(\frac{0.08}{0.92}\right) - 0.75, 0.675\right)^{\mathrm{T}}, \begin{bmatrix} 1.5 & 0 \\ 0 & 0.75 \end{bmatrix}\right)$$

are chosen to satisfy approximately two conditions on prior distribution of $\tilde{d}_0$ and $\tilde{d}_1$ formulated above. While these prior parameters are chosen to satisfy condition on two combination only, the difference can be found for the rest of combinations. Figure 2.2 shows the prior distributions of the toxicity probability on each combination imposed by the specified parameters' prior distributions for Emax, L2 and P1 models.

As expected the prior distributions of the toxicity probability associated with combinations $\tilde{d}_0$ and $\tilde{d}_1$ have similar characteristics for all models. At the same time, the differences can be found for other combinations. Emax and L2 being

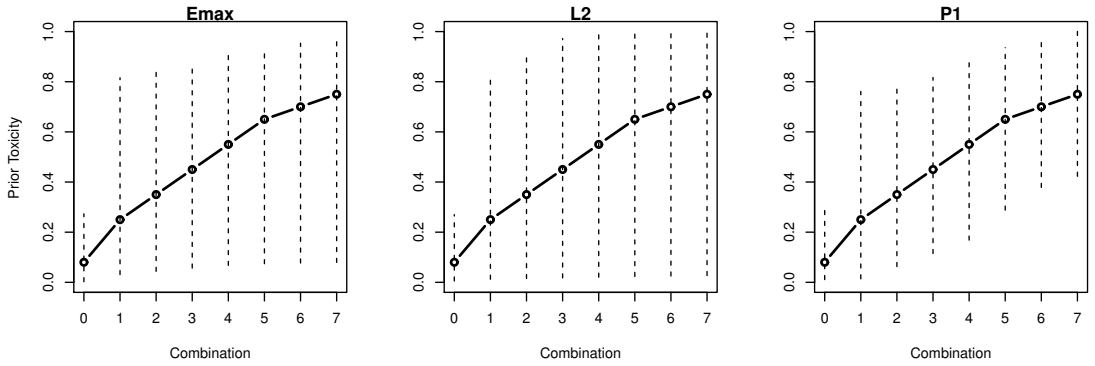

Figure 2.2: Mean values (black circles) and 95% credible intervals (dashed lines) of toxicity probabilities for each combination using Emax, L2 and P1 models.

more flexible models allow for a more informative prior distribution of the toxicity probability associated with the standard of care while leaving the distributions of the rest of combinations vague. At the same time, the prior by P1 cannot achieve the same as one parameter is used only and the distribution of the toxicity probability associated with the standard of care cannot be imposed independently of the other combinations.

The characteristics of all models compared are evaluated in `R` (R Core Team, 2015) using the `bcrm`-package by Sweeting et al. (2013). To accommodate the randomisation to the control arm and the Emax model corresponding modifications to the package were made. For all methods 10^4 replicated trials were used. The posterior distribution of parameters were found using JAGS (Hornik et al., 2003). The number of burn-in iterations is 2000, the number of production iterations is 10^4 and two chains were used.

2.2.5 Operating Characteristics

Proportion of Correct Selections The results of the comparison are summarized in Table 2.1 (scenarios 1-5) and Table 2.2 (scenario 6-9), with each figure representing the proportion of each combination selections. The last columns corresponds to the mean proportion of toxicity outcomes (DLTs).

Table 2.1: Operating characteristics of EmaxR, L2R, L2 and P1 models in scenarios 1-5: proportions of each combination selections and mean proportions of toxicity outcomes (DLTs). The MTC selection is in bold. Results are based on 10^4 replications.

	$\tilde{d}_0$	$\tilde{d}_1$	$\tilde{d}_2$	$\tilde{d}_3$	$\tilde{d}_4$	$\tilde{d}_5$	$\tilde{d}_6$	$\tilde{d}_7$	DLTs
Scenario 1									
Toxicity	0.08	0.10	0.12	0.15	0.25	0.40	0.45	0.47	
EmaxR	0.0	2.1	8.8	24.9	**43.1**	13.6	3.8	3.7	19.1
L2R	0.0	1.8	8.3	26.5	**44.6**	13.8	3.1	2.0	19.2
L2($N = 36$)	0.4	0.6	4.4	29.7	**45.4**	15.6	2.4	1.5	25.2
L2	0.4	0.1	3.5	24.9	**52.6**	16.4	2.1	1.3	25.1
P1	0.0	1.0	4.2	16.5	**51.4**	20.4	5.5	1.0	29.2
Scenario 2									
Toxicity	0.080	0.090	0.095	0.100	0.105	0.110	0.112	0.115	
EmaxR	0.0	0.3	0.5	0.9	0.5	1.6	1.9	**94.4**	10.2
L2R	0.0	0.6	0.8	1.5	1.8	1.6	2.5	**91.2**	10.1
L2($N = 36$)	0.1	0.3	0.8	0.3	0.9	1.1	2.0	**94.5**	11.0
L2	0.3	0.1	0.3	0.5	1.8	1.0	0.5	**95.6**	11.4
P1	0.0	0.0	0.0	0.0	0.0	0.6	0.8	**98.7**	11.3
Scenario 3									
Toxicity	0.08	0.25	0.26	0.265	0.27	0.27	0.27	0.27	
EmaxR	2.4	**44.3**	13.9	7.3	7.1	5.0	4.1	15.9	21.2
L2R	2.8	**44.7**	15.8	10.1	7.6	4.4	2.5	12.1	21.3
L2($N = 36$)	5.3	**25.5**	13.4	10.1	10.6	7.53	5.7	21.9	25.3
L2	5.1	**27.0**	13.0	9.4	11.5	7.1	6.1	20.7	25.3
P1	0.0	0.0	0.5	0.5	13.4	18.9	17.9	**44.7**	26.6
Scenario 4									
Toxicity	0.080	0.250	0.500	0.510	0.520	0.525	0.530	0.535	
EmaxR	5.1	**82.1**	11.5	0.9	0.3	0.1	0.0	0.0	24.4
L2R	4.1	**86.4**	8.4	0.9	0.3	0.0	0.0	0.0	24.6
L2($N = 36$)	20.7	**72.0**	6.1	0.7	0.3	0.1	0.0	0.0	28.4
L2	20.9	**75.0**	4.0	0.1	0.0	0.0	0.0	0.0	26.9
P1	20.2	**71.7**	7.4	0.7	0.0	0.0	0.0	0.0	31.0
Scenario 5									
Toxicity	0.080	0.085	0.090	0.100	0.50	0.50	0.50	0.50	
EmaxR	0.0	1.9	7.8	**57.1**	31.2	1.6	0.1	0.4	18.8
L2R	0.0	1.5	9.5	**56.4**	30.2	1.4	0.8	0.3	18.2
L2($N = 36$)	0.0	0.0	4.1	**59.7**	31.7	1.5	0.5	0.4	26.1
L2	0.1	0.1	1.4	**62.7**	35.5	0.2	0.1	0.0	25.1
P1	0.0	0.0	0.0	**54.9**	36.4	7.0	1.2	0.5	33.2

Overall, both randomised models EmaxR and L2R result in nearly the same PCS with differences no more than 4% across scenarios. The models that do not employ randomisation have a comparable performance in scenarios 1, 2, 4 and 6. In the rest of scenarios, P1 is less conservative and selects more toxic combinations more often than L2. The non-randomised designs select the MTC more often if the

Table 2.2: Operating characteristics of EmaxR, L2R, L2 and P1 models in scenarios 6-9: proportions of each combination selections and mean proportions of toxicity outcomes (DLTs). The MTC selection is in bold. Results are based on 10^4 replications.

	$\tilde{d}_0$	$\tilde{d}_1$	$\tilde{d}_2$	$\tilde{d}_3$	$\tilde{d}_4$	$\tilde{d}_5$	$\tilde{d}_6$	$\tilde{d}_7$	DLTs
Scenario 6									
Toxicity	0.020	0.030	0.050	0.080	0.090	0.100	0.110	0.115	
EmaxR	0.0	0.0	0.0	0.0	0.4	1.1	1.2	**97.2**	8.3
L2R	0.0	0.0	0.0	0.0	1.0	1.1	1.9	**96.1**	8.7
L2($N = 36$)	0.0	0.0	0.0	0.0	0.6	0.9	0.8	**97.5**	10.9
L2	0.0	0.0	0.0	0.1	0.1	0.2	0.5	**99.3**	11.3
P1	0.0	0.0	0.0	0.0	0.0	0.1	1.1	**98.8**	11.3
Scenario 7									
Toxicity	0.20	0.23	0.40	0.40	0.40	0.40	0.40	0.40	
EmaxR	12.7	**64.3**	15.1	4.7	1.7	0.7	0.5	0.4	25.8
L2R	9.4	**67.2**	15.7	5.3	1.4	0.3	0.4	0.3	25.8
L2($N = 36$)	19.0	**52.0**	18.3	4.5	2.0	1.7	1.4	1.1	30.1
L2	15.0	**62.5**	15.1	3.6	1.8	0.8	0.8	0.5	29.2
P1	0.2	**28.0**	29.2	23.5	14.1	3.87	0.8	0.4	35.2
Scenario 8									
Toxicity	0.17	0.20	0.25	0.35	0.40	0.40	0.40	0.40	
EmaxR	5.1	40.3	**29.5**	16.1	5.9	1.7	0.8	0.5	23.6
L2R	2.8	45.3	**31.5**	12.7	4.9	1.8	0.8	0.20	23.3
L2($N = 36$)	4.5	27.5	**34.3**	19.1	7.3	3.1	1.5	2.7	29.0
L2	4.5	24.9	**41.1**	19.1	6.5	1.9	1.1	0.9	28.6
P1	0.0	3.4	**30.2**	37.5	18.7	7.27	1.7	1.1	32.3
Scenario 9									
Toxicity	0.25	0.255	0.26	0.265	0.27	0.27	0.27	0.27	
EmaxR	**21.6**	37.6	10.0	7.9	4.7	4.3	2.3	11.6	25.5
L2R	**19.6**	43.3	12.9	7.2	5.2	2.4	2.1	7.3	26.1
L2($N = 36$)	**10.1**	20.7	12.5	10.5	10.1	6.3	5.8	24.0	26.0
L2	**10.1**	22.9	12.7	10.3	9.5	8.0	6.4	20.0	26.3
P1	**0.3**	6.5	15.0	26.7	28.1	12.8	6.9	3.7	26.6

MTC is located in the middle of the curve - scenarios 1, 5 and 8. For instance, L2 outperforms its randomised version L2R by 6-10% in these scenarios. The decrease in the PCS can be mainly explained by the reduced number of patients on the investigational arm as the differences between L2R and L2($N = 36$) vary in the range 1-3%. While randomisation entails a minor drop in these scenarios, it might also result in a notable increase in the PCS if the MTC is located at the beginning of the curve - scenarios 3, 4 and 7. Considering L2R and L2, randomisation leads to a more accurate MTC selection by 17%, 11% and 5%, respectively. These differences increase to 19%, 14% and 15% compared to L2($N =$

36). Note that the misspecified prior distribution for the control arm (scenarios 6-9) does not affect the MTC selection noticeably. Note that the inclusion of randomisation to the control leads to decrease the proportion of toxicity outcomes almost in all scenarios. For instance, it is decreased by up to 6% in scenario 1 that results in nearly two fewer patients experienced adverse events. The decrease in number of DLTs in the approaches with randomisation can be explained by the fact that the standard of care $\tilde{d}_0$ has the lowest DLT probability and at least 25% of patients are allocated to it. At the same time, the design without randomisation tends to allocate the majority of patient to the investigational arms having greater probabilities of toxicity.

In scenarios 2 and 6, the MTC is the highest combination and all models select it correctly in more than 90% of replications with L2R having the least PCS: 91.2% in scenario 2 and 96.1% in scenario 6. In scenario 9 where the control arm is already associated with the maximum acceptable toxicity, the randomised designs select the control arm by nearly 10% more often than L2. Generally, the randomised designs select combinations in the beginning of the curve ($\tilde{d}_1 - \tilde{d}_3$) more often with nearly 70% of $\tilde{d}_0, \tilde{d}_1, \tilde{d}_2$ selections against 45% for L2 and 22% for P1.

Probability to find the ET While both 4- and 2-parameter models with randomisation were shown to have comparable performances in terms of the PCS, major differences can be found in Table 2.3 in which the upper line represents the proportion of times ET is found ($\hat{\mathcal{P}}$) and the lower line shows the NMSE.

Comparing randomised designs, EmaxR results in greater proportions of correctly identified ETs in scenarios 1, 3, 4, 6 and 8 by 2-7%. In scenarios with no ET (2, 5, 7 and 9), EmaxR finds the ET less often than L2R by 2-9%. Note, however, that both approaches wrongly conclude the ET in the majority of trials under scenario 5. They struggle to capture the jump in toxicity risks between $\tilde{d}_3$ and $\tilde{d}_4$, and overestimate the MTC toxicity probability (see also Figure 2.3 and Figure 2.4)

Table 2.3: Operating characteristics of EmaxR, L2R, L2 and P1 models. The upper line: proportions of times the ET is found. The lower line: NMSE. Scenarios with no ET are underlined. Results are based on 10^4 replications.

		Sc 1	Sc 2	Sc 3	Sc 4	Sc 5	Sc 6	Sc 7	Sc 8	Sc 9
EmaxR	$\hat{\mathcal{P}}$	74.7	16.3	66.7	71.5	67.9	24.2	12.2	29.3	6.6
	NMSE	1.7	2.0	2.2	3.5	3.1	1.4	2.5	1.7	3.7
L2R	$\hat{\mathcal{P}}$	71.8	14.4	59.7	64.1	75.0	20.9	21.3	31.4	11.8
	NMSE	2.0	2.2	7.2	4.4	3.4	1.5	4.4	3.1	7.2
L2 ($N = 36$)	$\hat{\mathcal{P}}$	47.0	13.7	35.8	35.3	43.5	13.7	25.1	21.3	32.3
	NMSE	1.8	2.0	6.3	5.1	3.0	1.4	6.3	4.0	8.9
L2	$\hat{\mathcal{P}}$	61.5	15.8	50.1	50.2	64.7	18.1	37.5	30.2	43.5
	NMSE	2.0	2.5	7.7	5.8	3.6	1.6	7.6	4.5	9.0
P1	$\hat{\mathcal{P}}$	99.9	93.6	99.7	99.7	99.4	95.3	99.8	99.9	99.7
	NMSE	2.1	2.6	7.8	6.1	3.8	2.1	7.8	4.5	9.2

for the illustration of fitted curves). It is also challenging for both models to determine the ET in scenarios 6 and 8 when the prior distribution for the control arm is misspecified. EmaxR correctly identifies the ET in 24% and 30% of trials. At the same time, the prior misspecification is not being an issue in scenarios 7 and 9 with no ET as EmaxR finds the ET in 12% and 6% of trials, respectively.

Regarding the designs with no randomisation, P1 almost always finds the ET regardless the underlying scenario. As expected the strict monotonicity assumption imposed by the power model results in these (often) incorrect conclusions. The model L2 does not share the same pattern and it is able to find the ET with a low probability if it is not present (for example, scenario 2). At the same time, the inclusion of randomisation in L2 leads to a large increase in proportions of correctly identified ET by 10-14% in scenarios 1, 3 and 4. Again, comparing approaches with the same number of patients on the investigational arm these proportions differ by 25-30%. In the rest of scenarios (6 and 8) with the ET, L2 and L2R have similar characteristics and outperform L2 ($N = 36$) by 5-10%. In the majority of scenarios with no ET, using L2R instead of L2 leads to a decrease in the proportion of incorrect conclusions about the ET, for example, by 16% and 32% in scenarios 7 and 9, respectively. There is only one example when randomisation increases the probability to find the ET while it is not there - scenario 5.

Goodness-of-fit Comparing the ability of models to fit the combination-toxicity curve, EmaxR being the most flexible model results in the NMSE smallest values in all scenarios with the largest difference in scenario 3 − 2.2 against 7.2 for L2R, and the smallest difference in scenario 6 − 1.4 against 1.5 for L2R. Similarly, P1, the least flexible alternative, results in the greatest NMSE values in all scenarios. L2 shows a better fit than P1, which can be improved further by randomisation to the control. Low values of the NMSE for L2 with the reduced number of patients can be explained by the decreased accuracy of the non-parametric optimal benchmark.

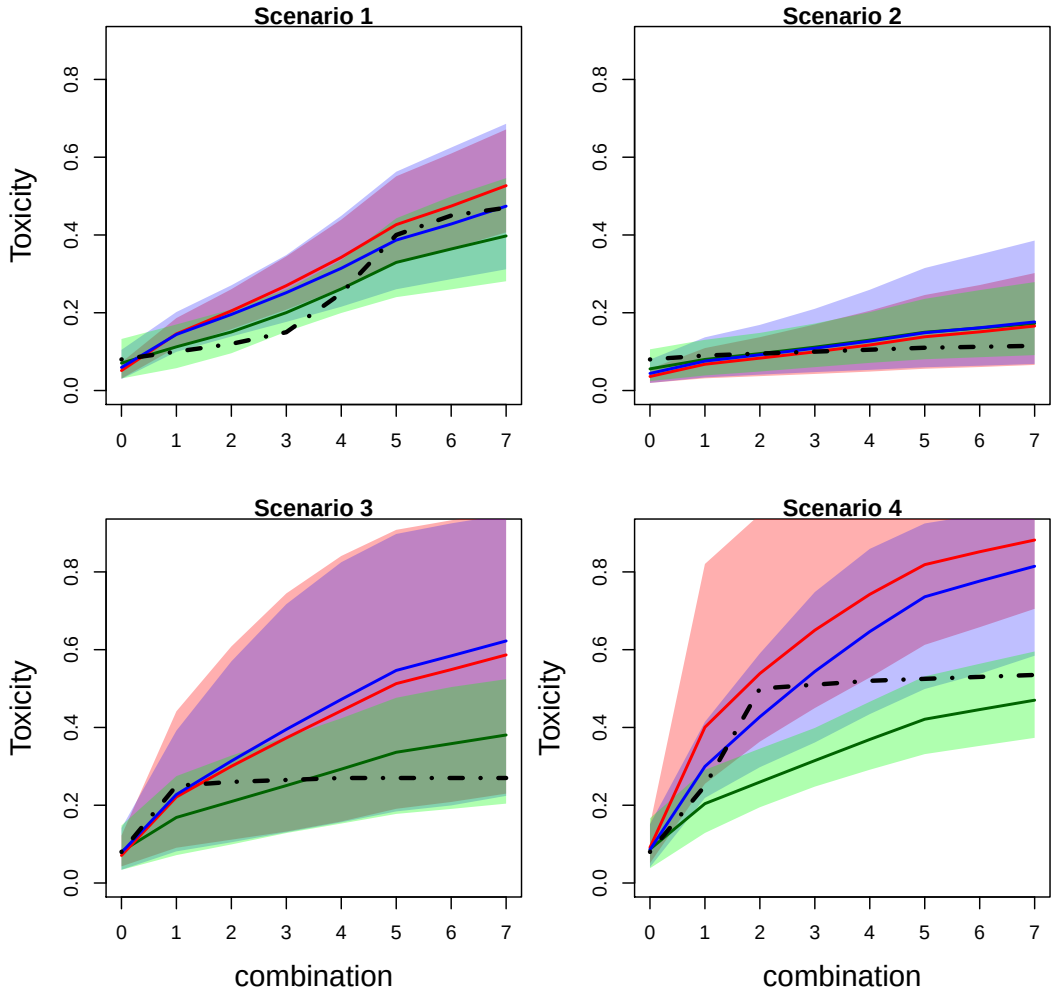

Figure 2.3: Mean values (solid lines) and 90% credible intervals (shadowed area) of toxicity probabilities for each combination obtained by fitted EmaxR (green), L2R (blue) and L2 (red) models in scenarios 1-4. True toxicity curves are marked by dashed-dotted lines. Results are based on 10^4 replications.

Further differences in the ability of these models to detect dose-effects are given

in Figure 2.3 and Figure 2.4 presenting mean values of toxicity probabilities and corresponding 90% credible interval. The mean values for estimated probabilities by EmaxR (green), L2R (blue) and L2 (red) are given by solid lines while the credible intervals are given by shadow areas of the corresponding colour. The true toxicity probabilities are marked by dashed-dotted lines.

EmaxR corresponds to the best fit of the toxicity curve and to the narrowest credible interval among all alternatives in all scenarios. The fitted curves of L2R and L2 are similar for the first combinations, but probability estimates obtained by L2R are more accurate due to the shift toward the true toxicity probability curve for the rest of combinations.

2.2.6 Sensitivity Analysis

Prior Distributions In this section the influence of different sets of prior distributions is studied. Since an investigator usually has a reasonable prior for E_0 and the detection of a plateau should be determined by the data alone (hence an uninformative prior on the $Emax$ parameter is imposed), we consider cases of less informative prior for ED_{50} and more informative one for λ

$$ED_{50} \sim \Gamma(0.25, 0.25),\ \lambda \sim \Gamma(3, 3) \tag{2.2.4}$$

and more informative prior for ED_{50} and less informative for λ

$$ED_{50} \sim \Gamma(1, 1),\ \lambda \sim \Gamma(0.05, 0.05). \tag{2.2.5}$$

We compare previously used prior distributions given in Equation (2.2.3) to the prior distributions given in Equation (2.2.4) denoted by EmaxR (λ) and the set given in Equation (2.2.5) denoted by EmaxR (ED_{50}). The summary of the operating characteristics using different prior distributions are given in Table 2.4.

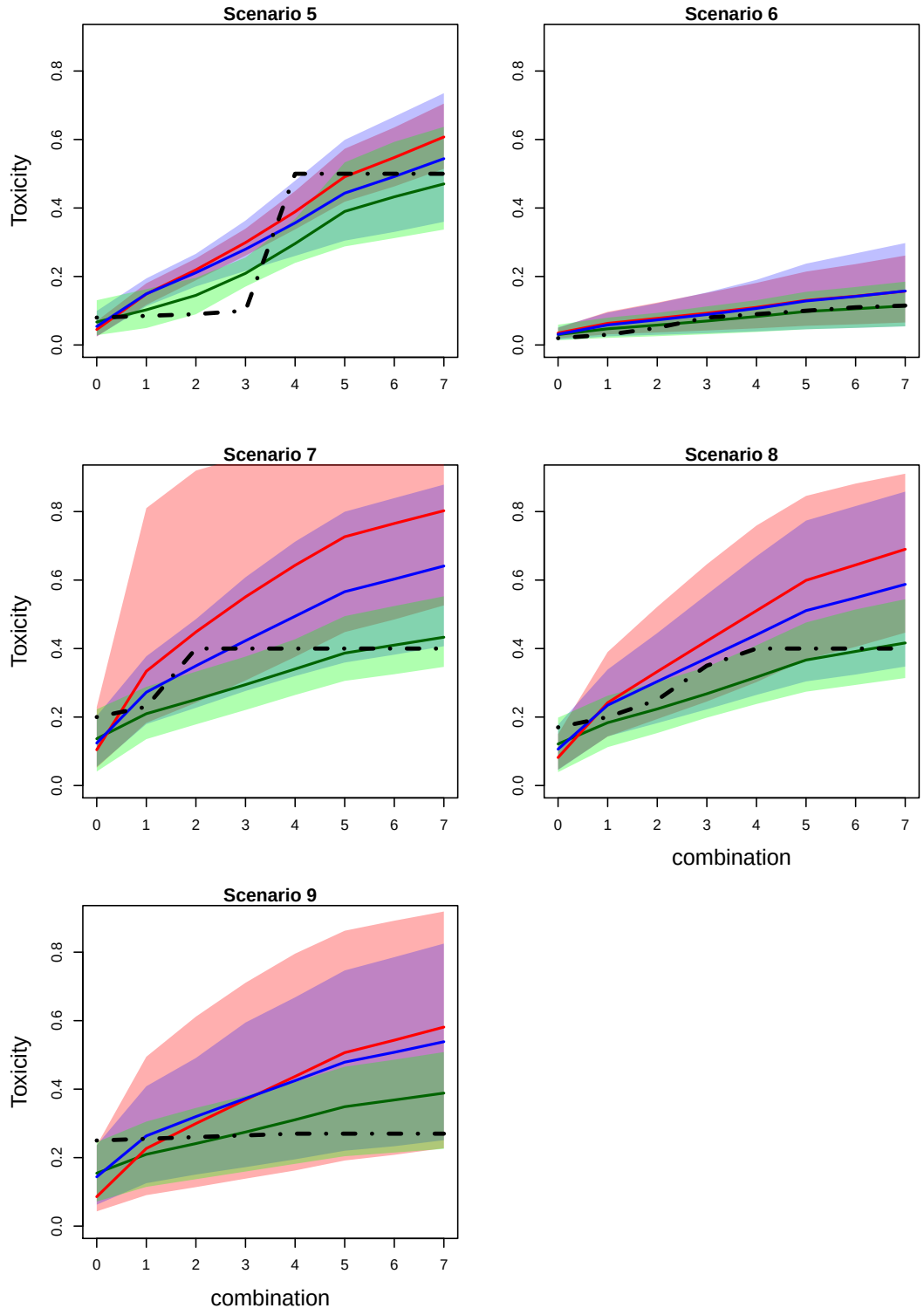

Figure 2.4: Mean values (solid lines) and 90% credible intervals (shadowed area) of toxicity probabilities for each combination obtained by fitted EmaxR (green), L2R (blue) and L2 (red) models in scenarios 5-9. True toxicity curves are marked by dashed-dotted lines. Results are based on 10^4 replications.

In the majority of scenarios the choice of prior did not influence the PCS by more than 4% with scenarios 1, 3 and 8 being exceptions. In cases where the MTC

Table 2.4: Operating characteristics of EmaxR using different prior distributions. Scenarios with no ET are underlined and the most noticeable differences across scenarios are in bold. Results are based on 10^4 replications.

	Sc 1	Sc 2	Sc 3	Sc 4	Sc 5	Sc 6	Sc 7	Sc 8	Sc 9
					Proportion of correct selections				
EmaxR	45.1	94.4	44.3	82.1	57.1	97.2	64.3	29.5	21.6
EmaxR (λ)	43.8	93.3	41.3	85.5	58.7	99.7	64.1	27.3	22.7
EmaxR (ED_{50})	**35.3**	96.5	**52.9**	85.5	54.0	99.8	63.4	**16.5**	24.8
					Proportion of toxicity outcomes				
EmaxR	19.1	10.2	21.2	24.4	18.8	8.3	25.8	23.6	25.5
EmaxR (λ)	19.2	10.3	21.1	24.7	19.1	8.2	25.9	23.5	25.7
EmaxR (ED_{50})	19.1	10.5	21.1	24.5	19.3	8.4	25.4	23.5	26.1
					$\hat{\mathcal{P}}$				
EmaxR	74.7	16.3	66.7	71.5	67.9	24.2	12.2	29.3	6.6
EmaxR (λ)	74.0	12.7	71.1	78.4	**80.0**	28.5	15.6	31.2	10.0
EmaxR (ED_{50})	68.1	9.2	66.3	70.5	61.5	23.6	16.2	**20.4**	11.2
					NMSE				
EmaxR	1.7	2.0	2.2	3.5	3.1	1.4	2.5	1.7	3.7
EmaxR (λ)	1.9	1.8	2.2	3.3	3.7	1.4	2.5	1.7	3.6
EmaxR (ED_{50})	2.0	2.2	2.2	**4.3**	3.2	1.9	3.0	2.3	3.1

lies in the middle of the curve (scenarios 1 and 8), the informative prior for ED_{50} results in nearly 10% loss in the PCS. Conversely, in scenario 3 where the MTC is located at the beginning of the curve, the PCS was greater. At the same time, the proportion of toxicity outcomes are not influenced by the prior choices in all scenarios.

The proportion of times the ET is found is again not affected by the choice of prior in scenarios 1-4, 6-7 and 9 with the difference between competing models below 10%. Major differences in $\hat{\mathcal{P}}$ can be found in scenario 5 in which the informative prior for λ leads to 20% more false conclusions than the informative prior for ED_{50}. At the same time, it is generally harder for EmaxR (ED_{50}) to detect the ET with the largest difference of 10% in scenario 8. While NMSE is not largely affected by the choice of prior distributions, EmaxR (ED_{50}) results in slightly greater values in scenarios 1-2, 4 and 6-8 with the greatest difference of 1 in scenario 4.

Overall, the choice of prior does not seem to have a noticeable impact in the majority of the scenarios, but EmaxR (λ) seems to be more robust and results

in better toxicity curve fittings. Therefore, an informative prior for λ and an uninformative prior for ED_{50} appear to be preferable choices.

Randomisation Ratio and No Skipping Constraint The randomisation was shown to be the key factor in improving the fit of the dose-toxicity curve and in establishing the ET. We now consider the impact of different randomisation ratios on the performance of the design. We also investigate the impact of the "no skipping constraint". The prior given in Equation (2.2.3) is used and the summary of the operating characteristics is given in Table 2.5.

Table 2.5: Operating characteristics of EmaxR using different randomisation ratios and no skipping constraint. Scenarios with no ET are underlined and the most noticeable differences across scenarios are in bold. Results are based on 10^4 replications.

	Sc 1	Sc 2	Sc 3	Sc 4	Sc 5	Sc 6	Sc 7	Sc 8	Sc 9
					Proportion of correct selections				
EmaxR(2:1)	43.5	91.6	44.1	84.2	58.5	97.2	62.7	28.2	**27.7**
EmaxR(2:1) No Skip	46.9	90.7	44.8	83.7	57.2	96.7	62.3	27.1	**29.2**
EmaxR(3:1)	45.1	94.4	44.3	82.1	57.1	96.8	64.3	29.5	21.6
EmaxR(3:1) No Skip	45.7	90.4	47.1	83.8	55.9	97.2	66.5	31.0	21.5
EmaxR(4:1)	43.3	94.6	45.1	86.8	57.1	98.1	66.5	29.1	19.1
EmaxR(4:1) No Skip	44.0	**88.4**	48.2	89.2	58.8	98.0	69.2	29.7	19.5
					Proportion of toxicity outcomes				
EmaxR(2:1)	17.5	9.9	19.7	22.9	18.2	8.1	25.2	22.5	25.8
EmaxR(2:1) No Skip	17.6	9.5	19.9	22.5	17.8	8.5	24.8	22.4	25.5
EmaxR(3:1)	19.1	10.2	21.2	24.4	18.8	8.3	25.8	23.6	25.5
EmaxR(3:1) No Skip	18.8	9.9	21.5	24.3	18.9	7.9	25.6	23.1	25.5
EmaxR(4:1)	20.4	10.7	23.2	26.8	20.9	8.6	27.6	25.1	26.9
EmaxR(4:1) No Skip	19.5	10.3	23.1	26.0	20.7	8.9	26.8	24.4	26.8
					$\hat{\mathcal{P}}$				
EmaxR(2:1)	73.5	11.4	63.1	71.9	68.3	19.9	7.8	**17.6**	4.8
EmaxR(2:1) No Skip	69.1	12.1	63.8	69.6	69.1	20.0	8.5	**16.7**	4.6
EmaxR(3:1)	74.7	16.3	66.7	71.5	68.0	24.2	12.2	29.3	6.6
EmaxR(3:1) No Skip	75.6	12.2	68.7	72.8	71.3	24.1	11.0	27.2	7.7
EmaxR(4:1)	76.3	11.6	64.5	72.1	72.3	20.5	11.9	24.4	9.9
EmaxR(4:1) No Skip	74.5	14.4	**69.2**	69.9	**66.1**	20.9	13.1	23.1	9.9
					NMSE				
EmaxR(2:1)	1.7	2.1	3.4	3.3	3.2	1.7	2.4	1.6	3.6
EmaxR(2:1) No Skip	1.7	2.1	3.3	3.5	3.5	1.7	2.5	1.6	3.5
EmaxR(3:1)	1.7	2.0	2.2	3.5	3.1	1.5	2.5	1.7	3.7
EmaxR(3:1) No Skip	1.6	2.0	3.4	3.3	3.2	1.7	2.4	1.6	3.7
EmaxR(4:1)	1.6	1.9	3.2	3.4	3.3	1.5	2.5	1.7	3.7
EmaxR(4:1) No Skip	1.7	2.0	3.1	3.6	3.4	1.6	2.6	1.6	3.6

Comparing models with the same randomisation ratios, no skipping constraint has a minor influence on characteristics of EmaxR in all scenarios. In the majority of cases, the difference in the PCS does not exceed 4% with scenario 2 being an exception. In this case, $N = 48$ is not always enough to test all combinations and reach the MTC given a large cohort size ($c = 5$). For EmaxR(4:1) no skipping constraint also increases $\hat{\mathcal{P}}$ by 5% in scenario 3 with the MTC being at the beginning of the curve and decreases it by 6% in scenario 5 with the MTC being the medium combination.

Considering different randomisation ratios with combinations skipping, all approaches result in the comparable PCS in all scenarios with EmaxR(2:1) in scenario 9 being an exception. The increased number of patients on the control (16 patients) leads to a higher PCS when the MTC is a standard therapy itself. At the same time, a smaller number of patients on the investigational arm might result in a substantial loss in the ability to detect the ET, for example, in scenario 8, $\hat{\mathcal{P}}$ decreased by 10% compared to EmaxR(3:1). Note that EmaxR(4:1) also results in a less accurate ET detection in this case as there are fewer patients on the control arm. Moreover, ratio 4:1 results in more toxicity outcomes as more patients are assigned to the standard of care. There is no noticeable difference in the NMSE across ratios. Overall, the operating characteristics remain unchanged for various randomisation ratios in the majority of scenarios, but the ratio 3:1 seems to be a reasonable trade-off between the MTC selection and the plateau detection.

2.3 A Benchmark for Studies with Complex Endpoints

To assess the goodness-of-fit across different scenarios in the previous section, the non-parametric optimal benchmark was used. It is, however, applicable to trials with a single binary endpoint only. Below, we generalise the benchmark to trials with (possibly multiple) continuous endpoints.

2.3.1 A Benchmark for a Continuous Endpoint

Consider a Phase I clinical trial with continuous outcome Y_{ij} at dose d_j for patient i having a cumulative distribution function (CDF) $F_j(y)$. The goal of the trial is find the target dose (TD) which minimises (or maximises as defined by an investigator) some decision criterion $T(\cdot, \gamma_t)$. We will further suppress the γ_t notation from $T(\boldsymbol{y}_j)$. In simulations, the CDF F_j is chosen by a statistician and specifies the distribution of outcomes for a given dose d_j, and the set of CDFs corresponding to doses $d_1, \ldots, d_m$ defines a simulation scenario. This simple fact is going to be a central part of our proposal. To illustrate the construction of the novel benchmark step-by-step, we use a setting studied by Wang and Ivanova (2015) throughout this section.

Example 1. *Wang and Ivanova (2015) considered a setting with $m = 6$ doses and a biomarker for toxicity measured on a continuous scale. In one of the simulation scenarios presented, it is assumed that toxicity outcome Y_{ij} given dose level d_j has normal distribution $\mathcal{N}(0.1j, (0.1j)^2)$, $j = 1, \ldots, 6$. Then, the CDF F_j is the CDF of a normal random variable with corresponding parameters*

$$\Phi(\cdot,\ \mu_j\ =\ 0.1j\ ,\ \sigma^2\ =\ (0.1j)^2).$$

These CDFs will be used to obtain the benchmark in this scenario.

Let us denote the quantile transformation as

$$F_j^{-1}(x) = \inf\{y | F_j(y) \geq x\},\ 0 < x < 1. \tag{2.3.1}$$

Then,

Probability Integral Transform. *If $U \sim \mathcal{U}(0, 1)$ is a uniform random variable on the unit interval, then F_j is the cumulative distribution function of a random variable $F_j^{-1}(U)$.*

This result is commonly used for the *inverse transform sampling* (e.g. see Bekele and Shen, 2005, for an example in dose finding) to generate random variables with CDF F_j.

Assume that the information about a patient's toxicity profile is summarised in a single value u_i drawn from $\mathcal{U}(0,1)$. For patient i with profile u_i, the quantile transformation $y_{ij} = F_j^{-1}(u_i)$ is applied to obtain a continuous outcome that this patient would have given dose d_j, $j = 1, \ldots, m$. Different dose levels are modelled by applying the corresponding quantile transformation. This results in a vector of responses $(y_{i1}, \ldots, y_{im})$, also called *the complete information* about patient i. The same procedure is repeated for all patients $i = 1, \ldots, N$ which results in the vector of responses for each dose level $\boldsymbol{y}_j = (y_{1j}, \ldots, y_{Nj})$, $j = 1, \ldots, m$.

Example 1 (Continued)**.** *Following the setting by Wang and Ivanova (2015), assume that the first patient has a toxicity profile* $u_1 = 0.40$*. The benchmark answers the question "how would patient* 1 *respond to dose level* d_j *with the response having distribution* $\mathcal{N}(0.1j, (0.1j)^2)$*". Applying the corresponding quantile transformation, the response of patient* 1 *given the dose level* d_1 *is equal to* $y_{11} = \Phi^{-1}(u_1 = 0.40, \mu_j = 0.1, \sigma^2 = 0.1^2) \approx 0.075$*. Subsequently, the complete information about patient* 1 *consists in the vector of responses at all dose levels* $d_1, \ldots, d_6$

$$(0.075, 0.149, 0.224, 0.299, 0.373, 0.448).$$

The complete information for 5 *patients with randomly generated profiles* $u_1, \ldots, u_5$ *is given in Table 2.6.*

Recalling the decision criterion $T(\boldsymbol{y}_j)$ on which the TD selection is based, the dose level d_j such that $T(\boldsymbol{y}_j)$ is minimised (or maximised) is selected as the TD in a single trial. For instance, if the goal of the trial is to find the dose having the average level of toxicity γ_t, the decision criterion given in Equation (2.1.7) can be used. The benchmark can be constructed for different decision criteria and then

Table 2.6: Complete information for 5 patients with randomly generated toxicity profiles.

Patient's profile u_i	Patient's response					
	d_1	d_2	d_3	d_4	d_5	d_6
$u_1 = 0.40$	0.075	0.149	0.224	0.299	0.373	0.448
$u_2 = 0.25$	0.033	0.065	0.098	0.130	0.163	0.195
$u_3 = 0.92$	0.241	0.481	0.722	0.962	1.203	1.443
$u_4 = 0.67$	0.144	0.288	0.432	0.576	0.720	0.864
$u_5 = 0.31$	0.050	0.101	0.151	0.202	0.252	0.302
Mean	0.109	0.217	0.325	0.434	0.542	0.650
Variance	0.007	0.029	0.065	0.116	0.181	0.261

can be used to evaluate any design under investigation.

Example 1 (Continued)**.** *The goal of the trial considered by Wang and Ivanova (2015) is to find the dose with the mean response closest to the target response γ_t. The criterion of choosing the dose which maximises the probability of the average level of toxicity μ_j to be in the ε neighbourhood of γ_t was considered. Let $g_j(\cdot|\boldsymbol{y}_j)$ be a probability density function of μ_j given the data $\boldsymbol{y}_j$. Then, the decision criterion takes the form*

$$T(\boldsymbol{y}_j) = \int_{\gamma_t-\varepsilon}^{\gamma_t+\varepsilon} g_j(v|\boldsymbol{y}_j)\mathrm{d}v. \tag{2.3.2}$$

The TD is the dose for which the criterion $T(\boldsymbol{y}_j)$ is maximised. Following the original framework, $\gamma_t = 0.1$ and $\varepsilon = 0.01$ are chosen. Using the complete information generated in Table 2.6 and the density function of Normal distribution with corresponding mean and variance parameters yields: $T(\boldsymbol{y}_1) = 0.09$; $T(\boldsymbol{y}_2) = 0.04$; $T(\boldsymbol{y}_3) = 0.02$; $T(\boldsymbol{y}_4) = 0.01$; $T(\boldsymbol{y}_5) = 0.01$ and $T(\boldsymbol{y}_6) = 0.01$. The value of the criterion is maximised for d_1 which is selected as the TD in this single trial. The procedure is repeated for $s = 1, \ldots, S$ simulated trials to obtain the proportion of correct selections. The evaluation of the method by Wang and Ivanova (2015) using the proposed benchmark is provided in Section 2.3.3.

Algorithm 1 provides the step-by-step guidance on how the benchmark can be constructed based on S simulated trials.

Algorithm 1 Computing a benchmark for a single continuous endpoint

1. Specify CDFs F_j for all doses d_j, $j = 1, \ldots, m$ and define the decision criterion $T(\cdot)$
2. Generate a sequence of patients' profiles $\{u_i\}_{i=1}^{N}$ from uniform distribution $\mathcal{U}(0, 1)$.
3. Transform u_i for dose level d_j using $y_{ij} = F_j^{-1}(u_i)$, $i = 1, \ldots, N$, $j = 1, \ldots, m$ and store $\boldsymbol{y}_j = (y_{1j}, \ldots, y_{Nj})$.
4. Compute $T(\boldsymbol{y}_j)$ for all $j = 1, \ldots, m$, find dose J for which $T(\boldsymbol{y}_J)$ is maximised (minimised) and set $Z_s = J$.
5. Repeat steps 2-4 for $s = 1, \ldots, S$ simulated trials
6. Use $\bar{Z}^{(j)} = \sum_{s=1}^{S} \mathbb{I}(Z_s = j)/S$ as the selection proportion of dose d_j, $j = 1, \ldots, m$

The proposed benchmark can be applied to a wide range of distributions as it requires the quantile information only, which is routinely available for many distributions in various statistical software (for example, `qbinom`, `qnorm`, `qexp` , etc in `R` (R Core Team, 2015)). Note that the Probability Integral Transform can be also applied to discrete random variables in which case the quantile transformation $F_j^{-1}(\cdot)$ is given explicitly. It is easy to see that $F_j^{-1}(\cdot)$ of a Bernoulli random variable in Algorithm 1 results in the original benchmark construction proposed by O'Quigley et al. (2002). Importantly, the benchmark employs the concept of the complete information implying that a statistician conducting simulations knows distributions at all dose levels. This means that the proposed benchmark can be applied to more complex regimen finding trials (e.g. combination or dose-schedule studies) straightforwardly as the complete information concept, proposed by O'Quigley et al. (2002), does not suffer from the uncertainty in the monotonicity ordering.

The novel benchmark can be also applied to clinical trials with multiple endpoints. This construction is provided below.

2.3.2 A Benchmark for Multiple Endpoints

In the setting with several endpoints, the correlation between them is important. Below, we describe the algorithm generating correlated outcomes in the bench-

mark framework. The approach described below has been known for a long time (Tate, 1955; Molenberghs et al., 2001). We apply it to an arbitrary distribution of outcomes to generate the complete information vector. We start with the case of binary toxicity and continuous efficacy responses that has attracted a lot of attention in the literature recently (see e.g. Hirakawa, 2012; Yeung et al., 2015, 2017).

Consider a Phase I/II clinical trial with toxicity outcome $Y_{ij}^{(1)}$ and efficacy outcome $Y_{ij}^{(2)}$ with CDFs $F_j^{(1)}$ and $F_j^{(2)}$, respectively, at dose level d_j for patient i. We will use the setting studied by Bekele and Shen (2005) to illustrate the construction of the benchmark for multiple endpoints throughout this section.

Example 2. *Bekele and Shen (2005) considered a setting with $m = 4$ dose levels, an efficacy outcome at dose d_j having Gamma distribution $\Gamma(\lambda_j\omega, \omega)$ where $\lambda_j\omega$ is the shape parameter, $\omega = 0.1$ is the rate parameter (i.e., the mean equals to λ_j), and a DLT outcome having probability p_j. In one of the simulation scenarios the following parameters are assumed $\lambda_1 = 25,\ \lambda_2 = 60,\ \lambda_3 = 115,\ \lambda_4 = 127$ and $p_1 = 0.01,\ p_2 = 0.10,\ p_3 = 0.25,\ p_4 = 0.60$. Then, $F_j^{(1)}$ is the CDF of a Bernoulli random variable with parameter p_j and $\mathrm{G}(\cdot, \lambda_j\omega, \omega)$ is the CDF of a Gamma random variable with parameter λ_j, $j = 1, 2, 3, 4$.*

The toxicity/efficacy profile of patient i is given by two characteristics: $u_i^{(1)} \in (0, 1)$ corresponding to toxicity and $u_i^{(2)} \in (0, 1)$ corresponding to efficacy. Firstly, we generate a bivariate standard normal vector $(x_i^{(1)}, x_i^{(2)})$ with the mean $\mu = (0, 0)$ and the covariance matrix

$$\Sigma = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \tag{2.3.3}$$

where ρ is the correlation coefficient. In simulation studies, the correlation coefficient, ρ, is specified by a statistician as part of the scenario. By applying the CDF

of the standard normal random variable

$$(u_i^{(1)}, u_i^{(2)}) = (\Phi(x_i^{(1)}), \Phi(x_i^{(2)})),$$

one can obtain two correlated random variables having uniform distributions. Then, the corresponding quantile transformations are applied to $u_i^{(1)}$ and $u_i^{(2)}$ marginally as described in Section 2.3.1 and responses for patient i given the dose level d_j are obtained as $y_{ij}^{(1)} = F_j^{-1^{(1)}}(u_i^{(1)})$, $y_{ij}^{(2)} = F_j^{-1^{(2)}}(u_i^{(2)})$. This results in the complete information vector of toxicity and efficacy outcomes at all dose level for the patient i. The procedure is repeated for N patients and pairs of vectors $\boldsymbol{y}_j^{(1)} = (y_{1j}^{(1)}, \ldots, y_{Nj}^{(1)})$ and $\boldsymbol{y}_j^{(2)} = (y_{1j}^{(2)}, \ldots, y_{Nj}^{(2)})$ are obtained for each dose level $d_j, j = 1, \ldots, m$.

Example 2 (Continued)**.** *The correlation coefficient considered by Bekele and Shen (2005) is* $\rho = 0.25$*. The bivariate normal vector with mean* $\mu = (0, 0)$ *and covariance matrix* Σ *is initially generated:* $(x_1, x_2) = (-0.892, 0.292)$*. Then, the first patient has a toxicity profile* $u_1^{(1)} = \Phi(-0.892) = 0.186$ *and an efficacy profile* $u_1^{(2)} = \Phi(0.292) = 0.615$ *which corresponds to the toxicity response* $F_1^{-1^{(1)}}(u_1^{(1)} = 0.186, p_1 = 0.01) = 0$ *and to the efficacy response* $G^{-1}(u_1^{(2)} = 0.615, \lambda_1\omega = 2.5, \omega = 0.1) = 26.3$ *. Subsequently, the vector of the complete toxicity information is* $(0, 0, 1, 1)$ *and the vector of the complete efficacy information is* $(26.3, 74.6, 121.8, 134.3)$*. The complete information for 5 patients with randomly generated profiles* $u_1^{(1)}, u_1^{(2)}, \ldots, u_5^{(1)}, u_5^{(2)}$ *is given in Table 2.7.*

Similar to the single endpoint case, the TD selection is based on a pre-specified decision criterion, $T(\boldsymbol{y}_j^{(1)}, \boldsymbol{y}_j^{(2)})$, which takes the minimum (maximum) value for the most desirable dose level. This would, however, involve the information for all endpoints of interest and can have more complicated structure. In the context of the Phase I/II clinical trial the decision criterion is also known as a trade-off function (see e.g. Thall and Cook, 2004).

Table 2.7: Complete information for 5 patients with randomly generated toxicity and efficacy profiles.

Patient's profile u_i	Patient's response			
	d_1	d_2	d_3	d_4
$u_1^{(1)} = 0.186$	0	0	1	1
$u_1^{(2)} = 0.615$	26.3	74.6	121.8	134.3
$u_2^{(1)} = 0.390$	0	0	0	1
$u_2^{(2)} = 0.214$	12.2	48.4	87.3	97.3
$u_3^{(1)} = 0.618$	0	0	0	0
$u_3^{(2)} = 0.898$	45.7	104.7	159.3	173.5
$u_4^{(1)} = 0.456$	0	0	0	1
$u_4^{(2)} = 0.545$	23.6	70.0	112.9	128.1
$u_5^{(1)} = 0.683$	0	0	0	0
$u_5^{(2)} = 0.869$	42.5	99.9	153.5	167.4
Number of toxicities	0	0	1	3
Mean (efficacy)	30.1	79.5	127.5	140.2
Standard Deviation (efficacy)	13.8	23.0	29.5	30.8

Example 2 (Continued)**.** *Bekele and Shen (2005) defined the TD as the dose with the highest expected efficacy while being safe ($p_j < 0.35$) and efficacious ($\lambda_j > 5$). This translates into the criterion*

$$T(\boldsymbol{y}_j^{(1)}, \boldsymbol{y}_j^{(2)}) = \frac{\sum_{i=1}^{n} y_{ij}^{(2)}}{n} \mathbb{I}\left(\int_0^5 g_j^{(2)}(v|\boldsymbol{y}_j^{(2)})\mathrm{d}v < \theta^{(2)},\ \int_0^{0.35} g_j^{(1)}(v|\boldsymbol{y}_j^{(1)})\mathrm{d}v > \theta^{(1)} \right) \tag{2.3.4}$$

where $g_j^{(1)}(\cdot|\boldsymbol{y}_j^{(1)})$ and $g_j^{(2)}(\cdot|\boldsymbol{y}_j^{(2)})$ are probability density functions of a toxicity probability and of an efficacy response given the data $\boldsymbol{y}_j^{(1)}, \boldsymbol{y}_j^{(2)}$, respectively, and $\theta^{(1)}$, $\theta^{(2)}$ are threshold probabilities. This decision criterion is used to construct the benchmark in this setting. Applied to the benchmark, the integrals in Equation (2.3.4) are computed using density functions of Beta and Normal random variables for toxicity and efficacy outcomes, respectively. Using summary statistics given in Table 2.7 and threshold probabilities $\theta^{(1)} = \theta^{(2)} = 0.50$, the values of the criterion are $T(\boldsymbol{y}_1^{(1)}, \boldsymbol{y}_1^{(2)}) = 0.30$; $T(\boldsymbol{y}_2^{(1)}, \boldsymbol{y}_2^{(2)}) = 0.79$; $T(\boldsymbol{y}_3^{(1)}, \boldsymbol{y}_3^{(2)}) = 1.28$; $T(\boldsymbol{y}_4^{(1)}, \boldsymbol{y}_4^{(2)}) = 0$. The criterion is maximised for dose level d_3 which is selected as the TD in this single trial. The procedure is repeated for $s = 1, \ldots, S$ simulated trials to obtain the proportion of correct selections. The evaluation of the method by Bekele and

Shen (2005) using the proposed benchmark is provided in Section 2.3.4.

Similarly, the benchmark can be applied to an arbitrary number of endpoints. For instance, consider a Phase I/II trial in which toxicity and efficacy are evaluated in four cycles. Then, the profile of patient i is given by $u_i^{(1)}, \ldots, u_i^{(8)}$ each drawn from $\mathcal{U}(0,1)$ and the rest of the construction remains unchanged. The procedure to generate the benchmark for K endpoints is given in Algorithm 2.

Algorithm 2 Computing a benchmark for multiple outcomes

1. Specify $K \times K$ covariance matrix Σ and define objective function $T(\cdot)$.
2. Generate $x_i = (x_i^{(1)}, \ldots, x_i^{(K)})$, $i = 1, \ldots, N$ from $\mathcal{N}(\mu, \Sigma)$ where $\mu = [0, \ldots, 0]_{1\times K}$.
3. Compute $u_i = (u_i^{(1)}, \ldots, u_i^{(K)})$, $i = 1, \ldots, N$ applying CDF Φ to each component of x_i.
4. Apply the quantile transformation $y_{ij}^{(k)} = F_j^{-1^{(k)}}(u_i^{(k)})$ for $k = 1, \ldots, K$ at each dose level d_j, $j = 1, \ldots, m$ and for $i = 1, \ldots, N$ as described in Algorithm 1 and store $\boldsymbol{y}_j^{(k)} = (y_{1j}^{(k)}, \ldots, y_{Nj}^{(k)})$
5. Compute $T(\boldsymbol{y}_j^{(1)}, \ldots, \boldsymbol{y}_j^{(K)})$, $j = 1, \ldots, m$, find dose level J for which $T(\cdot)$ is maximised (minimised) and set $Z_s = J$.
6. Repeat steps 2-5 for $s = 1, \ldots, S$ simulated trials
7. Use $\bar{Z}^{(j)} = \sum_{s=1}^{S} \mathbb{I}(Z_s = j)/S$ as the selection proportion of dose d_j, $j = 1, \ldots, m$

In the following section, we illustrate the implementation of Algorithm 1 (Section 2.3.3) and Algorithm 2 (Section 2.3.4) in different clinical contexts.

2.3.3 Application to a Phase I Trial with Continuous Toxicity

A dichotomization of the toxicity endpoint (DLT/no DLT) in Phase I clinical trials restricts the available information about the drug's toxicity. In fact, a continuous toxicity endpoint can provide more information about a drug's profile (Wang et al., 2000; Bekele and Thall, 2004; Wang and Ivanova, 2015).

Recently, Wang and Ivanova (2015) proposed the Bayesian Design for Continuous Outcomes (BDCO) which can be applied to clinical trials with a continuous toxicity

endpoint. In short, BDCO assumes that outcome Y_{ij} given dose d_j to patient i has normal distribution $\mathcal{N}(\mu_j, \sigma_j^2)$ where μ_j is considered as a random variable itself. Based on the posterior distributions of μ_j, BDCO is driven by the probability that μ_j is within ε of the target, γ_t:

$$\pi_j = \mathbb{P}\left(\gamma_t - \varepsilon \leq \mu_j \leq \gamma_t + \varepsilon\right). \tag{2.3.5}$$

The design targets the dose which maximizes the probability in Equation (2.3.5). Below, we apply the proposed benchmark to the setting considered in the original paper using this decision criterion and compare its performances to BDCO.

Recalling the setting by Wang and Ivanova (2015), we consider six scenarios with six dose levels $d_1, \ldots, d_6$, a sample size of $N = 36$, parameter $\varepsilon = 0.01$ and two cases: (i) the case of equal variances in which outcome Y_{ij} has normal distribution $\mathcal{N}(0.1j, 0.2^2)$ and (ii) the case of unequal variances corresponding to normal distribution $\mathcal{N}(0.1j, 0.1^2 j^2)$. In each of six scenarios the target values $\gamma_t = \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6\}$ are used, respectively. As a consequence, the target dose is dose d_1 in scenario 1, d_2 in scenario 2, and so on.

Table 2.8 shows the proportions of each dose selections corresponding to the BDCO against the respective benchmark. The results of the BDCO are extracted from Table 2 in the original article, and the benchmark is evaluated using $S = 10^6$ trial replications.

Under scenarios 2-5, the proportion of correct selections using the benchmark is 87%, which illustrates that they have the same level of "complexity". Conversely, the benchmark shows that it is easier to find the TD if it is either the first or the last dose. Under all scenarios with equal variances, the BDCO has the accuracy close to the benchmark. The ratio of the probability of correct selection of the BDCO relative to the benchmark ranges between 92% and 98% in these cases.

Table 2.8: Comparison of the BDCO design against the respective benchmark in six scenarios considered by Wang and Ivanova (2015).

Design	Variance	Proportion of selections					
		d_1	d_2	d_3	d_4	d_5	d_6
Scenario 1							
BDCO	Equal	**0.91**	0.10	0.00	0.00	0.00	0.00
Benchmark		**0.94**	0.06	0.00	0.00	0.00	0.00
BDCO	Unequal	**0.97**	0.03	0.00	0.00	0.00	0.00
Benchmark		**0.98**	0.02	0.00	0.00	0.00	0.00
Scenario 2							
BDCO	Equal	0.07	**0.86**	0.08	0.00	0.00	0.00
Benchmark		0.07	**0.87**	0.07	0.00	0.00	0.00
BDCO	Unequal	0.04	**0.84**	0.11	0.01	0.00	0.00
Benchmark		0.02	**0.86**	0.11	0.00	0.00	0.00
Scenario 3							
BDCO	Equal	0.00	0.07	**0.83**	0.09	0.00	0.00
Benchmark		0.00	0.07	**0.87**	0.07	0.00	0.00
BDCO	Unequal	0.00	0.16	**0.65**	0.16	0.02	0.00
Benchmark		0.00	0.11	**0.69**	0.17	0.02	0.00
Scenario 4							
BDCO	Equal	0.00	0.00	0.08	**0.81**	0.11	0.00
Benchmark		0.00	0.00	0.07	**0.87**	0.07	0.00
BDCO	Unequal	0.00	0.00	0.27	**0.50**	0.18	0.04
Benchmark		0.00	0.00	0.19	**0.55**	0.20	0.05
Scenario 5							
BDCO	Equal	0.00	0.00	0.00	0.09	**0.80**	0.11
Benchmark		0.00	0.00	0.00	0.07	**0.87**	0.07
BDCO	Unequal	0.00	0.00	0.02	0.34	**0.45**	0.20
Benchmark		0.00	0.00	0.00	0.25	**0.45**	0.29
Scenario 6							
BDCO	Equal	0.00	0.00	0.00	0.00	0.10	**0.90**
Benchmark		0.00	0.00	0.00	0.00	0.07	**0.93**
BDCO	Unequal	0.00	0.00	0.00	0.07	0.40	**0.54**
Benchmark		0.00	0.00	0.00	0.02	0.27	**0.71**

Under scenarios with unequal variances, the benchmark demonstrates that it is harder to find the TD if the corresponding variance is high. For example, the benchmark leads to 86% of correct selections under scenario 2 and 45% under scenario 5. Again, it appears that it is easier to find the TD if it is the first or the last dose for any methods. BDCO shows a high accuracy in scenario 1-5 with unequal variances. The correct probability ratios never go below 91% and even reach nearly 100% under scenario 5. In the former case, BDCO recommends the TD in 45% of replications (as well as the benchmark), but it recommends the highest dose d_6 systematically less often - 20% against 29% by the benchmark.

This implies that BDCO leads to more conservative decisions. Scenario 6 confirms this finding in which the correct probability ratio equals 76% which, however, is still high.

Overall, BDCO selects the correct dose uniformly less often than the benchmark in all scenarios (as expected), but the efficiency of the design is high. The minimum ratio of the probabilities of correct selections is 76% which corresponds to the scenario with highly variable outcomes. This indicates that parameters of the BDCO are adequately calibrated and the BDCO in the proposed form is able to find the MTD in many different scenarios.

2.3.4 Application to a Phase I/II Trial with Continuous Efficacy and Binary Toxicity

Similarly to a continuous toxicity outcome, the continuous efficacy endpoint can provide a better guidance on the TD selection than a dichotomized one. One of the first designs proposed for Phase I/II clinical trials considering continuous efficacy outcomes is by Bekele and Shen (2005) who developed a Bayesian approach to model toxicity and (continuous) biomarker of efficacy jointly. We denoted this design by BS.

Bekele and Shen (2005) introduced a latent normal random variable which is related to the observed binary toxicity. A bivariate normal distribution allows for different strengths of the dependence between toxicity and efficacy. Dose escalation/de-escalation decision rules are based on the posterior distributions of both toxicity and efficacy endpoints. The design was shown to have good operating characteristics in many scenarios. Therefore, the majority of subsequently proposed designs (e.g. see Hirakawa (2012) and Yeung et al. (2015)) were compared to it. Below, we provide the comparison of the design by Bekele and Shen (2005) against the respective benchmark.

Recalling the framework by Bekele and Shen (2005) we consider an efficacy outcome at dose d_j having Gamma distribution $\Gamma(\lambda_j\omega, \omega)$ with rate parameter $\omega = 0.1$ and a DLT outcome having probability p_j. A total of six scenarios and four dose levels per scenario are explored using the total sample size $N = 36$. The parameters of λ_j and toxicity probability p_j are given in Table 2.9. In each scenario a weak association, $\rho = 0.25$, between the toxicity endpoint and the efficacy biomarker is used. The TD is defined as the dose with the highest expected efficacy while being safe ($p_j < 0.35$) and efficacious ($\lambda_j > 5$).

Table 2.9 shows proportions of each dose selections of the BS design against the respective benchmark. The results for BS are extracted from Table 1 of the original work which uses 1000 replications, and the benchmark is evaluated using $S = 10^6$ trial replications.

Table 2.9: Comparison of the BS design against the respective benchmark in six scenarios considered by Bekele and Shen (2005).

Design	Proportion of selections				
	d_1	d_2	d_3	d_4	None
Scenario 1					
(λ_j, p_j)	(25,0.01)	(70,0.10)	**(115,0.25)**	(127,0.60)	
BS	0.00	0.03	**0.95**	0.02	0.00
Benchmark	0.00	0.08	**0.92**	0.00	0.00
Scenario 2					
(λ_j, p_j)	(5,0.50)	(70,0.70)	(90,0.80)	(135,0.85)	
BS	0.06	0.00	0.00	0.00	**0.94**
Benchmark	0.02	0.00	0.00	0.00	**0.98**
Scenario 3					
(λ_j, p_j)	(25,0.03)	(46,0.05)	(90,0.10)	(135,0.15)	
BS	0.00	0.00	0.02	**0.98**	0.00
Benchmark	0.00	0.00	0.01	**0.99**	0.00
Scenario 4					
(λ_j, p_j)	(20,0.05)	(75,0.05)	(75,0.35)	(75,0.65)	
BS	0.00	**0.83**	0.17	0.00	0.00
Benchmark	0.00	**1.00**	0.00	0.00	0.00
Scenario 5					
(λ_j, p_j)	(60,0.05)	(65,0.50)	(80,0.70)	(95,0.85)	
BS	**0.94**	0.06	0.00	0.00	0.00
Benchmark	**0.97**	0.03	0.00	0.00	0.00
Scenario 6					
(λ_j, p_j)	(2,0.03)	(2,0.03)	(2,0.03)	(2,0.03)	
BS	0.01	0.00	0.00	0.00	**0.99**
Benchmark	0.01	0.00	0.00	0.00	**0.99**

Under scenarios 1, 3 and 5 with increasing dose-efficacy relationships, the BS design performs with a high accuracy and the proportions of correct selections are close to the benchmark. Interestingly, the BS design recommends the TD d_3 3% more often than the benchmark under scenario 1. Given the number of replications for the BS and the benchmark, 3% difference is significant. This can be an indication that the prior distribution used by BS is in favour of d_3. It would also explain the relatively lower performance under scenario 4 in which the BS recommends the target dose d_2 in 83% of trials against 100% by the benchmark. The BS recommends the dose with the same efficacy, but noticeably greater toxicity in 17% of trials. An alternative explanation of the difference in proportions of selections under scenario 4 can be a plateau in the dose-efficacy relation that is not modelled by the BS. Nevertheless, the ratio of correct probabilities is 83% demonstrating good operating characteristics of the BS design.

Under unsafe scenario 2 and inefficacious scenario 6, the BS design comes to the correct conclusion in nearly the same proportions of trials as the benchmark. This shows the ability of the BS design to avoid unethical selections due to either high toxicity or low activity.

Overall, the benchmark confirmed that the BS design is flexible and can recommend the TD under many different scenarios. It also gives some possible clue to the super-efficient performance under scenario 1 and to potential challenges that the BS design can face in plateau dose-efficacy scenarios.

2.4 Discussion

In the first part of this chapter, the Bayesian model-based dose-escalation methods incorporating randomisation to the control arm into the Emax model were studied and compared to currently used model-based alternatives. The proposed approach shares the following features:

- The design can identify the ET between the toxicity probabilities at the MTC and the single agent more frequently than alternative approaches
- The randomisation step allows establishing whether effects observed for combinations are associated with the interaction of drugs or with one of the therapies in the combination only.
- The cost of randomisation to the control arm is presented by a small reduction in proportions of correct selections in some scenarios.
- The sensitivity analysis suggests that the proposed design is robust to reasonable choices of prior distributions and to allocation ratios between investigational and control arms.

Summarising, although the control arm is not routinely used in oncology early phase clinical trials, it is advocated to be included in combination trials of the considered types.

Importantly, the following limitation of the proposed design should be taken into account:

- As the classical Bayesian CRM design, the proposed design might have a large variance of number of patients on the MTC as indicated by Oron and Hoff (2013).

  Therefore, it is hard to guarantee the specific level of power in the frequentist hypothesis testing framework that can also be of the great usefulness. Consequently, the design that could maximise the power of such a test might be also of a practical interest.

- A risk of overfitting in trials with fewer combinations in the trials

  In the considered simulation study motivated by the actual trial, fitting the Emax model has not lead to overfitting as the number of combinations

is higher than the number of parameters. However, the considered Emax model requires four parameters to be estimated. In the context with fewer combinations studied in a trial, the proposed design can result in overfitting of the combination-toxicity curve.

In the second part of this chapter, we formulated the novel benchmark for dose finding studies. In essence, the novel benchmark is similar to the original proposal by O'Quigley et al. (2002) as the information about a patient is summarised in a single value, but can be also applied to studies with continuous outcomes. In the era of an increasing complexity of clinical trials, the procedure evaluating an adequacy of novel dose finding methods is crucial. It is shown that

- the proposed benchmark provides an accurate upper limit on the performance of model-based dose finding designs;
- the benchmark is able to reveal some inadequacy in the model/parameter/prior specifications or, alternatively, to confirm the design;
- the benchmark assesses the complexity of scenarios and can serve as a standardization of scenarios of various difficulty.

Therefore, it should be definitely recommended for the complete analysis of dose finding designs as it helps to evaluate them in a more comprehensive way.

Finally, it is important to mention that while the benchmark is a useful tool to assess the performance of any given dose finding methods, it does not capture all aspects of the evaluation. For instance, it does not provide information on a patient allocation distribution or a mean number of DLTs. Developments in this direction will be of great value for the complete design assessment.

Chapter 3

Loss Functions in Restricted Parameter Spaces

In many applications, for example, in clinical trials, when considering a problem of parameter estimation, the support of this parameter often is either naturally restricted (e.g. probability, variance, exponential distribution parameter) or an investigator can restrict it based on previously obtained knowledge (e.g. treatment effects on children given the data for adults). The knowledge about restricted space can carry important information that an experimenter has and can potentially improve the estimation (Mahmoudi and Zakerzadeh, 2011). There are also areas in which estimates on bounds of the restricted space are highly undesirable as they can lead to severe consequences. For instance, it can be worse to underestimate the potentiality of an event having life-threatening consequences than to overestimate it (Norstrom, 1996). An erroneously low estimated risk-level can lead to the absence of necessary initiative to reduce it.

Considering dose finding trials, once the estimate is obtained the decision to which dose the next patient should be allocated is commonly made based on the absolute

distance (2.1.4) or the squared distance between the estimate $\hat{p}_j$ and the target γ_t. These decision rules, however, do not take into account that both the probability estimate and the target γ_t lie in the restricted space (unit interval). We show why a simple logit transformation of these probabilities might not easily solve the problem. In this chapter, we consider a general Bayesian estimation problem in restricted parameter spaces and a problem of an appropriate loss function choice. Subsequently, we demonstrate how the proposed loss functions can be used as a criterion for allocation of patients in dose escalation trials.

3.1 Background

One of the ways to incorporate the information about a restricted space is to employ a Bayesian approach and to define a (uniform) prior distribution on the restricted space (Hartigan, 2004). However, once a posterior is obtained the squared error loss

$$L_q(\vartheta, d) = (\vartheta - d)^2 \tag{3.1.1}$$

where d is a decision the statistician has to take in order to approximate an unknown ϑ, called parameter, is often used to summarise a posterior distribution. The squared loss function (3.1.1) ignores the information about the restricted parameter space and is recognised to lead to suboptimal solutions (see e.g. Stein, 1964; Brown, 1968, for alternative loss functions for a scale parameter). Research on improving the Bayes estimator under the squared loss function (a posterior mean) for a scale parameter has subsequently attracted a great deal of attention in the literature (Arnold, 1970; Zidek, 1973; Brewster, 1974; Brewster and Zidek, 1974; Kubokawa, 1994).

To illustrate potential inconsistencies the squared error loss can lead to when estimating a parameter on a positive real line, consider two inferential procedures based on two independent experiments:

1. Estimate $\mu \in (-\infty, +\infty)$ given i.i.d. observations $X_i \sim \mathcal{N}(\mu, \sigma^2)$, where σ^2

is known;

2. Estimate $\sigma \in (0, +\infty)$ given i.i.d. observations $Y_i \sim \mathcal{N}(\mu, \sigma^2)$, where μ is known.

The unknown parameters of interest (denoted by ϑ in Equation (3.1.1)) are μ and γ, respectively. Using the squared error loss (3.1.1), the decision $\mu = 0$ in the first experiment and $\sigma = 0$ in the second are equally penalised, while this should not be the case. The claim of $\sigma = 0$ implies that the Y's are degenerate random variables, an extremely strong statement which should be penalised similarly to the decision $\mu = +\infty$ or $\sigma = +\infty$. The squared error loss function imposes an infinite penalty to a boundary decision in the first experiment and does not in the second one. While the decision $\sigma = 0$ is usually prevented by a proper choice of the prior, the squared loss function does not imply that it should be avoided and associated with a severe penalty. In contrast, an appropriate loss function imposes such a penalty and can be also used to prevent boundary decisions.

In addition, since the loss function (3.1.1) does not penalise boundary values, it was found to be unacceptable in many application areas: see Norstrom (1996); Soliman (2002) for examples in reliability analysis, Karimnezhad et al. (2014) in environmental sciences. To avoid boundary values of a scale parameter, Norstrom (1996) introduced the *precautionary loss function*

$$L_{sq}(\vartheta, d) = \frac{(d - \vartheta)^2}{d} \text{ where } \vartheta, d \in (0, +\infty) \tag{3.1.2}$$

which was used by many researchers (Kiapour and Nematollahi, 2011; Karimnezhad and Moradi, 2016).

The precautionary loss function covers the case of the scale parameter. There are, however, many applications in which the parameter of interest is restricted to an interval (a, b) and similar problems of severe consequences of boundary decisions can appear. The question of an appropriate loss function choice for a parameter ϑ

defined on the interval (a, b) has received less attentions in the statistical literature compared to a scale parameter. At the same time, its importance is acknowledged in many fields (see e.g. Aitchison, 1992; Pawlowsky-Glahn and Egozcue, 2001; Mateu-Figueras et al., 2013, for examples in compositional data analysis). Moreover, despite the variety of literature on families of loss functions for parameters restricted to an interval and on corresponding improved Bayes estimators (see e.g. Kubokawa, 1994; Marchand and Strawderman, 2005), they seem to be rarely applied in practice due to their complexity and a lack of closed-form solutions. The choice of loss functions for parameters defined on the interval and on the positive real line is yet under-represented area in the Bayesian literature and the usual mean still remains a common summary statistic.

The contribution of the first part of this chapter is twofold. Firstly, we provide a unified approach to define symmetry of a loss function when a parameter space is restricted to a particular open subset based on an appropriate definition of the distance. We outline that distances on their corresponding parameter spaces share the same property - an infinite penalisation of the bounds which is also known as the *balance property* (Mahmoudi and Zakerzadeh, 2011). We also recall two other desirable properties of loss functions: convexity and invariance. Secondly, we propose several loss functions which are as simple as the squared loss function (3.1.1), have explicit solutions for the corresponding Bayes estimator and include the information about the restricted parameter space in the corresponding Bayes estimator. In particular, we propose the scale invariant generalisation of the the precautionary loss function for a scale parameter $\vartheta \in (0, +\infty)$ and the interval squared loss function

$$L_{iq}(\vartheta, d) = \frac{(d - \vartheta)^2}{(d - a)(b - d)}$$

for the parameter $\vartheta \in (a, b)$. We show that the Bayes estimator corresponding to the interval squared loss function includes the Bayes estimator of the squared loss function (3.1.1) and of the precautionary loss function as limiting cases. It

is found that the interval squared and precautionary loss functions are both symmetric on corresponding parameter spaces and can be useful in application areas where conservative estimates are preferred. We generalise the approach for the multivariate parameter space and demonstrate how Bayes estimators obtained using the proposed loss functions behave in three classic problems of Bayesian estimation compared to standard approaches.

3.2 Scale Symmetry

3.2.1 A Historical Anecdote: Galileo on Scale Symmetry

In the Spring of 1627, a *peculiar controversy* [1] arose in one of Florence intellectual circles, where *noble gentlemen* used to entertain *erudite talks*:

Un cavallo, che vale veramente cento scudi, da uno è stimato mille scudi e da un altro dieci scudi: si domanda chi abbia di loro stimato meglio, e chi abbia fatto manco stravaganza nello stimare.

The problem translates into: "A horse, whose true worth is one hundred *scudi*[2], is estimated by someone to be one thousand *scudi* and by someone else to be ten *scudi*: the question is, who gave a better estimate, and who instead gave a more extravagant estimate?". It is formulated in a letter from Andrea Gerini to Nozzolini, an *erudite priest*. Gerini wanted Nozzolini's opinion on a sentence by Galilei (1627), according to whom

. . . li due stimatori abbiano egualmente esorbitato e commesse eguali stravaganze nello stimare l'uno mille e l'altro dieci quello che realmente val cento,

which translates to: "The two estimators have been equally exorbitant and are

[1] the italic is a translation of a commentary to Galilei (1627) appearing in the edition of Galilei's works mentioned in the bibliography, from which all of the quotes are taken, following Scàrdovi (1980).

[2] a monetary unit, literally, a shield

responsible for an equal extravagance by estimating, one thousand the former and ten the latter, what is really worth one hundred".

In the intense correspondence following the initial letters, Nozzolini argues that the estimates should be evaluated according to the *arithmetic proportion*, whereas Galileo insists that the correct method of judging is by the *geometric proportion*. The crux of the problem is that the estimand is a positive quantity, for which the *geometric proportion* seems more appropriate, as wittingly argued by Galileo in another letter:

Se uno stimasse alta dugento braccia una torre, che veramente fusse alta cento, con quale esorbitanza nel meno pareggerà il signor Nozzolini l'altra nel più ?

which translates as: "If one were to overestimate a one-hundred arm high tower as two-hundred arm high, what underestimate would Nozzolini consider as equally deviating?"

3.2.2 Scale Symmetry, Convexity and Scale Invariance

Let us start with the following definition for the parameter define on the whole real line.

Definition 1. *A loss function $L(\vartheta, d)$ is symmetric if, for every d_1, d_2 and $\vartheta \in \mathbb{R}^1$*

$$(\vartheta - d_1)^2 = (d_2 - \vartheta)^2 \tag{3.2.1}$$

implies $L(\vartheta, d_1) = L(\vartheta, d_2)$.

Definition 1 implies that two decisions defined on the real line should be equally penalised by a symmetric loss function $L(\cdot, \vartheta)$ if they stand on the same squared distance from ϑ. Note that for $d_1 < d_2$ Equation (3.2.1) can be rewritten as $\vartheta = \frac{d_1+d_2}{2}$. It follows that if ϑ is the *arithmetic mean* of d_1 and d_2 then these deci-

sions should be equally penalised. Clearly, the squared error loss of equation (3.1.1) is symmetric on the real line by definition.

Then, Galilei's claim of *eguali stravaganze* for a positive parameter ϑ can be expressed in modern terminology as the requirement of the *scale symmetric* loss function, as in the following definition.

Definition 2. *A loss function $L(\vartheta, d)$ is scale symmetric if, for every d_1, d_2 and $\vartheta \in \mathbb{R}_+$*

$$\frac{d_1}{\vartheta} = \frac{\vartheta}{d_2} \tag{3.2.2}$$

implies $L(\vartheta, d_1) = L(\vartheta, d_2)$.

Note that Equation (3.2.2) can be rewritten as

$$\vartheta = \sqrt{d_1 d_2}$$

or

$$\log(\vartheta) = \frac{\log d_1 + \log d_2}{2}.$$

In other words, if ϑ is the *geometric mean* of d_1 and d_2, then these decisions should be equally penalised by a scale symmetric loss function. As in Definition 1 of symmetry for the real line, two decisions are symmetric if the parameter ϑ is their appropriate mean – geometric in this case as opposed to arithmetic. This fact will be used for our proposal of the symmetry definition on the interval.

The definition of the squared distance on the positive real line, $\mathbb{R}_+$ (Mateu-Figueras et al., 2013)

$$\mathcal{D}_+(\vartheta, d) = (\log \vartheta - \log d)^2 \tag{3.2.3}$$

is known in Statistics as Brown's loss function (Brown, 1968). Its motivation is to rescale the positive real line to the whole real line via the log transformation and to use the squared error loss function. Here, the logarithm function is a natural

choice for a positive random variable.

Note that

$$\mathcal{D}_+(\vartheta, d_1) = \mathcal{D}_+(\vartheta, d_2) \tag{3.2.4}$$

implies either $d_1 = d_2$ or Equation (3.2.2). Therefore, we can also restate Definition 2 in terms of (3.2.4).

The Euclidean distance on the real line and $\mathcal{D}_+$ on the positive real line share the same property - they infinitely penalise boundary values on corresponding parameter space. In case of $\vartheta \in \mathbb{R}$, the squared distance $L_q(\vartheta, d)$ takes an infinite value when $d = \pm\infty$. For similar reasons, we require that an appropriate loss function for a scale parameter should go to infinity as the decision approaches the natural boundaries of the parameter space, to reproduce the behaviour at $\pm\infty$ of the squared error loss function. A loss function with this property are also called *balanced* according to Mahmoudi and Zakerzadeh (2011).

We recall one more property of loss functions for a parameter on the positive real line - the scale invariance.

Definition 3. *Loss function $L(\vartheta, d)$ is scale invariant if for every $c > 0$ and every pair (ϑ, d).*

$$L(\vartheta, d) = L(c\vartheta, cd)$$

It follows that the squared loss function (3.1.1) is not scale symmetric, is not scale invariant and does not penalise all boundaries for a scale parameter.

3.2.3 Symmetric Loss Functions on the Positive Real Line

The inadequacy of difference-based loss functions, like the squared error loss, for estimating certain positive quantities has often been recognized (James and Stein, 1961; Stein, 1964). Several alternative loss functions have been proposed, the

best-known being the normalized squared loss function proposed by Stein (1964)

$$L_{nq}(\vartheta, d) = \left(\frac{d}{\vartheta} - 1\right)^2, \tag{3.2.5}$$

Stein's loss (or an entropy loss function)

$$L_S(\vartheta, d) = \frac{d}{\vartheta} - 1 - \log\left(\frac{d}{\vartheta}\right) \tag{3.2.6}$$

and Brown's loss function (Brown, 1968) itself, $\mathcal{D}_+(\vartheta, d)$ given in Equation (3.2.3). One can check that all of functions above are scale invariant, but only Brown's loss function is scale symmetric and infinitely penalises the boundary decisions. Unfortunately, Brown's loss function is not *convex*, a feature of loss functions which is often required to represent risk aversion and for the sake of regularising the associated minimisation problems. An unpleasant consequences is that, the Bayes estimator associated with Brown's loss function is usually difficult to compute. Below we propose simple alternative loss functions which share the desirable properties of a loss function on the positive line and have explicit Bayes estimators.

We propose a family of loss functions defined for $k > 0$ as

$$L_k(\vartheta, d) = \left(\frac{d}{\vartheta}\right)^k + \left(\frac{\vartheta}{d}\right)^k - 2 \tag{3.2.7}$$

which are scale symmetric, scale invariant, convex, and which tend to infinity at the boundaries. In this chapter, we focus on the case $k = 1$

$$L_1(\vartheta, d) = \frac{(d - \vartheta)^2}{\vartheta d} \tag{3.2.8}$$

which can be considered as a modification of the squared error loss function. The numerator is again the squared distance, but the denominator guarantees the infinite penalisation for $d = 0$. It is easy to see that the loss function (3.2.8) is a scale invariant version of the precautionary loss function (3.1.2). The graphs

of loss functions $L_1(\vartheta, d)$, $L_{sq}(\vartheta, d)$ and $\mathcal{D}_+(\vartheta, d)$ as functions of d for fixed $\vartheta = 0.4$ are given in Figure 3.1.

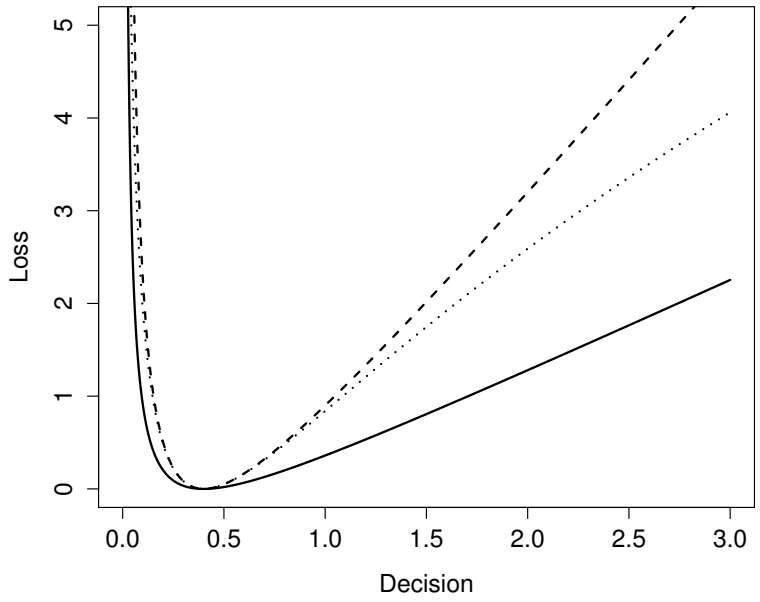

Figure 3.1: Scale symmetric loss functions L_1 (dashed), L_{sq} (solid) and $\mathcal{D}_+$ (dotted) for $d \in (0, 3)$ and $\vartheta = 0.4$.

Note that loss functions of similar-looking forms which have been considered in the literature, for example, the normalized squared loss L_{nq} (3.2.5) and the relative squared error loss (Savchuk and Tsokos, 2011) $L_{rq}(\theta, d) = \frac{(d-\vartheta)^2}{\vartheta}$, do not tend to infinity at the boundary and are not scale symmetric.

3.2.4 Scale Means and Scale Variances

Within the Bayesian approach, ϑ is a random parameter with a distribution which conveys the uncertainty in the given state. In such a case, a point summary of the distribution of ϑ minimising the risk (i.e. the expected loss) associated with the given loss function is often required. Such a minimiser of an expected loss is usually called a Bayes estimator. When a scale symmetric loss function is used, we propose to call such minimisers *scale means.* In case of convex loss functions, such as the novel loss function given in Equation (3.2.7) and the loss function in Equation (3.2.8) , the minimisation can be performed explicitly.

Theorem 1. *Let ϑ be a positive random variable with a posterior density function f and such that $\mathbb{E}(\vartheta^k) < \infty$ and $\mathbb{E}(\vartheta^{-k}) < \infty$, where $\mathbb{E}$ denotes the posterior mean with respect to f and $k > 0$. Then,*

(a) *Expectation of the loss function $L_k(\vartheta, d)$ in Equation (3.2.7) with respect to f is minimised by the Bayes estimator (scale mean)*

$$\hat{d}_k = \left(\frac{\mathbb{E}(\vartheta^k)}{\mathbb{E}(\vartheta^{-k})} \right)^{\frac{1}{2k}}. \tag{3.2.9}$$

In particular, the expectation of loss function $L_1(\vartheta, d)$ in Equation (3.2.8) is minimised by

$$\hat{d}_1 \;=\; \sqrt{\mathbb{E}(\vartheta)/\mathbb{E}(\vartheta^{-1})}.$$

(b) *Expectation of the precautionary loss function (3.1.2) is minimised by the Bayes estimator (scale mean)*

$$\hat{d}_{sq} = \sqrt{\mathbb{E}(\vartheta^2)}, \tag{3.2.10}$$

for which the following bound holds: $\hat{d}_{sq} \geq \mathbb{E}(\vartheta)$.

Proof. **(a)** The expectation of the loss function (3.2.7) with respect to the posterior density function f takes the form

$$\mathbb{E}(L_k(\vartheta, d)) = \mathbb{E}\left(\frac{d}{\vartheta}\right)^k + \mathbb{E}\left(\frac{\vartheta}{d}\right)^k - 2 = d^k \mathbb{E}\left(\vartheta^{-k}\right) + d^{k-1}\mathbb{E}\left(\vartheta^k\right) - 2.$$

Then, the decision d minimising the expected loss function is found solving

$$\frac{\partial \mathbb{E}(L_k(\vartheta, d))}{\partial d} = k d^{k-1}\mathbb{E}\left(\vartheta^{-k}\right) - k d^{-k-1}\mathbb{E}\left(\vartheta^k\right) = 0$$

This results in

$$\hat{d}_k = \left(\frac{\mathbb{E}(\vartheta^k)}{\mathbb{E}(\vartheta^{-k})} \right)^{\frac{1}{2k}},$$

and in the special case of $k = 1$, $\hat{d}_1 = \sqrt{\mathbb{E}(\vartheta)/\mathbb{E}(\vartheta^{-1})}$.

(b) Similarly to the previous point, the expectation of the precautionary loss function (3.1.2) with respect to the posterior density function f taken the form

$$\mathbb{E}(L_{sq}(\vartheta, d)) = \mathbb{E}\left(\frac{(d-\vartheta)^2}{d} \right) = \frac{d^2 - 2d\mathbb{E}(\vartheta) + \mathbb{E}(\vartheta^2)}{d}.$$

Then, the decision d minimising the expected loss function is found solving

$$\frac{\partial \mathbb{E}(L_{sq}(\vartheta, d))}{\partial d} = 0.$$

This results in $\hat{d}_{sq} = \sqrt{\mathbb{E}(\vartheta^2)}$. Using the Jensen inequality (Jensen, 1906) for ϑ^2 one can obtain that

$$\mathbb{E}\left(\vartheta^2\right) \geq \mathbb{E}^2\left(\vartheta\right).$$

Applying the squared root to the both sides of the inequality the result immediately follows.

□

In a more fundamental Bayesian approach, a Bayes estimator is regarded only as a convenient summary of the posterior, and a loss function as a way to prescribe what kind of summary is appropriate. Typically, the posterior expectation is used as the Bayes estimator, implying that the squared loss function is being used. A second step is usually taken to accompany the Bayes estimator with a measure of uncertainty of the posterior distribution. If the posterior mean is used, the posterior variance is often presented. However, if a scale symmetric loss function is considered to be appropriate then it is also reasonable to present the achieved minimum of the posterior expected loss as the second summary of the posterior

distribution. For the given loss functions (3.2.7) and (3.1.2), particularly simple expected posterior losses can be obtained. For the loss function (3.2.7), the *scale variance* of order k of the random variable ϑ in Theorem 1(a) can be written as

$$\varrho_k(\vartheta) := 2\sqrt{\mathbb{E}(\vartheta^k)\mathbb{E}(\vartheta^{-k})} - 2,$$

whereas the scale variance for the precautionary loss function (3.1.2) in Theorem 1(b) is

$$\varrho(\vartheta) = 2\left(\sqrt{\mathbb{E}(\vartheta^2)} - \mathbb{E}(\vartheta)\right).$$

3.3 Interval Symmetry

The approach used above for a positive parameter can be generalised to the parameter defined on the interval (a, b). The issue of a restricted parameter space is not usually discussed in the choice of the loss function and corresponding Bayes estimator: bounds are taken into account through the prior specification only, then the squared loss function and posterior mean (the corresponding Bayes estimator) are used. Such solutions can be suboptimal if boundary decisions are to be avoided. Below, we define the property of the symmetry on an interval and show that the novel definition generalises the cases of parameters on the whole real line and on the positive real line. We provide the loss function with desirable properties which is, again, a generalisation of the squared loss function and the precautionary loss function.

3.3.1 Symmetric Loss Functions on Interval

Let us consider an inferential problem for which the parameter of interest lies in the interval (a, b). Define the transformation

$$\operatorname{logit}_{(a,b)}(x) = \log\frac{x-a}{b-x} \tag{3.3.1}$$

where $a < x < b$. Note that, for $a = 0$ and $b = 1$, transformation (3.3.1) reduces to the logit transformation. Following the same line of reasoning as in Section 3.2, we introduce the following definition of the *symmetric on* (a, b) loss function (or simply the *interval symmetric* loss function).

Definition 4. *A loss function* $L(\vartheta, d)$ *is symmetric on the interval* (a, b) *if, for every choice of* $d_1, d_2 \in (a, b)$ *and* $\vartheta \in (a, b)$

$$\text{logit}_{(a,b)}(\vartheta) = \frac{\text{logit}_{(a,b)}(d_1) + \text{logit}_{(a,b)}(d_2)}{2}. \tag{3.3.2}$$

implies $L(\vartheta, d_1) = L(\vartheta, d_2)$.

In other words, two decisions d_1 and d_2 should be penalised equally if the arithmetic mean of their logit transformation is equal to the logit transformation of ϑ. Lemma 1 justifies the use of the logit transformation (3.3.1).

Lemma 1. *Definition 4 is equivalent to Definition 1 when* $a \to -\infty$ *and* $b \to +\infty$ *and equivalent to Definition 2 when* $a \to 0$ *and* $b \to +\infty$.

Proof. Condition (3.3.2) for $a < d_1 < d_2 < b$ can be rewritten

$$\vartheta = f(a, b, d_1, d_2) \equiv \frac{ab - d_1 d_2 + \sqrt{(d_1 - a)(b - d_1)(d_2 - a)(b - d_2)}}{a + b - d_1 - d_2}.$$

Obviously, ϑ is a symmetric function of d_1 and d_2. Considering two limits

$$\lim_{a \to -\infty,\ b \to +\infty} f(a, b, d_1, d_2) = \frac{d_1 + d_2}{2}$$

and

$$\lim_{a \to 0,\ b \to +\infty} f(a, b, d_1, d_2) = \sqrt{d_1 d_2},$$

it can be seen that the definitions are equivalent. □

It follows from Lemma 1 that Definition 4 is a convenient generalisation of the definition of the symmetry.

3.3.2 An Interval Symmetric Loss Function and Its Minimiser

As in the case of a positive parameter and scale symmetric loss functions, the approach by Brown (1968) of specifying the squared loss function after rescaling the interval (a, b) to the real line via, for example, the logit transformation (3.3.1) provides the loss function

$$L_{iB}(\vartheta, d) = \left(\text{logit}_{(a,b)}(d) - \text{logit}_{(a,b)}(\vartheta)\right)^2. \tag{3.3.3}$$

On the unit interval this loss function is equivalent to so-called Aitchison distance proposed by Aitchison (1992) for parameters defined on a simplex. However, the loss function (3.3.3) is not convex and its minimisation problem does not have an explicit solution. As an alternative, we propose the following loss function

$$L_{iq}(\vartheta, d) = \frac{(d-\vartheta)^2}{(d-a)(b-d)}. \tag{3.3.4}$$

which is interval symmetric and tends to infinity when the decision d tends to bounds a and b. In the special case $a = 0$ and $b = 1$ the loss function (3.3.4) takes the form

$$L_{iq}(\vartheta, d) = \frac{(d-\vartheta)^2}{d(1-d)}. \tag{3.3.5}$$

The graph of the corresponding loss function for the fixed $\vartheta = 0.5$ and for different values of d is given in Figure 3.2.

Note that the loss function (3.3.5) for $a = 0$ and $b = 1$ looks similar to the well-known loss function (Ferguson, 2014)

$$(d-\vartheta)^2/(\vartheta(1-\vartheta))$$

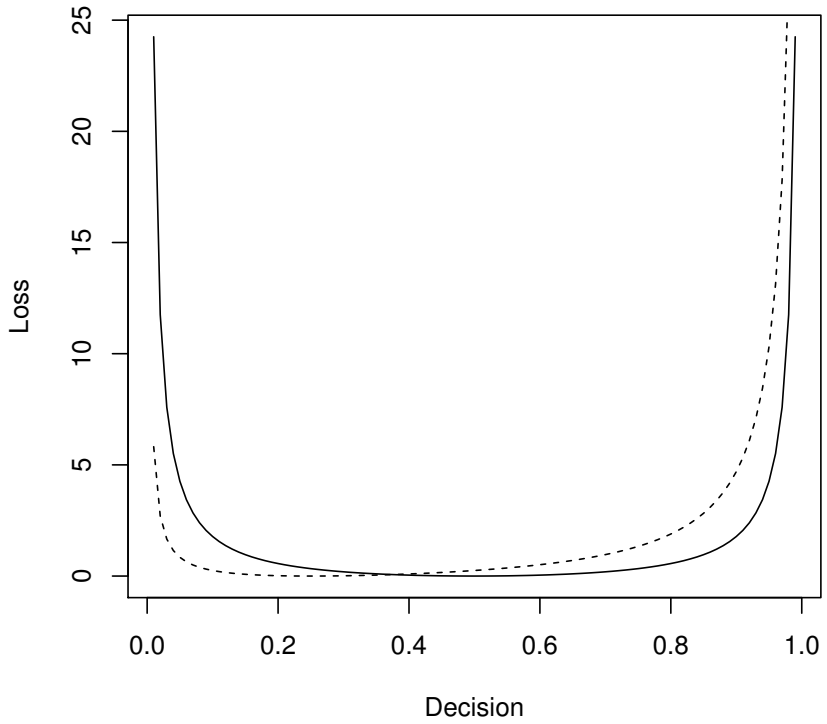

Figure 3.2: An example of the interval symmetric loss function (3.3.4) for $a = 0$, $b = 1$, $\vartheta = .5$ (solid line) and $\vartheta = .25$ (dashed line) and different values of $d \in (0, 1)$.

which, however, does not have the desirable property of assigning infinite penalties to boundary decisions.

The decision which minimises the interval symmetric loss function (3.3.4) is given explicitly in Theorem 2.

Theorem 2. *Let $\vartheta \in (a, b)$ be a random variable with a posterior density function f and $\mathbb{E}(\vartheta^2) < \infty$ where $\mathbb{E}(\cdot)$ denotes the expectation with respect to f. Then,*

***(a)** Expectation of the interval symmetric loss function L_{iq} in Equation (3.3.4) with respect to f is minimised by the Bayes estimator*

$$\hat{d}_{iq} = \frac{ab - \mathbb{E}(\vartheta^2) + \sqrt{(\mathbb{E}(\vartheta^2) - ab)^2 - (a + b - 2\mathbb{E}(\vartheta))(2ab\mathbb{E}(\vartheta) - (a + b)\mathbb{E}(\vartheta^2))}}{a + b - 2\mathbb{E}(\vartheta)} \tag{3.3.6}$$

(b) Estimator (3.3.6) minimises the expectation of squared loss function (3.1.1) in the limiting case $a \to -\infty$ and $b \to +\infty$, and the expectation of precautionary loss

function L_{sq} (3.1.2) in the limiting case $a \to 0$ and $b \to +\infty$.

Proof. (a) The expectation of the loss function L_{iq} in Equation (3.3.4) with respect to the posterior probability density function takes the form

$$\mathbb{E}\left(L_{iq}(\vartheta, d) = \frac{d^2 - 2d\mathbb{E}(\vartheta) + \mathbb{E}(\vartheta^2)}{(d-a)(b-d)}.\right.$$

The Bayes estimator minimising the expected loss function is found solving

$$\frac{\partial \mathbb{E}\left(L_{iq}(\vartheta, d)\right)}{\partial d} = 0,$$

which takes the form

$$\frac{b(d^2 - \mathbb{E}(\vartheta^2)) + a\left(d^2 - 2b(\mathbb{E}(\vartheta) - d) - \mathbb{E}(\vartheta^2)\right) + 2d\left(\mathbb{E}(\vartheta^2) - d\mathbb{E}(\vartheta)\right)}{(d-a)^2(b-d)^2} = 0.$$

Solving this equation results in the Bayes estimator given in Equation (3.3.6).

(b) Taking the limits

$$\lim_{a \to -\infty,\ b \to +\infty} \hat{d}_{iq}(a, b, \vartheta) = \mathbb{E}(\vartheta)$$

and

$$\lim_{a \to 0,\ b \to +\infty} \hat{d}_{iq}(a, b, \vartheta) = \sqrt{\mathbb{E}(\vartheta^2)},$$

it is easy to see that the obtained estimators are equivalent to the minimisers of the squared loss function L_q and of the precautionary loss function L_{sq}, respectively. Note that $d_{iq} \to \frac{a+b}{2}$ as $\mathbb{E}(\vartheta) \to \frac{a+b}{2}$. □

3.3.3 Multivariate Generalisations

The definition of symmetry can be generalised to cases of parameters belonging to subsets of $\mathbb{R}^m$ by applying the same ideas to different restricted parameter spaces as in Definition 5.

Definition 5. *Let* $\boldsymbol{\vartheta} = (\vartheta^{(1)}, \vartheta^{(2)}, \ldots, \vartheta^{(m)})^{\mathrm{T}}$ *be a parameter lying in one of the parameter spaces* $\vartheta \subset \mathbb{R}^m$ *listed below. Let* $\boldsymbol{d_i} = (d_i^{(1)}, d_i^{(2)}, \ldots, d_i^{(m)})^{\mathrm{T}}$, $i = 1, 2$ *be two vectors of decisions defined on the same parameter space. A loss function* $L(\boldsymbol{\vartheta}, \mathbf{d})$ *is a multivariate* $\vartheta-$*symmetric if the equality*

$$L(\boldsymbol{\vartheta}, \mathbf{d_1}) = L(\boldsymbol{\vartheta}, \mathbf{d_2})$$

is implied by each triple $\boldsymbol{\vartheta}, \mathbf{d_1}, \mathbf{d_2} \in \vartheta$ *satisfying the following respective equality of distances:*

(a) when $\vartheta = \mathbb{R}^m$ *(symmetry on* $\mathbb{R}^m$ *itself):*

$$\sqrt{\sum_{j=1}^{m} \left(d_1^{(j)} - \vartheta^{(j)}\right)^2} = \sqrt{\sum_{j=1}^{m} \left(d_2^{(j)} - \vartheta^{(j)}\right)^2};$$

(b) when $\vartheta = \mathbb{R}_+^m = \{\vartheta : \vartheta^{(i)} > 0, i = 1, \ldots, m\}$ *(scale symmetry on* $\mathbb{R}_+^m$*):*

$$\sqrt{\sum_{j=1}^{m} \log^2 \left(\frac{d_1^{(j)}}{\vartheta^{(j)}}\right)} = \sqrt{\sum_{j=1}^{m} \log^2 \left(\frac{d_2^{(j)}}{\vartheta_j}\right)};$$

(c) when $\vartheta = \{\vartheta : (a_1 < \vartheta^{(1)} < b_1), \ldots, (a_m < \vartheta^{(m)} < b_m)\}$ *(symmetry on an* $\mathbb{R}^m$*-rectangle):*

$$\sqrt{\sum_{j=1}^{m} \left(\mathrm{logit}_{(a_j,b_j)} d_1^{(j)} - \mathrm{logit}_{(a_j,b_j)} \vartheta^{(j)}\right)^2} =$$

$$\sqrt{\sum_{j=1}^{m} \left(\mathrm{logit}_{(a_j,b_j)} d_2^{(j)} - \mathrm{logit}_{(a_j,b_j)} \vartheta^{(j)}\right)^2};$$

(d) when $\vartheta = \{\vartheta : \vartheta^{(1)} > 0,\ \vartheta^{(2)} > 0,\ \ldots, \vartheta^{(m)} > 0; \sum_{i=1}^{m} \vartheta^{(i)} = 1\}$ *(symmetry*

on the unit simplex $\mathcal{S}^m$):

$$\sqrt{\frac{1}{m}\sum_{i<j}\left(\log\frac{d_1^{(i)}}{d_1^{(j)}}-\log\frac{\vartheta^{(i)}}{\vartheta^{(j)}}\right)^2}=\sqrt{\frac{1}{m}\sum_{i<j}\left(\log\frac{d_2^{(i)}}{d_2^{(j)}}-\log\frac{\vartheta^{(i)}}{\vartheta^{(j)}}\right)^2}.$$

The definition of the symmetric loss function in each case employs a distance corresponding to the particular restricted space. While the distances in (a)-(c) are natural extensions of the previously used, the definition in (d) is less straightforward. Definition 5(d) uses the Aitchison distance again as it is argued to have desirable properties of the metric on the simplex $\mathcal{S}^m$ (Billheimer et al., 2001). Following Definition 5, Lemma 2 similar to Lemma 1 holds.

Lemma 2. *Let $\boldsymbol{\vartheta}=(\vartheta^{(1)},\vartheta^{(2)},\ldots,\vartheta^{(m)})^{\mathrm{T}}$ be a vector of parameter of interest such that $\vartheta^{(1)}\in(a_1,b_1)$, $\vartheta^{(2)}\in(a_2,b_2),\ldots,\vartheta^{(m)}\in(a_m,b_m)$ and denote a vector of corresponding decisions lying in corresponding intervals by $\boldsymbol{d_i}=(d_i^{(1)},d_i^{(2)},\ldots,d_i^{(m)})^{\mathrm{T}}$. Then, Definition 5(c) is equivalent to Definition 5(a) when $a_i\to-\infty$ and $b_i\to\infty$ for all $i=1,\ldots,m$ and to Definition 5(b) when $a_i=0$ and $b_i\to\infty$ for all $i=1,\ldots,m$.*

Following Brown (1968), the distances in Definition 5 could be taken as corresponding symmetric loss functions. For example, in case *(b)* in Definition 5 one can define

$$\mathcal{D}_+^{(m)}(\boldsymbol{\vartheta},\boldsymbol{d})=\sum_{j=1}^{m}\log^2\left(\frac{d^{(j)}}{\vartheta^{(j)}}\right). \tag{3.3.7}$$

At the same time, some convex alternatives could be considered when leading to simple solutions of minimisation problems. However, the search of the symmetric multivariate generalisation of the proposed loss functions L_k, L_{sq} and L_{iq} seems to be non-trivial. We propose the following loss functions for parameters with non-negative components.

Proposition 1. *Let $\boldsymbol{\vartheta}=(\vartheta^{(1)},\vartheta^{(2)},\ldots,\vartheta^{(m)})^{\mathrm{T}}\in\mathbb{R}_+^m$ be a vector of unknown parameter and $\boldsymbol{d}=(d^{(1)},d^{(2)},\ldots,d^{(m)})^{\mathrm{T}}\in\mathbb{R}_+^m$ be a vector of decisions. Loss*

functions

$$L_1^{(m)}(\boldsymbol{\vartheta}, \boldsymbol{d}) = \sum_{j=1}^{m} \frac{(d^{(j)} - \vartheta^{(j)})^2}{d^{(j)}\vartheta^{(j)}} \tag{3.3.8}$$

$$L_{sq}^{(m)}(\boldsymbol{\vartheta}, \boldsymbol{d}) = \sum_{j=1}^{m} \frac{(d^{(j)} - \vartheta^{(j)})^2}{d^{(j)}} \tag{3.3.9}$$

are additive multivariate generalisations of the loss function (3.2.7) and the loss function (3.1.2),respectively, which infinitely penalise the boundary decisions.

Clearly, loss functions $L_1^{(m)}$ and $L_{sq}^{(m)}$ penalise the boundaries infinitely and a desirable performance of the corresponding estimators can be expected. However, the property of symmetry is not satisfied. One can find two decisions $\tilde{\boldsymbol{d}}_{\boldsymbol{1}}$ and $\tilde{\boldsymbol{d}}_{\boldsymbol{2}}$ for which $\mathcal{D}_+^{(m)}\left(\boldsymbol{\vartheta}, \tilde{\boldsymbol{d}}_{\boldsymbol{1}}\right) = \mathcal{D}_+^{(m)}\left(\boldsymbol{\vartheta}, \tilde{\boldsymbol{d}}_{\boldsymbol{2}}\right)$, but $L_1^{(m)}\left(\boldsymbol{\vartheta}, \tilde{\boldsymbol{d}}_{\boldsymbol{1}}\right) \neq L_1^{(m)}\left(\boldsymbol{\vartheta}, \tilde{\boldsymbol{d}}_{\boldsymbol{2}}\right)$. For instance, $\tilde{\boldsymbol{d}}_{\boldsymbol{1}} = (\frac{10471}{1280}, 2)^{\mathrm{T}}$, $\tilde{\boldsymbol{d}}_{\boldsymbol{2}} = (6, 6)^{\mathrm{T}}$ and $\boldsymbol{\vartheta} = (1, 2)^{\mathrm{T}}$. Then $\mathcal{D}_+^{(2)}\left(\boldsymbol{\vartheta}, \tilde{\boldsymbol{d}}_{\boldsymbol{1}}\right) = \mathcal{D}_+^{(2)}\left(\boldsymbol{\vartheta}, \tilde{\boldsymbol{d}}_{\boldsymbol{2}}\right) \approx 2.10$, but $L_1^{(2)}\left(\boldsymbol{\vartheta}, \tilde{\boldsymbol{d}}_{\boldsymbol{1}}\right) \approx 6.30$, $L_1^{(2)}\left(\boldsymbol{\vartheta}, \tilde{\boldsymbol{d}}_{\boldsymbol{2}}\right) = 5.5$. The same is true for $L_{sq}^{(m)}$. Even if the whole loss functions are not symmetric, they are "component-wise" symmetric as shown above. The comparison of loss functions $L_1^{(2)}$, $L_{sq}^{(2)}$ and $\mathcal{D}_+^{(2)}$ for different values of decision $\boldsymbol{d}_{\boldsymbol{1}} = \left(d_1^{(1)}, d_1^{(2)}\right)^{\mathrm{T}}$ and $\boldsymbol{d}_{\boldsymbol{2}} = \left(d_2^{(1)}, d_2^{(2)}\right)^{\mathrm{T}}$ and fixed $\boldsymbol{\vartheta} = (1, 2)^{\mathrm{T}}$ is given in Figure 3.3. The proposed loss functions demonstrate

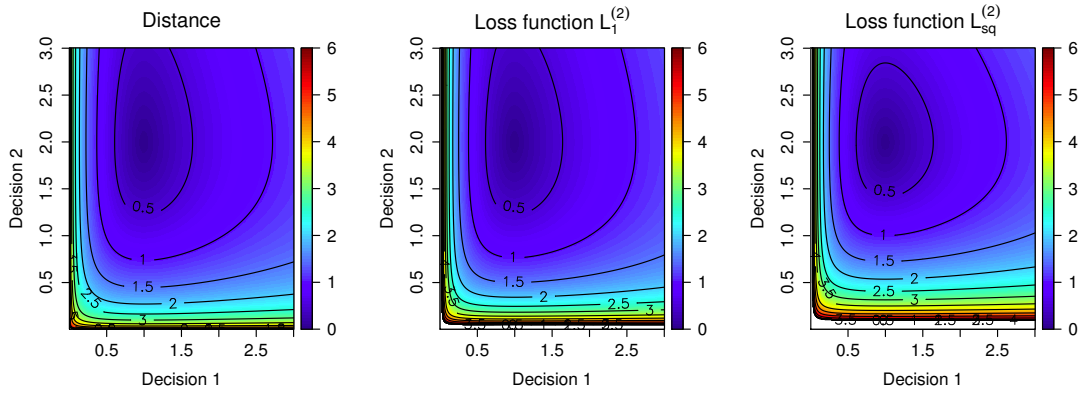

Figure 3.3: Contour plots of loss functions $\mathcal{D}_+^{(2)}$, $L_1^{(2)}$, $L_{sq}^{(2)}$ for the case $m = 2$ and $\boldsymbol{\vartheta} = (1, 2)^{\mathrm{T}}$.

the behaviour similar to the distance $\mathcal{D}_+^{(2)}$, and also have some favourable properties, like convexity.

3.4 Examples

Below, the performance of the Bayes estimators corresponding to the proposed loss functions in three classic examples is demonstrated. The mean squared error (MSE) is chosen to compare the different estimators on common grounds. Note that, in general, it is not favourable to the new proposals, since the MSE is derived from the squared error loss.

3.4.1 Estimation of a Probability

An important example of a parameter defined on the unit interval is a probability. In the presence of a binary random sample with an unknown probability of success, Beta prior distribution $\mathcal{B}(1,1)$ is often assumed. Having observed x successes out of n trials, the posterior distribution is conjugate Beta distribution with the posterior density function

$$f_n(p) = \frac{p^x(1-p)^{n-x}}{\mathbb{B}(x+1, n-x+1)}$$

where $\mathbb{B}(\cdot,\cdot)$ is Euler's Beta function. The estimator corresponding to the squared error loss function (posterior mean) has the form

$$\hat{p}_q = \frac{x+1}{n+2}. \tag{3.4.1}$$

Another widely used estimator is the so-called "add two successes and two failures" Agresti-Coull estimator (Agresti and Coull, 1998)

$$\hat{p}_{AC} = \frac{x+2}{n+4}.$$

Below we compare these approaches to the newly proposed estimator (3.3.6).

The symmetric optimal Bayes estimator (3.3.6) in the case $a = 0$ and $b = 1$ can be written as

$$\hat{d} = \frac{\mathbb{E}(\vartheta^2) - \sqrt{\mathbb{E}(\vartheta^2)(1 - 2\mathbb{E}(\vartheta) + \mathbb{E}(\vartheta^2))}}{2\mathbb{E}(\vartheta) - 1}$$

where ϑ is an unknown success probability, over which a posterior distribution is given. It is assumed that the extremes of the interval are not possible values for the parameter. The first and second moments of Beta distribution can be computed explicitly and plugged in formula (3.3.6) to obtain the interval symmetric optimal Bayes estimator

$$\hat{p}_{iq} = \left(1 + \sqrt{\frac{(n-x+1)(n-x+2)}{(x+1)(x+2)}}\right)^{-1}. \tag{3.4.2}$$

For the comparison, simulated trials with sample size $N = 15$ are considered. On a grid of values $\vartheta \in (0, 1)$, data from a binomial distribution with $N = 15$, and $p = \vartheta$ were simulated. For each value of ϑ, 10^4 simulated data sets were used. In each simulation the three estimators are computed; for each value of ϑ, the MSE is computed as

$$MSE_k \equiv \frac{1}{S}\sum_{s=1}^{S}(\hat{p}_k^{(s)} - \vartheta)^2,$$

where $\hat{p}_k^{(s)}$ corresponds to s^{th} simulation and $k = q, iq, AC$ corresponds to the estimation method. The results are given in Figure 3.4.

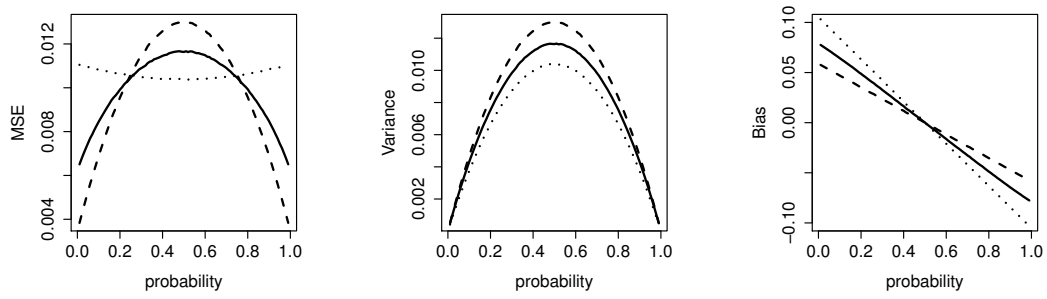

Figure 3.4: MSE, variance and bias for the restricted estimator $\hat{p}_{iq}$ (solid), the squared error loss function estimator $\hat{p}_q$ (dashed) and the Agresti-Coull estimator $\hat{p}_{AC}$ (dotted). Results are based on $N = 15$ observations and 10^4 replications.

The proposed estimator $\hat{p}_{iq}$ outperforms (in terms of the MSE) the Bayes estimator obtained with the squared error loss function $\hat{p}_q$ in the interval $\vartheta \in (0.2, 0.8)$. The cost of this advantage is the worse performance on the intervals close to the bounds

as the proposed form of the loss function penalises the boundary decisions and drives estimates away from them. However, the proposed estimator outperforms the Agresti-Coull estimator $\hat{p}_{AC}$ on the same intervals $\vartheta \in (0, 0.2)$ and $\vartheta \in (0.8, 1)$. Thus, the proposed estimator provides a reasonable trade-off between currently used estimators $\hat{p}_q$ and $\hat{p}_{AC}$ outperforming $\hat{p}_{AC}$ on bounds and $\hat{p}_q$ away from them.

In addition to the MSE, associated confidence intervals and coverage probabilities are extensively studied in the literature (see e.g. Brown et al., 2001). In particular, coverage probabilities were shown to have an erratic behaviour and often to go below their nominal level. Corrections were proposed by Agresti and Coull (1998). Confidence intervals can also be constructed around the newly proposed point estimator $\hat{p}_{iq}$. The following confidence intervals are compared via simulated coverage probabilities in Figure 3.5:

[1] Normal approximation of the confidence interval centred around $\hat{p}_k$, as suggested by Brown et al. (2001)

$$CI_N^{(k)} = \hat{p}_k \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}_k(1-\hat{p}_k)}{n}},$$

where $1-\alpha$ is the confidence level and $k = q, iq, AC$.

[2] Wilson confidence interval centred around $\hat{p}_{AC}$ (Wilson, 1927)

$$CI_W^{(AC)} = \frac{x+2}{n+4} \pm 2\frac{\sqrt{n}}{n+2}\sqrt{\frac{x(n-x)}{n^2} + \frac{1}{n}}.$$

[3] Approximate confidence interval using the delta-method (Resnick, 2013) centred around newly proposed $\hat{p}_{iq}$

$$CI_D^{(iq)} = \hat{p}_{iq} \pm 2\sqrt{\hat{V}_{iq}}$$

with $\hat{V}_{iq} = \left(\frac{\partial f(x)}{\partial x}\right)^2 \Big|_{x=n\hat{p}_{iq}} n\hat{p}_{iq}(1-\hat{p}_{iq})$ and $f(x) = \left(1 + \sqrt{\frac{(n-x+1)(n-x+2)}{(x+1)(x+2)}}\right)^{-1}$.

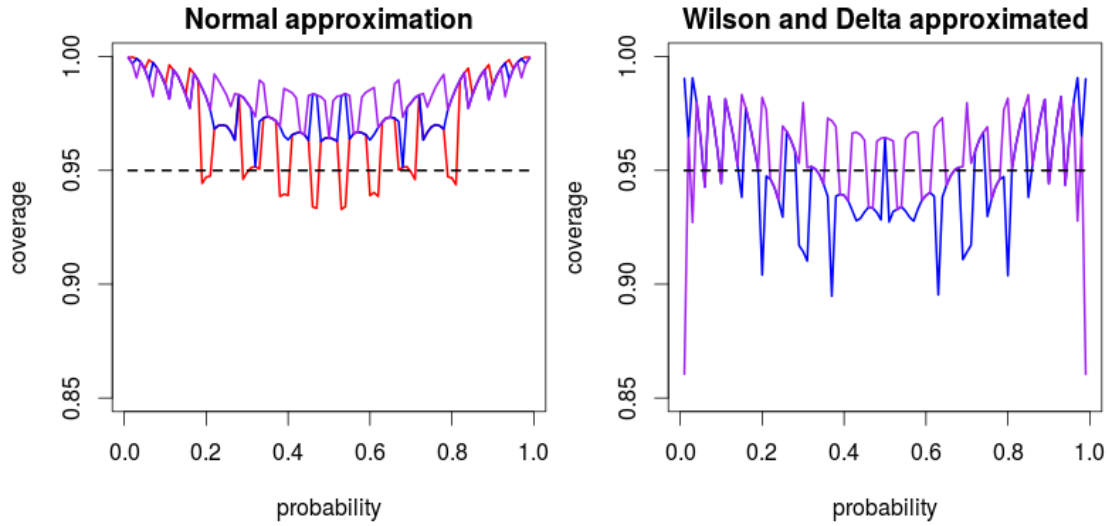

Figure 3.5: Left panel: Coverage probabilities of $CI_N^{(iq)}$ (blue line), $CI_N^{(q)}$ (red line) and $CI_N^{(AC)}$ (purple line) using [1], the normal approximation interval. Right panel: Coverage probabilities of $CI_D^{(iq)}$ (blue line) and $CI_W^{(AC)}$ (purple line) using [2], the Wilson, and [3], the delta-method, confidence intervals respectively. Results are based on $N = 15$ observations and 10^4 replications.

As it was originally found by Vollset (1993); Newcombe (1998), the coverage probabilities have an erratic behaviour for different values of the true probability ϑ. Using the normal approximation confidence interval, the coverage probability of $CI_N^{(q)}$ goes below the nominal value for several values of ϑ. The coverage probabilities of $CI_N^{(iq)}$ and $CI_N^{(AC)}$ also fluctuate but never go below 0.95 for $N = 15$. While one can find combinations of N and ϑ for which coverage probabilities $CI_N^{(iq)}$ and $CI_N^{(AC)}$ might be below 0.95 (Brown et al., 2001), it would be generally true that their coverage probabilities are greater than those of $CI_N^{(q)}$ for larger intervals of ϑ and are more robust. In addition, a comparison between the normal approximation method (left panel of Figure 3.5) and the Wilson and delta method intervals (right panel of Figure 3.5) supports the suggestion by Brown et al. (2001) that the normal approximation gives a simple way to construct confidence intervals with - on average - better coverage probabilities than alternative methods.

3.4.2 Restricted Estimation of a Normal Distribution Mean

In the following example, it is demonstrated what benefits the proposed form of loss function (3.3.4) can provide in a Bayesian framework in the presence of the additional information that the true parameter lies in an interval (a, b).

The problem of restricted mean estimation has been known for a long time (e.g. Bartholomew, 1965) and has been extensively studied in the literature (see e.g. Casella and Strawderman, 1981; Bickel, 1981; Gatsonis et al., 1987; Kumar and Tripathi, 2008). The previously proposed estimators were constructed using the squared error loss function and compared using a Bayes risk (Marchand et al., 2011). Interestingly, despite the variety of the literature on the problem and the fact that the Bayes estimator with respect to the uniform prior distribution on $(-a, a)$ outperforms uniformly the "unrestricted" Bayes estimator under the squared error loss function (Hartigan, 2004), the sample mean estimator is still widely used in practice. For this reason we propose our simple alternative and compare it to commonly used estimators.

Consider the problem of estimating the mean μ of a normal sample of i.i.d. $X_i \sim \mathcal{N}(\mu, \sigma^2)$, $i = 1, \ldots, n$ where σ^2 is known. Assume it is known that the parameter μ belongs to the interval $(-a, +a)$ where $a > 0$. An example is the estimation of the treatment effect in paediatric studies of a drug that has been already tested in the adult population. For instance, a clinician might be sure that a toxicity biomarker for the same doses cannot be less than for adults.

As stated above, the common way to incorporate this information in the Bayesian framework is to restrict the prior distribution of the parameter μ to the interval $(-a, +a)$ and then to use the squared error loss to obtain the Bayes estimator. Thus, the information about the restricted space is ignored when obtaining a

summary statistic, while the proposed form of loss function (3.3.4) allows to incorporate it together with the prior distribution. Furthermore, in practice the prior information can be often ignored as well and the sample mean estimator (corresponding to the Jeffrey's prior) is used. Therefore, a comparison of the proposed loss function and the currently used approaches (with and without incorporating the prior information) is of interest.

Consider samples of size $N = 15$ and Bayes estimators as follows:

- Bayes estimator under Jeffrey's prior $g(\mu) \propto k\sqrt{\frac{n}{\sigma^2}}$ for μ and the squared error loss function. Denoted by J;
- Bayes estimator under the uniform prior $\mathbb{U}(-a, +a)$ for μ and squared error loss function. Denoted by U_1.
- Bayes estimator in Equation (3.3.6) under the uniform prior $\mathbb{U}(-a, +a)$ for μ and the interval symmetric squared error loss function. Denoted by U_2.

Alternatively to U_2, the Bayes estimator U_2' defined by (3.3.6), but with a wider interval $(-1.25a, +1.25a)$ is also applied to investigate how a less strict penalisation of the bounds influences the estimation. This wider interval could be interpreted as a conservative way to incorporate information about the parameter location.

Two cases $a = 2$ and $a = 4$ are considered. The parameter $\sigma^2 = 4$ is assumed to be known in both cases. As before, the MSE is used as a measure of the comparison. The results for 10^4 replications for each value of μ are given in Figure 3.6.

Incorporating the interval information in the Bayes estimator allows for improvement if the true value of the parameter μ is not close to the bounds. In the case $a = 2$ estimator U_2 outperforms U_1 on the interval $\mu \in (-1, +1)$ and in the case $a = 4$ on the interval $\mu \in (-3, +3)$. The same holds for estimator U_2', however, its MSE never falls below the level of the MSE corresponding to J. Clearly, the wider interval improves estimation on the bounds, at the cost of the higher MSE

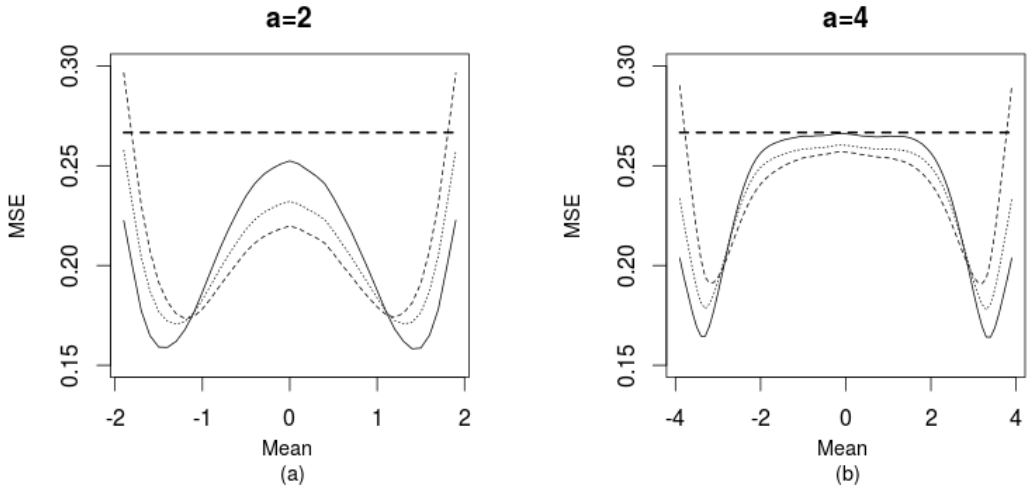

Figure 3.6: MSE corresponding to different values of the restricted mean parameter μ with (a) $a = 2$ and (b) $a = 4$ and Bayes estimator U_1 (solid), U_2 (dashed), U_2' (dotted) and the simple mean estimator J (solid dashed). Results are based on 10^4 replications.

in the middle. Note that the wider interval $a = 4$ corresponds to less amount of the additional information and to a smaller benefit in the MSE comparing U_1, U_2 and U_2' against J. Overall, the Bayes estimator corresponding to the interval symmetric loss function avoids the boundary decisions, improves the estimation if the parameter lies away from bounds and can be recommended for the application if boundary decisions lead to severe consequences.

3.4.3 Estimation of the Parameters of Gamma Distribution

An important example of the multidimensional restricted parameters estimation is Gamma distribution with positive shape and scale parameters $\alpha_1, \alpha_2 > 0$. A Bayesian inference for this problem has been studied, for example, in Miller (1980) and methods for approximate computation of Bayes estimators were proposed using the Lindley approximation (Lindley, 1980).

We consider the function $L_{sq}^{(2)}$ given in Equation (3.3.9) to obtain Bayes estimators for the parameters. As the loss function (3.3.9) is the sum of the univariate

precautionary loss functions (3.1.2), the following estimators are used

$$\hat{\alpha}_i = \sqrt{\mathbb{E}(\alpha_i^2)},\ i = 1, 2 \tag{3.4.3}$$

where the expectation $\mathbb{E}$ is taken with respect to the posterior density function.

Let us consider an experiment with sample size $N = 15$. The parameters of Gamma distribution are varied over a grid, $\alpha_1, \alpha_2 \in (0, 10)$ and the performance of the different approaches are compared by simulations. The two approaches compared use the same prior distribution for parameters α_1 and α_2 using the following Bayes estimators:

1) Bayes estimator under the squared loss function - the posterior mean

2) Bayes estimator under the multivariate precautionary loss function - the scale posterior mean (3.4.3).

We use non-informative Gamma prior distribution with parameter $(10^{-4}, 10^{-4})$ for both α_1 and α_2 and an approximate computation method proposed by Lindley (1980). We emphasise the choice of the prior distribution is out of scope and the goal is to compare two Bayes estimators under equal conditions. The differences between MSEs corresponding to the method 1) and to the method 2) for parameters α_1 and α_2 in 10^4 replicated trials is given in Figure 3.7.

The differences in the MSEs for both parameters are positive for all values of the true parameters α_1 and α_2. It means that the Bayes estimator from method 2) is associated with a smaller MSE than the Bayes estimator from method 1). The difference in the MSE increases as a true value of parameter increases. This result makes the proposed estimator and the associated loss function $L_{sq}^{(2)}$ good candidates for the further investigation in multidimensional estimation problems.

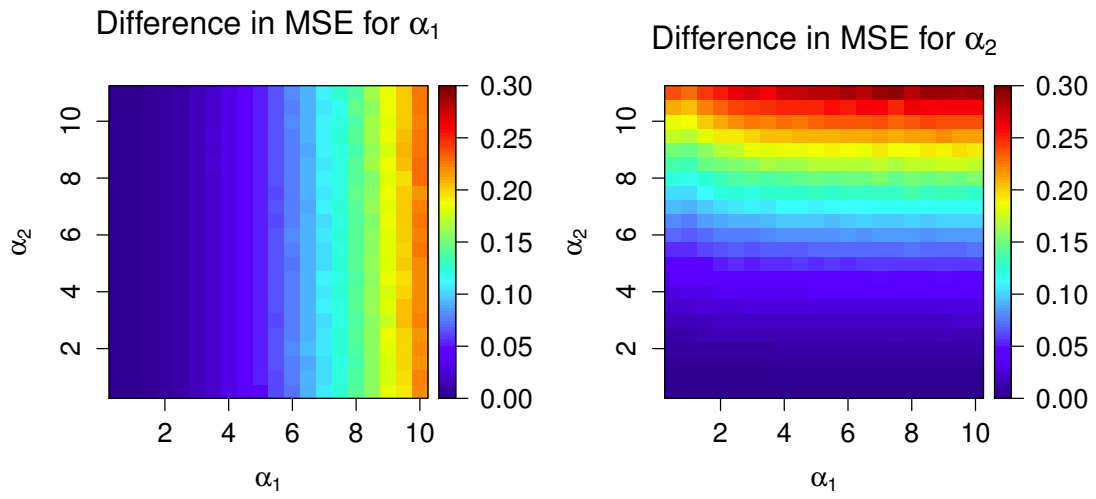

Figure 3.7: Differences in the MSEs for parameters α_1 and α_2 using Bayes estimator under the squared error loss function and the Bayes estimator under $L_{sq}^{(2)}$. Results are based on 10^4 replications.

3.5 A Modified Allocation Rule for the Continual Reassessment Method

The absolute distance criterion (2.1.4) used by the Bayesian CRM (Section 2.1.1) implies that each next patient should be allocated to the estimated MTD. While this is expected to result in the assignment of as much as possible patients to the MTD, it can also lead to not exploring a dose-toxicity relation enough and to "locking-in" (selecting regardless further outcomes) on suboptimal doses (Oron and Hoff, 2013). Furthermore, an investigator should take into account another ethical constraint and avoid the allocation of patients to highly toxic doses. Finally, following the arguments above, the absolute distance criterion might not be a reliable measure of the distance between objects ($\hat{p}_j$ and γ_t) defined on the unit interval. In this section, we show how both statistical and ethical counterparts can be taken into account in the allocation criterion using the proposed interval symmetric loss function (3.3.4).

3.5.1 Motivation

Following Section 2.1.1, consider a Phase I clinical trial with two doses (d_1, d_2) and a binary endpoint. The goal of the trial is to find the MTD which has a probability of DLT closest to the target, say $\gamma_t = 0.30$. Assume that 10 patients were assigned to each dose and 2 and 4 toxicities were observed, respectively. Then, a typical question in a sequential trial is: "which dose should be administered to the next patient". As described in Chapter 2, a conventional criterion for many dose-escalation designs is to assign the next patient to dose d_j corresponding to the point estimate $\hat{p}_j$ closest to γ_t using either the absolute distance (2.1.4) or the squared distance

$$(\hat{p}_j - \gamma_t)^2 . \tag{3.5.1}$$

Assume that in the example above, the probabilities p_1 and p_2 are considered as random variables having Beta distributions $\mathcal{B}(2, 8)$ and $\mathcal{B}(4, 6)$ and one uses the means as point estimates: $\hat{p}_1 = 0.2$ and $\hat{p}_2 = 0.4$. Following the criterion (3.5.1), the next patient can be assigned to either of doses as both probabilities are equally close to the target. At the same time, one can argue that these doses are not "equal" for at least two reasons. On the one hand, the criterion (3.5.1) ignores the uncertainty in the estimates. Indeed, the probability of being within 5% of γ_t is larger for p_2

$$\mathbb{P}\left(p_2 \in (0.25, 0.35)\right) > \mathbb{P}\left(p_1 \in (0.25, 0.35)\right) . \tag{3.5.2}$$

The larger variance of p_2 favours the decision to assign the next patient to d_2. On the other hand, assigning the patient to a dose with the estimated toxicity risk of 0.4 can be unethical as it potentially exposes a patient to an excessive toxicity. It is usually of interest to balance these two concerns - an investigator aims to decrease an uncertainty, but also tends to avoid exposing patients to doses far from the MTD. In the illustration above, the conventional absolute (or squared) distance criteria fails to address both of these concerns.

The question of safety was firstly addressed in the allocation rule by Babb et al. (1998) using the Escalation with Overdose Control (EWOC) design. The EWOC uses the following criterion for patients allocations (Cheung, 2011)

$$\mathbb{E}\left(\alpha(\gamma - p_i)^+ + (1-\alpha)(p_i - \gamma)^+\right) \tag{3.5.3}$$

where $(x)^+ = \max(0, x)$ and α is a parameter of asymmetry. Criterion (3.5.3) imposes that more toxic doses should have more severe penalty than less toxic doses. While the EWOC design was shown to result in a low mean number of DLTs, it also leads to the underestimation of the MTD in many scenarios (Wheeler et al., 2017). Therefore, some modifications were proposed. The main idea beyond the modifications is to use a varying parameter α_i in the criterion (3.5.3) for allocating patient i rather than a fixed value α. Particularly, Tighiouart et al. (2010) used

$$\alpha_2 = \ldots = \alpha_9 = 0.25, \quad \alpha_i = \min\left(\alpha_{i-1} + 0.05, 0.50\right). \tag{3.5.4}$$

We will denote the corresponding design by TAR. Wheeler et al. (2017) proposed toxicity-dependent feasibility bound design (TDFB) which uses

$$\alpha_{i+1} = \min\left(0.50, \alpha_{min} + (0.50 - \alpha_{min})\frac{N - 1 - \sum_{i=1}^{N} y_i}{S}\right) \tag{3.5.5}$$

where $\alpha_{\min}$ and S are parameters chosen by an investigator. These designs, however, prioritise the ethical issue (avoiding the overdosing of patients) and do not fully address the statistical issue (minimising the uncertainty in the MTD estimate).

Another design aimed to resolve both uncertainty and safety concerns is the Bayesian Logistic Regression Model (BLRM) proposed by Neuenschwander et al. (2008). To address the statistical part, BLRM uses the whole distribution of the DLT probability, while the ethical part is addressed via a penalty for overly toxic intervals. The allocation is determined by a loss function computed for each dose.

For instance, the loss function for $\gamma_t = 0.33$ takes the form

$$L = \begin{cases} 1 \text{ if } p \in (0.00, 0.26) \\ 0 \text{ if } p \in (0.26, 0.41) \\ 1 \text{ if } p \in (0.41, 0.66) \\ 2 \text{ if } p \in (0.66, 1.00) \end{cases} \tag{3.5.6}$$

Given the posterior distribution f_n after n patient have been observed, the expected value of the loss function for each dose is computed and the next patient is assigned to the dose corresponding to the minimum value of the loss. While this approach has been proven to be useful in practice, it requires several additional parameters (the values of the loss function and corresponding intervals) to be carefully calibrated.

In this section we propose a new criterion for allocation of patients in a dose escalation trial which is based on the interval symmetric loss function (3.3.4). The criterion takes into account both the variance of the probability distribution and the ethical concerns of an investigator. It requires only one parameter, which has a simple and intuitive interpretation, to be specified. This parameter controls the mean number of toxicity outcomes explicitly. We incorporate the proposed criterion into the Bayesian CRM using the 1-parameter power model (2.1.5) and compare operating characteristics to the standard CRM, EWOC type designs and BLRM.

3.5.2 Criterion

Derivation Following the construction of the interval symmetric loss function (3.3.4) we propose to use the following criterion for the allocation of patients during the trial

$$\delta(p, \gamma_t) = \frac{(p - \gamma_t)^2}{p(1-p)}. \tag{3.5.7}$$

The criterion (3.5.7) takes its minimum value $\delta(\cdot) = 0$ at $p = \gamma_t$, and, as shown above, its distinguishing property is an infinite penalisation of the boundary decisions. In the context of dose escalation trials, it means that doses corresponding to $p = 0$ and $p = 1$ would be never selected. The denominator of the criterion (3.5.7) also corresponds to the variance of a binary event probability. Therefore, it follows that the uncertainty is incorporated into the criterion's point estimate directly and the criterion (3.5.7) can be considered as a statistic. Applying it to the illustration example above helps to address the uncertainty issue: $\delta(\hat{p}_1 = 0.2, \gamma_t = 0.3) = 1/16$ and $\delta(\hat{p}_2 = 0.4, \gamma_t = 0.3) = 1/24$ leading to selecting d_2 for the next patient as follows from Inequality (3.5.2).

The target toxicity γ_t is always below 0.5 in Phase I clinical trials. Consequently, if one considers two point estimates which stand on the same squared distance $(\gamma_t - \nu)^2$ from the target γ_t (for $0 < \nu < \gamma_t$), the criterion (3.5.7) favours the higher probability estimate. This might conflict with the ethical constraint. Indeed, it was shown above that the novel criterion is *symmetric* on the unit interval implying that overly toxic and overly safe doses are equally penalised. To address needs of Phase I trial, we generalise the interval symmetric loss function to allow the asymmetric penalisation by including an asymmetry parameter a as follows

$$\delta(p, \gamma_t) = \frac{(p - \gamma_t)^2}{p^a (1 - p)^{2-a}}. \tag{3.5.8}$$

The asymmetry parameter $0 \leq a \leq 2$ corresponds to the penalisation of overly toxic doses and $(2 - a)$ to the penalisation of overly safe doses. The constant 2 is chosen to preserve the same rate of p in both nominator and denominator which guarantees that $\delta \to 0$ when $p \to \gamma_t$ for any value of γ_t. Clearly, $0 \leq a < 1$ implies a more severe penalty for selections of more toxic doses. Applying the criterion (3.5.8) using $a = 0.5$ one can obtain that

$$\delta(\hat{p}_1 = 0.2, \gamma_t = 0.3, a = 0.5) < \delta(\hat{p}_2 = 0.4, \gamma_t = 0.3, a = 0.5)$$

implying that dose d_1 should be selected due to the penalty for overly toxic dose selections. We refer to the proposed allocation criterion as to the *Convex Infinite Bounds Penalisation* (CIBP) approach. Illustrations of the standard criterion (3.5.1) and of CIBP using $a = 1$ and $a = 0.5$ are given in Figure 3.8.

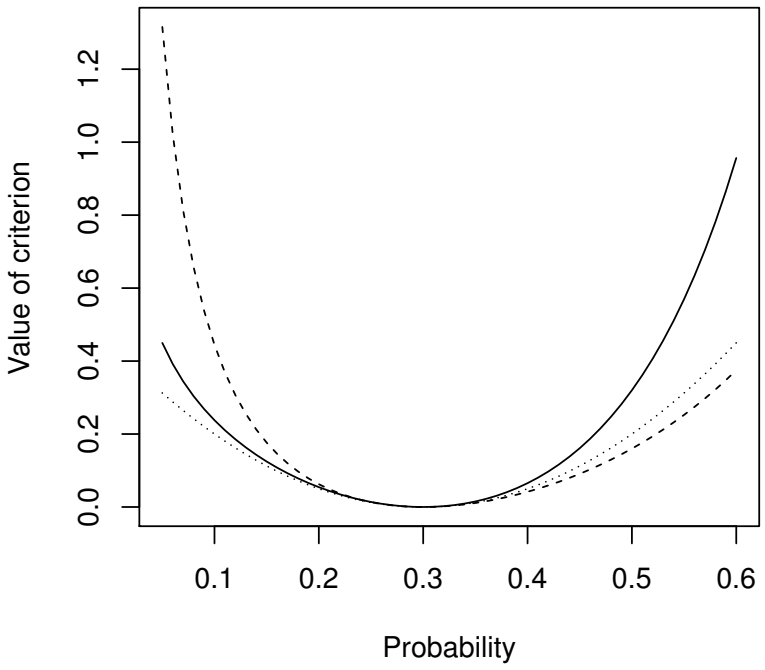

Figure 3.8: Squared distance (dotted line) and CIBP criteria using $a = 1$ (dashed line), $a = 0.5$ (solid line) for different values of $p \in (0.05, 0.6)$. The target probability is $\gamma_t = 0.3$.

The CIBP criteria using both $a = 1$ and $a = 0.5$ go to infinity faster than the squared distance as p approaches the lower bound. Moreover, for $a = 1$ the overly toxic doses are penalised less than overly safe, because corresponding values are located far from the upper bound. The asymmetric CIBP criterion solves this issue and penalises probabilities above $\gamma_t = 0.3$ more severely than alternatives. All criteria behave similarly in the neighbourhood of the target γ_t. Overall, the properties of the CIBP criterion allows to resolve the statistical and ethical concerns by the use of an appropriate value of the asymmetry parameter a. The guidance on the choice of a is given in the following section.

Choice of the parameter of asymmetry Firstly, note that the denominator alone is maximised at the point $p = a/2$. Then, if $\hat{p}$ is an estimator of p (depending on the approach, for instance, MLE or the Bayes optimal estimator) the "plug-in" estimator of the CIBP criterion

$$\delta(\hat{p}, \gamma_t) = \frac{(\hat{p} - \gamma_t)^2}{\hat{p}^a (1 - \hat{p})^{2-a}} \tag{3.5.9}$$

using $a = 2\gamma_t$ leads to the same allocation as the plug-in estimator of the squared distance criterion (3.5.1). Then, values $a < 2\gamma_t$ lead to a more conservative allocation of patients compared to the squared distance.

Secondly, the asymmetry parameter a represents the trade-off between the ethical and uncertainty concerns. Then, a sensible choice of a should satisfy the following conditions. Consider an interval $(\gamma_t - \nu, \gamma_t + \nu)$. Assume that given two probability estimates belonging to this interval and standing on the same squared distance from γ_t, one would like to select the lower estimate due to the safety reasons. Then, $(\gamma_t - \nu, \gamma_t + \nu)$ defines an interval in which the safety is prioritised. Similarly, given two estimates lying outside of the interval $(\gamma_t - \nu, \gamma_t + \nu)$ for which the squared distances are equal, one prefers decision with a greater uncertainty. Evidently, the estimates on the bounds of the interval should correspond to the same value of the CIBP criterion. Formally, solving

$$\frac{((\gamma_t - \nu) - \gamma_t)^2}{(\gamma_t - \nu)^a \left(1 - (\gamma_t - \nu)\right)^{2-a}} = \frac{((\gamma_t + \nu) - \nu)^2}{(\gamma_t + \nu)^a \left(1 - (\gamma_t + \nu)\right)^{2-a}}$$

one can obtain that

$$a = \frac{2}{1 + A}$$

where

$$A = \left(\log \frac{\gamma_t - \nu}{\gamma_t + \nu}\right) \Big/ \left(\log \frac{1 - \gamma_t - \nu}{1 - \gamma_t + \nu}\right)$$

For the fixed target values of the target γ_t and the half width ν, one can compute the corresponding value of the asymmetry parameter a. Figure 3.9 shows values of the asymmetry parameter for different half-widths ν and target values $\gamma_t = \{0.20, 0.25, 0.30\}$, and provides a useful insight on the intuitive interpretation of the parameter a.

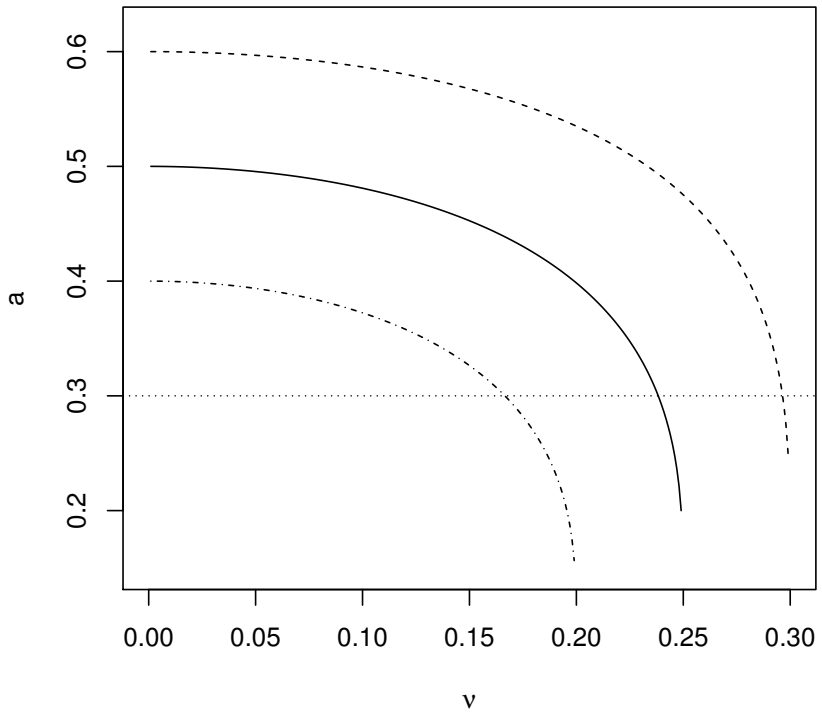

Figure 3.9: Values of the parameter of asymmetry a for $\gamma_t = 0.20$ (dashed-dotted line) , $\gamma_t = 0.25$ (solid line), $\gamma_t = 0.30$ (dashed line) and $\nu \in (0, 0.35)$. The horizontal line corresponds to $a = 0.30$.

As $\nu \to 0$, a tends to $2\gamma_t$ corresponding to the squared distance allocation rule. Increasing values of ν imply that an investigator prefers an interval in which a lower estimate is selected to be wider, and consequently, more conservative allocations and smaller values of a. Note that a corresponding to $\nu \approx \gamma_t$ guarantees that for any two estimates standing on the same squared distance from the target γ_t, the dose corresponding to the lower toxicity will be selected. For example, for $\gamma_t = 0.25$ the corresponding value of a is close to 0.3 marked by the dotted horizontal line.

Novel Criterion in the CRM We incorporate the novel criterion (3.5.8) in the CRM design. The design proceeds as described in Section 2.1.1, but replaces

the criterion (2.1.4) with the following. The dose minimising

$$\mathbb{E}\left(\frac{(\psi(d_j,\theta)-\gamma_t)^2}{\psi(d_j,\theta)^a(1-\psi(d_j,\theta))^{2-a}}\right) \tag{3.5.10}$$

among all $d_1,\ldots,d_m$ is selected for the next group of patients where the expectation is found with respect to the posterior probability f_n given in Equation (2.1.2). The procedure is repeated until the maximum number of patients, N, has been treated. As the uncertainty and the conservatism is important in the allocation only, we propose to use the conventional criterion (3.5.1) for the final MTD recommendation. Note that if the CIBP criterion (3.5.10) is used for the MTD recommendation then the design might lead to the MTD underestimation (Appendix A).

3.5.3 Illustration

To illustrate the impact of the proposed escalation criterion, we revisit the results of the actual clinical trial of Everolimus in patients with Metastatic Breast Cancer (NCT00426556). The study considers three regimens of Everolimus given together with Paclitaxel and Trastuzumab (PT):

1. Daily dosing of Everolimus 5mg plus PT (d_1)
2. Daily dosing of Everolimus 10mg plus PT (d_2)
3. Weekly dosing of Everolimus 30mg plus PT (d_3)

The goal is to find the regimen corresponding to the target toxicity $\gamma_t = 0.3$. Note that the amount of the complimentary drugs is fixed during the trial and a clinician is confident in the monotonicity of toxicity probabilities corresponding to d_1, d_2, d_3. Thus, the trial can be analysed using the tools for single-agent trials. The following aggregated data is available by the end of the trial.

We apply the CRM design using the 1-parameter power model (2.1.5) and the robust operational prior distribution $\theta \sim \mathcal{N}(0, 1.34)$ used by O'Quigley and Shen

Table 3.1: Aggregated data of the Everolimus trial

Dose	d_1	d_2	d_3
Number of Patients assigned	6	17	10
Number of DLTs	3	6	7

(1996) and the skeleton $(0.20, 0.30, 0.40)$ implying that the prior MTD is d_2. We restrict the design such that the dose skipping is not allowed and enforce the start from the lowest dose. Patients are enrolled in cohorts of three. Note that the parameters of the design are the same for both the original CRM and the proposed modification. The only difference is the allocation criterion: the original CRM uses the squared distance criterion (3.5.1) while the proposed design uses the CIBP criterion (3.5.10). Following the interpretation of the asymmetry parameter we fix $a = 0.3$ to favour less toxic selections (see Figure 3.9). The designs are implemented using the interactive functions of the `bcrm`-package by Sweeting et al. (2013). We use the aggregated data from Table 3.1 to generate responses in a single realisation of the trial. Clearly, DLTs indicated in Table 3.1 can appear in any sequence. Therefore, we generate a random sample (without replacement) for each dose to have a specific order of DLTs and fix it for both trials. The only exception is that the realisation for the first cohort is chosen by us. We consider its influence separately.

Allocation of Patients The first 3 patients are assigned to d_1 by the restriction. We begin by assuming that all 3 patients do not experience DLTs and consider the behaviour of the designs. The sequential allocations of patients for the CRM and CIBP designs in this case are given in Figure 3.10. Values of the criteria after responses for each cohort are given in Table 3.2.

After no DLTs were observed for the first cohort, CRM and CIBP allocate the second cohort of patients to d_2 for which one patient experiences a DLT. Given the toxic outcome, CRM recommends to stay at d_2 for the third cohort. In contrast, CIBP recommends to return to the previous dose level due to the conservatism of the criterion. As aggregated data are used, the sequence of patients are fixed for

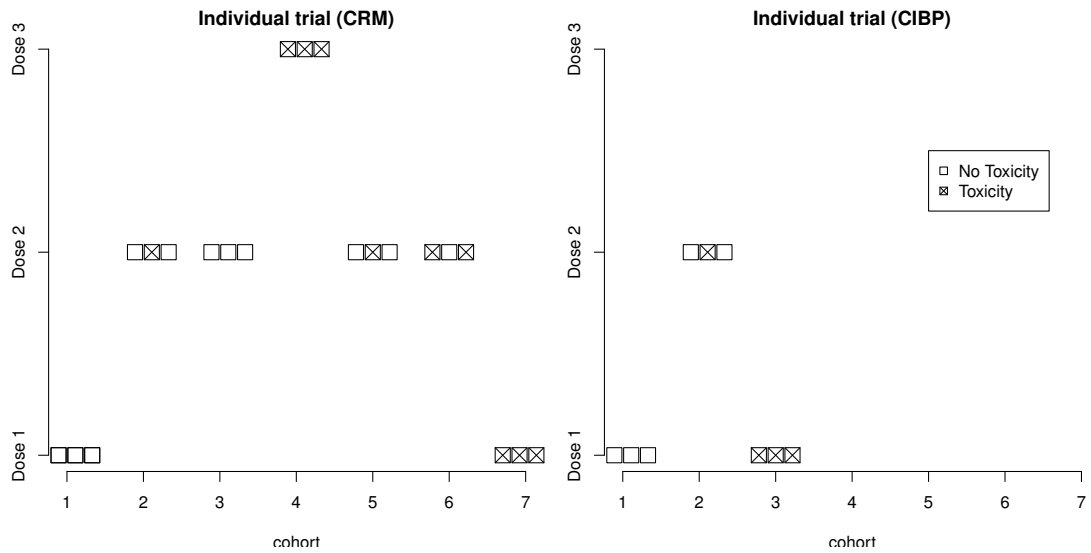

Figure 3.10: Allocations of 7 cohorts in the individual Everolimus trial.

both methods given a dose. Therefore, due to the construction of the illustration, all patients in cohort 3 (using CIBP) experienced DLTs to consistent with the aggregared data on d_1. Then, the trial would be terminated by a clinician. At the same time, the trial using the original CRM design proceeds as no DLTs were observed for cohort 3 at dose d_2. The dose d_3 is recommended for cohort 4 in which all patients have DLT. This leads to the de-escalation to d_2 and after 2 cohorts for which 3 patients (out of 6) had DLT and the further de-escalation to d_1. All 3 patients in cohort 7 experienced DLTs and a clinician terminates the trial. Overall, CRM assigned 21 patients and ten of them experienced DLTs which resulted in the same conclusion as CIBP did after 9 patients with 4 toxicity outcomes only.

Table 3.2: Values of the squared distance criterion (2.1.4) used by CRM and of the CIBP criterion (3.5.10) in the individual trial after responses for each cohort. Value of criteria corresponding to the dose selected for a next cohort by are in bold.

	CRM			CIBP		
Cohort 1	0.031	**0.012**	0.002	0.62	**0.47**	0.45
Cohort 2	0.01	**0.001**	0.005	**0.10**	0.12	0.21
Cohort 3	0.03	0.013	**0.001**	**0.30**	0.67	1.41
Cohort 4	0.003	**0.002**	0.003			
Cohort 5	0.003	**0.002**	0.024			
Cohort 6	**0.000**	0.009	0.039			
Cohort 7	**0.014**	0.048	0.096			

The illustration above demonstrates the allocation if no toxicity outcomes are observed in cohort 1. Regarding all other possible outcomes (1, 2 or 3 toxicities), one can find that both approaches result in the same conclusions, same numbers of DLTs and the total number of patients. Therefore, the proposed approach results in the same MTD recommendation, but in either a similar or safer allocation of patients. This motivates a comprehensive simulation study of the novel allocation criterion.

3.5.4 Comparison to the CRM

Setting Below we compare the performance of the proposed CIBP criterion against the squared distance criterion both applied to the 1-parameter power model. The single-agent Phase I trial with $m = 6$ doses and $N = 30$ patients is considered. The goal is to find the MTD corresponding to the target $\gamma_t = 0.25$. We consider 6 dose-toxicity scenarios with target doses located at positions corresponding to scenario's number. The dose-toxicity shapes are given in Figure 3.11. Toxicity scenarios were chosen "equally difficult" in terms of the optimal non-

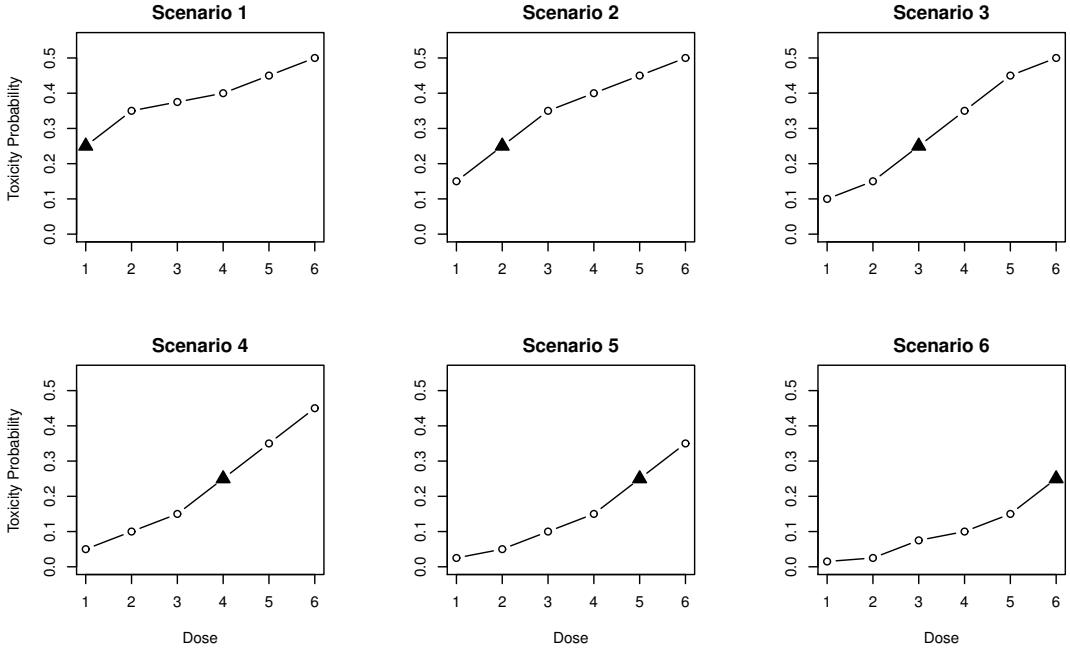

Figure 3.11: Six considered dose-toxicity scenarios for the comparison to the CRM. The MTD is marked by a triangle.

parametric benchmark (Section 2.1.2). It allows to compare the proportions of correct selections between scenarios. We use function `getprior` from the `dfcrm` R-package to specify the skeleton for the 1-parameter power choosing d_2 to be the prior MTD and the half width of the equivalent interval 0.05. The prior distribution of the parameter is chosen to be $\theta \sim \mathcal{N}(0, 1.34)$ (O'Quigley and Shen, 1996). Different skeletons corresponding to d_3 and d_4 being the MTD are also investigated and the corresponding (quantitatively similar) results are given in Appendix A. Again, both the standard CRM and its proposed modification use the same model parameters and the only difference is the allocation rule. This allows to connect observed differences to the choice of the allocation criterion only. We study (i) the proportion of correct selections and (ii) the proportion of patients experienced DLT. We consider different values of $a = \{0.3, 0.4, 0.5\}$ corresponding approximately to half widths of the interval $\nu \approx \{0.24, 0.20, 0.00\}$.

We denote CRM using the new allocation criterion with parameter a by CIBP(a). The characteristics of both methods are evaluated in `R` (R Core Team, 2015) using the `bcrm`-package by Sweeting et al. (2013). To accommodate the new criterion, corresponding modifications to the package were made.

Operating Characteristics The proportions of each dose selections and the proportion of patients experienced DLT for CRM, CIBP and the non-parametric optimal benchmark (Section 2.1.2) are given in Table 3.3. We use $40,000$ replications to declare any differences above 1% as significant ones.

All considered designs lead to the PCS close to the characteristics of the benchmark in scenarios 1-5 and have some difficulties in selecting the highest dose in scenario 6. Comparing the performance of the novel criteria for different values of the asymmetry parameter, as one can expect, more conservative selections correspond to CIBP(0.3). The greatest differences are in scenarios 1 and 6: the increase in a from 0.3 to 0.5 leads to an increase in the PCS by 5% in the toxic scenario 1

Table 3.3: Proportions of each dose selections and mean proportions of DLTs for CRM, CIBP using $a = \{0.3, 0.4, 0.5\}$, and the non-parametric optimal benchmark. Results are based on $40,000$ replications.

	d_1	d_2	d_3	d_4	d_5	d_6	DLTs
			Scenario 1				
Toxicity	**25.00**	35.00	37.50	40.00	45.00	50.00	
Benchmark	70.12	16.23	7.31	4.23	1.61	0.50	
CIBP(0.3)	69.18	21.65	6.18	2.27	0.61	0.11	28.31
CIBP(0.4)	66.40	22.20	7.25	3.08	0.91	0.16	29.46
CIBP(0.5)	64.12	22.25	8.49	3.80	1.15	0.18	30.58
CRM	65.59	21.16	8.22	3.79	1.07	0.17	30.17
			Scenario 2				
Toxicity	15.00	**25.00**	35.00	40.00	45.00	50.00	
Benchmark	21.01	50.18	19.91	6.25	2.10	0.55	
CIBP(0.3)	24.06	47.88	21.98	5.00	0.93	0.15	23.33
CIBP(0.4)	24.00	46.95	22.28	5.39	1.19	0.19	24.87
CIBP(0.5)	23.97	46.12	22.20	6.02	1.46	0.24	26.45
CRM	25.41	45.76	21.36	5.96	1.27	0.24	26.10
			Scenario 3				
Toxicity	10.00	15.00	**25.00**	35.00	45.00	50.00	
Benchmark	2.49	19.24	50.37	24.08	3.28	0.44	
CIBP(0.3)	4.26	25.61	46.48	20.20	3.16	0.28	20.91
CIBP(0.4)	4.08	25.53	46.25	20.57	3.23	0.34	22.56
CIBP(0.5)	3.77	25.64	46.49	20.46	3.31	0.32	24.21
CRM	3.91	26.66	45.62	20.37	3.06	0.37	23.97
			Scenario 4				
Toxicity	5.00	10.00	15.00	**25.00**	35.00	45.00	
Benchmark	0.10	2.58	19.13	50.37	23.96	3.85	
CIBP(0.3)	0.22	5.01	27.27	44.62	19.65	3.23	19.36
CIBP(0.4)	0.17	4.78	26.64	45.66	19.59	3.15	20.99
CIBP(0.5)	0.18	4.74	27.16	45.74	19.36	2.83	22.43
CRM	0.18	4.50	27.82	45.32	19.15	3.03	22.43
			Scenario 5				
Toxicity	2.50	5.00	10.00	15.00	**25.00**	35.00	
Benchmark	0.00	0.00	2.39	19.23	50.57	27.73	
CIBP(0.3)	0.00	0.34	6.54	27.67	43.34	22.11	17.71
CIBP(0.4)	0.00	0.31	5.89	27.77	44.12	21.89	19.24
CIBP(0.5)	0.00	0.33	5.50	28.06	44.84	21.28	20.73
CRM	0.01	0.27	5.46	28.89	44.10	21.28	20.56
			Scenario 6				
Toxicity	1.50	2.50	7.50	10.00	15.00	**25.00**	
Benchmark	0.00	0.00	0.57	2.67	20.00	76.86	
CIBP(0.3)	0.00	0.04	2.97	10.42	26.84	59.72	15.25
CIBP(0.4)	0.00	0.05	2.30	9.55	27.53	60.58	16.53
CIBP(0.5)	0.00	0.04	1.68	7.31	27.71	63.26	17.98
CRM	0.00	0.05	1.88	8.65	28.89	60.53	17.34

and to the decrease in the PCS by 3.5% in the flat scenario 6. The differences in the rest of scenarios are smaller, but still significant. CIBP(0.3) has a greater PCS

than CIBP(0.4) in scenarios 1 and 2, the same PCS in scenario 3 and a smaller in scenario 4, 5 and 6. Overall, greater values of a favour the selection of higher doses and lead to higher proportions of DLTs with the difference around 2-3% between CIBP(0.3) and CIBP(0.5) under all scenarios.

Regarding the comparison of the novel criterion to the original CRM, one can find that the CIBP(0.4) has a similar PCS, but smaller proportions of DLTs in all considered scenarios. The CIBP(0.5) performs comparably to the CRM (scenarios 2-5) or better (scenario 6) than the CRM at the cost of 1% decrease in the PCS in the toxic scenario 1. The CIBP(0.5) also results in a slight increase in the number of toxic responses. The most noticeable difference can be observed by comparing the CRM to the CIBP(0.3). In terms of the PCS, CIBP(0.3) outperforms the CRM by 4% and 2% in toxic scenarios 1, 2 and shows the comparable performance in the rest of scenarios. At the same time, CIBP(0.3) outperforms the CRM in terms of the proportion of patients experienced a DLT in all considered scenario by nearly 3% in scenarios 2-5 and by 2% in scenarios 1 and 6. While the margin of these differences might be seen to be negligibly small, this improvement results in nearly one fewer patient experienced DLT. Therefore, the novel allocation rule is more ethical as it exposes fewer patients to more toxic doses while leading to either no changes or an increase in the PCS.

Another valuable feature of the novel approach is an additional flexibility which allows controlling the mean number of DLTs directly. Clinicians can tailor the parameter a based on their conservatism. For instance, they might be ready to sacrifice the PCS in the flat scenario 6 for the sake of not giving overly toxic recommendation in the toxic scenario 1. The new criterion enables such an option. At the same time, the design preserves its simplicity and does not result in extra computational costs. The comparison of the proposed design to alternative methods is given in the following section.

3.5.5 Comparison to Alternative Methods

Setting In this section, we compare CRM incorporating the CIBP allocation criterion to alternative designs. We consider the setting by Wheeler et al. (2017) for discrete dose levels. There are $N = 40$ patients and $m = 6$ doses in the trial. The goal is to find the MTD corresponding to the target $\gamma_t = 0.33$. The dose-toxicity scenarios are given in Figure 3.12.

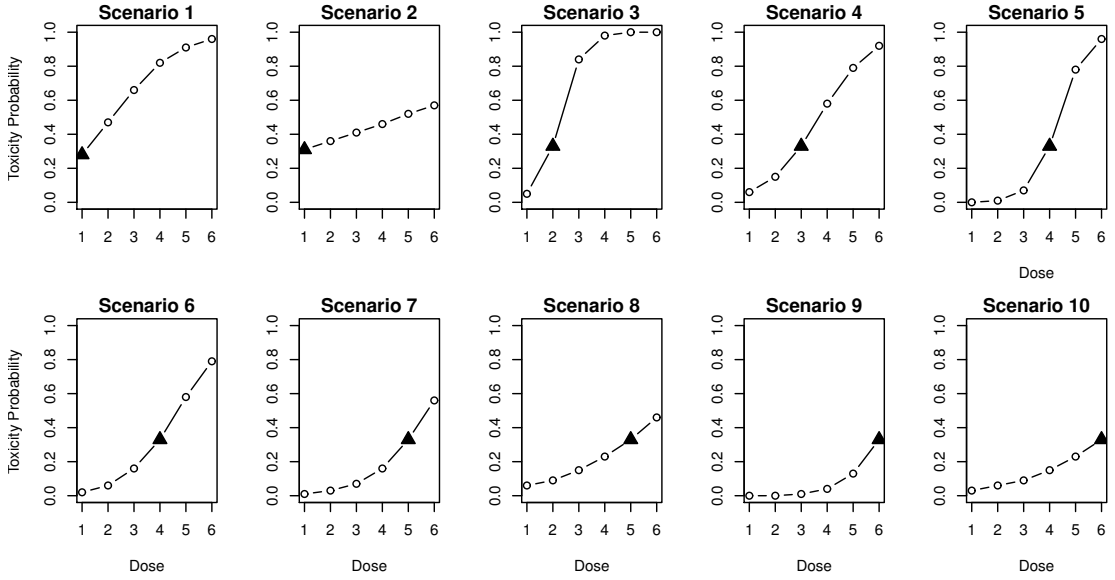

Figure 3.12: Ten considered dose-toxicity scenarios for the comparison to EWOC. The MTD is marked by a triangle.

We compare operating characteristics to the original EWOC design and its modification (TAR and TDFB with parameters $\alpha_{min} = 0.25$ and $S = 12\frac{2}{3}$) and to BLRM using the loss function (3.5.6). All of alternative methods use the 2-parameter logistic model (2.1.6). The bivariate normal prior distribution for parameters of BLRM suggested by Neuenschwander et al. (2008) is used. The parameters of EWOC, TAR and TDFB are specified as by Wheeler et al. (2017). For the novel design the prior distribution of θ as above is used. The only difference with the previous section is the skeleton which is now set using the same information the prior MTD as for the alternative methods - the skeleton is constructed using d_3 as the prior MTD. Assuming that the ethical issue is of the great interest in this trial we study different values of the asymmetry parameter $a = \{0.5, 0.25, 0.10\}$.

We study the performance of designs in terms of (i) Accuracy

$$\mathcal{A} = 1 - m\frac{\sum_{j=1}^{m}(p_j - \gamma_t)^2 \pi_j}{\sum_{j=1}^{m}(p_j - \gamma_t)^2}$$

where p_j is the true toxicity probability for d_j and π_j is the probability to select d_j, and in terms of (ii) mean number of patients experience a DLT. Note that the accuracy index $\mathcal{A}$ takes into account the selection at all dose levels. As many different scenarios are considered, one can expect that one design can have a better performance in one scenarios, but worse in others (Wages, 2015). Therefore, we focus on the average performance: the (geometric) mean accuracy and the mean number of DLTS across all scenarios.

Operating Characteristics Accuracy indices for the non-parametric optimal benchmark, CIBP with asymmetry parameter $a = \{0.5, 0.25, 0.10\}$, TDFB, EWOC, TAR and BLRM designs are given in Figure 3.13.

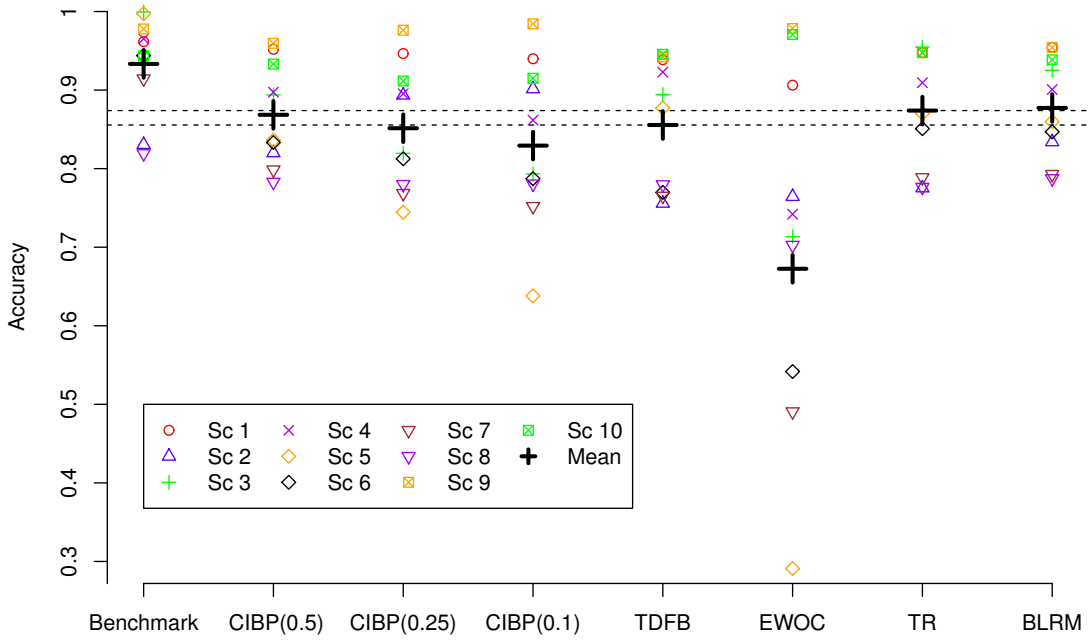

Figure 3.13: Accuracy indexes and mean accuracy indexes for the non-parametric optimal benchmark, the CIBP using $a = \{0.5, 0.25, 0.10\}$, TDFB, EWOC, TAR and BLRM designs. Upper dashed horizontal line corresponds to accuracy of the TAR and the lower one to the accuracy of the TDFB. Results are based on 2000 simulations.

All of the considered methods show the accuracy close to the accuracy of the non-parametric optimal benchmark with EWOC being an exception. Comparing CIBP approaches, one can see that the mean accuracy decreases as the parameter a decreases. A more conservative allocation of patients results in smaller accuracy indices in non-flat scenarios. The decrease in the mean accuracy index is, however, rather small - from 0.87 using $a = 0.5$ to 0.83 using $a = 0.1$. The most noticeable drop across scenarios can be found in scenario 5 - 0.20. Clearly, the variance of the accuracy indices increases as a decreases - a more conservative design leads to the better performance in toxic scenarios 1-3 for the cost of the less accurate performance in flat scenarios 8-10.

Comparing different approaches, TAR, BLRM and CIBP(0.5) correspond to the greatest mean accuracy. TDFB performs comparably to CIBP(0.25) both in terms of the mean accuracy and its variability. As expected, EWOC results in the least mean accuracy index due to the MTD underestimation in scenarios 5-7. The mean accuracy index associated with the most conservative CIBP(0.10) is greater than associate with the EWOC by 0.15.

The mean number of DLTs in all considered designs are given in Figure 3.14. Regarding CIBP approaches, while lower values of a result in a lower accuracy they also result in the fewer mean number of DLTs under all scenarios. The mean number of DLTs across all scenarios is decreased by nearly 2 comparing CIBP(0.5) and CIBP(0.10). Considering different designs, it is of interest to compare designs which have comparable mean accuracy indices. TAR and BLRM result in 0.5 more DLTs on average than CIBP(0.5). TDFB that has the accuracy similar to CIBP(0.25) results in nearly 1.5 more patients experienced DLT. Interestingly, the EWOC, having the least mean accuracy index, results in the same mean number of DLTs across all scenarios as CIBP(0.10).

Overall, in contrast to TDFB and TAR, the novel design does not change the

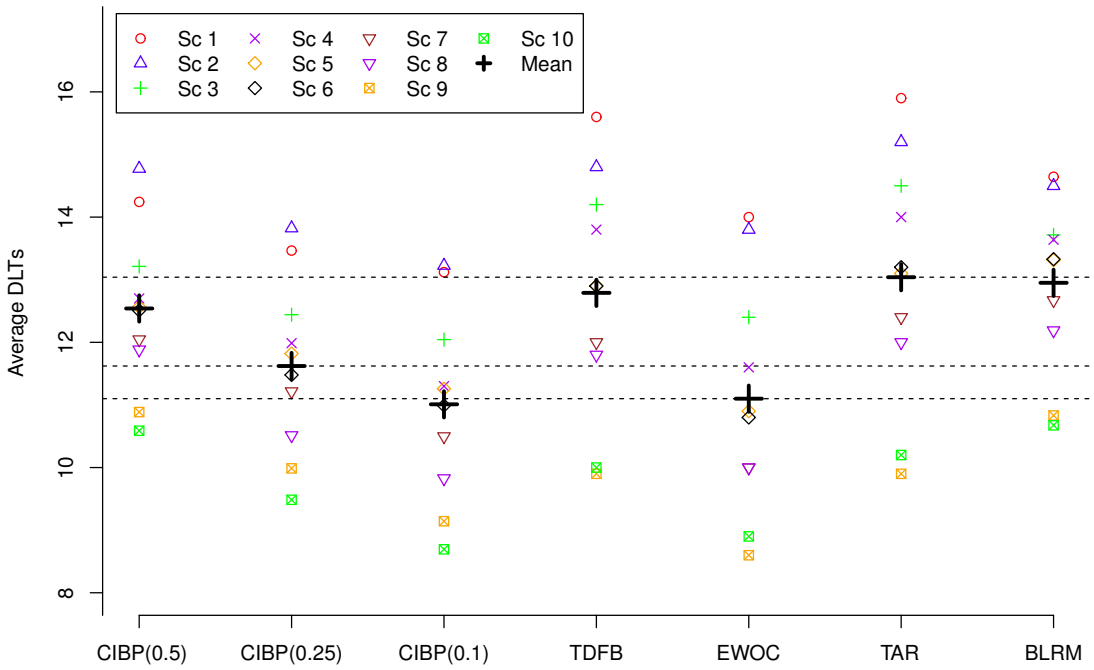

Figure 3.14: Mean number of DLTs for CIBP using $a = \{0.5, 0.25, 0.10\}$, TDFB, EWOC, TAR and BLRM designs. The upper, middle and lower dashed horizontal lines correspond to performance of TAR, CIBP(0.25) and EWOC, respectively. Results are based on 2000 replications.

conservatism parameter as the trial progress which results in only one extra parameter to be specified. One can find a value of parameter a that leading to a similar accuracy index, but a fewer mean number of DLTs. At the same time, the safest version of the CIBP design results in comparable to the EWOC mean number of DLTs across all scenarios, but also in the noticeably higher proportion of correct MTD selections.

3.6 Discussion

In this chapter, we have introduced the unified definition of a symmetric loss function in a restricted parameter space. Scale symmetric and interval symmetric loss functions which are as simple as the squared loss function and which share desirable properties are provided. On the basis of three examples, we show that the corresponding Bayes estimators

- perform well when compared to the squared error loss and other available estimators;
- improve the estimation if the parameter lies away from bounds;
- allow to avoid boundary decisions and can be recommended in applications where more conservative estimates are preferable.

We have subsequently shown how one can benefit from incorporating the novel interval symmetric loss function as the allocation criterion in dose escalation trials. The novel criterion

- requires only one additional parameter;
- has a clear intuitive interpretation;
- can be easily tuned according to the trial purposes.

A guideline on the choice of parameter is also given. We incorporated the criterion into the Bayesian CRM method using the 1-parameter power model. The proposed design results in a fewer mean number of patients experienced DLT with no loss in the probability of correct selections compared to the original CRM design. Comparing the proposed design to other alternatives, it was shown that one can find a value of the asymmetry parameter such that the novel design would result in either

- a similar accuracy and a lower mean number of DLTs, or
- a higher accuracy and a similar number of DLTs.

Therefore, the new allocation criterion makes a model-based design a more ethical option.

Importantly, the novel criterion proposed in this work can be applied to any parametric model and is not limited to the 1-parameter one. Furthermore, the application of the criterion was demonstrated in the context of a single-agent trial only, but it is also of interest what benefits its application to the CRM generalisation (e.g. to the time-to-event CRM by Cheung and Chappell (2000) or to the partial ordering CRM by Wages et al. (2011)) can have. These are the directions for a further research.

Chapter 4

Dose Finding Designs Which Do Not Require Monotonicity Assumptions

In the previous chapters, the contexts in which the monotonicity assumption is satisfied (to various extents) were considered. It was shown that in case of known monotonicity ordering, one can benefit from using a parametric model. There are, however, many regimen finding clinical trials in which a monotonic ordering of regimens with respect to increasing toxicity and (or) efficacy cannot be specified prior to a study. This includes Phase I and Phase I/II combination trials, dose-schedules trials and single drug studies of molecularly targeted agents (MTA). The ability to find the target regimen (TR) using methods employing monotonicity assumption is, however, rather limited if it is not satisfied (Shen and O'Quigley, 1996). There is also a broad class of clinical trials in which regimens are considered as being independent, for example, many Phase II trials. In this section, we propose a class of designs that do not require any parametric or monotonicity

assumptions and can be applied to different types of clinical trials.

4.1 Background

To overcome the issue of an unknown ordering in the context of Phase I trials, some specialised approaches for dual agents combination trials have been proposed (e.g. Thall et al., 2003; Wages et al., 2011; Riviere et al., 2015). The common feature of the majority of novel Phase I designs relaxing the monotonicity assumption is that they either rely on a complex parametric model or on an explicit set of orderings of toxicity. While such methods can borrow information between regimens, they might fail to find the TR if the model or orderings are misspecified. At the same time, there is a growing interest in advanced trials with a large number of potential orderings where specifying all of them (or a corresponding parametric model) is infeasible. In addition, more complex endpoints than a binary one are becoming more common in early phase trials (see e.g. Yuan et al., 2007; Lee et al., 2010, 2017) as they can carry more information about drug's mechanism of action. Despite this, methods for studies with complex outcomes that do not require the monotonicity assumption or a complex model are sparse to date.

Thinking more broadly about selecting one or more regimens during a trial (the main objective of many Phase II studies), different methods that consider regimens being independent have been proposed (see e.g. Stallard and Todd, 2003; Koenig et al., 2008; Whitehead and Jaki, 2009; Magirr et al., 2012). Williamson et al. (2017) have recently advocated designs maximising the expected number of responses in small population trials. As a result, adaptive randomisation methods and optimal Multi-Arm Bandit (MAB) approaches are starting to be considered in Phase II trials. Although MAB designs outperform other well-established methods of randomisation (e.g. a fixed equal randomisation) in terms of the expected number of successes, they can suffer a low statistical power for testing hypotheses (Villar et al., 2015). This problem corresponds to the "exploration versus

exploitation" (also known as "learn versus earn") trade-off (Azriel et al., 2011). Some rule-based modifications have been proposed to achieve a better balance of these two objectives (see e.g. Villar et al., 2015; Williamson et al., 2017). They, however, require calibration of additional parameters that should be carefully calibrated. Moreover, the majority of MAB approaches consider a univariate binary endpoint only and hone in on the regimen (also called arm) with the largest effect by default. They cannot be applied to studies which aim to select the TR corresponding to the target probability $\gamma_e \in (0.7, 1)$.

Regarding regimen finding designs for Phase I/II clinical trials, Gooley et al. (1994) was one of the first to consider two dose-outcome models simultaneously. Similar to later works on Phase I/II designs (e.g. Thall and Russell, 1998; O'Quigley et al., 2001; Braun, 2002; Thall and Cook, 2004; Yin et al., 2006, among others), it is based on the monotonicity assumption for both toxicity and efficacy. As discussed above, these assumptions can be violated for MTAs. Several designs for either single agent or combination therapy trials that relax this assumption for the dose-efficacy relationship have been proposed in the literature: see e.g. Cai et al. (2014); Wages and Tait (2015); Riviere et al. (2016) for a binary efficacy endpoint and Yeung et al. (2015, 2017) for a continuous efficacy endpoint. The majority of novel designs are model-based and the selection of the TR (defined as an optimal biologic regimen) is governed either by (i) a trade-off function (e.g. Thall and Cook, 2004; Yeung et al., 2015, 2017) or (ii) a two-stage procedure in which the safe subset is firstly defined and the most efficacious dose is then estimated (e.g. Thall and Russell, 1998; Yin et al., 2013; Wages and Tait, 2015). Regardless of the approach used, the number of parameters to be included in the underlying model increases considerably if the monotonicity assumption is relaxed. The design proposed by Wages and Tait (2015) uses the 1-parameter model, but employs the idea of different orderings (Wages et al., 2011) to overcome the uncertainty in efficacy ordering. This idea can be extended to a range of problems (see e.g. Wages and Conaway, 2014, for cytotoxic drugs combination

trials), although the extension to more complex settings, for instance, combination studies or combination-schedule trials involving MTAs, can also be challenging due to a large number of possible orderings. Consequently, designs for more complex Phase I/II trials have not been extensively studied yet.

Overall, the research problems described above can be considered as a general issue of correct selection of the TR whose response probability is the closest to the percentile $0 < \gamma < 1$ or, equivalently in the multidimensional case, whose characteristics are the closest to the vector $\boldsymbol{\gamma} \in \mathbb{S}^d$ where $\mathbb{S}^d$ is a d-dimensional unit simplex. In this chapter, we propose a general experimental design for studies with multinomial outcomes to solve this generic problem. Based on developments in the information theory of context-dependent measures of information (Belis and Guiasu, 1968; Kelbert and Mozgunov, 2015; Suhov et al., 2016; Kelbert and Mozgunov, 2017), we derive a criterion which governs regimen selection in the experiment. The criterion is based on the maximisation of the information gain when considering an experiment with a particular interest in regimens whose response probabilities are in the neighbourhood of $\boldsymbol{\gamma}$. Recently other designs using the information gain principle have been proposed (see e.g. Barrett, 2016; Kim and Gillen, 2016). In contrast to these methods, the novel approach allows for the incorporation of the context of outcomes (e.g. avoid high toxicity or low efficacy) in the information measures themselves. This is achieved by assigning a greater "weight" to the information obtained about regimens with characteristics close to $\boldsymbol{\gamma}$. Another difference to the majority of information-theoretic approaches is that the design is based on the so-called "patients' gain" and allocates each patient to the regimen considered to be the optimal one while taking into account the uncertainty. This leads to fulfilling of statistical goals of the experiment under ethical constraints.

The proposed criterion is not restricted to a particular model and can be used, for examples, to govern selection within traditional parametric designs. However,

motivated by relaxing parametric and monotonicity assumptions, we demonstrate that good operating characteristics of the design can be achieved without employing these assumptions. We study the asymptotic behaviour of the design and compare the performance to the currently used methods in contexts of Phase I, Phase II and complex combination-schedule Phase I/II clinical trials.

4.2 Information-Theoretic Criterion

The novel design is based on maximisation of the information gain in the experiment with an interval of the specific interest (the neighbourhood of $\boldsymbol{\gamma}$). Below, we derive an explicit formula for the information gain in such a trial with multinomial outcomes.

4.2.1 Derivation

Consider a discrete random variable taking one of d values and a corresponding random probability vector $\mathbf{Z} = \left[Z^{(1)}, Z^{(2)}, \ldots, Z^{(d)}\right] \in \mathbb{S}^d$ defined on a unit simplex

$$\mathbb{S}^d = \{\mathbf{Z} : Z^{(1)} > 0, Z^{(2)} > 0, \ldots, Z^{(d)} > 0; \sum_{i=1}^{d} Z^{(i)} = 1\}. \tag{4.2.1}$$

Assume that $\mathbf{Z}$ has a prior Dirichlet distribution $\mathrm{Dir}(\mathbf{v} + \mathbf{J})$ where parameters $\mathbf{v} = \left[v^{(1)}, \ldots, v^{(d)}\right]^{\mathrm{T}} \in \mathbb{R}_+^d$, $\sum_{i=1}^{d} v^{(i)} = \beta$ and $\mathbf{J}$ is a d-dimensional unit vector. After n realizations of a discrete random variable in which $x^{(i)}$ outcomes of i are observed, $i = 1, \ldots, d$, the random vector $\mathbf{Z}_n$ has a Dirichlet posterior distribution with a probability density function (pdf)

$$f_n(\mathbf{p}|\mathbf{x}) = \frac{1}{B(\mathbf{x}+\mathbf{v}+\mathbf{J})} \prod_{i=1}^{d} \left(p^{(i)}\right)^{x^{(i)}+v^{(i)}}, B(\mathbf{x}+\mathbf{v}+\mathbf{J}) = \frac{\prod_{i=1}^{d} \Gamma(x^{(i)}+v^{(i)}+1)}{\Gamma\left(\sum_{i=1}^{d}(x^{(i)}+v^{(i)}+1)\right)} \tag{4.2.2}$$

where $\mathbf{p} = \left[p^{(1)}, \ldots, p^{(d)}\right]^{\mathrm{T}}$, $\mathbf{x} = \left[x^{(1)}, \ldots, x^{(d)}\right]$, $\sum_{i=1}^{d} x^{(i)} = n$, $0 < p^{(i)} < 1$, $\sum_{i=1}^{d} p^{(i)} = 1$ and $B(\mathbf{x}+\mathbf{v}+\mathbf{J})$ is the Beta function and $\Gamma(x)$ is the Gamma function.

Let $\boldsymbol{\alpha} = \left[\alpha^{(1)}, \ldots, \alpha^{(d)}\right]^{\mathrm{T}} \in \mathbb{S}^d$ be the vector in the neighbourhood of which f_n concentrates as $n \to \infty$. A classic question of interest in this setting is to estimate the probability vector $\boldsymbol{\alpha}$. The information required to answer the estimation question can be measured by the Shannon differential entropy of f_n (Cover and Thomas, 2012)

$$h(f_n) = -\int_{\mathbb{S}^d} f_n(\mathbf{p}|\mathbf{x}) \log f_n(\mathbf{p}|\mathbf{x}) \mathrm{d}\mathbf{p} \tag{4.2.3}$$

with convention $0 \log 0 = 0$. The classic formulation of the estimation question, however, does not take into account the fact that an investigator would like to find the TR having characteristics $\boldsymbol{\gamma} = \left[\gamma^{(1)}, \ldots, \gamma^{(d)}\right] \in \mathbb{S}^d$. Nor does it reflect that one would like to have a more precise estimation of the vector $\boldsymbol{\alpha}$ if it is close to $\boldsymbol{\gamma}$ only. This is a consequence of the fact that the classic information measures do not depend on the nature of outcomes $\mathbf{p}$, but on their probabilities $f(\mathbf{p})$ and, therefore, are called *context-free* (Suhov et al., 2016). While it gives the notion of information a great flexibility which explains its successful applications in various fields, the context-free nature might be considered as a drawback in many application areas, e.g. in clinical trials, as it is demonstrated below.

To take into account the context of the experiment and the nature of outcomes $\mathbf{p}$, one can consider an estimation experiment with "sensitive" area (i.e. the neighbourhood of $\boldsymbol{\gamma}$). The information required in such an experiment can be measured by the *weighted* Shannon differential entropy (Belis and Guiasu, 1968; Clim, 2008; Kelbert and Mozgunov, 2015; Suhov et al., 2016; Kelbert and Mozgunov, 2017) of f_n with a positive weight function $\phi_n(\mathbf{p})$

$$h^{\phi_n}(f_n) = -\int_{\mathbb{S}^d} \phi_n(\mathbf{p}) f_n(\mathbf{p}|\mathbf{x}) \log f_n(\mathbf{p}|\mathbf{x}) \mathrm{d}\mathbf{p}. \tag{4.2.4}$$

The crucial difference between the information measures given in Equation (4.2.3) and Equation (4.2.4) is the weight function, $\phi_n(\mathbf{p})$, which emphasizes the interest in the neighbourhood of $\boldsymbol{\gamma}$ rather than on the whole simplex $\mathbb{S}^d$. It reflects that

the information about the probability vector which lies in the neighbourhood of $\boldsymbol{\gamma}$ is more valuable in the experiment.

Due to the limited sample size in actual studies, an investigator is typically interested in answering the question: which regimen has an associated probability vector closest to $\boldsymbol{\gamma}$ while ensuring an accurate estimation in the neighbourhood of the TR only. For this question, the information gain from considering the experiment with the sensitive area equals to

$$\Delta_n = h(f_n) - h^{\phi_n}(f_n). \tag{4.2.5}$$

Following the information gain approach, the first term in the equation above is the information in a classic experiment using the context-free measure, while the second (novel) term is the information when the context of events is taken into account. Alternatively, Δ_n can be considered as an average amount of the additional statistical information required when considering the context-dependent estimation problem instead of the traditional one.

The information gain in Equation (4.2.5) requires a weight function which defines the "value" of the information in different areas of the simplex $\mathbb{S}^d$ to be specified. To track the influence of the weight function explicitly we consider a weight function in the Dirichlet form:

$$\phi_n(\mathbf{p}) = C(\mathbf{x}, \boldsymbol{\gamma}, n) \prod_{i=1}^{d} \left(p^{(i)}\right)^{\gamma^{(i)} n^{\kappa}} \tag{4.2.6}$$

where $\kappa \in (0, 1)$ is a parameter and $C(\mathbf{x}, \boldsymbol{\gamma}, n)$ is a constant which is chosen to satisfy the normalization condition $\int_{\mathbb{S}^d} \phi_n(\mathbf{p}) f_n(\mathbf{p}|\mathbf{x}) \mathrm{d}\mathbf{p} = 1$. The parameter κ is restricted to the unit interval to ensure an asymptotically unbiased estimate of the vector $\boldsymbol{\alpha}$: $\lim_{n\to\infty} \int_{\mathbb{S}^d} \mathbf{p}\phi_n(\mathbf{p}) f_n(\mathbf{p}) \mathrm{d}\mathbf{p} = \boldsymbol{\alpha}$. This emphasises the interest in the identification of the TR for the small and moderate sample sizes typical for clinical

trials. The form of the weight function for $d = 2$ is demonstrated in Figure 4.1.

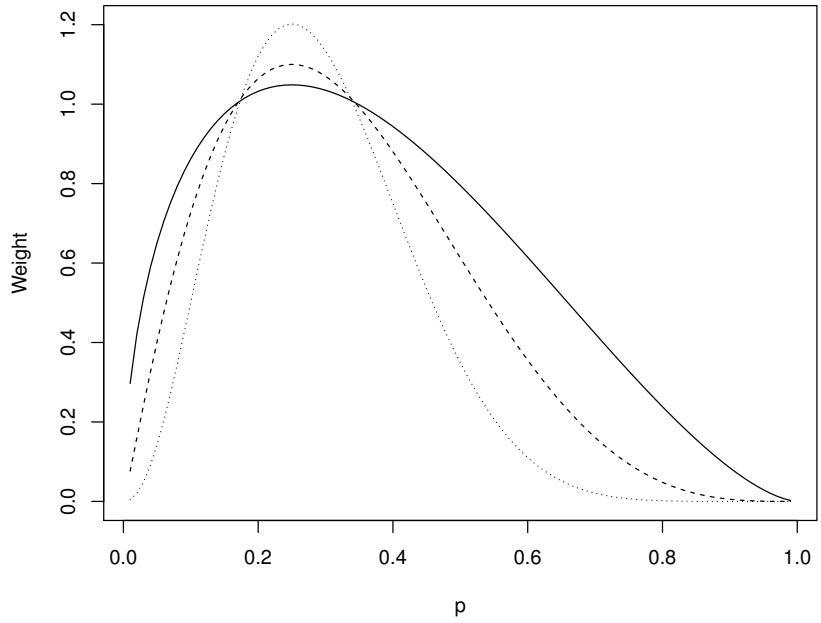

Figure 4.1: Weight function $\phi_n(p)$ given in Equation 4.2.6 in case $d = 2$ using $\kappa = 0.25$ (solid line), $\kappa = 0.50$ (dashed line), $\kappa = 0.75$ (dotted line), the target value $\gamma = 0.25$ and $n = 20$.

The asymptotic behaviour of the information gain, Δ_n, for the family of weight functions (4.2.6) is studied in Theorem 3.

Theorem 3. *Let $h(f_n)$ and $h^{\phi_n}(f_n)$ be the standard and weighted differential entropies of the pdf (4.2.2) with the weight function (4.2.6). Let $\lim_{n\to\infty} \frac{x^{(i)}(n)}{n} = \alpha^{(i)}$ for $i = 1, 2, \ldots, d$ and $\sum_{i=1}^{d} x^{(i)} = n$, then as $n \to \infty$*

$$\Delta_n = \begin{cases} O\left(\frac{1}{n^{1-2\kappa}}\right) \text{ if } \kappa < \frac{1}{2}; \\ -\frac{1}{2}\left(\sum_{i=1}^{d} \frac{\left(\gamma^{(i)}\right)^2}{\alpha^{(i)}} - 1\right) n^{2\kappa-1} + c(\boldsymbol{\alpha}, \boldsymbol{\gamma}, \kappa, n) + O\left(\frac{1}{n^{\eta(1-\kappa)-\kappa}}\right) \text{ if } \kappa \geq \frac{1}{2} \end{cases}$$

where

$$c(\boldsymbol{\alpha}, \boldsymbol{\gamma}, \kappa, n) = \sum_{j=3}^{\eta} \frac{(-1)^{j-1}}{j} n^{j\kappa-j+1} \left(\sum_{i=1}^{d} \frac{\left(\gamma^{(i)}\right)^j}{\left(\alpha^{(i)}\right)^{j-1}} - 1\right) \text{ and } \eta = \lfloor (1-\kappa)^{-1} \rfloor.$$

Proof. The differential entropy of the pdf f_n can found as

$$h(f_n) = -\int_{\mathbb{S}^d} f_n(\mathbf{p}) \log \left[\frac{1}{B(\mathbf{x}+\mathbf{v}+\mathbf{J})} \prod_{i=1}^{d} \left(p^{(i)}\right)^{x^{(i)}+v^{(i)}} \right] \mathrm{d}\mathbf{p}$$
$$= -\log \left[\frac{1}{B(\mathbf{x}+\mathbf{v}+\mathbf{J})} \right] - \sum_{i=1}^{d} \int_{\mathbb{S}^d} f_n(\mathbf{p}) \log \left[\left(p^{(i)}\right)^{x^{(i)}+v^{(i)}} \right] \mathrm{d}\mathbf{p}.$$

The last term is the sum of integrals having the same structure. For the sake of notations we assume $x^{(i)} + v^{(i)} = x^{(i)}$ which will not affect the asymptotic results as $v^{(i)}$ does not depend on n. Then, i^{th} integral can be written as

$$\int_{\mathbb{S}^d} \log \left(p^{(i)}\right)^{x^{(i)}} f_n(\mathbf{p}) \mathrm{d}\mathbf{p} = x^{(i)} \int_{\mathbb{S}^d} \log \left(p^{(i)}\right) f_n(\mathbf{p}) \mathrm{d}\mathbf{p}.$$

Note that integrating over $d-1$ variables does not change, the only difference is in the i^{th} term. The integral can be computed as follows (Gradshteyn and Ryzhik, 2014, Formula 4.253.1)

$$x^{(i)} \int_{\mathbb{S}^d} \log \left(p^{(i)}\right) f_n(\mathbf{p}) \mathrm{d}\mathbf{p} = x^{(i)} \left(\psi \left(x^{(i)} + 1 \right) - \psi \left(n + d \right) \right), \ i = 1, 2, \ldots, d,$$

where $\psi(x) = \frac{\mathrm{d}}{\mathrm{d}x} \log \Gamma(x)$ is the digamma function. Putting all terms together,

$$h\left(f_n\right) = \log B\left(\mathbf{x}+\mathbf{v}+\mathbf{J}\right) + n\psi\left(n+d\right) - \sum_{i=1}^{d} x^{(i)} \psi\left(x^{(i)}+1\right).$$

To obtain the asymptotics of $h(f_n)$ as $n \to \infty$, the asymptotics of digamma function

$$\psi(n) = \log(n) - \frac{1}{2n} + O\left(\frac{1}{n^2}\right),$$

and the Stirling formula (for the first term in the differential entropy)

$$n! = \sqrt{2\pi n} \left(\frac{n}{e}\right)^n \left(1 + \frac{1}{12n} + O\left(\frac{1}{n^2}\right)\right)$$

are used. Using that $x^{(i)} \simeq \alpha^{(i)} n$ and putting all terms together, one can obtain

that

$$h\left(f_n\right) = \frac{1}{2}\log\left(2\pi e\right)^{d-1} + \frac{1}{2}\log\frac{\prod_{i=1}^{d}\alpha^{(i)}}{n^{d-1}} + O\left(\frac{1}{n}\right), \text{ as } n \to \infty.$$

Similarly, the weighted differential entropy of f_n takes the form

$$\begin{aligned} h^{\psi_n}(f_n) &= -\int_{\mathbb{S}^d} \phi_n(\mathbf{p}) f_n(\mathbf{p}) \log\left[\frac{1}{B(\mathbf{x}+\mathbf{v}+\mathbf{J})}\prod_{i=1}^{d}\left(p^{(i)}\right)^{x^{(i)}+v^{(i)}}\right]\mathrm{d}\mathbf{p} \\ &= -\log\left[\frac{1}{B(\mathbf{x}+\mathbf{v}+\mathbf{J})}\right] - \sum_{i=1}^{d}\int_{\mathbb{S}^d}\phi_n(\mathbf{p})f_n(\mathbf{p})\log\left[\left(p^{(i)}\right)^{x^{(i)}+v^{(i)}}\right]\mathrm{d}\mathbf{p}. \end{aligned}$$

Then, each integral can be computed as

$$x^{(i)}\int_{\mathbb{S}^d}\log\left(p^{(i)}\right)\phi_n(\mathbf{p})f_n(\mathbf{p})\mathrm{d}\mathbf{p} = x^{(i)}\left(\psi\left(x^{(i)}+\gamma^{(i)}n^{\kappa}+1\right) - \psi\left(n+n^{\kappa}+d\right)\right).$$

Again, using the asymptotics of the digamma function one can obtain that

$$\begin{aligned} h^{\phi_n}\left(f_n\right) &= \frac{1}{2}\log\left(2\pi e\right)^{d-1} + \frac{1}{2}\log\frac{\prod_{i=1}^{d}\alpha^{(i)}}{n^{d-1}} + \\ &+ \sum_{j=1}^{+\infty}\frac{(-1)^{j-1}}{j}n^{j\kappa-j+1}\left(\sum_{i=1}^{d}\frac{\left(\gamma^{(i)}\right)^{j}}{\left(\alpha^{(i)}\right)^{j-1}} - 1\right) + O\left(\frac{1}{n^{\kappa}}\right). \end{aligned}$$

The first two terms are the same as in the differential entropy, and the third term, the sum, is to be considered in more details. For $j = 1$, the constant in brackets is cancelled out by the construction of the vector $\boldsymbol{\gamma}$

$$\sum_{i=1}^{d}\frac{\left(\gamma^{(i)}\right)^{j}}{\left(\alpha^{(i)}\right)^{j-1}} - 1 = \sum_{i=1}^{d}\gamma^{(i)} - 1 = 0.$$

Then, if $\kappa < \frac{1}{2}$ all terms decay to zero as $n \to +\infty$. If $\kappa \geq \frac{1}{2}$, first $\eta = \lfloor (1-\kappa)^{-1} \rfloor$ terms are non-decaying and the rest goes to zero as $n \to +\infty$. Finally, writing the term of the sum for $j = 2$ separately, the result immediately follows. □

The information gain, Δ_n, tends to 0 for $\kappa < 1/2$ which implies that assigning a value of information with a rate less than $1/2$ is insufficient to emphasise the

importance of the study context. However, the limit is non-zero for $\kappa \geq 1/2$. Following the conventional information gain approach, one would like to make a decision which maximises the statistical information in the experiment. The information gain Δ_n is always non-positive and for any fixed n its asymptotics achieves the maximum value 0 at the point $\alpha^{(i)} = \gamma^{(i)}$, $i = 1, \ldots, d$ (all constants are cancelled out). Indeed, Theorem 3 implies that when maximising the information gain Δ_n, one tends to collect more information about the regimen which has characteristics $\boldsymbol{\alpha}$ closer to the target $\boldsymbol{\gamma}$. To keep the solution easily interpretable in applications, we construct the regimen selection criterion using the leading term of the asymptotic expression for Δ_n in Theorem 3:

$$\delta^{(\kappa)}(\boldsymbol{\alpha}, \boldsymbol{\gamma}) := \frac{1}{2} \left(\sum_{i=1}^{d} \frac{\left(\gamma^{(i)}\right)^2}{\alpha^{(i)}} - 1 \right) n^{2\kappa - 1}. \tag{4.2.7}$$

Note that maximising the leading term of the information gain asymptotics is equivalent to minimising $\delta^{(\kappa)}(\boldsymbol{\alpha}, \boldsymbol{\gamma})$. Criterion (4.2.7) can be considered as the measure of the divergence between $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ and the criterion which governs the selection such that the information gain is maximised. The criterion (4.2.7) is intuitive as it reflects explicitly the fact that an investigator tends to collect more information about the regimen with the probability vector close to $\boldsymbol{\gamma}$, and also shares some desirable properties. Clearly, $\delta^{(\kappa)}(\cdot) \geq 0$ and $\delta^{(\kappa)}(\cdot) = 0$ iff $\boldsymbol{\alpha} = \boldsymbol{\gamma}$ for all κ and n. The boundary values $\alpha^{(i)} = 0$, $i = 1, \ldots, d$ correspond to infinite values of $\delta^{(\kappa)}(\boldsymbol{\alpha}, \boldsymbol{\gamma})$ which is one of the important properties for functions defined on simplex $\mathbb{S}^d$ as advocated by Aitchison (1992) and in Chapter 3. We construct the design based on the selection criterion (4.2.7) below.

4.2.2 Selection Criterion

Consider a discrete set of m regimens, $T_1, \ldots, T_m$, associated with probability vectors $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_m$ and $n_1, \ldots, n_m$ observations. The regimen T_j is optimal if it satisfies $\delta^{(\kappa)}(\boldsymbol{\alpha}_j, \boldsymbol{\gamma}) = \inf_{i=1,\ldots,m} \delta^{(\kappa)}(\boldsymbol{\alpha}_i, \boldsymbol{\gamma})$. To estimate $\delta^{(\kappa)}(\boldsymbol{\alpha}_i, \boldsymbol{\gamma})$ a random variable $\tilde{\delta}_{n_i}^{(\kappa)} \equiv \delta^{(\kappa)}(\mathbf{Z}_{n_i}, \gamma)$ is introduced where $\mathbf{Z}_{n_i}$ has the Dirichlet distribution

given in Equation (4.2.2). Let us fix a regimen $\mathbf{Z}_{n_i} \equiv \mathbf{Z}_n$ and denote $\tilde{\delta}_{n_i}^{(\kappa)} \equiv \tilde{\delta}_n^{(\kappa)}$. It is a known that a Dirichlet random variable (after appropriate normalisation) weakly converges to a multivariate normal distribution. In fact, a stronger result can be shown using the Kullback-Leibler distance

$$\mathbb{D}(f \mid\mid g) \;=\; \int_{\mathbb{R}} f(x) \log \frac{f(x)}{g(x)} \mathrm{d}x$$

where g and f are probability density functions.

Theorem 4. *Let $\widetilde{\boldsymbol{Z}}_n = \Sigma^{-1/2}\left(\boldsymbol{Z}_n - \boldsymbol{\alpha}\right)$ be a random variable with pdf $\widetilde{f}_n$ where pdf of $\boldsymbol{Z}_n$ is given in Equation (4.2.2) with $\lim_{n\to\infty} \frac{x^{(i)}(n)}{n} = \alpha^{(i)}$ for $i = 1, 2, \ldots, d$, $\sum_{i=1}^{d} x^{(i)} = n$ and where Σ is a d-dimensional square matrix with elements $\Sigma_{[ij]} = \frac{\alpha^{(i)}(1-\alpha^{(i)})}{n}$ if $i = j$ and $\Sigma_{[ij]} = -\frac{\alpha^{(i)}\alpha^{(j)}}{n}$ if $i \neq j$. Let $\overline{\boldsymbol{Z}}$ be the multivariate Gaussian random variable $\mathcal{MN}\left(0, I_{d-1}\right)$ (I_{d-1} is the $(d-1)$-dimensional identity matrix) with pdf φ and the differential entropy $h(\varphi) = \frac{1}{2}\log\left((2\pi e)^{d-1}\right)$. Then the Kullback-Leibler divergence of φ from $\tilde{f}_n$ tends to 0 as $n \to \infty$ which implies that $\widetilde{\boldsymbol{Z}}_n$ weakly converges to $\overline{\boldsymbol{Z}}$.*

Proof. Following the result in Theorem 3, as $n \to \infty$

$$h\left(f_n\right) = \frac{1}{2}\log\left(2\pi e\right)^{d-1} + \frac{1}{2}\log\frac{\prod_{i=1}^{d}\alpha^{(i)}}{n^{d-1}} + O\left(\frac{1}{n}\right).$$

Using that $h\left(\widetilde{f}_n\right) = h\left(f_n\right) + \log\left|det\left(\Sigma^{-1/2}\right)\right|$ (Cover and Thomas, 2012) where $det(A)$ is a determinant of the matrix A and

$$\log\left|det\left(\Sigma^{-1/2}\right)\right| = -\frac{1}{2}\log\frac{\prod_{i=1}^{d}\alpha^{(i)}}{n^{d-1}} + O\left(\frac{1}{n}\right) \text{ as } n \to \infty.$$

Then,

$$\mathbb{D}(\widetilde{f}_n \mid\mid \varphi) = -h\left(\widetilde{f}_n\right) - \int_{\mathbb{S}^d} \widetilde{f}_n(\mathbf{p}) \log\varphi(\mathbf{p}) \mathrm{d}\mathbf{p} =$$

$$-\frac{1}{2}\log\left(2\pi e\right)^{d-1} + \frac{1}{2}\log\left(2\pi\right)^{d-1} + \frac{1}{2}\int_{\mathbb{S}^d}\sum_{i=1}^{d-1}\left(p^{(i)}\right)^2 \widetilde{f}_n(\mathbf{p})\mathrm{d}\mathbf{p} + O\left(\frac{1}{n}\right) = O\left(\frac{1}{n}\right),$$

as $\int_{\mathbb{S}^d} \sum_{i=1}^{d-1} \left(p^{(i)}\right)^2 \widetilde{f}_n(\mathbf{p})\mathrm{d}\mathbf{p} = d - 1 + O\left(\frac{1}{n}\right)$ is the sum of the second moments. Pinsker's inequality (Csiszar and Körner, 2011) implies the convergence in the total variation (Cover and Thomas, 2012) which implies the weak convergence. □

Using Theorem 4 the following result can be obtained for the proposed criterion.

Theorem 5. *Let* $\nabla\delta^{(\kappa)}(\boldsymbol{z},\boldsymbol{\gamma}) = \left[\frac{\partial\delta^{(\kappa)}(\boldsymbol{z},\boldsymbol{\gamma})}{\partial z^{(1)}},\ldots,\frac{\partial\delta^{(\kappa)}(\boldsymbol{z},\boldsymbol{\gamma})}{\partial z^{(d)}}\right]^{\mathrm{T}}$, $\bar{\Sigma} = \nabla_{\boldsymbol{\alpha}}^{\mathrm{T}}\Sigma\nabla_{\boldsymbol{\alpha}}$, $\nabla_{\boldsymbol{\alpha}} \equiv \nabla\delta^{(\kappa)}(\boldsymbol{z},\boldsymbol{\gamma})$ *evaluated at* $\boldsymbol{z} = \boldsymbol{\alpha}$ *and* $\bar{\delta}_n^{(\kappa)} = \bar{\Sigma}^{-1/2}\left(\delta^{(\kappa)}(\boldsymbol{Z}_n,\boldsymbol{\gamma}) - \delta^{(\kappa)}(\boldsymbol{\alpha},\boldsymbol{\gamma})\right)$. *Let* $\bar{Z}$ *be a standard Gaussian RV. Then, under the assumptions of Theorem 4,*

$$\lim_{n\to\infty}\mathbb{E}\tilde{\delta}_n^{(\kappa)} = \delta^{(\kappa)}(\boldsymbol{\alpha},\boldsymbol{\gamma}),\ \lim_{n\to\infty}\mathbb{V}\tilde{\delta}_n^{(\kappa)} = 0,$$

and $\bar{\delta}_n^{(\kappa)}$ *weakly convergences to* $\bar{Z}$.

Proof. The statements about expectation and variance are straightforward to prove by computing the first and the second moments of $\delta^{(\kappa)}(\mathbf{Z}_n,\boldsymbol{\gamma})$. The weak convergence result follows from Theorem 4 and by applying the delta method (Resnick, 2013). □

A single summary statistic for $\delta^{(\kappa)}(\mathbf{Z}_n,\boldsymbol{\gamma})$ is needed to select the most promising regimen in the sequential experiment. We will focus on an intuitively clear and simple "plug-in" estimator, $\hat{\delta}^{(\kappa)}(\hat{\mathbf{p}}_n,\gamma) \equiv \hat{\delta}_n^{(\kappa)}$ with $\hat{\mathbf{p}}_n = [\hat{p}_n^{(1)},\ldots,\hat{p}_n^{(i)},\ldots,\hat{p}_n^{(d)}]$ and $\hat{p}_n^{(i)} = \frac{x^{(i)}+v^{(i)}}{n+\beta^{(i)}}$, $i = 1,\ldots,d$, the mode of the posterior Dirichlet distribution. The estimator for the regimen T_j takes the form

$$\hat{\delta}_{n_j}^{(\kappa)} = \delta^{(\kappa)}(\hat{\mathbf{p}}_{n_j},\boldsymbol{\gamma}) = \frac{1}{2}\left(\sum_{i=1}^{d}\frac{\left(\gamma^{(i)}\right)^2}{\hat{p}_{n_j}^{(i)}} - 1\right) n_j^{2\kappa-1},\ j = 1,2,\ldots,m. \tag{4.2.8}$$

By Theorem 5 for any $\varepsilon > 0$ $\lim_{n_j\to\infty}\mathbb{P}\left(\tilde{\delta}_{n_j}^{(\kappa)} \in [\hat{\delta}_{n_j}^{(\kappa)} - \varepsilon, \hat{\delta}_{n_j}^{(\kappa)} + \varepsilon]\right) = 1$. The statistic (4.2.8) is used to govern the selection among regimens during the experiment. Note that the estimator above requires a vector of prior parameters $\mathbf{v}_j$,

$j = 1, \ldots, m$ to start the experiment. This choice implies an initial ordering in which an investigator would like to test the regimens before the data is available.

4.2.3 Specific Assignment Rules

The criterion (4.2.7) summarizes the regimen's characteristics and can be applied to different types of sequential experiments. We consider two assignment rules: Rule I which randomises between regimens and Rule II which selects the "best" regimen. These rules follow the setting of the motivating clinical trials: Rule II is widely used in Phase I trials where the randomisation to all doses is not ethical or in the typical MAB setting where the primary goal is to maximise the number of successes. On the other hand, Rule I is often used in Phase II trials to learn more about all regimens allowing to decrease the probability of identifying a suboptimal regimen (Thall and Wathen, 2007).

Rule I: Randomisation

Under Rule I, the regimen selected next in the experiment is randomised with probabilities

$$\tilde{w}_j \equiv \frac{1/\tilde{\delta}_{n_j}^{(\kappa)}}{\sum_{i=1}^{m} 1/\tilde{\delta}_{n_i}^{(\kappa)}}, \quad j = 1, \ldots, m. \tag{4.2.9}$$

Note that $\mathbb{P}(\tilde{w}_j = 0) = 0$ meaning that there is a non-zero probability to assign a next patient to the regimen T_j. From Theorem 5 one can show that

$$w_j = \lim_{n_1, n_2, \ldots, n_m \to \infty} \mathbb{E}(\tilde{w}_j) = \frac{1/\delta_{n_j}^{(\kappa)}}{\sum_{i=1}^{m} 1/\delta_{n_i}^{(\kappa)}}. \tag{4.2.10}$$

When no observations have yet been collected, the procedure randomises according to the criterion based on the prior distribution alone, $\tilde{\delta}_{\beta_j}^{(\kappa)}$, $j = 1, \ldots, m$. Then, given n_j observations, $\mathbf{x}_j$ outcomes for regimen T_j, $j = 1, \ldots, m$ and using the "plug-in" estimator (4.2.8), the regimen T_j is selected with probability $\hat{w}_j = 1$ if

$\hat{\delta}_{n_j}^{(\kappa)} = 0$ and with probability

$$\hat{w}_j = \frac{1/\hat{\delta}_{n_j}^{(\kappa)}}{\sum_{i=1}^{m} 1/\hat{\delta}_{n_i}^{(\kappa)}} \text{ if } \hat{\delta}_{n_i}^{(\kappa)} > 0,\ i = 1, \ldots, m. \tag{4.2.11}$$

The method proceeds until N observations are assigned in the experiment. The regimen $T_{j^\star}$ satisfying

$$\hat{\delta}_{N_{j^\star}}^{(1/2)} = \inf_{j=1,\ldots,m} \hat{\delta}_{N_j}^{(0.5)}. \tag{4.2.12}$$

is adopted for the final recommendation, where N_j is a total number of observation on T_j, $j = 1, 2, \ldots, m$. The value $\kappa = 0.5$ in Equation (4.2.12) is used to ensure that the final recommendation is not penalised by the sample size (see Section 4.2.4).

Rule II: Select the best

Let N be a total sample size and begin with the experiment with the regimen that minimises $\tilde{\delta}_{\beta_j}^{(\kappa)}$, $j = 1, \ldots, m$. Given n_j observations, $\mathbf{x}_j$ outcomes for the regimen T_j, $j = 1, \ldots, m$ and using the "plug-in" estimator, T_j is selected if it satisfies $\hat{\delta}_{n_j}^{(\kappa)} = \inf_{i=1,\ldots,m} \hat{\delta}_{n_i}^{(\kappa)}$. The method proceeds until N observations are assigned in the experiment. Again, we adopt $T_{j^\star}$ as in Equation (4.2.12) for the final recommendation.

4.2.4 Criterion in the Context of Clinical Trials

In examples, we apply the novel selection criterion to Phase I and Phase II clinical trials. In these cases, the regimens are the different treatments (doses, combinations, schedules, etc.) and the goal is to find the regimen corresponding to the specific toxicity (efficacy) characteristics. As the proposed information gain and the corresponding selection criterion tends to assign the next patients to the best estimated TR during the trial, the criterion based on Δ_n is the *patient's gain* (also known as *best intention*) criterion as classified by Whitehead and Williamson (1998). The balance in the "exploration vs exploitation" trade-off is tuned by the term $n_j^{2\kappa-1}$ reflecting the penalty on the number of observations on the same regimen (i.e. many observations on one regimen would favour selection of other

regimens). This implies that the design will keep selecting a specific regimen only, if the corresponding estimate, $\boldsymbol{\alpha}$ is close to $\boldsymbol{\gamma}$. As the trial progresses the design requires an increasing level of confidence that the selected regimen is the TR. Clearly, $\kappa = 1/2$ corresponds to no penalty and is of particular interest in trials with small sample sizes while larger values of $\kappa > 1/2$ correspond to a greater interest in the statistical power of the experiment. Importantly, the first term in Equation (4.2.7) guarantees that the vast majority of patients will be assigned to the TR even for $\kappa > 1/2$. Clearly, one is not interested in the penalty terms for the final recommendations, therefore, $\kappa = 1/2$ is proposed in Equation (4.2.12).

As many Phase I and Phase II clinical trials consider a binary endpoint $d = 2$ (toxicity: yes/no or response: yes/no), we start with this case in applications of the proposed design. In this case, the pdf (4.2.2) reduces to the Beta distribution and criterion (4.2.8) takes the form

$$\delta_{n_j}^{(\kappa)}(\alpha, \gamma) = \frac{1}{2} \frac{(\alpha - \gamma)^2}{\alpha(1-\alpha)} n_j^{2\kappa - 1}$$

where α is the probability of event (toxicity/efficacy) and γ is the target. The corresponding estimator takes the form

$$\hat{\delta}_{n_j}^{(\kappa)}(\alpha, \gamma) = \frac{1}{2} \frac{(\hat{p}_{n_j} - \gamma)^2}{\hat{p}_{n_j}(1 - \hat{p}_{n_j})} n_j^{2\kappa - 1} \tag{4.2.13}$$

where

$$\hat{p}_{n_j} = \frac{x_j + v_j}{n_j + \beta_j} \tag{4.2.14}$$

is a plug-in estimate, x_j is number of toxicity/efficacy responses on regimen T_j after n_j patients were assigned to it and v_j, β_j are parameters of prior Beta distributions. Note that the first term in the criterion (4.2.13) is the unit interval symmetric loss function (3.3.4) which was shown to have desirable properties on a simplex. The Euclidean distance term in the numerator of the criterion tends to assign patients to the TR. At the same time, the denominator takes into account the uncertainty

in the denominator which is a variance of the probability of a binary event. In the context of regimen finding trials, the denominator can be also considered as a penalty term which "drives away" the allocation from the bounds ($\hat{p}_{n_j} = 0$ and $\hat{p}_{n_j} = 1$) - the boundary values correspond to infinite values of the criterion. As pointed out in Section 3.5, the maximum variance of the binary probability is achieved at $\hat{p}_{n_j} = 0.5$, so the criterion favours greater values of $\hat{p}_{n_j}$ which can be unethical if the objective of a study is to control the toxicity. We will study whether the construction of the criterion creates any practical limitation in the context of Phase I clinical trial without the monotonicity assumption in Section 4.3.

Further, we also consider the case of $d = 3$ in the context of Phase I/II regimen finding clinical trials with binary toxicity and binary efficacy endpoints. In this case, there are three outcomes: (i) "efficacy and no toxicity", (ii) "no efficacy and no toxicity" and (iii) "toxicity". Outcomes "toxicity and no efficacy" and "toxicity and efficacy" are combined assuming that an efficacy response can only be observed when no toxicity occurs. Due to the limited sample size, we will consider the case of no penalty for the sample size, $\kappa = 1/2$. Then, the selection criterion takes the form

$$\delta\left(\boldsymbol{\alpha}, \boldsymbol{\gamma}\right) := \frac{\gamma_1^2}{\alpha_1} + \frac{\gamma_2^2}{\alpha_2} + \frac{(1 - \gamma_1 - \gamma_2)^2}{1 - \alpha_1 - \alpha_2} - 1, \tag{4.2.15}$$

where $\alpha_1, \alpha_2, \alpha_3$ are the probabilities of the responding events and γ_1, γ_2 and γ_3 are the targets. The criterion (4.2.15) is the sum of three contributions corresponding to each of three events considered. Terms in the criterion (4.2.15) have the following interpretations:

- When α_1 tends to 0 the regimen is either inefficacious or (and) highly toxic. Then, the value of the trade-off function tends to infinity meaning that the treatment should be avoided.

- When α_2 tends 0 the regimen is either highly efficacious or (and) highly toxic. Then, this term penalises a high toxicity regardless of a high efficacy

as the trade-off function tends to infinity. This term prevents a too quick escalation to highly toxic regimens.

- When $1 - \alpha_1 - \alpha_2$ tends 0, the regimen is associated with nearly no toxicity. However, the TR is expected to be associated with a non-zero toxicity. Then, these terms drive the regimen allocation away from the under-dosing regimens.

Note that all terms are dependent - an increase in one leads to decreases in others and the optimal value is attained at the point of target characteristics only. Again, the denominators "drive away" the selection from inefficacious and highly toxic regimens and concentrate the allocation in the neighbourhood of the TR.

4.2.5 Asymptotic Behaviour

Considering the asymptotic behaviour of a procedure ensures that an experimental design becomes more accurate as a sample size grows (Azriel et al., 2011). Recall that the goal of the sequential experiment is to find the regimen j which corresponds to the minimum value $\delta_j \equiv \delta(\boldsymbol{\alpha}_j, \boldsymbol{\gamma})$ among all regimens using random variables $\tilde{\delta}_i^{(\kappa)} = \delta^{(\kappa)}(\mathbf{Z}_{n_i}, \boldsymbol{\gamma}),\ i = 1, \ldots, m$. Denote the index of the regimen to be selected by $\nu = \arg\min_{i=1,\ldots,m} \tilde{\delta}_i^{(\kappa)}$. Below, we consider the risk-adjusted average approach (Polley and Cheung, 2008) and the probability of correct selection (PCS) (Cheung, 2013)

$$A_N = \frac{1}{m} \sum_{j=1}^{m} \mathbb{P}_{\pi_j}(\nu = j)$$

where $\mathbb{P}_{\pi_j}$ is the probability computed under the vector $\pi_j = [\alpha_{1,j}, \ldots, \alpha_{m,j}]^T$, $1 \leq j \leq m$, which assumes that $\alpha_{j,j}$ corresponds to the TR. We would say that the design is consistent if $\lim_{N\to\infty} A_N = 1$. The main result of this section is formulated in Theorem 6.

Theorem 6. *Let us consider the experimental design with the selection criterion based on $\tilde{\delta}_{n_i}^{(\kappa)}$, m regimens and true probabilities vectors $\boldsymbol{\alpha}_i$, $i = 1, \ldots, m$. Then,*

(a) *The design is consistent under assignment Rule I for* $\kappa \geq 0.5$.
(b) *The design is consistent under assignment Rule II for* $\kappa > 0.5$.

Proof. **(a) Rule I.** Under Rule I proportions of observations on each regimen converges to a constant as follows from Equation (4.2.10). Therefore, one can assume that it is fixed and the probability measure below is conditional on a non-zero allocation proportion.

We start from $\kappa = 1/2$ and adopt notation $\tilde{\delta}_{n_i}^{(1/2)} \equiv \tilde{\delta}_i$. Denoting

$$\bar{C}_{k,k+1} \equiv \{\tilde{\delta}_k < \tilde{\delta}_{k+1}\} \text{ and } C_{k,k+1} \equiv \{\tilde{\delta}_k > \tilde{\delta}_{k+1}\}$$

we find that

$$\{\nu = k\} \Leftrightarrow \{\bar{C}_{k,1} \cap \bar{C}_{k,2} \cap \ldots \cap \bar{C}_{k,k-1} \cap \bar{C}_{k,k+1} \cap \ldots \cap \bar{C}_{k,m}\} \Leftrightarrow \{\cap_{i=1,i\neq k}^{m} \bar{C}_{k,i}\}.$$

Using DeMorgan's law and Boole's inequality (Resnick, 2013), one can obtain

$$\mathbb{P}(\nu = k) = 1 - \mathbb{P}\left(\cup_{i=1,i\neq k}^{m} C_{k,i}\right) \geq 1 - \sum_{i=1,i\neq k}^{m} \mathbb{P}\left(C_{k,i}\right) = 1 - \sum_{i=1,i\neq k}^{m} \mathbb{P}\left(\tilde{\delta}_k > \tilde{\delta}_i\right) \tag{4.2.16}$$

where $\mathbb{P}\left(\tilde{\delta}_k > \tilde{\delta}_i\right) =$

$$\mathbb{P}\left(\bar{\Sigma}_{k,i}^{-1/2}\left(\tilde{\delta}_i - \delta_i - \tilde{\delta}_k + \delta_k\right) < \bar{\Sigma}_{k,i}^{-1/2}\left(\delta_k - \delta_i\right)\right) \approx \Phi\left(\bar{\Sigma}_{k,i}^{-1/2}\left(\delta_k - \delta_i\right)\right) \tag{4.2.17}$$

where $\Phi(\cdot)$ is the distribution function of a standard normal random variable, $\bar{\Sigma}_{k,i} = \left(\bar{\Sigma}_k + \bar{\Sigma}_i\right)$ and $\bar{\Sigma}_k$ is the covariance matrix corresponding to the regimen k as in Theorem 5. As regimens k and i are independent, there are two independent random variables in the left-hand side of the second term in Equation (4.2.17) and each of them converges to normal random variables (Theorem 5). Therefore, the sum converges to a standard Gaussian random variable after an appropriate

normalisation. Consequently, for $j = 1, \ldots, m$

$$\mathbb{P}_{\pi_j}(\nu = j) \geq 1 - \sum_{i=1, i\neq j}^{m} \Phi\left(\bar{\Sigma}_{j,i}^{-1/2}(\delta_{j,j} - \delta_{i,j})\right) \tag{4.2.18}$$

By the construction of the vector π_j, $\delta_{j,j} - \delta_{i,j} < 0$. The number of observations on each regimen N_j is proportional to the total sample size N under Rule I: $N_j \simeq w_j N$ where $a_n \simeq b_n$ means that $\lim_{n\to\infty} \frac{a_n}{b_n} = 1$, $j = 1, \ldots, m$. Thus,

$$\bar{\Sigma}_{j,i}^{-1/2}(\delta_{j,j} - \delta_{i,j}) \simeq c\sqrt{N} \tag{4.2.19}$$

where c is a negative constant. Plugging-in terms in the accuracy formula, we obtain that $\lim_{N\to\infty} A_N \geq 1$.

For $\kappa > 1/2$, the probability of the final selection in the experiment is given by Equation (4.2.16) for $\kappa = 1/2$ as the penalty term is not taken into account for the final recommendation. Therefore, the Inequality (4.2.18) holds. The only difference is the number of observations on each regimen which is now proportional to the total number of patients $N_j \simeq l_j(N)N$ with l_j depending on N. This results in a different constant $c < 0$ in Equation (4.2.19), but in the unchanged rate $\sqrt{N}$ due to the same rate in both nominator and denominator of Equation (4.2.9) with respect to the total number of patients N.

Binary outcomes

While the asymptotic result in Equation (4.2.17) is given in general terms of $\bar{\Sigma}_{k,i}$, it can be written explicitly for $d = 2$. In the special case of binary outcomes, $\bar{\Sigma}_{k,i} \equiv \sigma_{k,i} = \sigma_k^2 + \sigma_i^2$ and $\sigma_i = |\delta_i{}'|\sqrt{\frac{\alpha_i(1-\alpha_i)}{n_i}}$, with

$$\left.\frac{\partial \delta_i^{(\kappa=1/2)}(z,\gamma)}{\partial z}\right|_{z=\alpha_i} = \delta_i{}' = \frac{(\gamma - \alpha_i)(\gamma(2\alpha_i - 1) - \alpha_i)}{\alpha_i^2(1-\alpha_i)^2}.$$

Therefore,

$$\Phi\left(\frac{\delta_k - \delta_i}{\sigma_{k,i}}\right) = \Phi\left(\frac{\sqrt{n_k n_i}\,(\delta_k - \delta_i)}{\sqrt{n_k \alpha_i (1-\alpha_i)(\delta_i{}')^2 + n_i \alpha_k (1-\alpha_k)(\delta_k{}')^2}}\right). \tag{4.2.20}$$

The number of patients n_k and n_i are proportional to the total number of patients by the construction of Rule I. Therefore, the result in Equation (4.2.19) becomes explicit.

(b) Rule II. Consider $\kappa = \frac{1}{2}$. The design based on this measure and on its point estimate is inconsistent as it does not guarantee an infinite number of patients on all regimens as required under the assumption of the regimens independence (Robbins, 1952). Let us consider a counterexample with two regimens and the regimen T_1 being the TR. Suppose that prior parameters are specified such that $\hat{\delta}_{\beta_2} \ll \hat{\delta}_{\beta_1}$ and $\delta_2 \ll \hat{\delta}_{\beta_1}$, so T_2 is selected initially. While the number of observations on T_2 increases and the estimate $\hat{\delta}_2$ approaches the true value δ_2 (Theorem 5), the estimate $\hat{\delta}_1$ remains unchanged. One can find prior values $\hat{\delta}_{\beta_1} \gg \delta_2$ such that T_1 is never selected, because the point estimate $\hat{\delta}_2$ would not go below $\hat{\delta}_{\beta_1}$. Consequently, the selection would "lock-in" at the suboptimal regimen regardless of further outcomes. Therefore, the number of patients on both regimens do not tend to infinity as $N \to \infty$.

For $\frac{1}{2} < \kappa < 1$, let $L_j(t)$ be the indicator function such that

$$L_j(t) = \begin{cases} 1 \text{ with probability } \mathbb{P}(\tilde{\delta}_j^{(\kappa)}(t) = \min_i \tilde{\delta}_i^{(\kappa)}(t)) \\ 0 \text{ with probability } 1 - \mathbb{P}(\tilde{\delta}_j^{(\kappa)}(t) = \min_i \tilde{\delta}_i^{(\kappa)}(t)) \end{cases}$$

where $\tilde{\delta}_j^{(\kappa)}(t)$ is a random variable corresponding to the posterior density function after t observations in the experiment. Let $n_j(t) = \sum_{i=1}^{t} L_j(i)$ be the number of

observations on T_j up to the moment t. We then obtain

$$\mathbb{E}(n_j(t)) = \sum_{u=1}^{t} \mathbb{E}L_j(u) = \sum_{u=1}^{t} \mathbb{P}(\tilde{\delta}_j^{(\kappa)}(u) = \min_i \tilde{\delta}_i^{(\kappa)}(u)).$$

Note, that $\mathbb{P}(\tilde{\delta}_j^{(\kappa)}(u) = \min_i \tilde{\delta}_i^{(\kappa)}(u))$ has already been studied in **(a)** (see e.g. Equation (4.2.20) for the expression in the binary case). The mean of $\tilde{\delta}_j^{(\kappa)}(u)$ associated with T_j to be selected is an increasing polynomial with respect to N. The probability to be the minimum decreases for j and increases for $i = 1, \ldots, j-1, j+1, \ldots, m$. It follows that the probability of being selected is not a monotonic function and

$$\lim_{t\to\infty} \sum_{u=1}^{t} \mathbb{P}(\tilde{\delta}_j^{(\kappa)}(u) = \min_i \tilde{\delta}_i^{(\kappa)}(u)) = \infty.$$

The final selection is the regimen satisfying (4.2.12). Consequently, the number of observations on each regimen tends to infinity and we obtain that $\lim_{N\to\infty} A_N = 1$ using the arguments of **(a)**. □

Note that there is no explicit general expression for $\mathbb{P}(\tilde{\delta}_j^{(\kappa)}(u) = \min_i \tilde{\delta}_i^{(\kappa)}(u))$ and, consequently, for the PCS as it depends on the whole history of events. To study the asymptotic behaviour of the novel design under Rule II for different choice of κ, we propose an approximation procedure in Appendix B.1-B.3. It demonstrates that the design is indeed inconsistent for $\kappa = 0.5$ and provides insights on the convergence for different values of the penalty parameter κ.

Below, the performance of the proposed experimental design is studied in contexts of Phase I (Section 4.3), Phase II (Section 4.4) and Phase I/II (Section 4.5) regimen finding clinical trials. We will refer to our proposal as the Weighted Entropy (WE) design (WE_I under Rule I and WE_II under Rule II) and compare it to well-established alternative approaches. All computations have been conducted using R (R Core Team, 2015).

4.3 Application to Phase I Clinical Trials

4.3.1 Setting

To study the WE design in the context of Phase I clinical trials, let us consider $m = 7$ regimens, $N = 20$ patients and the regimen selection allowed after each patient. The goal is to find the regimen (which could be combination, schedule or combination-schedule) with the toxicity probability closest to $\gamma_t = 0.25$. In these studies randomisation to all regimens is not ethical for safety reasons and therefore, Rule II is used. We would like to emphasise that we do not consider the classic dose escalation problem in which the doses can be ordered according to the increasing toxicity. We focus on the setting in which clinicians do not know the toxicity ordering. While clinicians will be able to provide a presumed ordering of the regimens, this order might be misspecified.

We consider scenarios in which the prior order chosen by a clinician is either correct or misspecified. Scenarios with correctly specified ordering have monotonic regimen-toxicity relationships and scenarios with misspecified ordering have non-monotonic relationships. The investigated scenarios are shown in Figure 4.2 and include a variety of monotonic and non-monotonic shapes as well as one scenario with highly toxic regimens only. Note that when studying regimens, it is likely that several of them would have similar toxicity. Therefore, scenarios with probabilities of toxicity outcome close to each other are considered.

It is assumed that limited information about regimens is available and a linear increase in toxicity probabilities is expected such that

$$\hat{p}_{\beta_1} < \hat{p}_{\beta_2} < \ldots < \hat{p}_{\beta_7}$$

where $\hat{p}$ is given in Equation (4.2.14). For safety reasons, the trial is required

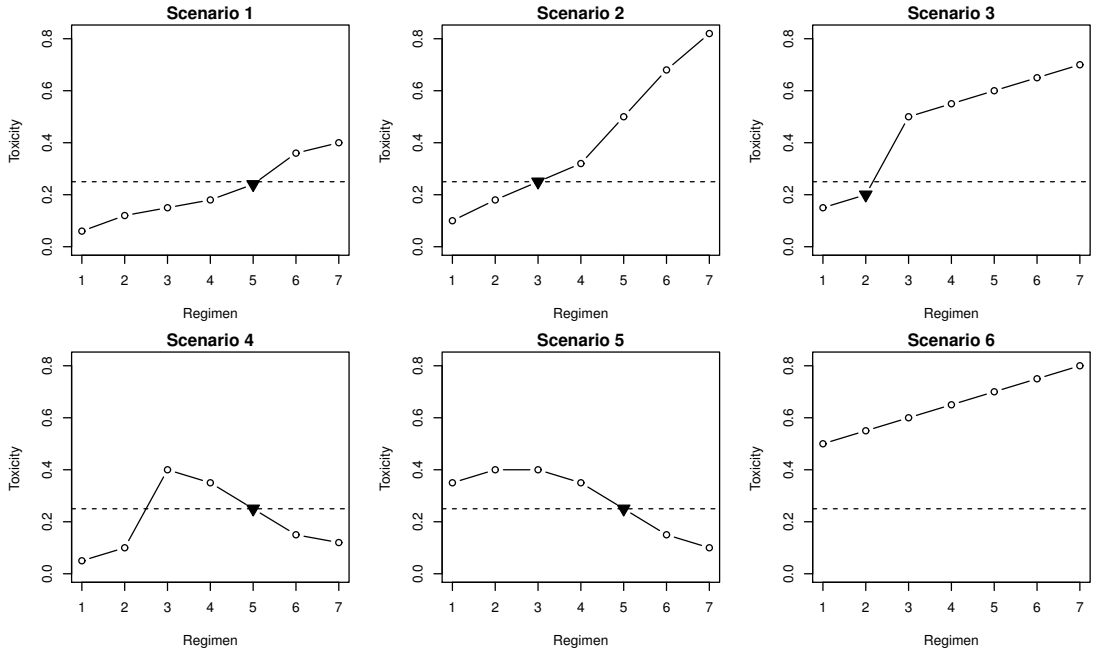

Figure 4.2: Toxicity scenarios. The target regimen is marked by a triangle and the maximum tolerated toxicity $\gamma_t = 0.25$ is marked by a dashed horizontal line. The monotonic scenarios 1-3, 6 correspond to a correctly prespecified ordering of regimens according to increased toxicity and non-monotonic scenarios 4 and 5 to a misspecified ordering of regimens.

to start at T_1. An "operational" prior, a prior that gives good operating characteristics under different scenarios, is calibrated over set of different scenarios and the details on the calibration are given in Appendix B.4. The prior uses $\beta_1 = \ldots = \beta_7 = 1$ and the prior toxicity probability modes of

$$\hat{\mathbf{p}} = [0.25, 0.30, 0.35, 0.40, 0.45, 0.50, 0.55]^{\mathrm{T}}.$$

The penalty parameter is fixed at $\kappa = 0.5$ due to the small sample size.

4.3.2 Comparators

The WE design is compared to common Phase I dose escalation designs. Specifically, we consider the Bayesian CRM (Section 2.1.1) and EWOC designs (Section 3.5) that assume a monotonic regimen-toxicity relationship. These designs are chosen to illustrate the consequences of the monotonicity assumption violation. We would also consider the partial ordering CRM (POCRM, Wages et al., 2011)

which relaxes the monotonicity assumption but uses the same model-based nature as CRM. The POCRM design is briefly summarised below.

If a single order of doses according to increasing toxicity cannot be specified, clinicians can often provide several orderings which are considered as feasible ones. Assume that there are R possible monotonic orderings. Under ordering r the DLT probability has a parametric form $p_j = \psi^{(r)}(d_j, \theta)$. Then, after n patients were assigned to dose levels $d(1), \ldots, d(n)$ and outcomes $y_1, \ldots, y_n$ were observed, the likelihood under ordering r can be computed as

$$\mathcal{L}_n^{(r)}(\theta) = \prod_{i=1}^{n} \phi^{(r)}(d(i), y_i, \theta)$$

as in the original CRM approach with the only difference in the model function

$$\phi^{(r)}(d(i), y_i, \theta) = \psi^{(r)}(d(i), \theta)^{y_i}(1 - \psi^{(r)}(d(i), \theta))^{1-y_i}.$$

Following the original CRM, the posterior distribution for parameter θ can be computed by Equation (2.1.2) under all orderings $r = 1, \ldots, R$. This results in R fitted CRM models. Given prior distribution of orderings $\{q^{(1)} \ldots, q^{(R)}\}$, where $0 \leq q^{(r)} \leq 1$ and $\sum_{r=1}^{R} q^{(r)} = 1$, the posterior probability of ordering r after n patients takes the form

$$\pi_n^{(r)} = \frac{q^{(r)} \int_{\mathbb{R}} \mathcal{L}_n^{(r)}(u) f_0(u) \mathrm{d}u}{\sum_{t=1}^{T} q^{(t)} \int_{\mathbb{R}} \mathcal{L}_n^{(t)}(u) f_0(u) \mathrm{d}u}. \tag{4.3.1}$$

The next group of patients is allocated based on ordering $r^\star$ corresponding to the maximum value of $\pi_n^{(r)}, r = 1, \ldots, R$. For chosen model $r^\star$, let $f_n^\star$ be the corresponding posterior distribution in Equation (2.1.2). Using estimate in Equation (2.1.3) for the toxicity probability, the next group of patients is assigned the dose level minimising the criterion (2.1.4). The procedure is repeated until the maximum number of patients N has been treated.

The POCRM requires ordering of regimens to be specified prior to a trial. In the comparison the correct orderings only are included to allow for the best possible performance of the POCRM under the evaluated scenarios. The orderings are given in Table 4.1. The same prior toxicity probabilities $\hat{\mathbf{p}}$ and rough prior distributions

Table 4.1: Orderings for POCRM.

	Order
1	(1,2,3,4,5,6,7)
2	(1,2,7,6,5,4,3)
3	(7,6,5,4,1,2,3)

of the model parameters were chosen for the model-based alternatives. Finally, we include the non-parametric optimal benchmark (Section 2.1.2).

The main characteristics to consider are: (i) the proportion of correct selections and (ii) the average number of toxic responses. The *bcrm* package (Sweeting et al., 2013) is used for the CRM and the EWOC and the *pocrm* package (Wages and Varhegyi, 2013) is used for POCRM. For all methods 10^4 simulations are used.

4.3.3 Safety Constraint

For ethical reasons an escalation procedure should be planned so that only a few patients are assigned to highly toxic regimens. This is typically achieved by the use of a safety constraint. The majority of existing safety constraints are based on the assumption of monotonicity and hence are not suitable for the proposed design. We adopt the following safety constraint instead. The regimen T_j is safe if after n patients

$$\int_{\gamma^*}^{1} f_{n_j}(p)\mathrm{d}p \leq \theta_{n_j}$$

where γ^* is an upper toxicity threshold, θ_{n_j} controls the overdosing probability and f_{n_j} is the posterior Beta distribution of the toxicity probability. Note that the overdosing threshold θ_{n_j} changes as the trial progresses. We require it to be a decreasing function of n with $\theta_0 = 1$ to give a possibility to test all regimens (if data suggests so) and $\theta_{final} \leq 0.3$ to ensure that the final recommendation is safe.

As an illustration, the linearly non-increasing

$$\theta_n = \max(1 - rn, \theta_{final})$$

is used with $r > 0$. We have calibrated the parameters of the safety constraints (details are given in Appendix B.4) and used $\gamma^* = 0.45$ and $r = 0.035$ in the simulations. Safety constraints of a similar form were incorporated in the model-based methods.

4.3.4 Operating Characteristics

The simulation results in monotonic scenarios 1-3 are given in Table 4.2.

Table 4.2: Operating characteristics of WE, CRM, POCRM and EWOC designs in scenarios 1-3. Term, DLT and $\bar{N}$ correspond to termination proportion, average number of toxic responses and average number of patients, respectively. The most likely recommendation is in bold, the TR is in italics. Results are based on 10^4 replications.

	T_1	T_2	T_3	T_4	T_5	T_6	T_7	Term	DLT	$\bar{N}$
Scenario 1. Linear response										
Scenario	0.06	0.12	0.15	0.18	*0.24*	0.36	0.40			
Optimal	0.92	9.12	10.60	14.44	**31.54**	27.50	5.87			
WE	7.03	14.72	23.33	**30.11**	*23.34*	1.39	0.05	0.03	3.3	20.0
CRM	2.66	7.21	14.17	20.58	***26.62***	15.95	12.53	0.3	4.38	19.9
POCRM	2.69	11.25	22.30	15.73	***22.60***	20.62	4.60	0.2	4.94	20.0
EWOC	8.06	13.80	20.30	**23.70**	*20.20*	9.62	3.88	0.4	3.76	18.8
Scenario 2. Logistic shape										
Scenario	0.10	0.18	*0.25*	0.32	0.50	0.68	0.82			
Optimal	6.05	29.03	**30.12**	28.27	6.48	0.05	0.00			
WE	17.78	27.43	***27.54***	22.51	3.76	0.11	0.00	0.9	5.23	20.0
CRM	17.24	25.88	***28.70***	19.37	6.24	0.56	0.04	1.9	4.84	19.7
POCRM	14.98	27.32	***27.89***	18.50	6.70	1.04	1.86	1.7	5.54	20.0
EWOC	**28.72**	**27.66**	*24.32*	13.65	4.21	0.00	0.00	1.4	3.27	18.0
Scenario 3. J shape										
Scenario	0.15	*0.20*	0.50	0.55	0.60	0.65	0.70			
Optimal	42.40	**45.79**	11.49	0.25	0.06	0.01	0.00			
WE	38.07	***44.65***	6.59	3.44	1.48	0.28	0.02	5.5	5.94	19.8
CRM	**37.47**	***37.85***	17.41	2.92	0.36	0.07	0.00	3.9	5.10	19.4
POCRM	**33.57**	***37.76***	13.27	2.55	0.54	1.33	6.04	4.9	6.06	19.8
EWOC	**51.00**	*26.11*	11.01	0.88	0.13	0.00	0.00	10.9	3.60	16.8

The WE design performs comparably to CRM and POCRM designs and selects the TR with the probability nearly 0.25 and 0.30 in scenarios 1-2. In scenario 1

the WE design underestimates the target treatment and recommends a less toxic treatment more often due to the safety constraint. Despite that, the performance of all methods is not far from the non-parametric optimal benchmark which shows that the selection of the TR is challenging in these scenarios. Proportions of terminations are close to 0 and the average number of toxic responses is largely the same for all considered methods. As expected, EWOC underestimates the TR in both scenarios.

In scenario 3, the WE design shows a better performance than the model-based alternatives with nearly 45% of correct selections against below 40% by CRM and POCRM, respectively. The safety constraint allows to prevent the recommendation of highly toxic regimens and controls the total number of toxic responses. Again, EWOC underestimates the TR but results only in 3 toxic responses compared to 5 for CRM and 6 toxicities for WE and POCRM. Clearly, the methods relaxing the monotonic assumption result in more toxic responses.

The results for non-monotonic scenarios 4-5 and unsafe scenario 6 are given in Table 4.3. As expected, the designs based on the monotonicity assumption are not able to find the TR in the non-monotonic settings. Comparing other designs, WE has a substantial advantage. It finds the TR with the probability nearly 0.28 comparing to 0.20 by POCRM while exposing nearly the same number of patients to toxic regimens. The time-varying safety constraint allows the selection of the TR even in non-monotonic scenarios where the TR lies beyond the toxic treatments ($T_3 - T_4$ in scenario 4 and $T_1 - T_4$ in scenario 5).

Considering scenario 6, the benchmark chooses the first regimen in 99.5% of replicated trials as the safety constraint is not incorporated in it. WE terminates earlier with probability 0.8 and performs similarly to POCRM and EWOC. It outperforms CRM which recommends a highly toxic regimen with a larger probability (32.56% against 19.16%). However, methods that relax the monotonicity

Table 4.3: Operating characteristics of WE, CRM, POCRM and EWOC designs. Term, DLT and $\bar{N}$ correspond to termination proportion, average number of toxic responses and average number of patients, respectively. The most likely recommendation is in bold, the TR is in intalics. Results are based on 10^4 replications.

	T_1	T_2	T_3	T_4	T_5	T_6	T_7	Term	DLT	$\bar{N}$
Scenario 4. Inverted-U shape										
Scenario	0.05	0.10	0.40	0.35	*0.25*	0.15	0.12			
Optimal	0.88	7.36	19.12	18.96	**38.47**	13.64	1.57			
WE	14.11	19.13	11.77	18.27	***27.90***	8.50	0.23	0.1	4.26	20.0
CRM	4.26	**19.90**	17.70	6.31	*2.84*	3.00	**46.10**	0.3	3.26	19.9
POCRM	2.87	11.39	11.75	9.32	*19.11*	**33.94**	11.62	0.2	4.29	20.0
EWOC	7.18	24.90	18.60	3.79	*2.52*	3.79	30.60	6.6	2.73	18.9
Scenario 5. Inverted-U shape										
Scenario	0.35	0.40	0.40	0.35	*0.25*	0.15	0.10			
Optimal	16.18	3.01	3.01	16.18	**39.46**	18.65	3.51			
WE	15.57	12.65	13.31	18.27	***27.92***	8.90	0.58	9.9	5.81	19.7
CRM	**47.41**	2.51	0.97	0.48	*0.72*	0.40	**30.10**	27.3	4.27	16.0
POCRM	16.81	5.98	5.66	12.42	*20.10*	**23.13**	10.23	9.7	5.14	19.5
EWOC	**30.75**	1.26	0.78	0.47	*0.47*	0.31	9.78	**56.2**	3.30	11.0
Scenario 6. Unsafe										
Scenario	0.50	0.55	0.60	0.65	0.70	0.75	0.80			
Optimal	99.53	0.35	0.10	0.02	0.00	0.00	0.00			
WE	13.63	5.53	2.45	0.88	0.27	0.06	0.00	***77.2***	8.02	14.2
CRM	32.24	0.32	0.08	0.00	0.00	0.00	0.00	***67.4***	5.33	10.3
POCRM	13.18	0.57	0.12	0.04	0.01	2.06	0.08	***83.9***	7.12	12.5
EWOC	16.17	0.00	0.12	0.00	0.00	0.00	0.00	***83.7***	3.07	6.1

assumption result in more toxic responses and require more patients on average to come to the termination conclusion. While CRM and EWOC require 5 and 6 patients only, it takes nearly 14 and 13 patients for WE and POCRM due to the consideration of non-monotonic orderings.

Summarising, the proposed design can perform comparably to the model-based approaches in trials with small sample size and with correctly specified ordering. At the same time, it can outperform them in cases when the ordering is misspecified. The time-varying safety constraint achieves the goals motivated by the ethical concerns while not preventing the TR selection in safe and non-monotonic scenarios. Therefore, one can conclude that the design is ethical and can be applied in practice if there is an uncertainty in the monotonicity ordering of regimens. Importantly, while a moderate number of regimens is considered in the example

above, similar results can be obtained for smaller number of regimens. As the proposed design does not fit any model to link the regimens, it will not experience the overfitting problem as model-based approaches, and can find the TR with fewer number of regimens.

4.4 Application to Phase II Clinical Trials

4.4.1 Setting

Consider a Phase II clinical trial investigating m regimens and in which the primary endpoint is a binary measure of efficacy (e.g. a response to the treatment). The goals of the study are (i) to find the most effective regimen (TR) and (ii) to treat as many patients as possible on it. Clearly, Rule I is preferable for the first goal and Rule II for the second one. We consider two hypothetical trials, each with $m = 4$ regimens, investigated by Villar et al. (2015) for Multi-Arm Bandit models (MAB). We compare the performance of the proposed approach to the MAB design based on the Gittins index (Gittins and Jones, 1979), which is the nearly optimal design in terms of maximising the expected number of successes (ENS), and to the fixed and equal randomisation (FR) which is expected to have a high statistical power.

Trial 1 investigates $N_1 = 423$ and efficacy probabilities are $(0.3, 0.3, 0.3, 0.5)$ while Trial 2 considers $N_2 = 80$ and the scenario is $(0.3, 0.4, 0.5, 0.6)$. Following the original application we consider the hypothesis $H_0 : \ p_0 \geq p_i$ for $i = 1, 2, 3$ with the family-wise error rate calculated at $p_0 = \ldots = p_3 = 0.3$, where p_0 corresponds to the control treatment efficacy probability. The Dunnett test (Dunnett, 1984) is used for the hypothesis testing in the FR setting. The hypothesis testing for MAB and WE designs is performed using an adjusted Fisher's exact test (Agresti, 1992). The adjustment chooses the cutoff values to achieve the same type-I error as FR. The Bonferroni correction is used for MAB and WE designs to correct for the multiple testing and the family-wise error rate is set to be less or equal to 0.05. Characteristics of interest are (i) the type-I error rate (α), (ii) statistical

power $(1-\eta)$, (iii) the expected number of successes (ENS) and (iv) the average proportion of patients assigned to the TR (p^*).

The WE design requires specification of the target value γ_e which can be defined in many trials by clinicians. This can be a maximum efficacy that they expect to see for the particular diseases. Below we consider the most challenging setting in which no target value is specified. Then, we set the target probability $\gamma_e = 1$ which corresponds to the aim "to find the regimen with the highest efficacy probability". The vector of the prior mode probabilities

$$p^{(0)} = [0.99, 0.99, 0.99, 0.99]^{\mathrm{T}}$$

is chosen to reflect no prior knowledge about which regimen has the highest success probability and that each treatment is considered as highly efficacious until data suggests otherwise. This choice of prior reflects the equipoise principle (Djulbegovic et al., 2000). We choose $\beta_0 = 5$ to ensure enough observations on the control regimen and $\beta_1 = \beta_2 = \beta_3 = 2$ to reflect no prior knowledge for competing regimens. We fix $\kappa = 0.5$ for $\mathrm{WE_I}$ and investigate the influence of different values of κ on $\mathrm{WE_{II}}$.

The trade-off between the expected number of successes (ENS) and the statistical power for different values of the penalty parameter κ under Rule II is illustrated in Figure 4.3.

In both trials, greater values of κ correspond to a greater power and a lower ENS as the increase in penalty tends to switch more often and leads to more spread allocation of patients. The exception is $\kappa \in (0.5, 0.55)$ in Trial 1 where the inconsistency for $\kappa = 0.5$ leads to "locking-in" on the suboptimal regimen. We choose two values of κ for the subsequent comparison to alternatives. These choices correspond to (i) high ENS, but unacceptable power (dashed line) and (ii)

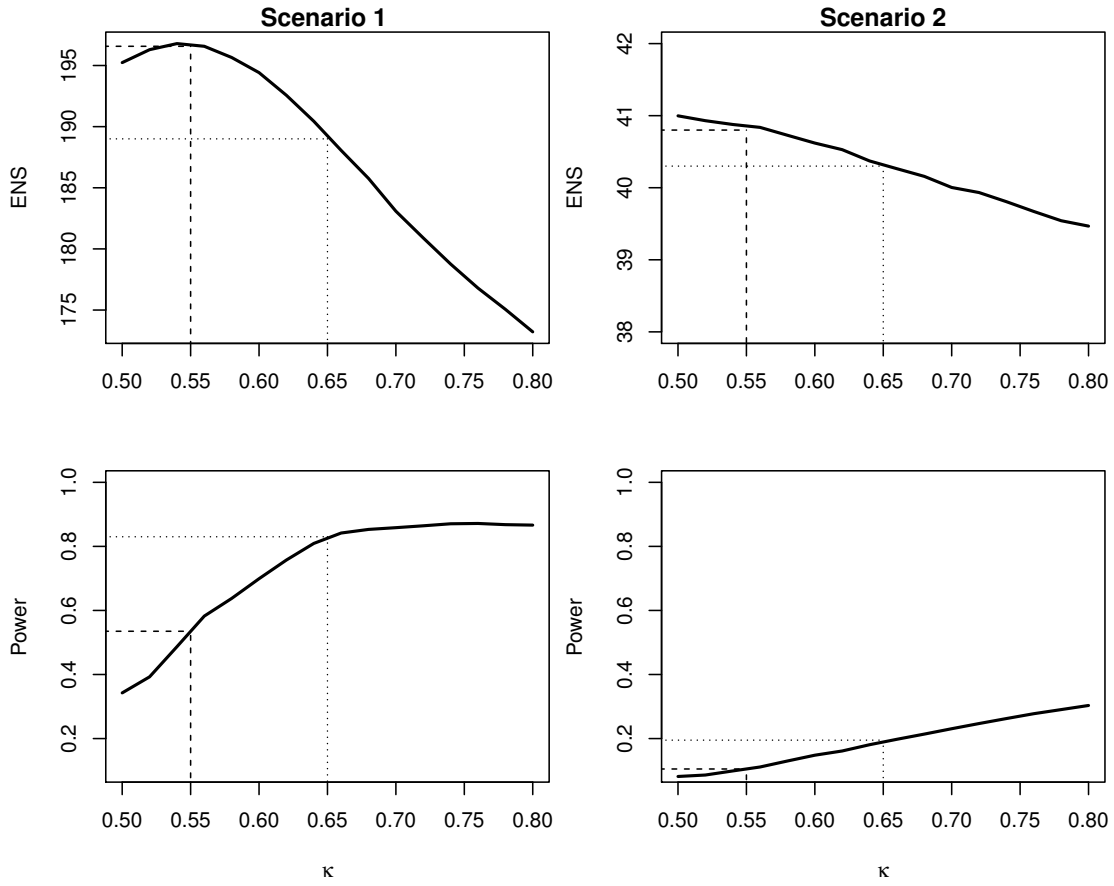

Figure 4.3: ENS and power (using a fixed cutoff value) for the WE design under the Rule II and different κ. Dashed lines correspond to the values chosen for the subsequent study. Results are based on 10^4 replications.

reduced ENS, but high power (dotted line).

4.4.2 Operating Characteristics

The operating characteristics of designs in Trial 1 are given in Table 4.4.

Table 4.4: Operating characteristics of the WE design under Rule I (WE_I), under Rule II (WE_II) for different κ (in brackets), MAB design and FR in Trial 1 with $N = 423$ under the null and alternative hypothesises. Results are based on 10^4 replications.

Method	$H_0 : p_0 = p_1 = p_2 = p_3 = 0.3$			$H_1 : p_0 = p_1 = p_2 = 0.3, p_3 = 0.5$		
	α	$p^*(s.e)$	ENS(s.e.)	$(1-\eta)$	$p^*(s.e.)$	ENS (s.e.)
MAB	0.05	0.25 (0.18)	126.68 (9.4)	0.43	0.83 (0.10)	198.25 (13.7)
FR	0.05	0.25 (0.02)	126.91 (9.4)	0.82	0.25 (0.02)	147.91 (9.6)
WE_I (0.50)	0.05	0.24 (0.05)	126.84 (9.5)	0.88	0.39 (0.06)	159.90 (11.0)
WE_II (0.55)	0.05	0.21 (0.20)	126.89 (9.4)	0.55	0.83 (0.18)	197.13 (17.8)
WE_II (0.65)	0.05	0.23 (0.13)	126.86 (9.4)	0.87	0.74 (0.10)	189.26 (13.7)

Under the null hypothesis, the performance of all methods is similar and the

type-I error is controlled. Under the alternative hypotheses, the WE_{II} design with $\kappa = 0.55$ performs comparably to the MAB in terms of the ENS, but yields almost 10% points increase in the power. Nevertheless, it has an unacceptable low statistical power which, however, can be increased by using higher values of the penalty parameter ($\kappa = 0.65$). It leads to an increase in the power from 0.53 to 0.86 at the cost of the slight ($\approx 4\%$) decrease in the ENS. In fact WE_{II} then has comparable power to FR, while resulting in 42 more patients responded to the treatment. Another way to increase the statistical power is to use WE_I for which both the associated power and the ENS is higher than for FR.

The operating characteristics of the designs in Trial 2 with fewer patients and a linearly increasing trend is given in Table 4.5. Under the null hypothesis, all

Table 4.5: Operating characteristics of the WE design under the Rule I (WE_I), under the Rule II (WE_{II}) for different κ (in brackets), MAB design and FR in Trial 2 with $N = 80$ under the null and alternative hypothesises. Results are based on 10^4 replication.

Method	$H_0 : p_0 = p_1 = p_2 = p_3 = 0.3$			$H_1 : p_i = 0.3 + 0.1i, i = 0, 1, 2, 3$		
	α	$p^*(s.e)$	ENS(s.e.)	$(1 - \eta)$	$p^*(s.e.)$	ENS (s.e.)
MAB	0.00	0.25 (0.13)	23.97 (4.10)	0.01	0.49 (0.21)	41.60 (5.4)
FR	0.05	0.25 (0.04)	24.02 (4.10)	0.50	0.25 (0.04)	35.98 (4.3)
WE_I (0.50)	0.05	0.23 (0.07)	23.92 (4.11)	0.59	0.33 (0.10)	37.55 (4.7)
WE_{II} (0.55)	0.01	0.20 (0.15)	24.01 (4.10)	0.11	0.50 (0.27)	40.72 (5.9)
WE_{II} (0.65)	0.05	0.22 (0.12)	23.96 (4.08)	0.52	0.47 (0.21)	40.19 (5.4)

designs perform similarly and type-I errors are controlled at the 5% level. Under the alternative hypothesis, MAB and WE_{II} with $\kappa = 0.55$, again, yield the highest (and similar) ENS among all alternatives, but also a low statistical power. WE_I or increased κ for WE_{II} result in a noticeable power increase - 0.59 and 0.52, respectively. Comparing to FR, both designs have a greater (or similar) power and result in more ENS.

Overall, WE designs can perform comparably to the nearly optimal MAB design in terms of the ENS, but with a greater statistical power for both large and small sample sizes. They have similar statistical power to FR, but with the considerably

greater ENS. The ENS and power trade-off can be tuned via the built-in parameter κ. Although, some modification to the MAB designs were proposed (e.g. see Villar et al., 2015) to prevent the lower statistical power, the majority of those are ruled-based. The proposed approach allows to avoid any algorithm-based rules and keeps the procedure fully adaptive. Additionally, the computation of the Gittens index for the MAB design is not trivial, requires more attention and is widely discussed in the literature (e.g. see Villar et al., 2015, and reference therein). Some of them require calibration and can be computationally intensive. The proposed criterion, in contrast, is extremely simple and easy to compute. While the proposed designs are compared for the target $\gamma_e = 1$, a similar performance can be expected for the problem of seeking the regimen associated with $\gamma_e \in (0.7, 1)$.

4.5 Application to Phase I/II Clinical Trials

4.5.1 Motivating Trial

This section is motivated by an ongoing Phase I/II clinical trial at the Hospital Gustave Roussy for patients with high-risk neuroblastoma. Both binary toxicity and efficacy endpoints are evaluated during the trial. The goal is to find the optimal biological regimen (OBR) defined as the safest regimen with a toxicity probability below the upper toxicity bound ϕ and a maximum efficacy above the lowest efficacy bound ψ. It is, however, expected to be challenging to find a single regimen having the safest toxicity and maximum efficacy simultaneously. Therefore, clinicians are also interested in finding "correct" regimens defined as regimens with the highest efficacy rate while still safeguarding patients.

Neuroblastoma is the most frequent individual type of a solid tumour in children (Steliarova-Foucher et al., 2017). Although different chemotherapy regimens of increasing intensities have been evaluated, the 5-years overall survival remains around 50% for the high risk group (Kreissman et al., 2013). A recent pre-clinical study has suggested that the use of a particular immunotherapy targeting the

disialoganglioside GD2 which is expressed in all neuroblastoma cells, in combination with conventional chemotherapy (etoposide and cisplatin) can improve the induction treatment. The first part of the high-risk neuroblastoma treatment aims to reduce the tumour burden to facilitate surgery and subsequent treatments. Combinations of the newly developed immunotherapy and chemotherapy are given under different schedules:

- Immunotherapy for 2 days after the chemotherapy (S_1)
- Immunotherapy for 3 days after the chemotherapy (S_2)
- Immunotherapy for 4 days with the chemotherapy. Overlap 1 day (S_3)
- Immunotherapy for 4 days with the chemotherapy. Overlap 2 days (S_4)

The combination of treatments is given for two cycles (three weeks). In each cycle the combination is given under one of schedules $S_1, \ldots, S_4$. Six different regimens are considered in the study and are given in Table 4.6. The toxicity outcome is evaluated by the end of the second cycle (after three weeks from the treatment start). The efficacy data is available after two more cycles (after six weeks from the start of the treatment) only. If a patient has experienced toxicity he will be treated off the protocol and the efficacy outcome cannot be observed. A maximum of 40 patients will be recruited to the study and it is anticipated that two patients can be recruited each month. Consequently, it is expected that the next two patients are assigned to a regimen before efficacy outcomes for the previous two patients are observed.

Table 4.6: The range of considered regimens in the motivating trial.

Regimen	T_1	T_2	T_3	T_4	T_5	T_6
Cycle 1		S_1	S_2	S_3	S_3	S_4
Cycle 2	S_1	S_2	S_2	S_3	S_4	S_4

A clinician is certain that toxicity probabilities of T_3, T_4 and T_5 are greater than of T_1, T_2 and smaller than of T_6. However, toxicity probabilities of T_3, T_4 and T_5

cannot be ordered according to increasing toxicity levels before the trial. Therefore, there are 6 possibilities how regimens $T_1, \ldots, T_6$ can be ordered with respect to the toxicity probabilities

$$\begin{aligned} &1)\ T_1, T_2, T_3, T_4, T_5, T_6 && 4)\ T_1, T_2, T_4, T_5, T_3, T_6 \\ &2)\ T_1, T_2, T_3, T_5, T_4, T_6 && 5)\ T_1, T_2, T_5, T_3, T_4, T_6 \\ &3)\ T_1, T_2, T_4, T_3, T_5, T_6 && 6)\ T_1, T_2, T_5, T_4, T_3, T_6. \end{aligned} \tag{4.5.1}$$

Additionally, a plateau or umbrella shape for the efficacy is plausible due to the action mechanism of the immunotherapy. This results in 48 efficacy orderings to be considered so that traditional designs cannot be applied directly. We illustrate how the proposed design can be applied in the context of the complex combination-schedule clinical trial. We also discuss practical and ethical issues such as coherence, delayed and missing efficacy responses, ethical constraints.

4.5.2 Practical Considerations

Re-parametrisation The selection criterion in Equation (4.2.15) depends on probabilities α_1 and α_2 corresponding to "efficacy and no toxicity" and "no efficacy and no toxicity" events, respectively. However, the goal of Phase I/II clinical trials is conventionally formulated in terms of toxicity (p_t) and efficacy (p_e) probabilities and corresponding targets γ_t and γ_e. Thus, we re-parametrise the criterion (4.2.15) using $\alpha_1 = (1-p_t)p_e$, $\alpha_2 = (1-p_t)(1-p_e)$, $\gamma_1 = (1-\gamma_t)\gamma_e$ and $\gamma_2 = (1-\gamma_t)(1-\gamma_e)$ and denote the selection criterion (trade-off function) by $\delta\left(p_t, p_e, \gamma_t, \gamma_e\right)$. This measure can be computed for each of m regimens with toxicity probabilities $p_{t,1}, \ldots, p_{t,m}$ and efficacy probabilities $p_{e,1}, \ldots, p_{e,m}$, respectively. Given $\delta(p_{t,j}, p_{e,j}, \gamma_t, \gamma_e)$, $j = 1, \ldots, m$, the TR $j^\star$ satisfies

$$\delta(p_{t,j^\star}, p_{e,j^\star}, \gamma_t, \gamma_e) = \min_{j=1,\ldots,m} \delta(p_{t,j}, p_{e,j}, \gamma_t, \gamma_e). \tag{4.5.2}$$

Contours of the efficacy-toxicity trade-off function for different combinations of toxicity and efficacy probabilities and $\gamma_t = 0.01$, $\gamma_e = 0.99$ are given in Figure 4.4. The most desirable point $p_t = 0.01$, $p_e = 0.99$ (right bottom corner) corresponds

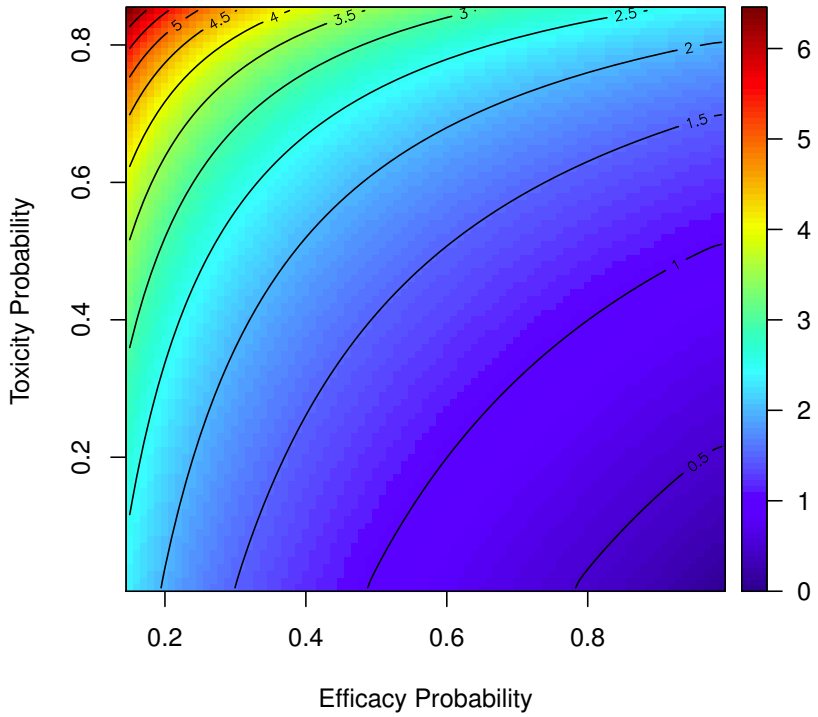

Figure 4.4: Contours of the efficacy-toxicity trade-off function $\delta(p_t, p_e, \gamma_t, \gamma_e)$ for $p_t \in (0, 0.85)$ and $p_e \in (0.15, 1)$, $\gamma_t = 0.01$, $\gamma_e = 0.99$.

to the minimum of the trade-off function. As probabilities move away from this point the trade-off function grows with an increasing rate that can be seen by contours located closer to each other.

Estimation The proposed trade-off function for regimen j depends on unknown parameters $p_{t,j}$ and $p_{e,j}$. While these could be estimated using model-based approaches such as the CRM, we can also consider each toxicity and efficacy probability independently as a Beta random variable estimated directly from the observed number of toxicities/efficacies. This allows the design to avoid the monotonicity assumption.

Let us consider an estimator for regimen j. The toxicity probability has prior distribution $\mathcal{B}(\nu_{t,j} + 1, \beta_{t,j} - \nu_{t,j} + 1)$, $\nu_{t,j}, \beta_{t,j} > 0$. After n_j patients and $x_{t,j}$

toxicities, we obtain Beta posterior $\mathcal{B}(x_{t,j} + \nu_{t,j} + 1, n_j - x_{t,j} + \beta_{t,j} - \nu_{t,j} + 1)$ which concentrates in the neighbourhood of $0 < p_{t,j} < 1$. Similarly, given prior Beta parameters $\nu_{e,j}, \beta_{e,j} > 0$, after n_j patients and $x_{e,j}$ efficacy responses the Beta posterior, $\mathcal{B}(x_{e,j} + \nu_{e,j} + 1, n_j - x_{e,j} + \beta_{e,j} - \nu_{e,j} + 1)$, for the efficacy probability can be found. The posterior modes of these Beta distributions are

$$\hat{p}_{t,j}^{(n_j)} = \frac{x_{t,j} + \nu_{t,j}}{n_i + \beta_{t,j}} \qquad \text{and} \qquad \hat{p}_{e,j}^{(n_j)} = \frac{x_{e,j} + \nu_{e,j\cdot}}{n_j + \beta_{e,j}}, \tag{4.5.3}$$

respectively. We then use the "plug-in" estimator $\delta(\hat{p}_{t,j}^{(n_j)}, \hat{p}_{e,j}^{(n_j)}, \gamma_t, \gamma_e) = \hat{\delta}_j^{(n_j)}$ as the criterion that governs the selection of a subsequent regimen to be studied in Phase I/II regimen finding trials.

Note that the trade-off function (4.2.15) is derived in the general form without assuming independence between toxicity and efficacy. While the estimators (4.5.3) does not account for the interaction between toxicity and efficacy, the derived trade-off function does.

Hybrid Randomised design Given unknown toxicity rates for regimens and a limited sample size, the assignment Rule I randomising between all regimens is not ethical to apply. Therefore, we consider the non-randomised version of the design under the assignment Rule II. At the same time, under this assignment rule the design can "lock-in" which means that one regimen would be tested regardless of further outcomes and the true optimal regimen can be never tested. In this case the design can benefit from the assignment rule based on a randomisation (Thall and Wathen, 2007; Wages and Tait, 2015).

We propose the randomisation within a safety set with probabilities proportional to the inverse of the trade-off function. For ethical considerations we restrict randomisation to the two best regimens only. Formally, assume that regimen $j^\star$ is the estimated best regimen (has the minimum value of $\hat{\delta}_{j^\star}^{(n_{j^\star})}$) and l is the second best regimen (i.e. has the second smallest value of $\hat{\delta}_l^{(n_l)}$). Then, randomisation

probabilities $\hat{w}_j^{(n_j)}$ for a regimen $j = 1, \ldots, m$ are

$$\hat{w}_j^{(n_j)} = \begin{cases} \frac{1/\hat{\delta}_j^{(n_j)}}{1/\hat{\delta}_{j^\star}^{(n_{j^\star})} + 1/\hat{\delta}_l^{(n_l)}}, \text{ if } j = j^\star, l \text{ and } \hat{\delta}_{j^\star}^{(n_j^\star)}, \hat{\delta}_l^{(n_l)} \neq 0. \\ 1, \text{ if } \hat{\delta}_j^{(n_j)} = 0. \\ 0, \text{ otherwise.} \end{cases} \tag{4.5.4}$$

The method proceeds until the maximum number of patients, N, have been treated. The regimen $j^\star$ satisfying criterion (4.5.2) is adopted as the final recommendation. We refer to this design as "WE(R)". This randomisation technique allows to get more spread allocation, while assigning only few patients to suboptimal regimens. We will focus on these two allocation rules only although other alternatives are possible. For instance, one might assign the first patients using randomisation and the rest using allocation to the "best" as suggested by Wages and Tait (2015).

Delayed responses So far, it was implied that efficacy responses are available at the same time as the toxicity information is. However, this is unlikely to be true in practice. While toxicity is usually quickly ascertainable, an efficacy endpoint may take longer to be observed (Riviere et al., 2016) and waiting for both endpoints increases the length of a trial substantially. However, the proposed criterion still can be applied in the trial with delayed responses as $\hat{p}_{t,i}^{(n_i)}$ and $\hat{p}_{e,i}^{(n_i)}$ can be estimated based on a different number of observations. Consequently, the design can proceed before the full response is observed and would only use the information available at the time the regimen for the next cohort is selected.

For instance, it takes twice as long to observe the efficacy outcome than the toxicity outcome (six weeks versus three weeks) in the motivating trial. To conduct the trial in a timely manner, the next cohort of patients is expected to be assigned based on the toxicity data only for the previous cohort and both toxicity and efficacy data for earlier patients. For instance, the recommendation for the third

cohort is based on both toxicity and efficacy outcomes for cohort 1 but only the toxicity outcomes for cohort 2. The proposed design accommodates this by computing toxicity and efficacy estimates based on different numbers of observations but on all the available information up to the time of the next patient allocation. Note that if an efficacy outcome is available earlier, it can (and should) be included in the design. Doing so can improve the performance of the design as illustrated in Appendix C.4. Note also that there may be situations in which auxiliary information about efficacy is available (e.g. through a short-term endpoint). Accounting for this information is beyond the scope of this section.

Coherence In practice, a clinician might be very cautious about further escalation if a DLT was observed in the previous cohort. For this reason, we force the WE design to satisfy principles of the coherent escalation/de-escalation (Cheung, 2005) with respect to known orderings. Assume that there are S known monotonic partial orderings. Denote the position of a regimen for cohort i in the partial ordering s by ${}_{[s]}T^{(i)}$ and the sum of the corresponding toxicity outcomes by $Q^{(i)}$. Then the coherent escalation means that,

$$\mathbb{P}\left({}_{[s]}T^{(i)} - {}_{[s]}T^{(i-1)} > 0 | Q^{(i-1)} \geq q\right) = 0, \ s \in S \tag{4.5.5}$$

where q is a threshold number of toxicities after which the escalation should be prohibited. The coherent de-escalation means that

$$\mathbb{P}\left({}_{[s]}T^{(i)} - {}_{[s]}T^{(i-1)} < 0 | Q^{(i-1)} < q\right) = 0, \ s \in S. \tag{4.5.6}$$

For instance, in the motivating example the following partial orderings are known

$$p_{t,1} \leq p_{t,2} \leq p_{t,3} \leq p_{t,6}$$

$$p_{t,1} \leq p_{t,2} \leq p_{t,4} \leq p_{t,6} \tag{4.5.7}$$

$$p_{t,1} \leq p_{t,2} \leq p_{t,5} \leq p_{t,6}.$$

It follows that if more than q toxicity outcomes are observed for T_3 the next cohort can be still allocated to T_4. Moreover, these coherence principles are used to incorporate the information about pairs of regimens that can be ordered (for example, T_1 and T_2). While it is not taken into account in the estimation step of the original design, it is reflected in allocation restrictions.

4.5.3 Illustration

Below, we demonstrate the performance of the non-randomised WE design in the context of the motivating trial. The major challenge of the motivating trial is the uncertainty in the regimen-toxicity relation for T_3, T_4, T_5 which results in six possible toxicity orderings given in Equation (4.5.1). Furthermore, for each of these toxicity orderings either a monotonic, plateau or umbrella regimen-efficacy relationship can be expected. Despite this complex setting, the WE design can be applied as both toxicity and efficacy endpoints are binary.

We consider the regimen-finding clinical trial with $m = 6$ regimens, $T_1, \ldots, T_6$, $N = 36$ patients and cohort size $c = 2$. The regimens are ordered (on the basis of clinician's beliefs) with increasing toxicity and efficacy. The trial is to be started at regimen T_1 and no regimen-skipping is allowed. The coherence parameter is fixed to be $q = 1$ and the escalation/de-escalation is required to be coherent with respect to known partial orderings (4.5.7). Following the motivating trial, toxicity is evaluated after two cycles of treatment (three weeks) and efficacy data is available after four cycles (six weeks) only. Since it is expected to recruit one patient per month, we assume that the next patient is assigned after the toxicity outcome is available for the previous patient. Moreover, efficacy is only observed for patients without toxicity.

True probabilities of toxicity and efficacy are $p_t = [0.05, 0.10, 0.45, 0.15, 0.30, 0.55]^{\mathrm{T}}$

and $p_e = [0.10, 0.40, 0.70, 0.70, 0.70, 0.70]^{\mathrm{T}}$. This scenario corresponds to a plateau in the regimen-efficacy relationship starting at T_3 and to the misspecified ordering of toxicities − regimen T_3 is more toxic than regimens T_4 and T_5. We study the ability of the WE design to recommend the *optimal* and *correct* regimens which are define as follows. The safest regimen with a toxicity probability below the upper toxicity bound ϕ and a maximum efficacy above the lowest efficacy bound ψ is called *optimal* and the regimen with the highest efficacy rate while still safeguarding patients (irrespective of it also having lowest toxicity) is called *correct*. Then, for a maximum toxicity probability bound $\phi = 0.35$ and a minimum efficacy probability bound $\psi = 0.20$, regimen T_4 is the optimal one and regimens T_4 and T_5 are correct ones. To ensure that a regimen with the same efficacy but a lower toxicity is preferred over one with a higher toxicity we set $\gamma_t = 0.01$ and, similarly, we set $\gamma_e = 0.99$ to prefer a regimen with a higher efficacy if the toxicity is the same. To see that T_4 is indeed the OBR using these targets, one can compute the true values of the trade-off function [9.48,1.77,1.59,**0.67**,1.03,2.16]$^{\mathrm{T}}$. The minimum among all regimes corresponds to T_4 as desired and the second smallest value corresponds to the correct regimen T_5.

Prior parameters of toxicity $\hat{p}_t^{(0)} = [0.10, 0.175, 0.25, 0.325, 0.40, 0.475]^{\mathrm{T}}$, efficacy $\hat{p}_e^{(0)} = [0.60, 0.65, 0.70, 0.75, 0.80, 0.85]^{\mathrm{T}}$ and $\beta_{t,j} = \beta_{e,j} = 1$, $j = 1, \ldots, 6$ are chosen. Note that the purpose of this priors is to specify in which order the regimens are trialled. The allocation of 18 cohorts in a single simulated trial is given in Figure 4.5. Note that "no toxicity and efficacy" and "no toxicity and no efficacy" outcomes in Figure 4.5 can be observed after four cycles only. Until an efficacy outcome is available the design uses information about "no toxicity" only.

The allocation starts at T_1 with no toxicities. As no efficacy responses have been observed yet and T_1 has the "promising" prior efficacy probability, it is selected again. Following no efficacies for cohort 1, cohort 3 is assigned to T_2 at which no patients experienced DLTs. This leads to selecting T_2 again until the efficacy

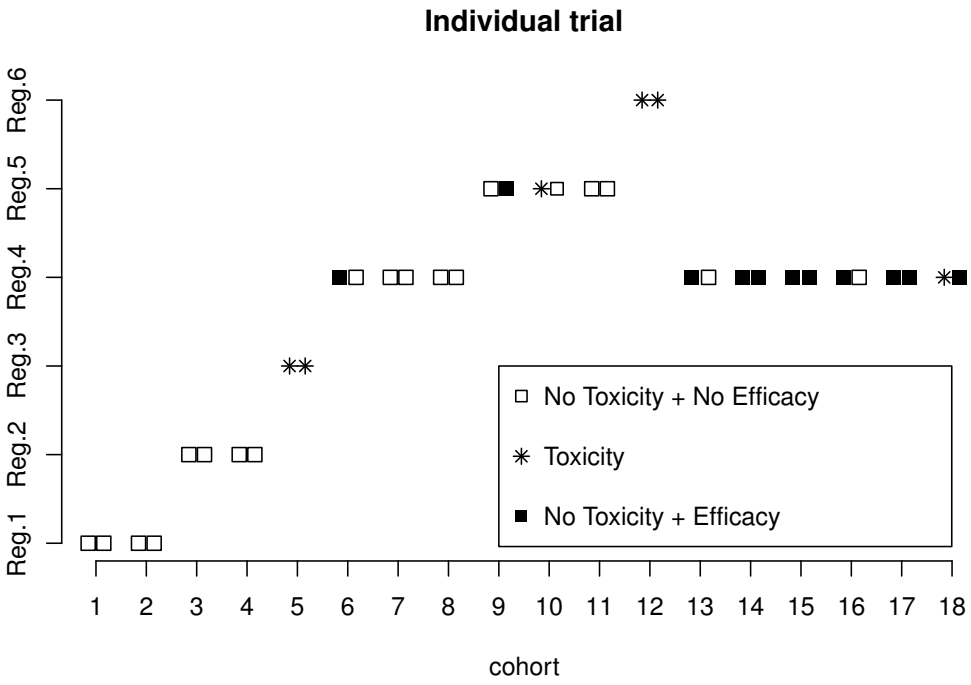

Figure 4.5: Allocation of 18 cohorts in the invididual trial.

data are available. After no toxicity (cohort 4) and no efficacy (cohort 3), cohort 5 is assigned to T_3 for which both patients experience toxicities. As there is the uncertainty whether T_3 is more toxic than T_4 and (or) T_5, cohort 6 is assigned to T_4. After no toxicity outcomes are observed, cohort 7 is allocated to regimen T_4 again. Due to no toxicities (cohort 7) and one efficacy (cohort 6), regimen T_4 is chosen for cohort 8 as well. However, after no efficacy for cohort 7, the design escalates to regimen T_5. Again, as no toxicity outcomes are observed for cohort 9, cohort 10 is assigned to T_5 too at which one patient experiences a toxicity. However, by the time cohort 11 is allocated, efficacy outcomes for cohort 9 become available and the allocation remains at regimen T_5. As no further efficacies have been observed for regimen T_5, design escalates further to regimen T_6 for which two toxicities are observed. Then, the design de-escalates to the optimal regimen T_4 for which one efficacy and no toxicity has been previously observed (against 1 efficacy and 1 toxicity for regimen T_5). All the consequent patients (up to cohort 18) are assigned to the optimal regimen T_4 which is finally selected in the trial. Clearly, a delayed efficacy response requires two cohorts to be assigned to each dose

conditionally on 'no toxicity'. It leads to fewer patients at the optimal regimen but also to more reliable selection due to a better exploration of regimes.

While the regimen-finding algorithm in the individual trial is considered above, allocation probabilities for each cohort in 10^6 replicated trials are given in Figure 4.6. Using the WE design, first and second cohorts are to be assigned to the

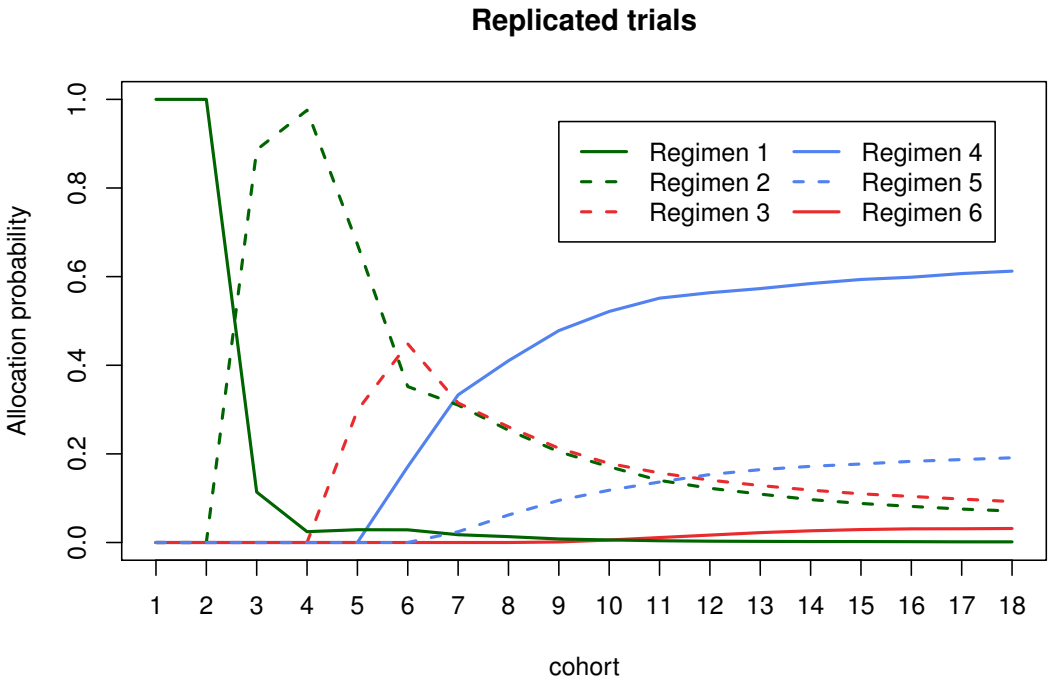

Figure 4.6: Probabilities to allocate each of 18 cohorts to T_1 (solid green), T_2 (dashed green), T_3 (dashed red), T_4 (solid blue), T_5 (dashed blue) and T_6 (solid red) in the motivating trial setting. The optimal regimen is Regimen T_4 and the correct regimens are Regimens T_4 and T_5). Results are based on 10^6 replications.

first regimen with probability 1. As illustrated above, the design stays at T_1 if no toxicity was observed (due to a "promising" efficacy) or if a toxicity is observed (due to the coherence principle). The probability to allocate a cohort to the optimal regimen T_4 starts to increase after cohort 5 and reaches nearly 60% for cohort 18. Considering the probability of regimen recommendations, we compare the performance to an equal allocation of 6 patients per each regimen. For the equal allocation, the trade-off functions for each regimen are estimated at the end of the trial and the regimen corresponding to the smallest value is selected. The

optimal regimen T_4 is recommended in 62.5% of trials and the correct regimen T_5 in 18.6% of trials by the WE design against 31% and 29% by the equal allocation, respectively. It means that the proposed design recommends one of the correct regimens in more than 80% of the trials and clearly favours the optimal one.

Overall, the proposed design appears to be able to recommend the optimal regimen with a high probability and the escalation/de-escalation algorithm in the individual trial is intuitive. A comprehensive study comparing the proposed method to alternative approaches and across different scenarios is given below.

4.5.4 Ethical Constraints

Not borrowing information across regimens is a key feature of the proposed design. However, some regimens might have high toxicity and/or low efficacy. Then, a design can result in a high (small) number of toxicity (efficacy) responses or in unsafe/inefficacious selections. For ethical reasons it is required to control the number of patients exposed to such regimens and two time-varying constraints are introduced.

Safety constraint An absence of the monotonicity assumption for toxicity makes the problem of the highly toxic regimen selection even more crucial. As stated above, a conventional (constant) safety constraint (e.g. as in Riviere et al., 2016) cannot be applied because no parametric model is used. On the one hand, a reliable safety constraint should give the hypothetical possibility to test all regimens if data suggests so. On the other hand, the final recommendation should be made with a high confidence in safety. Following the argument for Phase I clinical trials above, a time-varying safety constraint meets both of these requirements. A regimen j is safe if after n_j patients

$$\int_{\phi^*}^{1} f_{t,j}^{(n_j)}(p)\mathrm{d}p \leq \zeta^{(n_j)} \tag{4.5.8}$$

where $f_{t,j}^{(n_j)}$ is the Beta posterior density function of the toxicity probability, ϕ^* is the toxicity threshold and ζ_{n_j} is the probability that controls overdosing. As information increases, we gain confidence about a regimen's safety and hence consider the constraint that becomes more strict as the trial progresses. We therefore use a non-increasing function of n_j for $\zeta^{(n_j)}$. We choose $\zeta^{(0)} = 1$ initially to allow all regimens to be tested while the final recommendation is made with probability ζ_N. Subsequently we use a linearly decreasing function $\zeta^{(n_i)} = max(1 - r_t n_j, \zeta_N)$ where $r_t > 0$. These safety constraint parameters can either be specified by experts or alternatively calibrated with respect to trial's goals using simulations.

Futility Constraint The same reasoning is applied to a time-varying futility constraint. The regimen j is efficacious if after n_j patients

$$\int_{\psi^*}^{1} f_{e,j}^{(n_j)}(p)\mathrm{d}p \geq \xi^{(n_j)} \tag{4.5.9}$$

where $f_{e,j}^{(n_j)}$ is the Beta posterior density function of the efficacy probability, ψ^* is the efficacy threshold and $\xi^{(n_j)}$ is the controlling probability. This probability is an increasing function of n_j and the recommendation is made with probability ξ_N. We use a linearly increasing function $\xi^{(n_j)} = min(r_e n_j, \xi_N)$ where $r_e > 0$.

4.5.5 Simulation Setting

Scenarios In this section, we study the performance of the proposed design in the context of the motivating trial under various different scenarios. Following the motivating trial, we consider $m = 6$ regimens and $N = 36$ patients which are enrolled in cohorts of $c = 2$. The setting as stated in Section 4.5.3 remains unchanged. The upper toxicity and the lowest efficacy bounds are $\phi = 0.35$ and $\psi = 0.20$, respectively. The goal is to study the ability of the WE design to identify the *optimal* and *correct* regimens as defined above.

The major challenge of the motivating trial is the uncertainty in both regimen-

Table 4.7: Six permutations of scenario 1. The optimal regimen is in bold and correct regimens are underlined.

	T_1	T_2	T_3	T_4	T_5	T_6
Sc 1.1	(.005;.01)	(.01;.10)	(.02;.30)	(.05;.50)	**(.10;.80)**	(.15;.80)
Sc 1.2	(.005;.01)	(.01;.10)	(.02;.30)	**(.10;.80)**	(.05;.50)	(.15;.80)
Sc 1.3	(.005;.01)	(.01;.10)	(.05;.50)	(.02;.30)	**(.10;.80)**	(.15;.80)
Sc1.4	(.005;.01)	(.01;.10)	**(.10;.80)**	(.02;.30)	(.05;.50)	(.15;.80)
Sc 1.5	(.005;.01)	(.01;.10)	(.05;.50)	**(.10;.80)**	(.02;.30)	(.15;.80)
Sc 1.6	(.005;.01)	(.01;.10)	**(.10;.80)**	(.05;.50)	(.02;.30)	(.15;.80)

toxicity and regimen-efficacy relations. This increases the number of possible scenarios to be investigated enormously. Therefore, we start by defining scenarios on which the assessment will be based. Firstly, we specify 14 scenarios with increasing regimen-toxicity relations and different shapes of regimen-efficacy curves given in Figure 4.7: eight plateau regimen-efficacy scenarios (1-8) by Riviere et al. (2016), four umbrella shaped scenarios (9-12) by Wages and Tait (2015), and two scenarios with no optimal and correct regimens (13-14, due to inefficacy and toxicity, respectively). These scenarios were initially used to test the performance of the WE design in a context of Phase I/II single MTA trial and compare it to the methods by Wages and Tait (2015) and Riviere et al. (2016). The results of this evaluation are given in Appendix C. Secondly, to allow for the uncertainty in the toxicity ordering, we consider six permutations of each scenario with respect to toxicity orderings (4.5.1). For instance, six permutations of scenario 1 are given in Table 4.7. Overall, this results in 84 scenarios that cover a large variety of possibilities and allows the proposed design to be assessed in the setting of the motivating trial adequately. In the analysis we focus on (i) the proportion of optimal/correct selection, (ii) the average number of toxic responses, (iii) the average number of efficacy responses. The study is performed using `R` (R Core Team, 2015) and 10,000 replications for each scenario.

Design Specification As in the illustrative example above, we use the target toxicity of $\gamma_t = 0.01$ and the target efficacy of $\gamma_e = 0.99$. This implies that an investigator targets the most efficacious yet the least toxic regimen. The design

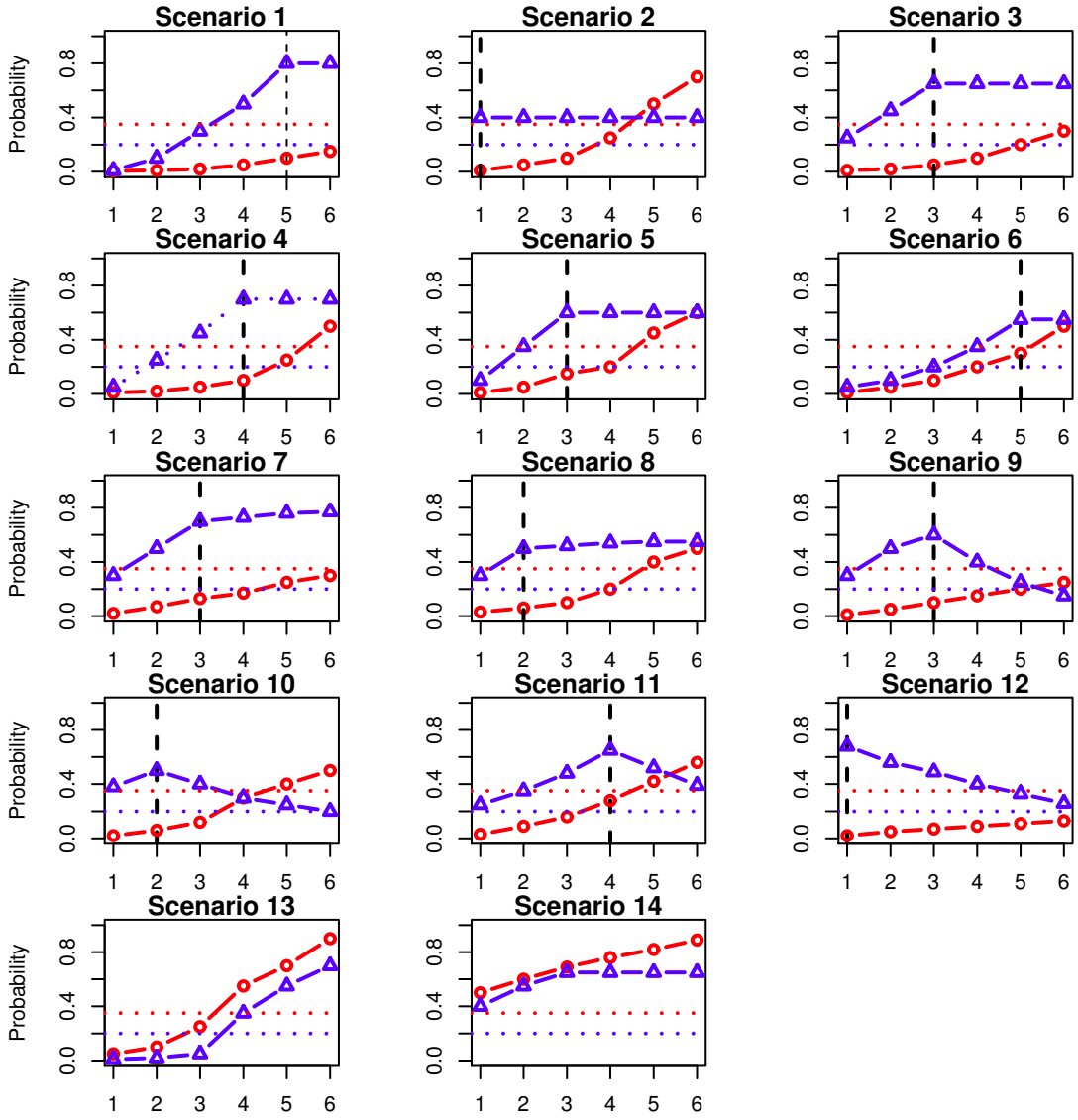

Figure 4.7: Eight plateau regimen-efficacy (blue triangles) scenarios (1-8), four umbrella regimen-efficacy scenarios (9-12) and two scenarios with no correct regimens (13-14) in the trial with $m = 6$ regimens. Toxicity and efficacy probabilities are marked by red circles and blue triangles, respectively. The red dotted horizontal line corresponds to the upper toxicity bound $\phi = 0.35$ and the blue dotted horizontal line corresponds to the lowest efficacy bound $\psi = 0.20$. A dashed black vertical line corresponds to the optimal regimen.

was restricted to satisfy the coherence conditions (4.5.5) - (4.5.6) with respect to partial orderings and $q = 1$. The randomised design is presented here as it has been shown in the evaluations of single agent trials (see Appendix C) to have more potential benefits when more than one correct regimen is expected in the trial.

Parameters $\beta_{t,j} = \beta_{e,j} = 1$ of the prior Beta distributions (4.5.3) are chosen for

all regimens $j = 1, \ldots, 6$ to emphasize a limited amount of information available. Parameters $\nu_{t,j}$ and $\nu_{e,j}$ (which coincide with prior toxicity efficacy probabilities for chosen values of $\beta_{t,j}, \beta_{e,j}$) are calibrated such that the optimal regimen can be found in various different scenarios with high probability. Note that the design is fully driven by the values of the trade-off function and, therefore, there are two requirements for the prior parameters. The prior should dictate to start the trial at the lowest regimen T_1 and the design should follow the escalation order of regimens specified by clinicians (with no regimen-skipping). To preserve a gradual escalation, the prior should assume that the higher regimens have greater efficacy but also greater toxicity. To restrict the number of prior parameters to be calibrated, we would assume a linear increase in prior toxicity and efficacy probabilities. Through extensive calibration, prior vectors of toxicity probabilities $\hat{\mathbf{p}}_t^{(0)} = [0.10, 0.14, 0.18, 0.22, 0.26, 0.30]^{\mathrm{T}}$ and efficacy probabilities $\hat{\mathbf{p}}_e^{(0)} = [0.60, 0.62, 0.64, 0.66, 0.68, 0.70]^{\mathrm{T}}$ were chosen for subsequently analysis. Note that despite the increasing prior probabilities, the ordering of the regimens is not fixed and can change as the trial progresses. Regarding the ethical constraint, parameters of the safety constraint $\zeta_N = 0.30, r_t = 0.02$, $\phi^* = 0.4$ and of the futility constraint $\xi_N = 0.50, r_e = 0.05, \psi^* = 0.35$ were calibrated. A further guideline on the choice of the operational prior and the parameters of safety and futility constraints are given in Appendix C.

Competing method We compare the performance of the novel approach to the extension of the design by Wages and Tait (2015) which was originally proposed for the fixed toxicity ordering. The original design employs the idea of partial orderings as in the POCRM by Wages et al. (2011) with respect to the various efficacy orderings. We extend the original design to allow for several toxicity orderings. As before, the dose-toxicity relationship is modelled by the 1-parameter power model

$$p_{t_k} = d_{k,w}^{exp(\theta_t)}$$

for a class of working dose-toxicity models where $d_{k,w}$ $w = 1, \ldots, W$ is a skeleton for model w, W is a number of working models and θ_t is an unknown parameter. One parameter power model is also used for the dose-efficacy relation

$$p_{e_k} = q_{k,s}^{exp(\theta_e)}, \ s = 1, \ldots, S,$$

for a class of working dose-efficacy models where $q_{k,s}, \ s = 1, \ldots, S$ is a skeleton for model s, S is a number of working models and θ_e is an unknown parameter. The working model is selected using the posterior model probabilities as for the POCRM (Section 4.3.2) - the model corresponding to the highest posterior probabilities is selected.

Following Wages and Tait (2015), a trial involving MTA with a monotonic regimen-toxicity relation and 6 regimens is associated with 11 efficacy orderings: one strictly monotonic, five cases of a plateau location and 5 cases of umbrella peaks. Similarly, one can deduce all possible efficacy orderings associated with each toxicity ordering. For instance, for the toxicity ordering $T_1, T_2, T_3, \boldsymbol{T_5}, \boldsymbol{T_4}, T_6$ we specify the following ordering (Figure 4.8)

1. 0.10, 0.20, 0.30, 0.50, 0.40, 0.60 (monotonic with respect to regimen-toxicity)
2. 0.20, 0.30, 0.40, 0.60, 0.50, 0.60 (plateau starting at T_4)
3. 0.30, 0.40, 0.50, 0.50, 0.60, 0.40 (peak at T_5)
4. 0.20, 0.30, 0.40, 0.60, 0.50, 0.50 (peak at T_4)
5. 0.40, 0.50, 0.60, 0.40, 0.50, 0.30 (peak at T_3)
6. 0.50, 0.60, 0.50, 0.30, 0.40, 0.20 (peak at T_2)
7. 0.60, 0.50, 0.40, 0.20, 0.30, 0.10 (peak at T_1)

Note that other orderings as a "plateau starting at $T_1/T_2/T_3/T_5$" are already included in the first regimen-toxicity case. Applying the same procedure to the rest of toxicity orderings (4.5.1) leads to 48 unique efficacy orderings. The design pro-

ceeds as follows. Firstly, given the available data, one of 6 toxicity orderings is selected as proposed by Wages et al. (2011). Secondly, one of 48 efficacy orderings is chosen as in the original design. The parameters of the designs are chosen as in the original specification by Wages and Tait (2015) with an exception of using cohort size $c = 2$ and 80% confidence intervals for stopping rules.The design waits for both toxicity and efficacy responses to allocated the next cohort and assumes that an efficacy outcome is observed regardless the toxicity outcome.We refer to this design as WT.

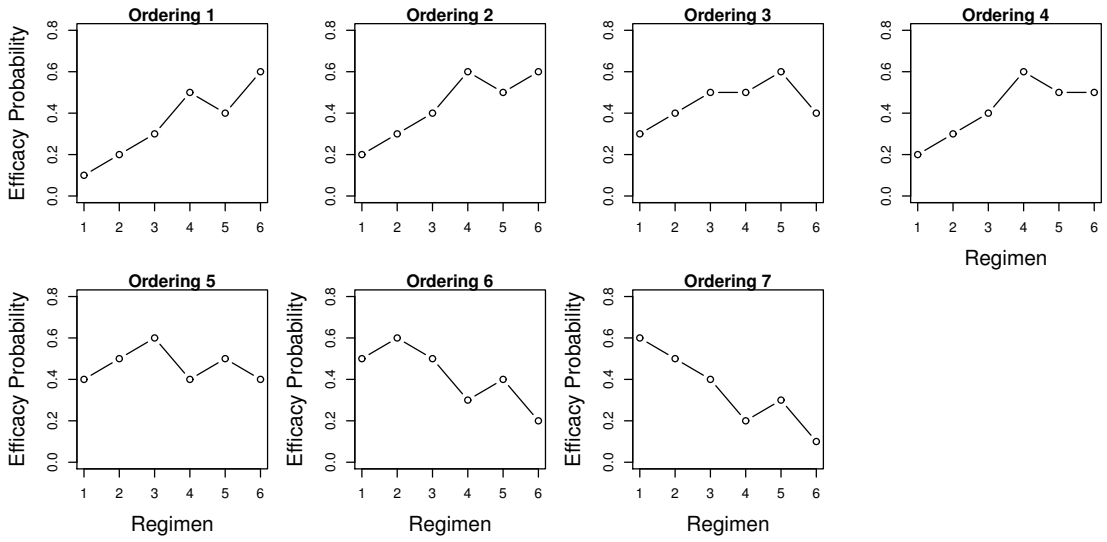

Figure 4.8: Seven efficacy orderings corresponds to the the toxicity ordering $T_1, T_2, T_3, T_5, T_4, T_6$.

Different variations of the extended WT design were also explored. A reduced number of efficacy orderings (plateau cases only) were investigated, but no significant difference was found, and the specification with the full number of orderings was found to be more robust. Additionally, the 'conditional' choice of the efficacy orderings was also studied. Using this approach, once the toxicity ordering is selected, that choice of the efficacy orderings is restricted to 11 orderings with respect to the toxicity profile. However, it was found to result in less accurate optimal and correct regimen recommendations across all scenarios.

4.5.6 Operating Characteristics

The results of the comparison in scenarios 1-12 using 6 permutations are summarized in Figure 4.9. The length of the bar corresponds to the proportion of the

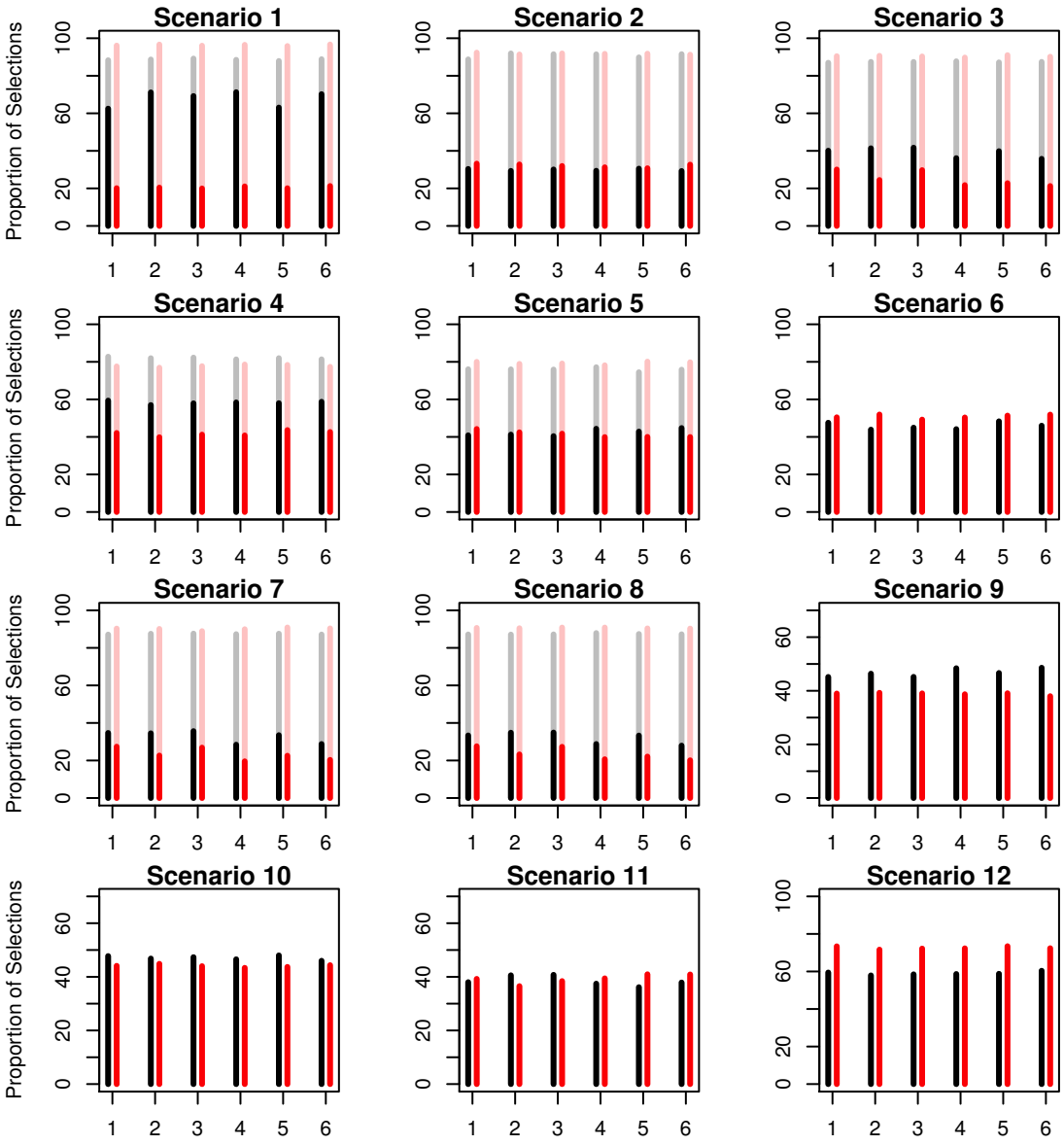

Figure 4.9: Proportion of optimal (bright) and correct (transparent) selections by WE(R) (black) and WT (red) designs in scenarios 1-12 across six permutations. Results are based on 10^4 replications.

optimal (solid) and correct (transparent) regimen selections by WE(R) (black) and WT (red). Overall, both designs are robust to toxicity ordering permutations under all scenarios. The difference of minimum and maximum proportions of the optimal and correct selections within one scenario does not exceed 5% and 8% for

WE and WT designs, respectively. Regarding the proportion of *correct* selections, the designs perform comparably (no more than 5% difference) in scenarios with several correct regimens (scenarios 2-5, 7 and 8) with scenario 1 being an exception (8% difference in favour to WT design). Regardless the large number of orderings, WT design preserves it accuracy in terms of the correct regimen selections.

Comparing *optimal* regimen selections, WE(R) favours the optimal regimen over other correct regimens systematically. It results in superior characteristics in half of all scenarios: 1, 3, 4, 7-9 with the minimum difference across permutations ranging from 7% (scenario 7) up to 45% (scenario 1) and the maximum difference varying between 11% and 52%, respectively. Generally, WT is less conservative as it favours safe regimens with higher toxicity that results in a lower proportion of optimal selections in these scenarios. Both designs perform comparably in scenarios 2, 5, 10 and 11 within the maximum difference of 2-4%. At the same time, a less conservative nature of WT allows outperforming WE(R) by 3-8% under scenario 6 in which the optimal regimen is the highest safe one. WT also shows a better performance in scenario 12 with the difference in the proportion of optimal selections ranging in 10-13%. The regimen 1 is optimal and 36 patients are not enough for WE(R) to investigate all regimens and come back to T_1 but it is enough for the WT to identify the decreasing ordering.

Investigating ethical aspects of the designs, the average numbers of toxicity and efficacy responses across all permutations in scenarios 1-12 are given in Table 4.8.

The WE(R) design is generally safer and results in at least nearly one less toxic response in scenarios 1, 3-8 with the largest difference of 2.7 in scenario 4. WT design results in fewer toxicities in scenarios where the target regimen is among the first ones (scenarios 2, 10, 12). While the prior toxicity vector chosen for WE(R) suggests proceeding escalation, the model-based approach is able to identify the

Table 4.8: Mean number of toxicity and efficacy responses in scenarios $1-12$ across six permutations using $N = 36$ patients and $m = 6$ regimens. Results are based on 10^4 replications.

Sc	1	2	3	4	5	6	7	8	9	10	11	12
						Toxicity responses						
WE(R)	2.5	6.4	3.2	4.4	7.0	7.7	5.1	5.1	3.9	5.9	7.8	2.4
WT	4.1	5.0	4.5	7.1	7.9	8.7	5.9	6.0	3.3	4.2	7.5	1.5
						Efficacy responses						
WE(R)	19.8	14.4	20.8	19.5	18.2	12.5	22.8	22.8	15.4	13.7	16.7	18.1
WT	24.5	14.4	21.0	21.4	19.0	13.8	23.4	23.5	15.8	14.4	16.7	21.5

correct ordering faster. The average number of efficacy outcomes does not differ by more than one in the majority of scenarios (2, 3, 5, 7-11). Due to the fact that WT waits for the complete efficacy response and due to model-based nature, it results in 1-5 more average efficacy responses in scenarios 1, 4, 6 and 12. At the same time, WE(R) waits for the efficacy outcome and escalates slower as it was shown in Section 4.5.3.

Regarding scenarios 13 and 14 with no correct and no optimal regimens due to futility and high toxicity, WE(R) terminates the trial earlier in nearly 72% and 85% of the trials against 68% and 79%, respectively. WE(R) results in 10.5 and 10.3 toxic responses on average in scenarios 13 and 14 against 10.4 and 10.5 by WT, respectively. Regarding the efficacy outcomes, WE(R) results in 7.0 and 8.7 responses versus 7.4 and 9.0 by WT. This follows that both designs are able to terminate the trial with high probability and prevent unethical patient allocations.

Summarising, the proposed design is robust to the possible true toxicity and efficacy orderings. WE(R) is found to be a good and comparative alternative to the model-based design when a large number of orderings is to be considered. While all possible toxicity and efficacy orderings are still feasible to specify, it can be challenging to convince a clinician to randomise between them given the limited sample size of $N = 36$ patients. WE(R) is able to identify optimal and correct regimens with high probability while being safer than model-based design and leads to a comparable number of efficacy responses in the majority of scenarios.

Table 4.9: Operating characteristics of WE (R) design in scenarios 1-12: proportion of optimal and correct rselections for different correlation values $r = \{-0.8, 0.0, 0.8\}$. The largest deviations are in bold. Results are based on 10^4 replications.

Scenario	1	2	3	4	5	6	7	8	9	10	11	12
WE (R)	Proportion of optimal selections											
$\rho = -0.8$	70.7	**22.5**	37.7	59.8	40.0	52.8	36.2	**24.7**	51.1	45.9	45.1	56.0
$\rho = 0.0$	66.5	31.0	39.5	59.3	40.5	46.8	33.2	32.8	44.8	47.7	38.5	59.1
$\rho = 0.8$	65.5	**41.2**	42.4	61.4	42.7	47.5	33.5	**38.3**	43.6	51.4	37.3	63.6
	Proportion of correct selections											
$\rho = -0.8$	90.9	**81.2**	91.4	91.8	78.0	52.8	91.6	75.4	51.1	45.9	48.1	56.0
$\rho = 0.0$	88.1	89.7	87.2	86.7	73.2	46.8	86.7	77.5	44.8	47.7	38.5	59.1
$\rho = 0.8$	88.4	**96.6**	84.8	84.5	73.6	47.5	86.2	80.9	43.6	51.4	37.3	63.6

4.5.7 Sensitivity Analysis

In the simulation study above, the toxicity and efficacy outcomes were assumed to be uncorrelated which may not hold in an actual trial. In this section, we investigate the robustness of the WE(R) design to the correlation in toxicity and efficacy under scenarios given in Figure 4.7. The study was also conducted for different toxicity orderings permutations and the same qualitative results were obtained (not shown).

We follow the procedure proposed by Tate (1955) to generate correlated binary toxicity and efficacy outcomes. The procedure generates a binary normal vector with unit variances and pre-specified correlation coefficient ρ. The generated random variable is then transformed to a binary response by applying the cumulative distribution function and a quantile transformation, subsequently.

The results of WE(R) performance in cases of high negative correlation ($\rho = -0.8$), absence of correlation ($\rho = 0.0$) and high positive correlation ($\rho = 0.8$) under scenarios 1-12 are summarised in Table 4.9.

In the majority of scenarios, WE(R) is robust to both negative and positive correlations. The proportions of optimal selection in correlated cases never differ by more than 10% compared to the uncorrelated case. Moreover, the differences be-

tween the proportions of optimal selections in positively and negatively correlated cases do not exceed 8% in 10 out of 12 scenarios with scenarios 2 and 8 being exceptions. A noticeable difference in the performance can be seen in plateau scenarios 2 and 8 in which the optimal regimens are among low regimens (T_1 and T_2, respectively). In these scenarios, efficacy probabilities are nearly the same and the negative correlation biases the selection to higher regimens which worsen the proportion of optimal selections by nearly 9% in scenarios 2 and 8. In contrast, the positive correlation biases the selection to lower regimens and leads to an increase by 10% and 6% in the proportion of optimal selection in these scenarios, respectively. There are no noticeable changes in the rest of scenarios as the bias caused by the correlation is smaller than the difference in toxicity and probability estimates.

Overall, the proposed design is robust to the highly correlated toxicity and efficacy outcomes. The proportion of optimal and correct selections in the correlated cases never differs by more than 10% compared to the uncorrelated case and mainly unchanged in the majority of scenarios. Despite assumed independence in the estimates, WE(R) is able to find the optimal and correct regimens in highly correlated cases.

4.6 Discussion

In this chapter, we have proposed a family of criteria for selecting the target regimen in experiments with multinomial outcomes. The novel criterion

- has an intuitively clear interpretation and can be easily computed by non-statisticians;
- leads to an accurate selection without the need for parametric or monotonicity assumptions;
- drives the allocation away from the bounds of spaces to the neighbourhood

of the target regimen;

- preserves flexibility and allows to tailor the design parameters in a light of the investigation goal.

The consistency conditions of the design based on the novel criterion are obtained. The application of the novel design was also demonstrated in the context of the Phase I, Phase II and Phase I/II clinical trials. The simulation results demonstrate that the proposal

- has comparable or better characteristics than other alternatives in Phase I and Phase II clinical trials with binary outcomes;
- can identify the optimal and correct regimens with high probability and leads to an ethical allocation of patients in the context of Phase I/II trials;
- is a good alternative to model-based designs when the ordering specification is challenging.

It is important to emphasize that the derived selection criterion can be also applied in conjunction with a parametric model which also expands its possible applications. This, in fact, was already demonstrated in Section 3.5 where the novel information-theoretic criterion in case of binary outcomes $d = 2$ and $\kappa = 0.5$ was incorporated into the one-parameter CRM model to govern the allocation of patients. This expands potential applications of the novel criteria. Moreover, the same form of the information-theoretic criterion and unit interval symmetric loss function suggests a link in the properties of the information gain and loss function on restricted parameters space which is subject of further research.

Despite clinical trials being used as the main motivation throughout, the design can be applied to a wide range of problems of a similar nature. For example,

applications where the MAB approach has found application: online advertising, portfolio design, queuing and communication networks, etc. (see Gittins et al., 2011, and references there in). On top of that, the proposed design can be used in more general problems of a percentile estimation rather than the identification of the highest success probability.

Chapter 5

Conclusions

In this work, we have considered various settings of complex regimen finding clinical trials which involve novel compounds. It was found in Section 2.2 that one can benefit from the monotonicity assumption in specific type of Phase I combination trials, but existing models should be appropriately adapted for the combination setting. It is also emphasized that the goals beyond the standard maximum tolerated regimen identification are of interest for molecularly targeted compounds and the designs should be tailored for them as well. To address objectives specific to combination trials with immunotherapies, we have proposed to include a control arm (standard of care) in the escalation trial and randomise patients between the control and the estimated MTC. We have also advocated the use of more complex combination-toxicity model in such trials. The proposed design, however, can face the same problem of "locking-in" on suboptimal combinations as other model-based designs (Azriel et al., 2011). While being one of the main concerns regarding the model-based designs, this problem has received limited attention in the literature and is the direction of the future research.

Motivated by a growing interest in more complex designs, we have proposed the

extension of the non-parametric benchmark in Section 2.3 which can be applied in the vast majority of novel early phase contexts. The novel tool allows for the assessment of the performance of many designs and two recently proposed methods were compared to the respective benchmarks as examples. Importantly, the novel tool provides the information on the proportion of correct selections, but not on other important characteristics in the regimen finding trials, for example, the average number of toxic responses and allocation proportions. These aspects remain unstudied and provide a valuable direction for the future research in terms of the non-parametric optimal designs.

In Chapter 3, we have considered a regimen finding trial as a generic problem of estimation on restricted parameter spaces in a Bayesian setting. While the problem has appeared to have quite a long historical record, it has received limited attention in the applications. Motivated by this, simple loss functions were proposed. It was found that they can lead to particular improvements when dealing with estimation on restricted spaces. A further consideration in the setting of the multivariate estimation is required and is an topic for further work. In Section 3.5, the new loss function was used as the allocation criterion incorporated into the 1-parameter CRM design. It has led to the same proportion of correct selections as the original allocation criterion, but to a lower average number of toxicities and imposing fewer patients to more toxic doses. Comparison to alternative methods has also revealed that the proposed criterion can lead either to the same accuracy and fewer toxicities or to a greater accuracy and the same number of toxicities. Therefore, such a design should be recommended for the application as leading to a more ethical patient allocation. This opens a door for further applications of the proposed loss function beyond the regimen finding studies.

In Chapter 4 we have considered a more general problem of the regimen finding in complex trials with the violated monotonicity assumption. The information-theoretic consideration of the problem lead to the class of novel criteria for studies

with multinomial outcomes. The proposal was demonstrated to have comparable or better characteristics than alternative methods and to preserve a flexibility by being applied in various setting with unknown monotonicity ordering. There are, however, important practical considerations that should be studied further with respect to the novel designs. Particularly, the influence of various implementation techniques of the missing and delayed outcomes which are common in regimen finding trials can be of interest in the setting with no parametric model. Furthermore, the proposed design assumes that outcomes of interest have a discrete distribution, but there is a growing interest in continuous endpoints. As an example, a continuous efficacy endpoint becomes a more common choice in a clinical practice (Hirakawa, 2012; Yeung et al., 2015, 2017). An extension of the proposed design for these cases is the subject of future research.

Appendices

Appendix A

A Modified Allocation Rule for the CRM

In Section 3.5.4, the comparison of the CRM and CIBP designs is presented. For the simulation results in the main text the skeleton with equivalent interval 0.05 and the prior MTD d_2 is used. In fact, the same qualitative results can be obtained for other skeletons. The performance of the CRM and CIBP if the skeleton is constructed using the prior MTD d_3 and d_4 are given in Table A.1 and Table A.2, respectively. The operating characteristics in case if the proposed allocation criterion (3.5.10) is used for the final selection are given in Table A.3. One can find that in this case the CIBP underestimates the MTD in the majority of the scenarios.

Table A.1: Proportions of dose selections and mean proportions of DLTs for CRM and the CIBP using $a = \{0.3, 0.4, 0.5\}$ and the prior distribution as described in Section 3.5, but the prior MTD d_3 is used for the skeleton construction. Results are based on 40000 replications.

	d_1	d_2	d_3	d_4	d_5	d_6	DLTs
			Scenario 1				
Toxicity	**25.00**	35.00	37.50	40.00	45.00	50.00	
CIBP(0.3)	67.59	22.50	6.44	2.60	0.74	0.13	28.73
CIBP(0.4)	64.12	22.53	8.29	3.77	1.08	0.21	29.98
CIBP(0.5)	61.78	22.23	9.61	4.80	1.30	0.28	31.31
CRM	63.27	21.28	9.45	4.52	1.28	0.21	31.02
			Scenario 2				
Toxicity	15.00	**25.00**	35.00	40.00	45.00	50.00	
CIBP(0.3)	22.86	47.00	23.37	5.45	1.15	0.17	23.94
CIBP(0.4)	22.89	45.59	23.64	6.25	1.40	0.23	25.64
CIBP(0.5)	23.47	44.65	23.12	6.97	1.49	0.29	27.14
CRM	24.38	44.26	22.64	6.81	1.64	0.27	27.23
			Scenario 3				
Toxicity	10.00	15.00	**25.00**	35.00	45.00	50.00	
CIBP(0.3)	3.76	23.83	46.95	21.62	3.51	0.33	21.74
CIBP(0.4)	3.84	23.95	46.45	21.98	3.47	0.32	23.25
CIBP(0.5)	3.53	23.99	46.36	22.17	3.56	0.40	25.09
CRM	3.52	24.57	46.51	21.54	3.50	0.37	25.17
			Scenario 4				
Toxicity	5.00	10.00	15.00	**25.00**	35.00	45.00	
CIBP(0.3)	0.17	4.45	25.60	46.06	20.42	3.29	20.10
CIBP(0.4)	0.16	4.15	25.51	46.38	20.60	3.20	21.62
CIBP(0.5)	0.15	4.00	25.66	46.66	20.38	3.14	23.37
CRM	0.18	4.08	26.25	46.39	20.24	2.85	23.40
			Scenario 5				
Toxicity	2.50	5.00	10.00	15.00	**25.00**	35.00	
CIBP(0.3)	0.00	0.27	5.71	27.24	44.13	22.64	18.39
CIBP(0.4)	0.00	0.24	5.06	26.41	45.88	22.41	19.96
CIBP(0.5)	0.00	0.27	4.97	27.23	45.48	22.05	21.41
CRM	0.00	0.25	4.58	27.98	45.88	21.31	21.56
			Scenario 6				
Toxicity	1.50	2.50	10.00	10.00	15.00	**25.00**	
CIBP(0.3)	0.00	0.02	2.42	9.37	26.83	61.36	15.83
CIBP(0.4)	0.00	0.02	1.74	7.62	27.18	63.44	17.20
CIBP(0.5)	0.00	0.05	1.74	7.75	27.89	62.58	18.20
CRM	0.00	0.04	1.29	6.63	28.55	63.49	18.36

Table A.2: Proportions of dose selections and mean proportions of DLTs for CRM and CIBP using $a = \{0.3, 0.4, 0.5\}$ and the prior distribution as described in Section 3.5, but the prior MTD d_4 is used for the skeleton construction. Results are based on 40000 replications.

	d_1	d_2	d_3	d_4	d_5	d_6	DLTs
				Scenario 1			
Toxicity	**25.00**	35.00	37.50	40.00	45.00	50.00	
CIBP(0.3)	64.38	23.48	7.74	3.21	1.03	0.15	29.25
CIBP(0.4)	62.80	22.43	9.04	4.15	1.35	0.24	30.56
CIBP(0.5)	59.16	22.61	10.51	5.68	1.73	0.31	31.93
CRM	60.09	21.50	10.73	5.63	1.74	0.30	31.96
				Scenario 2			
Toxicity	15.00	**25.00**	35.00	40.00	45.00	50.00	
CIBP(0.3)	21.23	46.23	24.93	6.07	1.36	0.18	24.61
CIBP(0.4)	21.78	44.21	25.04	6.94	1.74	0.29	26.46
CIBP(0.5)	22.46	43.10	24.44	7.82	1.87	0.31	27.91
CRM	23.79	42.60	23.60	7.78	1.92	0.32	28.17
				Scenario 3			
Toxicity	10.00	15.00	**25.00**	35.00	45.00	50.00	
CIBP(0.3)	3.50	22.07	47.11	23.09	3.92	0.31	22.38
CIBP(0.4)	3.37	21.87	46.68	23.73	3.97	0.38	24.31
CIBP(0.5)	3.54	22.51	46.14	23.47	3.93	0.41	25.67
CRM	3.53	23.31	45.96	22.88	3.89	0.44	26.19
				Scenario 4			
Toxicity	5.00	10.00	15.00	**25.00**	35.00	45.00	
CIBP(0.3)	0.13	3.62	23.72	47.04	22.04	3.44	20.85
CIBP(0.4)	0.16	3.71	23.79	47.30	21.75	3.29	22.57
CIBP(0.5)	0.15	3.67	24.02	47.06	21.78	3.31	23.99
CRM	0.15	3.56	24.48	47.12	21.39	3.30	24.57
				Scenario 5			
Toxicity	2.50	5.00	10.00	15.00	**25.00**	35.00	
CIBP(0.3)	0.00	0.20	4.69	25.61	45.86	23.64	19.11
CIBP(0.4)	0.00	0.22	4.35	25.54	46.44	23.44	20.82
CIBP(0.5)	0.00	0.21	4.14	25.68	46.36	23.61	22.10
CRM	0.00	0.23	4.09	26.50	46.24	22.94	22.45
				Scenario 6			
Toxicity	1.50	2.50	10.00	10.00	15.00	**25.00**	
CIBP(0.3)	0.00	0.03	1.74	8.63	26.63	62.97	16.33
CIBP(0.4)	0.00	0.03	1.33	6.53	25.96	66.16	18.01
CIBP(0.5)	0.00	0.02	1.24	6.20	26.59	65.94	18.88
CRM	0.00	0.05	1.17	6.44	27.95	64.38	18.88

Table A.3: Proportions of dose selections and mean proportions of DLTs for CRM and CIBP using $a = \{0.3, 0.4, 0.5, 0.6\}$ and the prior distribution as described in Section 3.5, but the prior MTD d_3 is used for the skeleton construction and the MTD is selected using the novel criterion (3.5.10). Results are based on 40000 replications.

	d_1	d_2	d_3	d_4	d_5	d_6	DLTs
				Scenario 1			
Toxicity	**25.00**	35.00	37.50	40.00	45.00	50.00	
CIBP(0.3)	74.85	17.07	5.36	2.07	0.57	0.08	28.73
CIBP(0.4)	69.02	19.12	7.49	3.33	0.87	0.17	29.98
CIBP(0.5)	64.28	20.65	9.12	4.52	1.19	0.23	31.31
CIBP(0.6)	58.83	22.17	10.90	5.97	1.82	0.30	32.63
CRM	63.27	21.28	9.45	4.52	1.28	0.21	31.02
				Scenario 2			
Toxicity	15.00	**25.00**	35.00	40.00	45.00	50.00	
CIBP(0.3)	32.69	44.56	17.77	4.03	0.83	0.10	23.94
CIBP(0.4)	28.78	44.97	19.72	5.25	1.10	0.18	25.64
CIBP(0.5)	26.61	44.34	21.09	6.35	1.35	0.24	27.14
CIBP(0.6)	23.71	42.91	23.30	7.82	1.94	0.31	28.82
CRM	24.38	44.26	22.64	6.81	1.64	0.27	27.23
				Scenario 3			
Toxicity	10.00	15.00	**25.00**	35.00	45.00	50.00	
CIBP(0.3)	6.25	30.85	44.27	16.23	2.22	0.18	21.74
CIBP(0.4)	5.29	29.01	45.06	17.95	2.45	0.25	23.25
CIBP(0.5)	4.21	26.71	45.62	20.13	3.02	0.32	25.09
CIBP(0.6)	3.81	24.55	45.84	21.84	3.60	0.36	26.39
CRM	3.52	24.57	46.51	21.54	3.50	0.37	25.17
				Scenario 4			
Toxicity	5.00	10.00	15.00	**25.00**	35.00	45.00	
CIBP(0.3)	0.43	6.82	32.95	42.94	15.05	1.80	20.10
CIBP(0.4)	0.28	5.75	30.69	44.37	16.70	2.21	21.62
CIBP(0.5)	0.21	4.72	28.28	45.80	18.49	2.49	23.37
CIBP(0.6)	0.19	3.86	26.35	46.51	20.24	2.85	24.61
CRM	0.18	4.08	26.25	46.39	20.24	2.85	23.40
				Scenario 5			
Toxicity	2.50	5.00	10.00	15.00	**25.00**	35.00	
CIBP(0.3)	0.00	0.60	8.54	34.20	41.13	15.53	18.39
CIBP(0.4)	0.01	0.43	6.88	31.60	43.16	17.93	19.96
CIBP(0.5)	0.00	0.39	5.67	30.01	44.74	19.19	21.41
CIBP(0.6)	0.00	0.22	4.63	27.74	45.66	21.73	22.62
CRM	0.00	0.25	4.58	27.98	45.88	21.31	21.56
				Scenario 6			
Toxicity	1.50	2.50	10.00	10.00	15.00	**25.00**	
CIBP(0.3)	0.00	0.12	3.31	12.32	33.03	51.22	15.83
CIBP(0.4)	0.00	0.08	2.27	9.30	31.25	57.11	17.20
CIBP(0.5)	0.00	0.10	1.96	8.59	30.63	58.72	18.20
CIBP(0.6)	0.00	0.03	1.29	6.25	27.43	65.00	19.36
CRM	0.00	0.04	1.29	6.63	28.55	63.49	18.36

Appendix B

Weighted Entropy design for trials with a binary endpoint

It was shown in Section 4.2.5 that the probability of the correct selection $\mathbb{P}(\nu = j)$ under Rule II has no analytical expression as it depends on the whole history of events. Instead, the exact characteristics can be found computationally as described in Appendix B.1. It will be shown that the computational approach is time-consuming even for moderate sample sizes and cannot be used to investigate the asymptotic behaviour. An approximation procedure for computing the expected sample size at each regimen is subsequently introduced in Appendix B.2. We use this approximation to investigate asymptotic properties of the design under Rule II in Appendix B.3. Below, we focus on the case of binary outcomes as in the examples for Phase I and Phase II clinical trials.

B.1 Computing Exact Operating Characteristics

B.1.1 Exact Algorithm

Below we describe an algorithm for computing proportions of each regimen selection for the WE design under Rule II using the "plug-in" estimator (4.2.14). As before we consider the sequential experiment with m regimens, N patients and a binary endpoint in which the regimen selection is allowed to change after each observation.

The "plug-in" estimator $\delta^{\kappa}(\hat{p}_{n_j}, \gamma) = \hat{\delta}^{\kappa}_{n_j}$ for regimen T_j given in Equation (4.2.13) after t observations enrolled in the experiment can be written as a function of the number of responses $x_j(t)$ and observations $n_j(t)$. We denote it by

$$\hat{\delta}^{\kappa}(x_j(t), n_j(t), \gamma) \;=\; \hat{\delta}^{\kappa}(x_j(t), n_j(t)).$$

The selection probability is then determined by the number of responses and observations on all regimens. As the method does not use the information borrowing between regimen, there are only two possibilities after an outcome of a current patient is observed: to stay at the estimated target regimen T_j or to switch to regimen T_i (which was the "second best" before). Either of these events can happen with probabilities $\{0, 1, \alpha_j, 1-\alpha_j\}$ where α_j is the probability of response associated with regimen T_j. For instance, probability 1 corresponds to regimen T_j being selected for the next subject regardless the previous observation. Assume that t subjects are already assigned in the experiment. Before the $t+1^{th}$ selection

is made, the transition matrix takes the form

$$\mathbb{P}_t = \begin{bmatrix} 0_{1,1} & \dots & 1_{1,j} & \dots & 0_{1,i} \dots & 0_{1,m} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0_{j,1} & \dots & \mathcal{P}_{j,j} & \dots & (1-\mathcal{P}_{j,j})_{j,i} \dots & 0_{1,m} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0_{m,1} & \dots & 1_{m,j} & \dots & 0_{m,i} \dots & 0_{m,m} \end{bmatrix}$$

where it is assumed that $j < i$ and $j, i \neq 1$, the double lower index corresponds to the position of the element,

$$\mathcal{P}_{j,j} = \mathbb{P}\left(\hat{\delta}^{(\kappa)}(x_j(t) + X_j, n_j(t) + 1) < \hat{\delta}^{(\kappa)}(x_i(t), n_i(t))\right)$$

and X_j is a binary random variable that takes value 0 with probability $1 - \alpha_j$ and 1 with probability α_j. The transition probability then is $\mathcal{P}_{j,j} =$

$$\begin{cases} 0, \text{ if } \hat{\delta}^{(\kappa)}(x_j, n_j + 1) > \hat{\delta}^{(\kappa)}(x_i, n_i) \text{ and } \hat{\delta}^{(\kappa)}(x_j + 1, n_j + 1) > \hat{\delta}^{(\kappa)}(x_i, n_i) \\ 1, \text{ if } \hat{\delta}^{(\kappa)}(x_j, n_j + 1) < \hat{\delta}^{(\kappa)}(x_i, n_i) \text{ and } \hat{\delta}^{(\kappa)}(x_j + 1, n_j + 1) < \hat{\delta}^{(\kappa)}(x_i, n_i) \\ \alpha_j, \text{ if } \hat{\delta}^{(\kappa)}(x_j, n_j + 1) > \hat{\delta}^{(\kappa)}(x_i, n_i) \text{ and } \hat{\delta}^{(\kappa)}(x_j + 1, n_j + 1) < \hat{\delta}^{(\kappa)}(x_i, n_i) \\ 1 - \alpha_j, \text{ if } \hat{\delta}^{(\kappa)}(x_j, n_j + 1) < \hat{\delta}^{(\kappa)}(x_i, n_i) \text{ and } \hat{\delta}^{(\kappa)}(x_j + 1, n_j + 1) > \hat{\delta}^{(\kappa)}(x_i, n_i) \end{cases}$$

Using the transition matrix, probabilities of all paths (all possible outcomes in the trial) can be computed. There are, however, 2^N possibilities that makes the procedure computationally costly for moderate values of N. At the same time, paths that reach the same number of outcomes $x_1, \dots, x_m$ and observations $n_1, \dots, n_m$, can be united to reduce the computational cost. However, the search and union of these paths can be costly as well. A trade-off is sought between merging paths often, in which case a little computational gain is expected and merging paths infrequently which means that many paths need to be searched over increasing computational complexity. We provide the illustration of the exact procedure in

the following section.

B.1.2 Illustration

Table B.1 provides the comparison of proportions of each regimen selections using the exact algorithm against 10^6 replicated trials in scenario 1 (Section 4.3) with $N = 20$.

Table B.1: Proportions of each regimen selections in scenario 1 (Section 4.3) using the exact computation and 10^6 replications.

Toxicity	0.06	0.12	0.15	0.18	0.24	0.36	0.40
Prior	0.25	0.30	0.35	0.40	0.45	0.50	0.55
Exact	4.04	10.09	18.85	22.09	26.52	12.97	5.42
Simulations	4.14	10.20	19.09	22.65	26.37	12.60	4.95

As expected, proportions of selections obtained by the exact algorithm are close to those obtained by simulations. The difference are minor and the simulation setting gives a good understanding of the design behaviour. The computational time for different sample sizes $N = \{10, 20, 25, 30\}$ using the paths unions is given in Table B.2. The paths are not united for $N = 10$ and united after selections 11, $11 - 15$ and $11 - 23$ for $N = 20, 25, 30$, respectively.

Table B.2: Average computational time (min) for the exact algorithm and the simulations approach in a scenario with 7 regimens and total sample sizes $N = \{10, 20, 25, 30\}$.

Sample size	All paths	United paths	10^6 replications
$N = 10$	0.001	-	16.33
$N = 20$	0.92	0.64	27.95
$N = 25$	24.44	10.60	32.58
$N = 30$	942.08	372.78	38.43

The exact algorithm becomes infeasible for $N > 25$ while it takes much less time for smaller sample sizes. The union procedure decreases the computation time for all sample sizes and makes the computation for $N = 25$ three times faster than the simulation approach and more than 2 times faster than the exact algorithm without united paths. For $N = 30$ the exact computation takes nearly 10 more

time than the simulation study. Therefore, the former one seems to be a reasonable substitution for larger sample sizes.

B.2 Approximation Procedure for Expected Number of Observations

B.2.1 Approximation

Both exact algorithm and simulation approach are computationally infeasible for large sample sizes. Therefore, these approaches cannot be effectively used to study the large samples behaviours, for example, using different values of parameter κ in many different scenarios. The main challenge of the exact procedure is to trace all possible trajectories and to estimate the expected sample size at each regimen. We propose the following procedure that approximates the expected number of patients at each regimen and allows computing the selection proportions.

The core idea of the approximate is to replace the exact values $x_j(t)$ and $n_j(t)$ with their expectations. Following the notations above

$$\mathbb{E}\left(n_j^{(t)}\right) = \sum_{u=1}^{t} \mathbb{P}\left((\hat{\delta}_j^{(\kappa)}(n_j^{(u-1)}, x_j^{(u-1)}) = \min_i \hat{\delta}_i^{(\kappa)}(n_i^{(u-1)}, x_i^{(u-1)}))\right) := \sum_{u=1}^{t} \phi_j(u).$$

Using the law of the total probability and the transition matrix

$$[\phi_1(t) \dots \phi_m(t)] = [\phi_1(t-1) \dots \phi_m(t-1)]\, \mathbb{P}_t$$

where the only non-trivial quantity of transition matrix $\mathbb{P}_t$

$$\mathcal{P}_{j,j} = \mathbb{P}\left(\hat{\delta}^{(\kappa)}(\mathbb{E}\left(x_j(t)\right) + X_j, \mathbb{E}\left(n_j(t)\right) + 1) < \hat{\delta}^{(\kappa)}(\mathbb{E}\left(x_i(t)\right), \mathbb{E}\left(n_i(t)\right))\right)$$

and $\mathbb{E}\left(x_j^{(t)}\right) = \alpha_j \mathbb{E}\left(n_j^{(t)}\right)$. This allows computing the expected number of observations and expected number of responses at all regimens.It is demonstrated below

that the approximate procedure can be effectively used to study the behaviour of the WE design under Rule II.

B.2.2 Illustration

We consider two scenarios with two regimens only. In scenario 1, the target percentile is $\gamma = 0.20$ and the prior probabilities of response are $\hat{\mathbf{p}} = [0.20, 0.20]^{\mathrm{T}}$. In scenario 2, the target is $\gamma = 0.99$ and the prior is $\hat{\mathbf{p}} = [0.99, 0.99]^{\mathrm{T}}$. In both examples we fix $\beta = 5$, the probability associated with regimen 2, $\alpha_2 = 0.40$, and $N = 50$. We study how the expected number of observations on regimen 1 depends on parameter α_1. The results for the approximation and the simulations approach are given in Figure B.1.

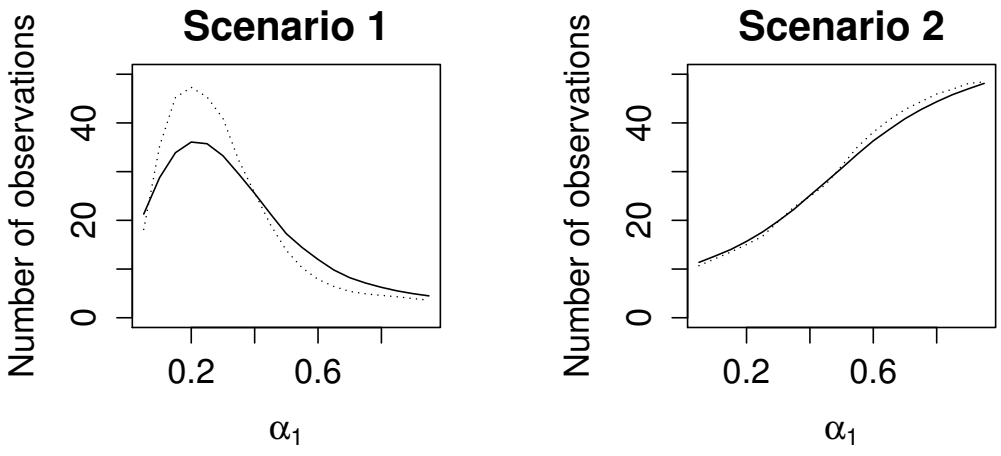

Figure B.1: Expected number of observations at regimen 1 using different values of α_1 obtained by the approximation (dashed line) and by 10^6 simulations (solid line). Left panel: the target probability $\gamma = 0.20$. Right panel: the target probability $\gamma = 0.99$.

In scenario 1, both the approximation and the simulation approach have a peak at the point $\alpha_1 = \gamma = 0.2$ and the overall dependence on the probability α_1 is represented correctly by the approximation. However, the exact values obtained by the approximation are largely overestimated in the neighbourhood of γ. The number of observations by both approaches is the same for $\alpha_1 = \alpha_2 = 0.4$ and the approximation is accurate for $\alpha_1 > 0.4$. Regarding scenario 2, the expected number of observations by both methods reaches the peak at $\alpha_1 = 0.99$ and has

the half total sample size for $\alpha_1 = 0.4$. The expected number of observations by the approximation is very close to ones obtained by simulations for all values of α_1. Importantly, the computational time for the simulated results in one scenario takes nearly 18 hours (10^6 replications for values α_1 on a grid $(0.05, 0.95)$ with step 0.05), while the approximation procedure takes less than one second. Due to this gain, the approximation is used to demonstrate the asymptotic behaviour of the PCS in the following section.

B.3 Asymptotic Behaviour of the Weighted Entropy Design Using the Approximation

We use the approximation of the expected number of observations under Rule II and Equation (4.2.20) to find the lower bound of the PCS in Equation (4.2.18) in Theorem 6 and to investigate the asymptotic properties of the design for different values of the parameter κ.

To demonstrate the large sample behaviour an illustrative trial with two regimens is considered. The true probabilities of responses are $(0.3, 0.5)$ for regimen 1 and regimen 2, respectively. The goal of the trial is to find the regimen corresponding to the maximum probability of response $\gamma = 1$. Following Section 4.4, we fix prior parameters $\beta_1 = \beta_2 = 2$ and $v_1 = v_2 = 0.99 \times 2$. The expected number of observations on the suboptimal arm, the proportion of observations on the suboptimal arm and the lower bounds of the PCS for different values of N and κ are given in Figure B.2.

As it was shown in Theorem 6, the total number of observations on the suboptimal regimen locks-in for $\kappa = 1/2$ which results in the design's inconsistency. For $\kappa > 1/2$ the total number of observations on the suboptimal regimen is a non-decreasing function of the number of observations. "Jumps" in the graph of the expected number of observations correspond to the sample size when the subop-

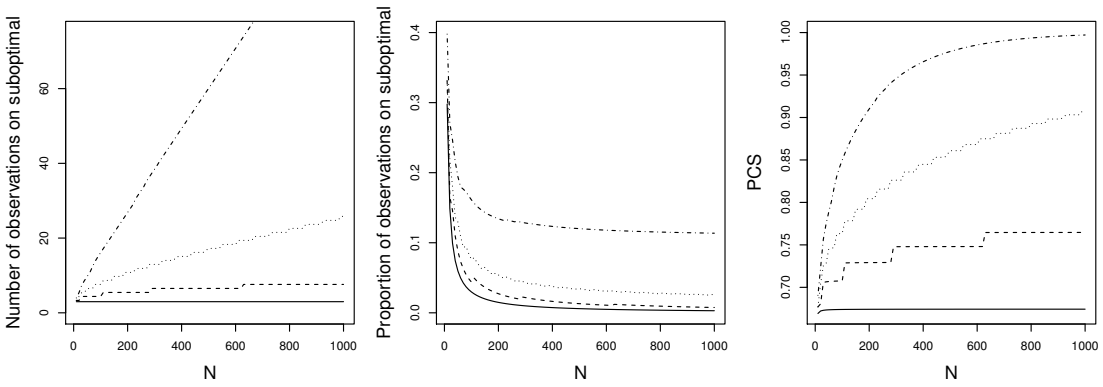

Figure B.2: Operating characteristics of the approximation: (i) Expected number of observations on suboptimal regimen, (ii) expected proportion of observations on suboptimal regimen, and (iii) the normal approximation for the PCS lower bound (4.2.18) for different sample sizes N and $\kappa = 0.5$ (solid), $\kappa = 0.55$ (dashed), $\kappa = 0.60$ (dotted), $\kappa = 0.70$ (dotted-dashed).

timal regimen is selected. The jumps become more regular as κ increases. For a given κ the number of observations between jumps increases with a sample size due to the decreasing rate of the penalty term.

B.4 Calibration of the Design Parameters for Phase I Clinical Trials

The WE design applied in the context of Phase I clinical trials (Section 4.3) requires specification of the prior distributions and the parameters of safety constraint. We would like to find the *operational* parameters, the parameters for which the WE design selects the TR with a high probability in many different scenarios. We provide a guideline how these parameters can be calibrated below.

B.4.1 Operational Prior

A prior distribution for regimen j can be specified through the mode of the prior distribution, $\hat{p}_{\beta_j} = \frac{\nu_j}{\beta_j}$. It is assumed that an investigator has the same (limited) knowledge about all regimens and $\beta_1 = \ldots = \beta_7$ are chosen. Then, the problem reduces to the choice of $\nu_1, \ldots, \nu_7$ which are equal to $\hat{p}_{\beta_1}, \ldots, \hat{p}_{\beta_1}$ in this case, respectively. To guarantee the procedure to start from T_1 we set $\hat{p}_{\beta_1} = \gamma$. Although

the initial ordering (with respect to increasing toxicities) might be misspecified, a clinician would like to test the regimen in the specific order. Therefore, increasing values of $\hat{p}_{\beta_1} < \ldots < \hat{p}_{\beta_7}$ are required. For simplicity, we would consider the prior parameters which assume a linear increase in toxicity probabilities.

To calibrate the operational parameters of prior distributions, we consider six scenarios with different locations of the TR given in Figure B.3. Given $m = 7$, we set the difference between prior toxicities at the first regimen T_1 and the last regimen T_7 and, then, interpolate the linear curve for the rest. We would define $step = \hat{p}_{\beta_7} - \hat{p}_{\beta_1}$. Then, we vary values of $step$ and β in each scenario. Larger values of β and the $step$ correspond to more conservative escalation schemes as an investigator needs more observations on each particular regimen to escalate. Therefore, one can expect a set of prior parameters that lead to similar proportions of correct selections. The proportion of correct selections (PCS) using different values of $step$ and β is given in Figure B.3.

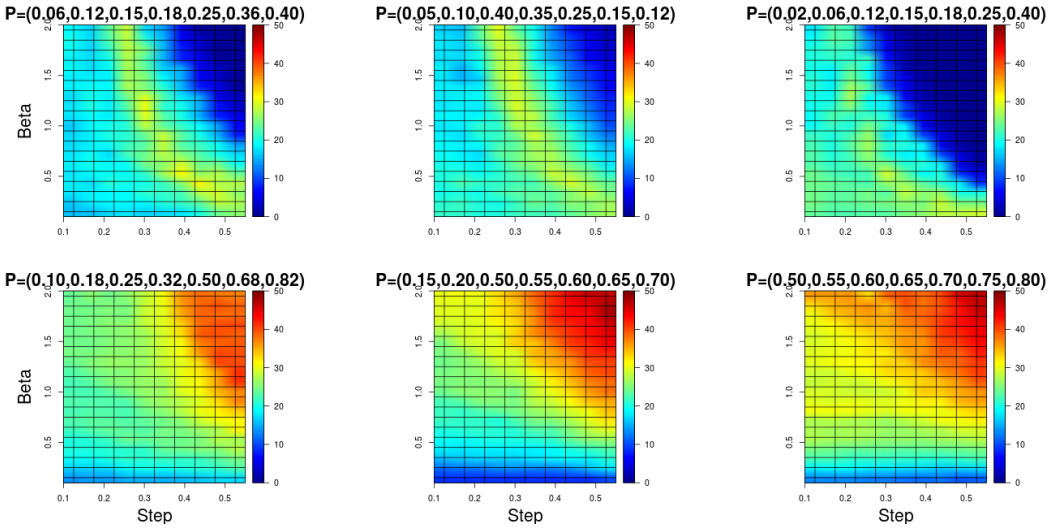

Figure B.3: Proportion of correct selections by WE design using different values of β (vertical axis) and $step$ (horizontal axis). Results are based on 10^6 replications.

Brighter colours correspond to higher values of the PCS. A conservative prior (top right corner of the grid) prevents the WE design from correct selections in upper line graphs scenarios as higher regimens can be hardly reached. At the same time,

it leads to accurate selections in scenarios with highly toxic regimens (lower line scenarios). In contrast, less conservative choices result in higher PCS in upper line scenarios and lower PCS in lower line ones. There is a clear trade-off between the ability to investigate higher regimen and the desire to prevent the high number of toxic selections. Therefore, the geometric mean of the PCS over all scenarios is chosen as the criterion for the operational prior choice. The geometric mean for different values of parameters is given in the Figure B.4.

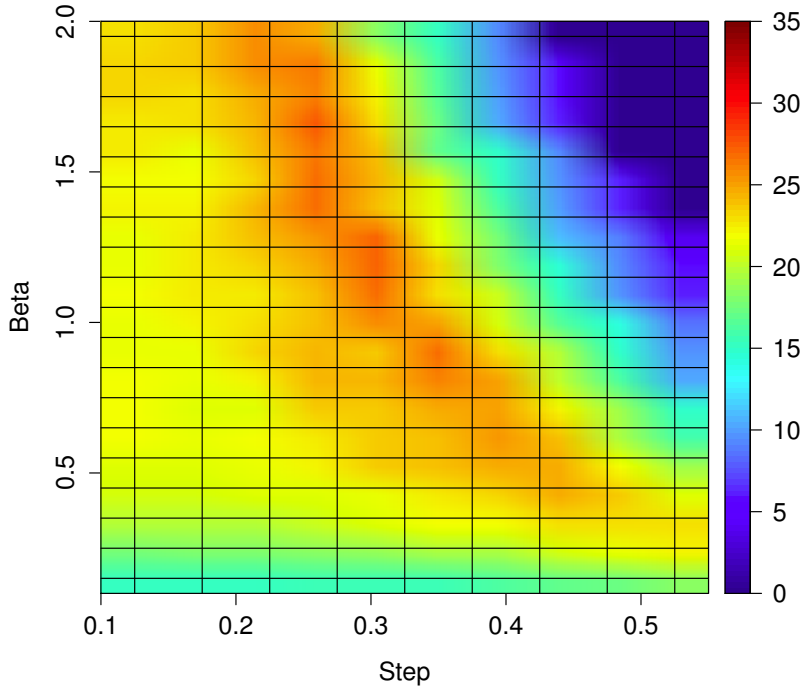

Figure B.4: Geometric mean of proportions of correct selections by WE design using different values of parameters β (vertical axis) and *step* (horizontal axis) in six scenarios. Results are based on 10^6 replications.

There is a set of the prior parameter leading to the same geometric mean of the PCS. We choose a prior that carries a limited information $\beta = 1$ and seek for the rate maximising the geometric mean. Thus, the following vector of modes is chosen

$$\hat{\mathbf{p}} = [0.25, 0.30, 0.35, 0.40, 0.45, 0.50, 0.55]^{\mathrm{T}}.$$

We also consider the case of the larger sample size $N = 25$ using the same sce-

narios to illustrate the choice of prior in trials with a slightly greater sample. The geometric mean of proportions of correct selection using $N = 25$ is given in Figure B.5. There is a similar dependence pattern on β and *step*. As one would expect, the set of the equivalent operational prior parameters is now wider. This means that the importance of the prior distribution choice decreases as the sample size increases.

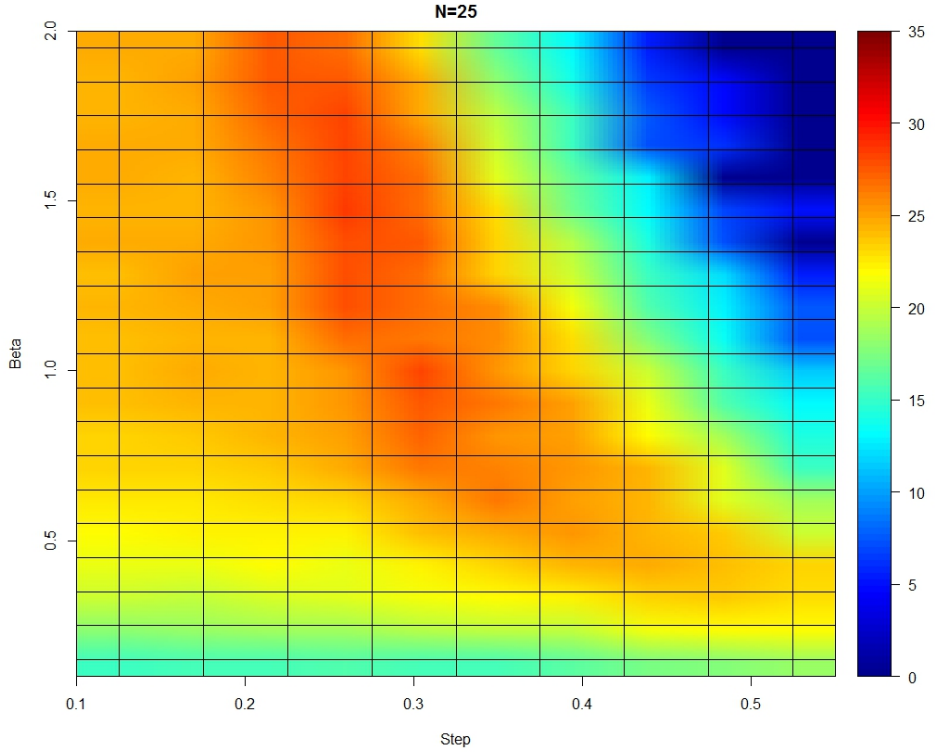

Figure B.5: Geometric mean of proportions of correct selections by WE design using different values of β (vertical axis) and *step* (horizontal axis). The total sample size $N = 25$ is used. Results are based on 10^6 replications.

B.4.2 Safety Constraint

Parameters r and γ^* define the strictness of the safety constraint. Greater values of r and smaller values of γ^* would lead to a more conservative escalation. While it would help to avoid high risk in unsafe scenarios, it would also prevent the correct

selections in flat safe scenarios. There is a clear trade-off in the choice of these parameters that is precisely studied below.

Let us consider two extreme scenarios: a linear flat regimen response shape (scenario 1 in Section 4.3) with the TR far from the bottom (T_5) and the scenario with no safe regimens at all (scenario 6 in Section 4.3). The proportion of correct selections (scenarios 1) and the proportion of early terminations (scenario 6) by WE design using different values of γ^* and r are given in Table B.3.

Table B.3: Operating characteristics of WE design in linear and unsafe scenarios using different parameters of the safety constraint. Figures in the upper line correspond to termination proportions in the unsafe scenario 6. Figures in the lower cells correspond to the PCS in the flat scenario 1. Subsequently used parameters of the safety constraint are in bold. Results are based on 10^6 replications.

	r							
	0.010	0.015	0.020	0.025	0.030	0.035	0.040	0.045
$\gamma^* = 0.55$	0.00	0.32	4.32	18.47	36.15	49.06	61.49	75.70
	26.47	26.65	26.40	26.05	26.85	25.03	24.10	20.23
$\gamma^* = 0.50$	0.15	2.50	17.76	38.75	52.74	63.06	74.94	87.22
	26.27	26.22	26.53	27.24	25.46	23.30	20.35	17.10
$\gamma^* = 0.45$	1.13	12.72	35.72	56.49	67.16	**77.55**	86.53	93.49
	26.15	26.02	26.81	25.18	24.26	**23.15**	18.16	11.05
$\gamma^* = 0.40$	7.47	37.95	59.49	70.52	80.53	88.32	94.18	97.63
	26.04	25.91	24.90	21.98	17.66	17.47	11.05	3.51
$\gamma^* = 0.35$	33.98	58.22	74.42	84.14	90.52	94.86	97.90	99.20
	25.65	24.54	20.45	15.55	13.77	9.21	6.25	0.70
$\gamma^* = 0.30$	55.51	77.02	87.21	92.99	96.50	98.55	99.37	99.83
	24.21	18.09	14.40	11.42	7.13	0.95	0.08	0.04

The upper line corresponds to the proportion of times the WE design terminates the trial early in the unsafe scenario 6. The lower line corresponds to the proportion of trials the TR is selected in the linear scenario 1. The most relaxed safety constraint corresponds to the left upper corner: no trials are terminated in the highly toxic scenario and the PCS in the linear scenario is relatively high. The right lower corner corresponds to the strictest safety constraint: nearly all trials are terminated in the unsafe scenario, but the method will often not find TR in the linear scenario. Therefore, the trade-off is to sacrifice the accuracy of the method when the TR is far from the bottom in order to prevent the selection of a highly

toxic regimen. Consequently, $\gamma^* = 0.45$ and $r = 0.035$ are used.

Appendix C

Weighted Entropy Design for Single Agent Phase I/II Trials with trinary outcomes

The proposed WE design can be applied to a wide range of Phase I/II clinical trials. While the performance of the WE design was demonstrated in the context of the motivating combination-schedule trial, it can be also applied to single agent dose finding Phase I/II trials of molecularly targeted agents for which several model-based designs were recently proposed (see e.g. Wages and Tait, 2015; Riviere et al., 2016). Here we show the comparison of the proposed design to the currently used and provide a step-by-step algorithm how the parameters of the novel design can be calibrated.

C.1 Simulation Setting

We consider $m = 6$ doses and $N = 60$ patients. The dose-toxicity relationship is known to be a non-decreasing function, but a clinician expects either a plateau or

an umbrella shape for the dose-efficacy curve. As before we would assume that it takes twice longer to evaluate an efficacy response than a toxicity response. To conduct the trial in a timely manner, the next cohort of patients is allocated after the toxicity data for the previous cohort are available. The upper toxicity and the lowest efficacy bounds are $\phi = 0.35$ and $\psi = 0.20$. The goal is to study the ability of the WE design to identify the *optimal* and *correct* doses. A dose is called *optimal* if it is safe, has maximal efficacy and minimal toxicity while a safe dose with maximum efficacy (irrespective of it also having lowest toxicity) is called *correct*.

We consider 14 scenarios (Figure 4.7) that were used for the motivating trial simulations: eight plateau scenarios (1-8) suggested by Riviere et al. (2016), 4 umbrella shaped scenarios (9-12) studied by Wages and Tait (2015) and two scenarios with no correct doses (13-14, due to inefficacy and toxicity, respectively)

In the analysis we focus on (i) the proportion of optimal/correct selections, (ii) the average number of toxic responses, (iii) the average number of efficacy responses. The study is performed using `R` (R Core Team, 2015) and 10,000 replications for each scenario. We compare the characteristics with the "RIV" design proposed by Riviere et al. (2016) and the "WT" design developed by Wages and Tait (2015). Parameters of the designs are chosen as in the original proposals with an exception of using cohort size $c = 3$ and 80% confidence intervals for stopping rules for the WT design.

C.2 Design Specifications

As before, we use the target toxicity $\gamma_t = 0.01$ and the target efficacy $\gamma_e = 0.99$. Due to the known toxicity ordering, the design is restricted to satisfy the coherence principals given in Equation (4.5.5) and Equation (4.5.6) with $q = 1$. We consider non-randomised and hybrid randomised versions of the WE design to study an allocation rule impact. As design specifications for both variants of the design are

similar, we provide the algorithm for the non-randomised WE design only.

C.2.1 Operational Prior

Parameters $\beta_{t,i} = \beta_{e,i} = 1$ of the prior Beta distribution in (8) are chosen for all dose levels $i = 1, \ldots, m$ to emphasize a limited available information. Parameters $\nu_{t,i}$ and $\nu_{e,i}$ (which coincide the prior probabilities of toxicity and efficacy for $\beta_{t,i} = \beta_{e,i} = 1$) are specified such that the WE design leads to accurate optimal dose selection in many different scenarios. The prior values of $\nu_{t,i}$ and $\nu_{e,i}$ are calibrated over scenarios 1-8 with different locations of the optimal and correct doses. There are two restrictions on the prior parameters: the escalation should start at the first dose and no dose skipping is allowed. To restrict number of possible parameters to be calibrated over, we assume that prior efficacy and toxicity probability increases linearly as $\nu_{t,i} = start_t + w_t \times i$ and $\nu_{e,i} = start_e + w_e \times i$. Then, we search for the values of $start_t, start_e, w_t, w_e$ such that the geometric mean of the proportion of optimal selection over all scenarios is maximised.

Prior vectors of toxicity probabilities

$$\hat{\mathbf{p}}_t^{(0)} = [0.05, 0.14, 0.23, 0.32, 0.41, 0.50]^{\mathrm{T}}$$

and efficacy probabilities

$$\hat{\mathbf{p}}_e^{(0)} = [0.55, 0.58, 0.61, 0.64, 0.67, 0.70]^{\mathrm{T}}$$

are subsequently used for the non-randomised WE design.

Similarly, vectors of prior toxicity

$$\hat{\mathbf{p}}_t^{(0)} = [0.25, 0.35, 0.45, 0.55, 0.65, 0.75]^{\mathrm{T}}$$

and

$$\hat{\mathbf{p}}_e^{(0)} = [0.65, 0.69, 0.73, 0.77, 0.81, 0.85]^{\mathrm{T}}$$

efficacy probabilities are used for the randomised WE(R) design. It was found that the randomised WE(R) design is more robust to the choice of the prior parameter than non-randomised WE.

C.2.2 Safety Constraint

To set the time-varying safety constraint, we use $\zeta_N = 0.30$ and calibrate ϕ^*, r_t using the highly toxic scenario 14 and the flat scenario 6. These two scenarios are chosen to represent the trade-off in the safety constraint. The proportion of correct selections (terminations) and the mean number of patients involved in a trial for different parameters values are given in Figure C.1. The mean number

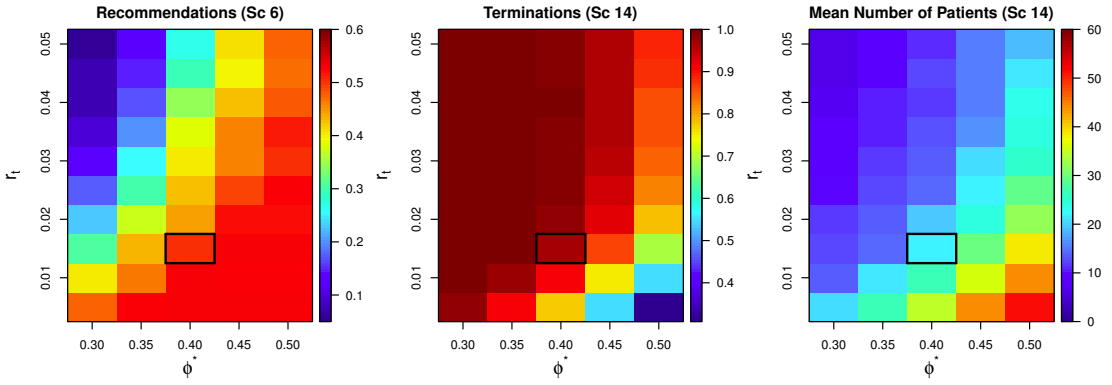

Figure C.1: Safety constraint parameters calibration: $\phi^* \in (0.3, 0.5)$, $r_t \in (0, 0.05)$ in scenarios 6 and 14. The proportion of correct selections (terminations) and the mean number of patients in a trial (scenario 14). The final choice is marked by a black frame. Results are based on 10^4 replications.

of patients in scenario 6 does not vary a lot and the corresponding graph is not shown. In scenario 6 the highest proportion of the optimal selections corresponds to the least strict safety constraint (right bottom corner), but only 35% of trials in scenario 14 are then terminated. At the same time, the most strict rule (left top corner) results in 100% of terminations in scenario 14, but only in 5% of correct selection in scenario 6. Parameters $r_t = 0.0125$ and $\phi^* = 0.4$ are chosen for subsequent study as a reasonable trade-off. The same parameters of the safety

constraints are used for the randomised design.

C.2.3 Futility Constraint

We calibrate the futility constraint by fixing $\xi_N = 0.50$ and tuning ψ^* and r_e using two opposite scenarios - 2 and 13. In scenario 2 all doses have the same efficacy probability. In scenario 13 there are no correct doses as all efficacious doses have unacceptable toxicity. The proportion of correct selections (terminations) and the mean number of patients are given in Figure C.2. Since the mean number of

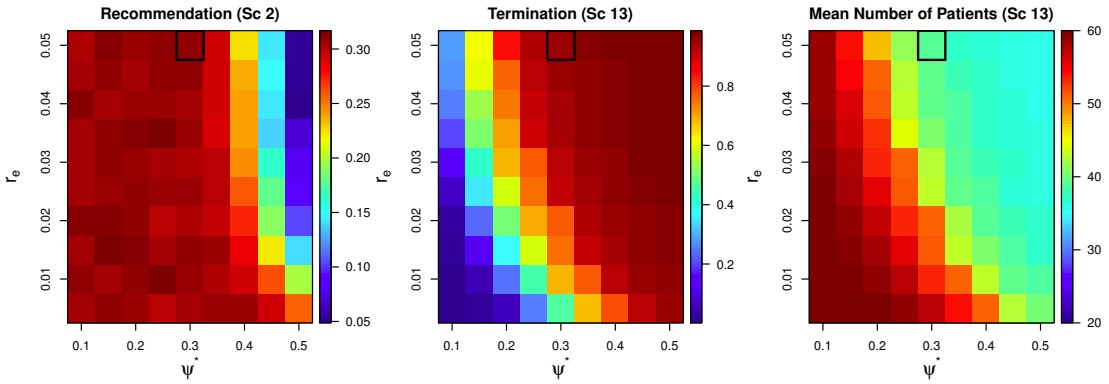

Figure C.2: Futility constraint parameters calibration: $\psi^* \in (0.1, 0.5)$ and $r_e \in (0, 0.05)$ in scenarios 2 and 13. The proportion of correct selections (terminations) and the mean number of patients in a trial (scenario 13). The final choice is in the black frame. Results are based on 10^4 replications.

patients in scenario 2 does not vary this graph is not shown. A stricter constraint is favourable in scenario 13 and less favourable in scenario 2 while the opposite is true for less strict constraints. Subsequently, parameters $\psi^* = 0.3$ and $r_e = 0.05$ are used for both non-randomised and randomised designs.

C.3 Operating Characteristics

The results of the comparison in scenarios 1-8 with plateau dose-efficacy relation and in scenarios 13-14 with no correct doses are summarised in Table C.1 and Table C.2. Each figure in Table C.1 corresponds to proportions of optimal or correct dose selections. The detailed results, such as the selection proportions and mean number of patients on each dose, are given in Table C.4 and Table C.5.

Table C.1: Proportion of optimal and correct dose selections in scenarios 1-8 and 13-14 using $N = 60$ patients and $m = 6$ doses. Results are based on 10^4 replications.

Scenario	1	2	3	4	5	6	7	8	13	14
Proportion of the optimal dose selections										
WE	58.8	30.3	65.8	71.1	60.3	53.5	60.0	37.8	95.2	96.9
WE (R)	72.0	35.0	51.0	69.9	54.5	56.7	48.5	36.4	93.2	97.4
RIV	57.0	60.2	48.4	53.7	55.3	55.9	37.9	43.0	91.9	91.0
WT	19.6	41.9	29.3	25.3	27.0	65.2	27.1	26.1	91.5	90.4
Proportion of the correct dose selections										
WE	60.3	90.0	87.5	79.8	89.7	53.5	77.8	91.5	-	-
WE (R)	89.3	97.0	92.8	89.1	87.5	56.7	88.3	91.1	-	-
RIV	94.1	96.6	83.8	82.3	80.6	55.9	89.9	77.4	-	-
WT	98.1	97.6	93.6	86.5	80.0	65.2	93.3	81.1	-	-

With respect to the *optimal* dose selections, both versions of the proposed design perform comparably or better than model-based designs in the majority of scenarios. The WE design without randomisation leads to a considerable improvement in scenarios 3-5, 7 and 13-14 and outperform the best model-based alternative by up to 20%. While the randomised WE(R) shows the comparable to the best model-based alternative performance in scenarios 3,5 and 6, it also results in more accurate optimal dose selections in scenarios 1, 4, 7, 13-14. However, both WE and WE(R) are outperformed by RIV in scenarios 2 and 8 in which dose-toxicity and dose-efficacy curves are flat in the neighbourhood of the optimal dose. While RIV recommends the lowest dose by default, such small differences in toxicity and efficacy probabilities are difficult to be found using the small sample size. At the same time, the absence of a parametric model is not found to be a problem in any other cases. WT design outperforms all other designs in scenario 6 with the optimal dose being the highest safe one. Generally, WT is less conservative as it favours safe doses with higher toxicities that results in a low proportion of optimal selections if the optimal dose is not the highest safe one (see also Table C.4 and Table C.5).

Considering the proportion of the *correct* selections, WT outperforms RIV in all scenarios and has the best performance among all alternatives in scenarios 1-3

Table C.2: Mean number of toxicity and efficacy responses in scenarios $1-8$ and $13-14$ using $N = 60$ patients and $m = 6$ doses. Results are based on 10^4 replications.

Scenario	1	2	3	4	5	6	7	8	13	14
Toxicity responses										
WE	3.1	6.0	2.8	4.5	7.9	11.0	5.9	5.6	11.7	11.0
WE (R)	4.0	6.9	4.3	6.0	8.7	10.8	6.9	6.9	13.0	10.9
RIV	5.5	8.1	6.0	10.0	12.1	13.2	9.6	9.3	11.0	**11.5**
WT	6.8	6.7	7.3	13.2	13.5	14.7	10.0	9.1	11.2	12.1
Efficacy responses										
WE	28.5	24.0	33.0	30.0	29.8	19.4	34.6	28.9	6.0	9.1
WE (R)	27.4	24.0	34.5	32.2	29.8	19.2	36.5	29.2	7.1	8.9
RIV	38.0	24.0	34.6	35.0	29.4	21.4	39.1	29.3	6.2	9.6
WT	41.5	24.0	35.5	37.0	32.3	24.4	39.9	29.2	5.4	9.7

and 6-7. In the rest of scenarios, WT has either comparable or worse performance than the randomised WE(R). Comparing WE and WE(R), the randomised design is more robust in terms of correct selections with the largest difference in scenario 1. Here, the chosen prior would not escalate to dose 6 once the optimal is already found at dose 5 if no randomisation is used.

In terms of toxicities we find that the non-randomised WE design results in a considerably lower number of toxicities in almost all scenarios with the largest difference observed in scenario 4. As the WT approach is less conservative, it results in a greater number of toxicities, but also leads to the highest average number of efficacies in all scenarios. In contrast, the cost of the WE's lowest number of toxicities is a smaller number of efficacies. In scenarios 13 and 14 with no optimal and correct doses, all alternatives result in nearly the same average number of toxicities and efficacies.

The results of the comparison in scenarios 9-12 with an umbrella-shaped dose-efficacy relationship and only one correct dose are given in Table C.3. Overall, WE designs lead to more robust optimal dose selections in non-monotonic scenarios. The WE design with no randomisation outperforms RIV by up to 35% and WT by up to 6%. WT has the highest proportion of the optimal dose selection in scenarios

Table C.3: Proportions of optimal dose selections, mean number of toxicity and efficacy responses in scenarios 9-12 using $N = 60$ patients and $m = 6$ doses. Results are based on 10^4 replications.

Scenario	9	10	11	12	9	10	11	12	9	10	11	12
	Optimal selection				Toxicity responses				Efficacy responses			
WE	54.7	55.9	46.5	80.1	4.5	5.7	10.0	1.8	29.7	25.6	27.4	35.4
WE(R)	56.7	56.2	47.9	70.9	5.5	6.7	10.0	3.3	28.4	24.6	27.1	32.7
RIV	20.3	35.3	46.0	96.1	5.0	6.3	12.8	2.6	26.3	23.9	28.7	35.5
WT	50.1	49.3	56.9	75.9	5.5	5.9	12.2	2.4	27.9	24.8	29.4	36.3

11 with nearly 10% difference compared to the non-randomised WE. The RIV design is more conservative and recommends d_1 with the highest probability that results in the best performance in scenario 12, but poor performance in other cases. The non-randomised WE is more favourable compared to the randomised version due to a single correct dose in each scenario. The average number of toxicities of the WE design is again the safest alternative. In contrast to the scenarios with a plateau, it can now also result in a larger number of efficacy responses (e.g. in scenario 9) due to the non-monotonic shape of the dose-efficacy curve.

Overall, the proposed approaches have better or comparable operating characteristics in 9 out of 14 considered scenarios even with less information used in a trial. Comparing two assignment rule of the WE design, the non-randomised WE is always less accurate in terms of the correct dose selection. As the result, the WE design without randomisation should be preferred if only one correct dose is expected or a clinician is cautious about the toxicity profile, while the randomised WE is a robust choice if multiple *correct* doses are expected.

Table C.4: Operating characteristics of WE, WE(R), RIV and WT designs in scenarios 1-4: selection proportions, mean number of patients assigned to a dose (in brackets), termination proportion (Term), mean number of toxicity (T) and efficacy (E) responses. The optimal dose is in bold and correct doses are underlined. Results are based on 10^4 replications.

	d_1	d_2	d_3	d_4	d_5	d_6	Term	T	E
				Scenario 1					
	(.005;.01)	(.01;.10)	(.02;.30)	(.05;.50)	**(.10;.80)**	(.15;.80)			
WE	0.0	0.1	2.3	37.4	58.8	1.5	0.0	3.1	28.5
	(6.1)	(6.3)	(9.5)	(20.5)	(17.4)	(0.3)			
WE(R)	0.0	0.2	1.0	9.5	72.0	17.3	0.0	4.0	27.4
	(5.1)	(5.0)	(8.2)	(13.8)	(21.7)	(6.2)			
RIV	0.0	0.0	0.7	4.5	57.0	37.1	0.7	5.5	38.0
	(3.3)	(3.7)	(4.8)	(7.7)	(21.5)	(19.0)			
WT	0.0	0.1	0.4	1.3	19.6	78.5	0.1	6.8	41.5
	(3.9)	(1.6)	(2.4)	(3.6)	(11.4)	(37.1)			
				Scenario 2					
	(.01;.40)	(.04;.40)	(.10;.40)	(.25;.40)	(.50;.40)	(.70;.40)			
WE	30.3	26.3	20.8	12.5	6.3	3.7	0.2	6.0	24.0
	(18.3)	(17.0)	(13.7)	(7.9)	(2.7)	(0.4)			
WE(R)	35.0	29.0	21.5	11.5	2.8	0.2	0.1	6.9	24.0
	(15.0)	(15.7)	(15.7)	(10.0)	(3.3)	(0.4)			
RIV	60.2	20.3	8.8	7.3	2.6	0.3	0.6	8.1	24.0
	(18.8)	(13.0)	(10.7)	(10.7)	(5.8)	(0.8)			
WT	41.9	24.5	16.9	14.2	2.4	0.0	0.0	6.7	24.0
	(23.1)	(13.0)	(10.7)	(8.6)	(3.4)	(1.3)			
				Scenario 3					
	(.01;.25)	(.02;.45)	**(.05;.65)**	(.10;.65)	(.20;.65)	(.30;.65)			
WE	0.6	12.6	65.8	18.2	2.8	0.1	0.0	2.8	33.0
	(7.2)	(15.9)	(29.2)	(6.7)	(0.9)	(0.0)			
WE(R)	0.7	6.6	51.0	30.5	10.6	1.2	0.0	4.3	34.5
	(6.4)	(9.6)	(21.3)	(16.1)	(5.8)	(0.9)			
RIV	2.0	14.3	48.4	19.2	9.8	6.4	0.0	6.0	34.6
	(6.3)	(9.8)	(15.5)	(12.9)	(10.4)	(5.1)			
WT	1.4	4.9	29.3	29.9	22.6	11.8	0.0	7.3	35.5
	(6.1)	(5.0)	(13.8)	(14.8)	(12.0)	(8.3)			
				Scenario 4					
	(.01;.05)	(.02;.25)	(.05;.45)	**(.10;.70)**	(.25;.70)	(.50;.70)			
WE	0.0	0.5	19.6	71.1	8.5	0.3	0.0	4.5	30.0
	(6.2)	(7.9)	(17.8)	(25.2)	(2.9)	(0.1)			
WE(R)	0.0	1.3	9.2	69.9	18.5	1.0	0.0	6.0	32.2
	(5.4)	(6.9)	(13.4)	(24.1)	(9.2)	(0.8)			
RIV	0.0	0.7	8.2	53.7	28.6	8.5	0.4	10.0	35.0
	(3.8)	(5.0)	(9.1)	(19.0)	(15.9)	(7.0)			
WT	0.0	0.4	2.1	25.3	61.2	10.8	0.2	13.2	37.0
	(4.5)	(2.4)	(4.0)	(13.5)	(24.7)	(10.9)			

Table C.5: Operating characteristics of WE, WE(R), RIV and WT designs in scenarios 5-8: selection proportions, mean number of patients assigned to a dose (in brackets), termination proportion (Term), mean number of toxicity (T) and efficacy (E) responses. The optimal dose is in bold and correct doses are underlined. Results are based on 10^4 replications.

	d_1	d_2	d_3	d_4	d_5	d_6	Term	T	E
				Scenario 5					
	(.01;.10)	(.05;.35)	**(.15;.60)**	(.20;.60)	(.45;.60)	(.60;.60)			
WE	0.1	6.2	60.3	28.9	3.6	0.8	0.1	7.9	29.8
	(6.4)	(12.3)	(29.4)	(10.5)	(1.3)	(0.1)			
WE(R)	0.1	7.3	54.5	35.4	4.2	0.2	0.3	8.7	29.8
	(6.1)	(12.9)	(23.0)	(14.8)	(3.0)	(0.2)			
RIV	0.0	8.5	55.3	25.3	9.7	1.2	0.1	12.1	29.4
	(5.0)	(8.7)	(18.5)	(15.6)	(10.1)	(2.0)			
WT	0.1	2.7	27.0	53.0	16.9	0.2	0.1	13.5	32.3
	(5.2)	(4.4)	(14.3)	(22.3)	(11.0)	(2.8)			
				Scenario 6					
	(.01;.05)	(.05;.10)	(.10;.20)	(.20;.35)	**(.30;.55)**	(.50;.55)			
WE	0.4	0.8	3.7	18.9	53.5	18.9	4.7	11.0	19.4
	(6.6)	(7.4)	(10.0)	(15.7)	(16.8)	(2.8)			
WE(R)	0.4	0.8	4.8	25.4	56.7	7.1	4.4	10.8	19.2
	(6.3)	(7.7)	(11.1)	(16.3)	(15.2)	(2.9)			
RIV	0.1	0.7	4.5	17.0	55.9	13.7	8.3	13.2	21.4
	(4.5)	(5.5)	(7.9)	(12.4)	(19.0)	(7.8)			
WT	0.2	0.9	3.8	21.4	65.2	5.8	2.7	14.7	24.4
	(5.0)	(3.0)	(5.3)	(13.0)	(25.1)	(7.9)			
				Scenario 7					
	(.02;.30)	(.07;.50)	**(.13;.70)**	(.17;.73)	(.25;.76)	(.30;.77)			
WE	0.5	21.4	60.0	16.2	1.7	0.0	0.0	5.9	34.6
	(8.7)	(20.6)	(25.3)	(5.1)	(0.4)	(0.0)			
WE(R)	0.8	10.9	48.5	31.9	7.2	0.6	0.0	6.9	36.5
	(8.6)	(13.8)	(22.0)	(12.5)	(2.9)	(0.3)			
RIV	1.4	8.7	37.9	24.5	16.4	11.1	0.0	9.6	39.1
	(6.2)	(8.9)	(14.6)	(14.0)	(11.3)	(5.1)			
WT	1.6	5.2	27.1	29.8	24.7	11.7	0.0	10.0	39.9
	(6.7)	(5.4)	(13.7)	(14.8)	(12.1)	(7.3)			
				Scenario 8					
	(.03;.30)	**(.06;.50)**	(.10;.52)	(.20;.54)	(.40;.55)	(.50;.55)			
WE	3.2	37.8	34.7	19.1	4.2	1.0	0.0	5.6	28.9
	(9.4)	(23.9)	(17.8)	(7.1)	(1.5)	(0.1)			
WE(R)	3.9	36.4	33.8	20.9	4.4	0.4	0.0	6.9	29.2
	(9.2)	(16.9)	(18.7)	(11.7)	(3.2)	(0.3)			
RIV	12.8	43.0	21.7	12.7	8.2	1.7	0.1	9.3	29.3
	(10.1)	(14.3)	(12.8)	(12.3)	(8.5)	(2.0)			
WT	7.1	26.1	27.0	28.0	11.3	0.4	0.0	9.1	29.2
	(9.9)	(13.0)	(13.9)	(13.8)	(7.1)	(2.2)			

Table C.6: Operating characteristics of WE, WE(R), RIV and WT designs in scenarios 9-12: selection proportions, mean number of patients assigned to a dose (in brackets), termination proportion (Term), mean number of toxicity (T) and efficacy (E) responses. The optimal dose is in bold and correct doses are underlined. Results are based on 10^4 replications.

	d_1	d_2	d_3	d_4	d_5	d_6	Term	T	E
				Scenario 9					
	(.01;.30)	(.05;.50)	**(.10;.60)**	(.15;.40)	(.20;.25)	(.25;.15)			
WE	3.0	34.7	54.7	5.8	1.3	0.5	0.0	4.4	29.7
	(8.9)	(23.1)	(22.5)	(3.9)	(1.4)	(0.4)			
WE(R)	4.1	31.4	56.7	6.5	1.1	0.1	0.0	5.5	28.4
	(8.5)	(15.6)	(22.0)	(9.1)	(3.6)	(1.2)			
RIV	24.2	54.7	20.3	0.6	0.0	0.0	0.2	5.0	26.3
	(13.2)	(18.3)	(14.1)	(8.5)	(4.7)	(1.2)			
WT	7.3	30.9	50.1	8.8	2.1	0.7	0.2	5.5	27.9
	(9.8)	(15.4)	(22.1)	(7.1)	(3.3)	(2.2)			
				Scenario 10					
	(.02;.38)	**(.06;.50)**	(.12;.40)	(.30;.30)	(.40;.25)	(.50;.20)			
WE	18.4	55.9	15.9	3.9	3.2	2.7	0.1	5.7	25.6
	(14.7)	(27.4)	(11.0)	(4.3)	(2.1)	(0.5)			
WE(R)	22.7	56.2	17.0	2.5	1.2	0.2	0.1	6.7	24.6
	(13.4)	(21.4)	(14.9)	(6.6)	(3.0)	(0.6)			
RIV	60.4	35.3	2.8	0.4	0.1	0.1	1.0	6.3	23.9
	(20.2)	(17.6)	(10.1)	(8.0)	(3.3)	(0.5)			
WT	29.0	49.3	15.8	4.9	0.9	0.1	0.1	5.9	24.8
	(19.5)	(21.6)	(10.5)	(4.8)	(2.3)	(1.2)			
				Scenario 11					
	(.03;.25)	(.09;.35)	(.16;.48)	**(.28;.65)**	(.42;.52)	(.56;.39)			
WE	2.2	9.8	30.4	46.5	8.6	2.8	0.2	10.0	27.4
	(9.3)	(13.5)	(19.4)	(15.1)	(2.4)	(0.3)			
WE(R)	3.6	12.0	30.9	47.9	5.3	0.5	0.2	10.0	27.1
	(10.2)	(13.9)	(18.0)	(14.5)	(3.0)	(0.2)			
RIV	6.7	14.0	27.0	46.0	5.6	0.3	0.3	12.8	28.7
	(8.1)	(10.2)	(14.1)	(17.0)	(8.7)	(1.7)			
WT	6.9	9.4	23.2	56.9	3.5	0.0	0.1	12.2	29.4
	(9.9)	(7.9)	(13.8)	(22.2)	(4.6)	(1.6)			
				Scenario 12					
	(.02;.68)	(.05;.56)	(.07;.49)	(.09;.40)	(.11;.33)	(.13;.26)			
WE	80.1	14.5	3.9	1.1	0.3	0.1	0.0	1.8	35.4
	(44.8)	(10.0)	(3.4)	(1.1)	(0.5)	(0.1)			
WE(R)	70.9	18.7	7.1	2.4	0.9	0.2	0.0	3.3	32.7
	(20.5)	(15.1)	(11.8)	(7.0)	(3.6)	(1.9)			
RIV	96.1	3.2	0.6	0.1	0.0	0.0	0.0	2.6	35.5
	(34.6)	(10.1)	(6.4)	(4.8)	(2.8)	(1.4)			
WT	75.9	17.1	5.1	1.2	0.6	0.2	0.0	2.4	36.3
	(37.0)	(11.0)	(5.4)	(3.0)	(2.1)	(1.4)			

Table C.7: Operating characteristics of WE, WE(R), RIV and WT designs in scenarios 13-14: selection proportions, mean number of patients assigned to a dose (in brackets), termination proportion (Term), mean number of toxicity (T) and efficacy (E) responses. The optimal dose is in bold and correct doses are underlined. Results are based on 10^4 replications.

	d_1	d_2	d_3	d_4	d_5	d_6	Term	T	E
				Scenario 13					
	(.05;.01)	(.10;.02)	(.25;.05)	(.55;.35)	(.70;.55)	(.90;.70)			
WE	0.0	0.0	0.2	2.9	0.6	1.1	95.2	11.7	6.0
	(6.7)	(7.6)	(10.9)	(12.2)	(1.6)	(0.1)			
WE(R)	0.1	0.0	0.2	4.4	1.1	0.2	93.9	13.0	7.1
	(6.9)	(7.7)	(10.6)	(12.1)	(3.2)	(0.4)			
RIV	0.0	0.0	2.3	5.8	0.0	0.0	91.9	11.0	6.2
	(5.8)	(5.9)	(7.7)	(11.0)	(2.7)	(0.3)			
WT	0.0	0.1	5.5	3.1	0.0	0.0	91.5	11.2	5.4
	(6.3)	(6.5)	(14.8)	(7.9)	(1.5)	(1.2)			
				Scenario 14					
	(.50;.40)	(.60;.55)	(.69;.65)	(.76;.65)	(.82;.65)	(.89;.65)			
WE	2.2	0.2	0.1	0.3	0.2	0.0	96.9	11.0	9.1
	(17.6)	(2.8)	(0.6)	(0.1)	(0.0)	(0.0)			
WE(R)	1.9	0.5	0.3	0.2	0.0	0.0	97.4	10.9	8.9
	(17.6)	(2.4)	(0.8)	(0.1)	(0.0)	(0.0)			
RIV	8.8	0.2	0.0	0.0	0.0	0.0	91.0	11.5	9.6
	(16.0)	(4.3)	(1.1)	(0.2)	(0.0)	(0.0)			
WT	9.5	0.0	0.0	0.0	0.0	0.0	90.4	12.1	9.7
	(23.0)	(0.4)	(0.1)	(0.1)	(0.1)	(0.1)			

C.4 Early Efficacy Data

In the setting above, it is assumed that it takes twice as long to observe the efficacy outcome than the toxicity endpoint. It is, however, possible that an efficacy (or lack of efficacy) can be observed at the time of the interim analysis for some of the patients. As the proposed design includes all available information, it can also accommodate earlier efficacy (no efficacy) data. This section we study how the operating characteristics of the non-randomised WE design are affected if a certain proportion of "no efficacy" responses can be observed earlier.

The setting above remains unchanged with the following exception: if the patient has observed no DLT and will have "no efficacy", it is assumed that the outcome can be observed at the time of toxicity evaluation with probability π. If observed earlier, the WE design uses this information for the next patient allocation. We consider two cases: $\pi = 0$ (the original setting) and $\pi = 1/2$ (half of "no efficacy" responses can be observed earlier). The results are given in Table C.8.

As expected, the availability of some of the efficacy information earlier leads to a less conservative design that allows a more rapid escalation. Earlier "no efficacy" data even in half of the patients lead to more ethical patient allocation. This can be seen by increased numbers of efficacy responses almost in all scenarios with the cost of a reasonable increase in the average number of toxicity responses. The largest increase can be seen in scenario 1 where the average number of efficacy response increase by nearly 7, while toxicity increases only by 1. The information about earlier efficacy also improves the proportion of optimal selections in the scenarios where the target dose is high - by 6% in scenario 1 and by 4% in scenario 6. As the design being less conservative it favours higher doses among correct ones. This decreases the proportion of optimal selections in scenario 3, 5, 7 and 12 by 3-7%. At the same time, the proportion of correct selections is either unchanged (scenario 5 and 8) or increased by at least 5% (all the rest plateau scenarios). This confirms

Table C.8: Operating characteristics of WE in scenarios 1-12 with no early efficacy data available ($\pi = 0$) and with half "no efficacy" outcomes ($\pi = 1/2$) available at the time of toxicity evaluation: selection proportions, mean number of toxicity (T) and efficacy (E) responses . The optimal dose is in bold and correct doses are underlined. Results are based on 10^4 replications.

WE	d_1	d_2	d_3	d_4	d_5	d_6	T	E
Sc.1	(.005;.01)	(.01;.10)	(.02;.30)	(.05;.50)	**(.10;.80)**	(.15;.80)		
$\pi = 0$	0.0	0.1	2.3	37.4	**58.8**	1.5	3.1	28.5
$\pi = 1/2$	0.0	0.0	1.5	23.5	**64.9**	10.0	4.3	34.8
Sc. 2	**(.01;.40)**	(.04;.40)	(.10;.40)	(.25;.40)	(.50;.40)	(.70;.40)		
$\pi = 0$	**30.3**	26.3	20.8	12.5	6.3	3.7	6.0	24.0
$\pi = 1/2$	**33.2**	27.1	21.9	12.2	3.7	1.6	7.2	24.0
Sc. 3	(.01;.25)	(.02;.45)	**(.05;.65)**	(.10;.65)	(.20;.65)	(.30;.65)		
$\pi = 0$	0.6	12.6	**65.8**	18.2	2.8	0.1	2.8	33.0
$\pi = 1/2$	0.9	9.4	**57.7**	24.0	6.8	1.1	3.7	34.8
Sc. 4	(.01;.05)	(.02;.25)	(.05;.45)	**(.10;.70)**	(.25;.70)	(.50;.70)		
$\pi = 0$	0.0	0.5	19.6	**71.1**	8.5	0.3	4.5	30.0
$\pi = 1/2$	0.0	0.9	14.6	**68.0**	15.8	0.7	5.9	33.9
Sc. 5	(.01;.10)	(.05;.35)	**(.15;.60)**	(.20;.60)	(.45;.60)	(.60;.60)		
$\pi = 0$	0.1	6.2	**60.3**	28.9	3.6	0.8	7.9	29.8
$\pi = 1/2$	0.1	6.3	**56.8**	32.8	3.3	0.7	9.2	31.5
Sc. 6	(.01;.05)	(.05;.10)	(.10;.20)	(.20;.35)	**(.30;.55)**	(.50;.55)		
$\pi = 0$	0.4	0.8	3.7	18.9	**53.5**	18.9	11.0	19.4
$\pi = 1/2$	1.4	1.4	4.9	22.2	**57.2**	10.0	13.1	22.2
Sc. 7	(.02;.30)	(.07;.50)	**(.13;.70)**	(.17;.73)	(.25;.76)	(.30;.77)		
$\pi = 0$	0.5	21.4	**60.0**	16.2	1.7	0.0	5.9	34.6
$\pi = 1/2$	0.9	16.4	**53.4**	23.2	5.5	0.6	6.9	37.0
Sc. 8	(.03;.30)	**(.06;.50)**	(.10;.52)	(.20;.54)	(.40;.55)	(.50;.55)		
$\pi = 0$	3.2	**37.8**	34.7	19.1	4.2	1.0	5.6	28.9
$\pi = 1/2$	3.0	**37.8**	34.0	20.1	4.0	1.1	6.9	29.5
Sc. 9	(.01;.30)	(.05;.50)	**(.10;.60)**	(.15;.40)	(.20;.25)	(.25;.15)		
$\pi = 0$	3.0	34.7	**54.7**	5.8	1.3	0.5	4.4	29.7
$\pi = 1/2$	3.8	34.2	**54.6**	6.	1.1	0.2	5.1	29.7
Sc. 10	(.02;.38)	**(.06;.50)**	(.12;.40)	(.30;.30)	(.40;.25)	(.50;.20)		
$\pi = 0$	18.4	**55.9**	15.9	3.9	3.2	2.7	5.7	25.6
$\pi = 1/2$	20.2	**57.1**	17.2	2.8	1.6	1.0	6.3	25.5
Sc. 11	(.03;.25)	(.09;.35)	(.16;.48)	**(.28;.65)**	(.42;.52)	(.56;.39)		
$\pi = 0$	2.2	9.8	30.4	**46.5**	8.6	2.8	10.0	27.4
$\pi = 1/2$	3.5	10.8	31.7	**45.5**	6.3	1.8	11.4	28.8
Sc. 12	**(.02;.68)**	(.05;.56)	(.07;.49)	(.09;.40)	(.11;.33)	(.13;.26)		
$\pi = 0$	80.1	14.5	3.9	1.1	0.3	0.1	1.8	35.4
$\pi = 1/2$	**74.1**	17.5	5.8	1.8	0.6	0.1	2.2	36.8

that the WE design in the setting with no earlier efficacy information is more conservative, but the difference in proportions of correct selections is relatively small.

Bibliography

Aglietta, M., Barone, C., Sawyer, M., Moore, M., Miller Jr, W., Bagalà, C., Colombi, F., Cagnazzo, C., Gioeni, L., Wang, E. et al. (2014) A Phase I dose escalation trial of tremelimumab (cp-675,206) in combination with gemcitabine in chemotherapy-naive patients with metastatic pancreatic cancer. *Annals of Oncology*, **25**, 1750–1755.

Agresti, A. (1992) A survey of exact inference for contingency tables. *Statistical Science*, **7**, 131–153.

Agresti, A. and Coull, B. A. (1998) Approximate is better than "exact" for interval estimation of binomial proportions. *The American Statistician*, **52**, 119–126.

Aitchison, J. (1992) On criteria for measures of compositional difference. *Mathematical Geology*, **24**, 365–379.

Arnold, B. C. (1970) Inadmissibility of the usual scale estimate for a shifted exponential distribution. *Journal of the American Statistical Association*, **65**, 1260–1264.

Azriel, D., Mandel, M. and Rinott, Y. (2011) The treatment versus experimentation dilemma in dose finding studies. *Journal of Statistical Planning and Inference*, **141**, 2759–2768.

Babb, J., Rogatko, A. and Zacks, S. (1998) Cancer Phase I clinical trials: efficient dose escalation with overdose control. *Statistics in Medicine*, **17**, 1103–1120.

Barrett, J. E. (2016) Information-adaptive clinical trials: a selective recruitment design. *Journal of the Royal Statistical Society: Series C*, **65**, 797–808.

Bartholomew, D. (1965) A comparison of some bayesian and frequentist inferences. *Biometrika*, **52**, 19–36.

Bekele, B. and Shen, Y. (2005) A bayesian approach to jointly modeling toxicity and biomarker expression in a Phase I/II dose-finding trial. *Biometrics*, **61**, 343–354.

Bekele, B. N., Li, Y. and Ji, Y. (2010) Risk-group-specific dose finding based on an average toxicity score. *Biometrics*, **66**, 541–548.

Bekele, B. N. and Thall, P. F. (2004) Dose-finding based on multiple toxicities in a soft tissue sarcoma trial. *Journal of the American Statistical Association*, **99**, 26–35.

Belis, M. and Guiasu, S. (1968) A quantitative-qualitative measure of information in cybernetic systems (corresp.). *IEEE Transactions on Information Theory*, **14**, 593–594.

Bickel, P. (1981) Minimax estimation of the mean of a normal distribution when the parameter space is restricted. *The Annals of Statistics*, **9**, 1301–1309.

Billheimer, D., Guttorp, P. and Fagan, W. F. (2001) Statistical interpretation of species composition. *Journal of the American Statistical Association*, **96**, 1205–1214.

Bornkamp, B. (2017) Dose-finding studies in Phase II: Introduction and overview. In *Handbook of Methods for Designing, Monitoring, and Analyzing Dose-Finding Trials* (eds. J. O'Quigley, A. Iasonos and B. Bornkamp), chap. 11, 189–204. CRC Press, Taylor and Francis Group.

Braun, T. M. (2002) The bivariate continual reassessment method: extending the

crm to Phase I trials of two competing outcomes. *Controlled Clinical Trials*, **23**, 240–256.

Brewster, J. F. (1974) Alternative estimators for the scale parameter of the exponential distribution with unknown location. *The Annals of Statistics*, **2**, 553–557.

Brewster, J.-F. and Zidek, J. (1974) Improving on equivariant estimators. *The Annals of Statistics*, **2**, 21–38.

Brown, L. (1968) Inadmissibility of the usual estimators of scale parameters in problems with unknown location and scale parameters. *The Annals of Mathematical Statistics*, **39**, 29–48.

Brown, L. D., Cai, T. T. and DasGupta, A. (2001) Interval estimation for a binomial proportion. *Statistical Science*, **12**, 101–117.

Buchbinder, E. I. and Desai, A. (2016) Ctla-4 and pd-1 pathways: similarities, differences, and implications of their inhibition. *American Journal of Clinical Oncology*, **39**, 98.

Cai, C., Yuan, Y. and Ji, Y. (2014) A bayesian Phase I/II design for oncology clinical trials of combining biological agents. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, **63**, 159–173.

Carter, S. K. (1973) Study design principles for the clinical evaluation of new drugs as developed by the chemotherapy programme of the national cancer institute. In *The Design of Clinical Trials in Cancer Therapy* (ed. M. Staquet), 242–289. Éditions Scientifiques Europeenes, Brussels, Belgium.

Casella, G. and Strawderman, W. E. (1981) Estimating a bounded normal mean. *The Annals of Statistics*, **9**, 870–878.

Cheung, Y. K. (2005) Coherence principles in dose-finding studies. *Biometrika*, **92**, 863–873.

— (2011) *Dose finding by the continual reassessment method*. CRC Press.

— (2013) Sample size formulae for the bayesian continual reassessment method. *Clinical Trials*, **10**, 852–861.

— (2014) Simple benchmark for complex dose finding studies. *Biometrics*, **70**, 389–397.

Cheung, Y. K. and Chappell, R. (2000) Sequential designs for Phase I clinical trials with late-onset toxicities. *Biometrics*, **56**, 1177–1182.

— (2002) A simple technique to evaluate model sensitivity in the continual reassessment method. *Biometrics*, **58**, 671–674.

Chevret, S. (2006) *Statistical methods for dose-finding experiments*, vol. 24. Wiley Online Library.

Chiuzan, C., Shtaynberger, J., Manji, G. A., Duong, J. K., Schwartz, G. K., Ivanova, A. and Lee, S. M. (2017) Dose-finding designs for trials of molecularly targeted agents and immunotherapies. *Journal of Biopharmaceutical Statistics*, **27**, 477–494.

Chow, S. and Liu, J. (1999) *Design and Analysis of Bioavailability and Bioequivalence Studies, Second Edition*. Chapman & Hall/CRC Biostatistics Series. Taylor & Francis.

Clertant, M. and O'Quigley, J. (2017) Semiparametric dose finding methods. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **79**, 1487–1508.

Clim, A. (2008) Weighted entropy with application. *Analele Universitatii Bucuresti, Matematica, Anul*, **57**, 223–231.

Conolly, R. B. and Lutz, W. K. (2004) Nonmonotonic dose-response relationships: mechanistic basis, kinetic modeling, and implications for risk assessment. *Toxicological Sciences*, **77**, 151–157.

Cover, T. M. and Thomas, J. A. (2012) *Elements of information theory*. John Wiley & Sons.

Crane, C. H., Abbruzzese, J. L., Evans, D. B., Wolff, R. A., Ballo, M. T., Delclos, M., Milas, L., Mason, K., Charnsangavej, C., Pisters, P. et al. (2002) Is the therapeutic index better with gemcitabine-based chemoradiation than with 5-fluorouracil-based chemoradiation in locally advanced pancreatic cancer? *International Journal of Radiation Oncology Biology Physics*, **52**, 1293–1302.

Csiszar, I. and Körner, J. (2011) *Information theory: coding theorems for discrete memoryless systems*. Cambridge University Press.

Disis, M. L. (2010) Immune regulation of cancer. *Journal of Clinical Oncology*, **28**, 4531–4538.

Djulbegovic, B., Lacevic, M., Cantor, A., Fields, K. K., Bennett, C. L., Adams, J. R., Kuderer, N. M. and Lyman, G. H. (2000) The uncertainty principle and industry-sponsored research. *The Lancet*, **356**, 635–638.

Dunnett, C. W. (1984) Selection of the best treatment in comparison to a control with an application to a medical trial. In *Design of Experiments: Ranking and Selection* (eds. T. Santer and A. Tamhane), 47–66. Marcel Dekker: New York.

Ezzalfani, M., Zohar, S., Qin, R., Mandrekar, S. J. and Deley, M.-C. L. (2013) Dose-finding designs using a novel quasi-continuous endpoint for multiple toxicities. *Statistics in Medicine*, **32**, 2728–2746.

Ferguson, T. S. (2014) *Mathematical statistics: A decision theoretic approach*, vol. 1. Academic press.

Finney, D. (1952) *Statistical method in biological assay*. Charles Griffin: London.

Friedman, L. M., Furberg, C., DeMets, D. L., Reboussin, D. and Granger, C. B. (2015) *Fundamentals of clinical trials*. Springer.

Galilei, G. (1627) Lettera (intorno la stima di un cavallo). *Le opere di Galileo Galilei*, Prima edizione completa. Societa editrice fiorentina. Firenze.

Gatsonis, C., MacGibbon, B. and Strawderman, W. (1987) On the estimation of a restricted normal mean. *Statistics & Probability Letters*, **6**, 21–30.

Gittins, J., Glazebrook, K. and Weber, R. (2011) *Multi-armed bandit allocation indices*. John Wiley & Sons.

Gittins, J. C. and Jones, D. M. (1979) A dynamic allocation index for the discounted multiarmed bandit problem. *Biometrika*, **66**, 561–565.

Gooley, T. A., Martin, P. J., Fisher, L. D. and Pettinger, M. (1994) Simulation as a design tool for Phase I/II clinical trials: an example from bone marrow transplantation. *Controlled Clinical Trials*, **15**, 450–462.

Gradshteyn, I. S. and Ryzhik, I. M. (2014) *Table of integrals, series, and products*. Academic press.

Hartigan, J. (2004) Uniform priors on convex sets improve risk. *Statistics & Probability Letters*, **67**, 285–288.

Hirakawa, A. (2012) An adaptive dose-finding approach for correlated bivariate binary and continuous outcomes in Phase I oncology trials. *Statistics in Medicine*, **31**, 516–532.

Hornik, K., Leisch, F. and Zeileis, A. (2003) Jags: A program for analysis of bayesian graphical models using gibbs sampling. In *Proceedings of DSC*, vol. 2, 1–10.

Iasonos, A., Wages, N. A., Conaway, M. R., Cheung, K., Yuan, Y. and O'Quigley, J. (2016) Dimension of model parameter space and operating characteristics in adaptive dose-finding studies. *Statistics in Medicine*, **35**, 3760–3775.

Ivanova, A. and Kim, S. H. (2009) Dose finding for continuous and ordinal outcomes with a monotone objective function: a unified approach. *Biometrics*, **65**, 307–315.

James, W. and Stein, C. (1961) Estimation with quadratic loss. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, 361–379.

Jensen, J. L. W. V. (1906) Sur les fonctions convexes et les inégalités entre les valeurs moyennes. *Acta mathematica*, **30**, 175–193.

Kang, S.-H. and Ahn, C. (2001) The expected toxicity rate at the maximum tolerated dose in the standard Phase I cancer clinical trial design. *Drug Information Journal*, **35**, 1189–1199.

Kang, S.-H. and Ahn, C. W. (2002) An investigation of the traditional algorithm-based designs for Phase I cancer clinical trials. *Drug Information Journal*, **36**, 865–873.

Karimnezhad, A. and Moradi, F. (2016) Bayes, e-bayes and robust bayes prediction of a future observation under precautionary prediction loss functions with applications. *Applied Mathematical Modelling*, **40**, 7051–7061.

Karimnezhad, A., Niazi, S. and Parsian, A. (2014) Bayes and robust bayes prediction with an application to a rainfall prediction problem. *Journal of the Korean Statistical Society*, **43**, 275–291.

Kelbert, M. and Mozgunov, P. (2015) Asymptotic behaviour of the weighted Renyi, Tsallis and Fisher entropies in a bayesian problem. *Eurasian Mathematical Journal*, **6**, 6–17.

— (2017) Generalization of Cramér-Rao and Bhattacharyya inequalities for the weighted covariance matrix. *Mathematical Communications*, **22**, 25–40.

Kerantzas, C. A. and Jacobs, W. R. (2017) Origins of combination therapy for tuberculosis: lessons for future antimicrobial development and application. *MBio*, **8**, e01586–16.

Khalil, D. N., Smith, E. L., Brentjens, R. J. and Wolchok, J. D. (2016) The future of cancer treatment: immunomodulation, cars and combination immunotherapy. *Nature Reviews Clinical oncology*, **13**, 273.

Kiapour, A. and Nematollahi, N. (2011) Robust bayesian prediction and estimation under a squared log error loss function. *Statistics & Probability Letters*, **81**, 1717–1724.

Kim, S. B. and Gillen, D. L. (2016) A bayesian adaptive dose-finding algorithm for balancing individual-and population-level ethics in Phase I clinical trials. *Sequential Analysis*, **35**, 423–439.

Kodama, A., To, H., Kinoshita, T., Ieiri, I. and Higuchi, S. (2009) Influence of dosing schedules on toxicity and antitumour effects of combined cisplatin and docetaxel treatment in mice. *Journal of Pharmacy and Pharmacology*, **61**, 615–621.

Koenig, F., Brannath, W., Bretz, F. and Posch, M. (2008) Adaptive dunnett tests for treatment selection. *Statistics in Medicine*, **27**, 1612–1625.

Korn, E. L., Midthune, D., Chen, T. T., Rubinstein, L. V., Christian, M. C. and Simon, R. M. (1994) A comparison of two Phase I trial designs. *Statistics in Medicine*, **13**, 1799–1806.

Kreissman, S. G., Seeger, R. C., Matthay, K. K., London, W. B., Sposto, R., Grupp, S. A., Haas-Kogan, D. A., LaQuaglia, M. P., Alice, L. Y., Diller, L. et al. (2013) Purged versus non-purged peripheral blood stem-cell transplantation for high-risk neuroblastoma (cog a3973): a randomised Phase III trial. *The Lancet Oncology*, **14**, 999–1008.

Kubokawa, T. (1994) A unified approach to improving equivariant estimators. *The Annals of Statistics*, **22**, 290–299.

Kumar, S. and Tripathi, Y. M. (2008) Estimating a restricted normal mean. *Metrika*, **68**, 271–288.

Kyi, C. and Postow, M. A. (2014) Checkpoint blocking antibodies in cancer immunotherapy. *FEBS letters*, **588**, 368–376.

Lagarde, F., Beausoleil, C., Belcher, S. M., Belzunces, L. P., Emond, C., Guerbet, M. and Rousselle, C. (2015) Non-monotonic dose-response relationships and endocrine disruptors: a qualitative method of assessment. *Environmental Health*, **14**, 13.

Le Tourneau, C., Lee, J. J. and Siu, L. L. (2009) Dose escalation methods in Phase I cancer clinical trials. *JNCI: Journal of the National Cancer Institute*, **101**, 708–720.

Lee, S., Hershman, D., Martin, P., Leonard, J. and Cheung, Y. (2011) Toxicity burden score: a novel approach to summarize multiple toxic effects. *Annals of Oncology*, **23**, 537–541.

Lee, S. M., Cheng, B. and Cheung, Y. K. (2010) Continual reassessment method with multiple toxicity constraints. *Biostatistics*, **12**, 386–398.

Lee, S. M. and Cheung, Y. K. (2009) Model calibration in the continual reassessment method. *Clinical Trials*, **6**, 227–238.

Lee, S. M., Ursino, M., Cheung, Y. K. and Zohar, S. (2017) Dose-finding designs for cumulative toxicities using multiple constraints. *Biostatistics*, kxx059. URL: `http://dx.doi.org/10.1093/biostatistics/kxx059`.

Lindley, D. V. (1980) Approximate bayesian methods. *Trabajos de estadística y de investigación operativa*, **31**, 223–245.

Magirr, D., Jaki, T., Whitehead, J. et al. (2012) A generalized dunnett test for multi-arm multi-stage clinical studies with treatment selection. *Biometrika*, **99**, 494.

Mahmoudi, E. and Zakerzadeh, H. (2011) An admissible estimator of a lower-bounded scale parameter under squared-log error loss function. *Kybernetika*, **47**, 595–611.

Marchand, É., Najafabadi, A. T. P. et al. (2011) Bayesian improvements of a mre estimator of a bounded location parameter. *Electronic Journal of Statistics*, **5**, 1495–1502.

Marchand, É. and Strawderman, W. E. (2005) On improving on the minimum risk equivariant estimator of a scale parameter under a lower-bound constraint. *Journal of Statistical Planning and Inference*, **134**, 90–101.

Mateu-Figueras, G., Pawlowsky-Glahn, V. and Juan José, E. (2013) The normal distribution in some constrained sample spaces. *SORT: Statistics and Operations Research Transactions*, **37**, 29–56.

Miller, R. B. (1980) Bayesian analysis of the two-parameter gamma distribution. *Technometrics*, **22**, 65–69.

Molenberghs, G., Geys, H. and Buyse, M. (2001) Evaluation of surrogate end-points in randomized experiments with mixed discrete and continuous outcomes. *Statistics in Medicine*, **20**, 3023–3038.

Morgan, B., Thomas, A. L., Drevs, J., Hennig, J., Buchert, M., Jivan, A., Horsfield, M. A., Mross, K., Ball, H. A., Lee, L. et al. (2003) Dynamic contrast-enhanced magnetic resonance imaging as a biomarker for the pharmacological response of ptk787/zk 222584, an inhibitor of the vascular endothelial growth factor receptor tyrosine kinases, in patients with advanced colorectal cancer and liver metastases: results from two Phase I studies. *Journal of Clinical Oncology*, **21**, 3955–3964.

Neuenschwander, B., Branson, M. and Gsponer, T. (2008) Critical aspects of the bayesian approach to Phase I cancer trials. *Statistics in Medicine*, **27**, 2420–2439.

Newcombe, R. G. (1998) Two-sided confidence intervals for the single proportion: comparison of seven methods. *Statistics in Medicine*, **17**, 857–872.

Norstrom, J. G. (1996) The use of precautionary loss functions in risk analysis. *IEEE Transactions on Reliability*, **45**, 400–403.

O'Quigley, J. (2002) Continual reassessment designs with early termination. *Biostatistics*, **3**, 87–99.

O'Quigley, J., Hughes, M. D. and Fenton, T. (2001) Dose-finding designs for HIV studies. *Biometrics*, **57**, 1018–1029.

O'Quigley, J., Paoletti, X. and Maccario, J. (2002) Non-parametric optimal design in dose finding studies. *Biostatistics*, **3**, 51–56.

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

O'Quigley, J. and Shen, L. Z. (1996) Continual reassessment method: a likelihood approach. *Biometrics*, **52**, 673–684.

O'Quigley, J. and Zohar, S. (2010) Retrospective robustness of the continual reassessment method. *Journal of Biopharmaceutical Statistics*, **20**, 1013–1025.

Oron, A. P. and Hoff, P. D. (2013) Small-sample behavior of novel Phase I cancer trial designs. *Clinical Trials*, **10**, 63–80.

Paller, C. J., Bradbury, P. A., Ivy, S. P., Seymour, L., LoRusso, P. M., Baker, L., Rubinstein, L., Huang, E., Collyar, D., Groshen, S. et al. (2014) Design of Phase I combination trials: recommendations of the clinical trial design task force of the nci investigational drug steering committee. *Clinical Cancer Research*, **20**, 4210–4217.

Paoletti, X. and Kramar, A. (2009) A comparison of model choices for the continual reassessment method in Phase I cancer trials. *Statistics in Medicine*, **28**, 3012–3028.

Paoletti, X., Le Tourneau, C., Verweij, J., Siu, L. L., Seymour, L., Postel-Vinay, S., Collette, L., Rizzo, E., Ivy, P., Olmos, D. et al. (2014) Defining dose-limiting toxicity for Phase I trials of molecularly targeted agents: results of a dlt-targett international survey. *European Journal of Cancer*, **50**, 2050–2056.

Pardoll, D. M. (2012) The blockade of immune checkpoints in cancer immunotherapy. *Nature Reviews. Cancer*, **12**, 252.

Patnaik, A., Kang, S. P., Rasco, D., Papadopoulos, K. P., Elassaiss-Schaap, J., Beeram, M., Drengler, R., Chen, C., Smith, L. S., Espino, G. et al. (2015) Phase I study of pembrolizumab (mk-3475; anti-pd-1 monoclonal antibody) in patients with advanced solid tumors. *Clinical Cancer Research*, **21**, 4286–93.

Pawlowsky-Glahn, V. and Egozcue, J. J. (2001) Geometric approach to statistical analysis on the simplex. *Stochastic Environmental Research and Risk Assessment*, **15**, 384–398.

Polley, M.-Y. and Cheung, Y. K. (2008) Two-stage designs for dose-finding trials with a biologic endpoint using stepwise tests. *Biometrics*, **64**, 232–241.

Postel-Vinay, S., Gomez-Roca, C., Molife, L. R., Anghan, B., Levy, A., Judson, I., De Bono, J., Soria, J.-C., Kaye, S. and Paoletti, X. (2011) Phase I trials of molecularly targeted agents: should we pay more attention to late toxicities? *Journal of Clinical Oncology*, **29**, 1728–1735.

R Core Team (2015) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL: `https://www.R-project.org/`.

Ratain, M. J., Mick, R., Schilsky, R. L. and Siegler, M. (1993) Statistical and

ethical issues in the design and conduct of Phase I and II clinical trials of new anticancer agents. *Journal of the National Cancer Institute*, **85**, 1637–1643.

Reiner, E., Paoletti, X. and O'Quigley, J. (1999) Operating characteristics of the standard Phase I clinical trial design. *Computational Statistics & Data Analysis*, **30**, 303–315.

Resnick, S. I. (2013) *A probability path*. Springer Science & Business Media.

Riviere, M.-K., Dubois, F. and Zohar, S. (2015) Competing designs for drug combination in Phase I dose-finding clinical trials. *Statistics in Medicine*, **34**, 1–12.

Riviere, M.-K., Yuan, Y., Jourdan, J.-H., Dubois, F. and Zohar, S. (2016) Phase I/II dose-finding design for molecularly targeted agent: Plateau determination using adaptive randomization. *Statistical Methods in Medical Research*, **27**, 466–479.

Robbins, H. (1952) Some aspects of the sequential design of experiments. *Bull. Amer. Math. Soc.*, **58**, 527–535.

Robert, C., Ribas, A., Wolchok, J. D., Hodi, F. S., Hamid, O., Kefford, R., Weber, J. S., Joshua, A. M., Hwu, W.-J., Gangadhar, T. C. et al. (2014) Anti-programmed-death-receptor-1 treatment with pembrolizumab in ipilimumab-refractory advanced melanoma: a randomised dose-comparison cohort of a Phase I trial. *The Lancet*, **384**, 1109–1117.

Rogatko, A., Schoeneck, D., Jonas, W., Tighiouart, M., Khuri, F. R. and Porter, A. (2007) Translation of innovative designs into Phase I trials. *Journal of Clinical Oncology*, **25**, 4982–4986.

Saad, E. D., Paoletti, X., Burzykowski, T. and Buyse, M. (2017) Precision medicine needs randomized clinical trials. *Nature Reviews Clinical Oncology*, **14**, 317–323.

Savchuk, V. and Tsokos, C. P. (2011) *Bayesian theory and methods with applications*, vol. 1. Springer Science & Business Media.

Scàrdovi, I. (1980) *Appunti di statistica*, vol. 1. Pàtron, Bologna.

Sharma, P. and Allison, J. P. (2015) The future of immune checkpoint therapy. *Science*, **348**, 56–61.

Shen, L. Z. and O'Quigley, J. (1996) Consistency of continual reassessment method under model misspecification. *Biometrika*, **83**, 395–405.

Soliman, A. A. (2002) Reliability estimation in a generalized life-model with application to the burr-xii. *IEEE Transactions on Reliability*, **51**, 337–343.

Stallard, N. and Todd, S. (2003) Sequential designs for Phase III clinical trials incorporating treatment selection. *Statistics in Medicine*, **22**, 689–703.

Stein, C. (1964) Inadmissibility of the usual estimator for the variance of a normal distribution with unknown mean. *Annals of the Institute of Statistical Mathematics*, **16**, 155–160.

Steliarova-Foucher, E., Colombet, M., Ries, L. A., Moreno, F., Dolya, A., Bray, F., Hesseling, P., Shin, H. Y., Stiller, C. A. et al. (2017) International incidence of childhood cancer, 2001–10: a population-based registry study. *The Lancet Oncology*, **18**, 719–731.

Storer, B. E. (1989) Design and analysis of Phase I clinical trials. *Biometrics*, **65**, 925–937.

Suhov, Y., Stuhl, I., Sekeh, S. Y. and Kelbert, M. (2016) Basic inequalities for weighted entropies. *Aequationes Mathematicae*, **90**, 817–848.

Sweeting, M., Mander, A., Sabin, T. et al. (2013) Bcrm: Bayesian continual reassessment method designs for Phase I dose-finding trials. *Journal of Statistical Software*, **54**, 1–26.

Tate, R. F. (1955) The theory of correlation between two continuous variables when one is dichotomized. *Biometrika*, **42**, 205–216.

Thall, P. F. and Cook, J. D. (2004) Dose-finding based on efficacy–toxicity trade-offs. *Biometrics*, **60**, 684–693.

Thall, P. F., Millikan, R. E., Mueller, P. and Lee, S.-J. (2003) Dose-finding with two agents in Phase I oncology trials. *Biometrics*, **59**, 487–496.

Thall, P. F. and Russell, K. E. (1998) A strategy for dose-finding and safety monitoring based on efficacy and adverse outcomes in Phase I/II clinical trials. *Biometrics*, **54**, 251–264.

Thall, P. F. and Wathen, J. K. (2007) Practical bayesian adaptive randomisation in clinical trials. *European Journal of Cancer*, **43**, 859–866.

Tighiouart, M., Rogatko, A. et al. (2010) Dose finding with escalation with overdose control (ewoc) in cancer clinical trials. *Statistical Science*, **25**, 217–226.

Villar, S. S., Bowden, J. and Wason, J. (2015) Multi-armed bandit models for the optimal design of clinical trials: benefits and challenges. *Statistical Science*, **30**, 199–215.

Vollset, S. E. (1993) Confidence intervals for a binomial proportion. *Statistics in Medicine*, **12**, 809–824.

Wages, N. A. (2015) Comments on 'Competing designs for drug combination in Phase I dose-finding clinical trials' by MK. Riviere, F. Dubois, S. Zohar. *Statistics in Medicine*, **34**, 18.

Wages, N. A. and Conaway, M. R. (2014) Phase I/II adaptive design for drug combination oncology trials. *Statistics in Medicine*, **33**, 1990–2003.

Wages, N. A., Conaway, M. R. and O'Quigley, J. (2011) Continual reassessment method for partial ordering. *Biometrics*, **67**, 1555–1563.

Wages, N. A. and Tait, C. (2015) Seamless Phase I/II adaptive design for oncology trials of molecularly targeted agents. *Journal of Biopharmaceutical Statistics*, **25**, 903–920.

Wages, N. A. and Varhegyi, N. (2013) pocrm: An r-package for Phase I trials of combinations of agents. *Computer Methods and Programs in Biomedicine*, **112**, 211–218.

— (2017) A web application for evaluating Phase I methods using a non-parametric optimal benchmark. *Clinical Trials*, **14**, 553–557.

Wang, C., Chen, T. T. and Tyan, I. (2000) Designs for Phase I cancer clinical trials with differentiation of graded toxicity. *Communications in Statistics-Theory and Methods*, **29**, 975–987.

Wang, Y. and Ivanova, A. (2015) Dose finding with continuous outcome in Phase I oncology trials. *Pharmaceutical Statistics*, **14**, 102–107.

Wheeler, G. M. (2017) Overview of Phase I designs. In *Handbook of Methods for Designing, Monitoring, and Analyzing Dose-Finding Trials* (eds. J. O'Quigley, A. Iasonos and B. Bornkamp), chap. 1, 3–26. CRC Press, Taylor and Francis Group.

Wheeler, G. M., Sweeting, M. J. and Mander, A. P. (2017) Toxicity-dependent feasibility bounds for the escalation with overdose control approach in Phase I cancer trials. *Statistics in Medicine*, **36**, 2499–2513.

Whitehead, J. and Jaki, T. (2009) One-and two-stage design proposals for a Phase II trial comparing three active treatments with control using an ordered categorical endpoint. *Statistics in Medicine*, **28**, 828–847.

Whitehead, J. and Williamson, D. (1998) Bayesian decision procedures based on logistic regression models for dose-finding studies. *Journal of Biopharmaceutical Statistics*, **8**, 445–467.

Wilhelm, S., Carter, C., Lynch, M., Lowinger, T., Dumas, J., Smith, R. A., Schwartz, B., Simantov, R. and Kelley, S. (2006) Discovery and development of sorafenib: a multikinase inhibitor for treating cancer. *Nature Reviews. Drug Discovery*, **5**, 835–844.

Williamson, S. F., Jacko, P., Villar, S. S. and Jaki, T. (2017) A bayesian adaptive design for clinical trials in rare diseases. *Computational Statistics & Data Analysis*, **113**, 136–153.

Wilson, E. B. (1927) Probable inference, the law of succession, and statistical inference. *Journal of the American Statistical Association*, **22**, 209–212.

Yamamoto, N., Nokihara, H., Yamada, Y., Shibata, T., Tamura, Y., Seki, Y., Honda, K., Tanabe, Y., Wakui, H. and Tamura, T. (2017) Phase I study of nivolumab, an anti-pd-1 antibody, in patients with malignant solid tumors. *Investigational New Drugs*, **35**, 207–216.

Yeung, W. Y., Reigner, B., Beyer, U., Diack, C., Sabanés bové, D., Palermo, G. and Jaki, T. (2017) Bayesian adaptive dose-escalation designs for simultaneously estimating the optimal and maximum safe dose based on safety and efficacy. *Pharmaceutical Statistics*, **16**, 396–413.

Yeung, W. Y., Whitehead, J., Reigner, B., Beyer, U., Diack, C. and Jaki, T. (2015) Bayesian adaptive dose-escalation procedures for binary and continuous responses utilizing a gain function. *Pharmaceutical Statistics*, **14**, 479–487.

Yin, G. (2013) *Clinical trial design: Bayesian and frequentist adaptive methods*, vol. 876. John Wiley & Sons.

Yin, G., Li, Y. and Ji, Y. (2006) Bayesian dose-finding in Phase I/II clinical trials using toxicity and efficacy odds ratios. *Biometrics*, **62**, 777–787.

Yin, G. and Yuan, Y. (2009) Bayesian dose finding in oncology for drug combinations by copula regression. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, **58**, 211–224.

Yin, G., Zheng, S. and Xu, J. (2013) Two-stage dose finding for cytostatic agents in Phase I oncology trials. *Statistics in Medicine*, **32**, 644–660.

Yuan, Z., Chappell, R. and Bailey, H. (2007) The continual reassessment method for multiple toxicity grades: A bayesian quasi-likelihood approach. *Biometrics*, **63**, 173–179.

Zidek, J. V. (1973) Estimating the scale parameter of the exponential distribution with unknown location. *The Annals of Statistics*, **1**, 264–278.

Zohar, S. and Chevret, S. (2001) The continual reassessment method: comparison of bayesian stopping rules for dose-ranging studies. *Statistics in Medicine*, **20**, 2827–2843.