DOCTORAL THESIS

Developments to established dose-finding methodologies for application in trials with complex and innovative designs

---

Author:

Amit PATEL

Supervisor:

Prof. Lucinda BILLINGHAM

Dr. Kristian BROCK

A thesis submitted in fulfillment of the requirements

for the degree of Doctor of Philosophy

Institute of Cancer and Genomic Sciences

College of Medical and Dental Sciences

University of Birmingham

August 14, 2020

Abstract

Developments to established dose-finding methodologies for application in trials with complex and innovative designs

by Amit PATEL

Insert abstract here...

Acknowledgements

Acknowledge people here . . .

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Chapter 2

Implementing the PO-TITE-CRM trial design into ADePT-DDR

2.1 Introduction

Worldwide there are approximately 600,000 new cases of Head and Neck Squamous Cell Carcinoma (HNSCC) each year [1]. Of which, 12,000 occur in the UK with the most common forms of treatment being surgery, radiotherapy and/or chemotherapy [2]. Radiotherapy is essential for the treatment of cancer, it has been estimated that more than 40% of patients will receive radiotherapy at some point in their treatment [3]. However, despite recent advancements in radiation techniques and the use of concomitant chemo radiotherapy, patients with solid tumours such as head and neck cancer have suboptimal cure rates [2], [4]. For those with advanced HNSCC primary radiotherapy with concurrent chemotherapy is often offered but, it has not been shown to improve survival in patients aged over 70 compared to radiotherapy alone [5]. Therefore, any strategy to improve the efficacy of radiotherapy without increasing toxicity would have a significant impact on patient outcomes.

DNA damage repair (DDR) inhibition is a potential technique which could

be utilised as it potentiates the therapeutic effects of ionising radiation in cancer cells without a substantial increase in acute and late toxicity. Combining radiotherapy with DDR inhibition could improve clinical outcomes for these patients [6].

The ADePT-DDR trial is a platform trial which aims to evaluate the safety and efficacy of different DDR agents, or different immunotherapy agents and/or DDR and immunotherapy combinations, together with radiotherapy in patients with HNSCC. The initial component of this trial is a single-arm dose-finding phase Ib/IIa trial investigating the Ataxia Telangiectasis and Rad3 Related (ATR) inhibitor AZD6738 in combination with radiotherapy. ATR inhibitors not only stop DNA repair but impairs the mechanism that allows for repairs to take place. Preclinical models have shown this double blocking to be effective in killing cancer cells [7].

Traditionally dose-finding trials aim to determine the maximum tolerated dose (MTD) of a treatment based on the cytotoxic assumption that the most toxic dose is the most efficacious. Rule-based or 'up and down' designs achieve this by escalating and de-escalating doses dependent on the observation of severe toxicity due to the drug, commonly referred to as a dose-limiting toxicity (DLT). In the case of the 3+3 design escalation continues until at least two patients in a cohort of three or six experience a DLT. More explicitly, the MTD is the dose at which $\geq 33\%$ of patients experience a DLT [8]. Model-based designs such as the continual reassessment method (CRM) [9] work on a slightly different principle which assumes that the probability of toxicity monotonically increases with dose. One key difference with the CRM is that it iteratively changes dose, seeking some acceptable target probability of toxicity also referred to as the MTD.

Due to the historical use of rule-based designs, the majority of the terminology used to describe them, and the ambiguity they raise, have been inherited

by modern designs such as the CRM. The MTD in the context of a CRM is not the 'maximum' dose patients could tolerate but rather a dose in which there would be an acceptable target probability of a DLT occurring. For example, if the target is set at 25% the MTD would be the dose at which there is a 25% probability of experiencing a DLT. Rather than using the term MTD, the dose to be found will be referred to as the target dose (TD%%, where the %'s are replaced by the target probability), i.e. TD25 would be the dose expected to be toxic in 25% of patients.

The investigation of multiple-agent treatments, where the monotonicity assumption may not hold, is increasing in early phase trials. Finding the TD in combinations of treatments, compared to single-agents, presents methodological challenges. Each drug individually may obey the monotonicity assumption, however, when combined, the ordering of doses in terms of toxicity may not be fully apparent. An order for a subset of the combined doses could be identified resulting in a partial order. Without a fully understood ordering it is uncertain which dose should be chosen in decisions of escalation and de-escalation and ultimately as the TD. This issue is not exclusively reserved for trials with multiple-agents. The monotonicity assumption may not hold for certain drugs in single-agent studies leading to partial orders of dose toxicity. This can occur in scenarios where multiple parameters of the treatment schedule are altered for each dose level. For example, either dose or treatment duration could be increased and even if patients receive an equal dose it would remain unclear as to if prolonged exposure to a lower dose is more toxic than short exposure to a higher dose, which implies a partial ordering of toxicity probabilities.

Further methodological challenges revolve around the issue of late-onset toxicities. Typically, early phase trials implement a short window, post-treatment,

to observe DLTs. This works well in situations where toxicities are likely to occur rapidly after treatment. However, this is not optimal for treatments that could cause late-onset toxicities such as radiotherapy. The aim here would be to incorporate a larger observation window to account for potential late-onset toxicities whilst also minimising the trial duration.

Cheung and Chappel [10] introduced an extension to the CRM to deal with the issues of treatments that may cause late-onset toxicity. This design referred to as the time-to-event CRM (TITE-CRM), uses a weighted dose-response model to incorporate the time it takes for a DLT to occur in a patient. There have also been published trial designs to deal with the issues that arise from investigating combinations of treatments. Thall et al. [11] proposed an adaptive two-stage Bayesian design which utilises a parametric model of toxicity as a function of two doses. Yin and Yaun [12] present a Bayesian design that uses a copula regression model to evaluate the joint toxicity probabilities of combined drugs. The continual reassessment method for partial orders (PO-CRM) developed by Wages et al. [13] extends the CRM design by relaxing the assumption of monotonicity and by modelling different potential orders.

Wages et al. [13], [14] further developed their work on the PO-CRM to deal with late-onset toxicities by implementing a TITE component. This trial design, referred to as the time-to-event continual reassessment method in the presence of partial orders (PO-TITE-CRM) by the authors, was chosen to be used in ADePT-DDR. A search of PubMed, conducted on the 25th of July 2020, found six articles that had cited the PO-TITE-CRM design by Wages et al. [14]. Five of these papers were methodological in nature, two of which only include the PO-TITE-CRM design in a brief introduction to current methodology before going on to present new Bayesian trial designs [15], [16]. The other three papers were authored by Wages. The first of which details practical considerations and specifications for the PO-CRM design, the TITE variant is only

cited as the source of an example which is being used [17]. One paper presents an R package 'pocrm' [18], [19], the package is only capable of analysing the PO-CRM design, similarly, the TITE variant is only referenced as it illustrates the issue of partial ordering. The last methodological paper by Wages et al. [20] presents three different methods for phase I studies of drug combinations one of which is the PO-CRM however, PO-TITE-CRM is only mentioned as an extension to this design. A key message in this paper is the fact that novel methodologies are constantly emerging but are rarely implemented in practice. The last paper is a protocol paper that is investigating a combination of treatments however, it is only escalating dose in one of the drugs [21]. The PO-TITE-CRM paper is cited but it appears as if another variant of the TITE-CRM is being used. This is just a brief review of the current literature but it seems that the PO-TITE-CRM has rarely been used or discussed since its inception.

Section 2.2 will detail how the PO-TITE-CRM works. Section 2.3 discusses the justification for implementing the design into the ADePT-DDR trial and our experiences doing so.

2.2 The PO-TITE-CRM Design

Wages et al. [14] introduced the PO-TITE-CRM design which builds directly upon the PO-CRM design by incorporating a TITE component into the dose toxicity model. The aim of which is to determine the target dose for combinations of drugs where the monotonicity assumption does not hold, in a setting where late-onset toxicities are possible.

To help understand partial ordering consider an example of an early phase trial investigating the combination of two agents. Drug A which consists of three doses (0.25, 1.0, 1.5 mg/day) and drug B which consists of two doses (1.0, 1.5 mg/day), for a total of six drug combinations d_1, ..., d_6 (Table 2.1).

For each drug independently we assume they have a monotonic dose-toxicity curve however, the ordering of toxicity probabilities for some of the treatment combinations is unknown. Specifically, we can say d_1 is less toxic than d_2 even the dose of drug B is the same as the dose of drug A increased. The same can be said for d_2 and d_3. The order between d_4, d_5 and d_3 is not known because the dose of drug A decreases whilst the dose of drug B increases. Assessing all these potential order toxicity relationships leaves 5 possible orderings.

1. $d_1 \to d_2 \to d_3 \to d_4 \to d_5 \to d_6$

2. $d_1 \to d_2 \to d_4 \to d_3 \to d_5 \to d_6$

3. $d_1 \to d_2 \to d_4 \to d_5 \to d_3 \to d_6$

4. $d_1 \to d_4 \to d_2 \to d_3 \to d_5 \to d_6$

5. $d_1 \to d_4 \to d_2 \to d_5 \to d_3 \to d_6$

TABLE 2.1: Example of drug combinations for a trial investigating two agents.

| | **Drug combinations** | | | | | |
Agent	d_1	d_2	d_3	d_4	d_5	d_6
A (mg/day)	0.25	0.5	1.0	0.25	0.5	1.0
B (mg/day)	1.0	1.0	1.0	1.5	1.5	1.5

Using the notation of Wages et al. [13], [14], let M denote the number of possible orders and Y be an indicator of a toxicity event. Then for a trial investigating k combinations, $d_1,...,d_k$, the dose for the jth patient, X_j, $j = 1,...,n$ can be thought of as random $x_j \in (d_1, ..., d_k)$. For a specific ordering m, $m = 1, ..., M$ the toxicity probability $R(x_j)$ is modelled by

$$R(x_j) = P(Y_j = 1 | X_j = x_j) = E(Y_j | x_j) \doteq \phi_m(d_i, w, a) = w \psi_m(d_i, a) \qquad (2.1)$$

for a weighted dose response model $\phi_m(d_i, w, a)$ and $a \in A$, where $A = (-\infty, \infty)$. The weight, w as defined by Cheung and Chappel [10], is a function of the time-to-event of each patient and is incorporated linearly with the dose toxicity model ψ so that $0 \leq w \leq 1$. Each patient is followed for a fixed amount of time T. Let U_j represent the time-to-toxicity of patient j. Then for $u \leq T$,

$$P(U_j \leq u) = P(U_j \leq u | U_j \leq T) P(U_j \leq T) \equiv w(u; T) \psi_m(d_i, a). \qquad (2.2)$$

For simplicity we will refer to the weight function $w(u; T)$ as w. The weight function will have to be decided upon by the trials team, dependent on the scenario, a simple linear function or a more complex adaptive weights function could be utilised. There are also several working dose models which could be used for ψ, Wages et al. [14] present their design with the power parameter model given by

$$\psi_m(d_i, a) = \alpha_{mi}^{exp(a)} \quad i = 1, ..., k; m = 1, ..., M. \qquad (2.3)$$

Here $0 < \alpha_{m1} < ... < \alpha_{mk} < 1$ are the prior estimates of toxicity probabilities, or skeleton, for each potential ordering. Furthermore, prior probabilities are assigned to each order M to account for any prior information regarding the plausibility of each model such that, $p(m) = \{p(1), ..., p(M)\}$, where $p(m) \geq 0$ and $\sum_m p(m) = 1$. When all orders are equally likely or there is no prior information available on possible orderings the prior is discretely uniform and would be $p(m) = 1/M$.

A Bayesian framework is used and a prior probability distribution $g(a)$ is assigned to the parameter a. The ordering with the largest prior probability is selected as the starting ordering, in the scenario where all priors are equal an ordering is selected at random, subsequently a starting dose is also chosen.

After j patients have been entered into the trial data is collected in the form of $\Omega_j = \{x_1, y_1, ..., x_j, y_j\}$. A weighted likelihood for the parameter a is used to establish running probabilities of toxicity for each treatment combination. The weighted likelihood under ordering m, is given by

$$\tilde{L}_m(a|\Omega_j) = \prod_{l=1}^{j} \phi_m^{y_l}(d_l, w_l, a)\{1 - \phi_m(d_l, w_l, a)\}^{(1-y_l)} \tag{2.4}$$

which can be used to generate a summary value $\hat{a}_{mj}$, for a, for each ordering. With the likelihood and the data Ω_j, the posterior density for a can be calculated using

$$\tilde{f}_m(a|\Omega_j) = \frac{\tilde{L}_m(a|\Omega_j)g(a)}{\int_A \tilde{L}_m(a|\Omega_j)g(a)da} \tag{2.5}$$

This can then be used to establish posterior probabilities of the orderings given the data as

$$\tilde{\pi}(m|\Omega_j) = \frac{p(m) \int_A \tilde{L}_m(a|\Omega_j)g(a)da}{\sum_{m=1}^{M} p(m) \int_A \tilde{L}_m(a|\Omega_j)g(a)da}. \tag{2.6}$$

We select the single ordering, h, with the largest posterior probability along with its associated working model $\psi_h(d_i, a)$ and generate toxicity probabilities for each dose level. Once the jth patient has been included the posterior probability of DLT can be calculated for d_i so that

$$\hat{R}(d_i) = \psi_h(d_i, \hat{a}_{hj}); \; \hat{a}_h = \int_A a\tilde{f}_h(a|\Omega_j)da. \tag{2.7}$$

In turn, the dose level $x_j \in \{d_1, ..., d_k\}$ assigned to the $(j+1)$th patient is the dose, d_i, which minimises

$$\triangle(\hat{R}(d_i), \theta) = |\hat{R}(d_i) - \theta|, \; i = 1, ..., k \tag{2.8}$$

where θ is the target toxicity rate. Similarly, once all patients have been recruited and observed and the trial ends, the target dose (TDθ) is the dose, d_i,

which minimises (2.8).

2.3 PO-TITE-CRM in ADePT-DDR

The decision to implement PO-TITE-CRM into ADePT-DDR was made by Piers Gaunt (PG) after discussions with other statisticians Kristian Brock (KB) and Daniel Slade (DS), as well as the chief investigator and other co-investigators. The design was chosen as the toxicity probabilities of the dose levels weren't monotonically increasing which restricts the use of most early phase designs such as the CRM. Additionally, the design also handles late-onset toxicities which would be an issue in ADePT-DDR due to the treatment involving radiotherapy. The availability of software to conduct the trial was also a factor that was considered. The R package 'pocrm' [18] only provides a means for implementing the PO-CRM design but the easy accessibility to this code meant that it could be extended to include the TITE component.

The intended use of this design is for dose-finding in combinations of therapies, as this is the source of the partial ordering issue. ADePT-DDR however, is a unique implementation of the design as even though it involves a combination of therapies (radiotherapy and AZD6783) the dose of radiotherapy is fixed and dose-finding is only planned for AZD6783. PO-TITE-CRM is still applicable in this case as the dose levels for AZD6783 are partially ordered.

A two-stage PO-TITE-CRM will be used to find the TD25 of AZD6783. This will be determined by dose-limiting toxicities evaluated by Common Terminology Criteria for Adverse Events (CTCAE) v5.0 and Radiation Therapy Oncology Group (RTOG) late toxicity score. The binary DLT events are predefined by a variety of grade 3-4 adverse events notably, haematological, cardiovascular and gastrointestinal/hepatic toxicities as well as significant non-haematological events and specific treatment-related toxicities. DLTs will be

monitored for the duration of treatment (seven weeks) and up to a year post-treatment.

A maximum of 60 patients will be recruited for the dose-finding aspect of this trial and up to 20 patients as controls. Controls will be utilised to make comparisons for secondary outcomes such as survival and efficacy. Cohorts of three patients will be recruited and assigned to dose levels chosen by the PO-TITE-CRM. Controls will be recruited in the interim period between the recruitment of the third patient in a cohort and the completion of the minimum follow-up period.

2.3.1 Partial Ordering in Practice

Each patient entered into ADePT-DDR will receive fixed dose radiation, totalling 70 Gy in 35 fractions over seven weeks. For the dose-finding aspect we investigate six doses of AZD6783 detailed in table 2.2. Treatment dose and duration to be selected for dose level 3 will be determined based on a combination of data observed, adverse events and compliance. The issue of partial ordering stems from dose levels 2a and 2b, which can be seen in Figure 2.1. The increased duration of 2a and increased dose of 2b mean they are both more toxic than dose level 1. However, when comparing 2a and 2b it is unknown whether the increase in duration or dose will be more toxic. Hence there are two possible orderings for ADePT-DDR.

1. $d_{-1} \rightarrow d_0 \rightarrow d_1 \rightarrow d_{2a} \rightarrow d_{2b} \rightarrow d_3$

2. $d_{-1} \rightarrow d_0 \rightarrow d_1 \rightarrow d_{2b} \rightarrow d_{2a} \rightarrow d_3$

Preliminary designs of the trial included only five dose levels and planned to use dose level 0 as the starting dose. During the trial design phase it was decided a new lower dose (dose level -1) would be introduced to allow for

TABLE 2.2: ADePT-DDR dose levels and duration of treatment
for AZD6783.

Dose Level	AZD6783 Daily dose (mg BD)	Weeks	Duration (days)	Radiotherapy
-1	20	1	5	70 Gy / 35 F
0	20	1&4	10	70 Gy / 35 F
1	40	1&4	10	70 Gy / 35 F
2a	40	1,2,4&5	20	70 Gy / 35 F
2b	80	1&4	10	70 Gy / 35 F
3	120	1&4	10	70 Gy / 35 F
	80	1,2,4&5	20	70 Gy / 35 F

FIGURE 2.1: ADePT-DDR dose levels across dose and duration.
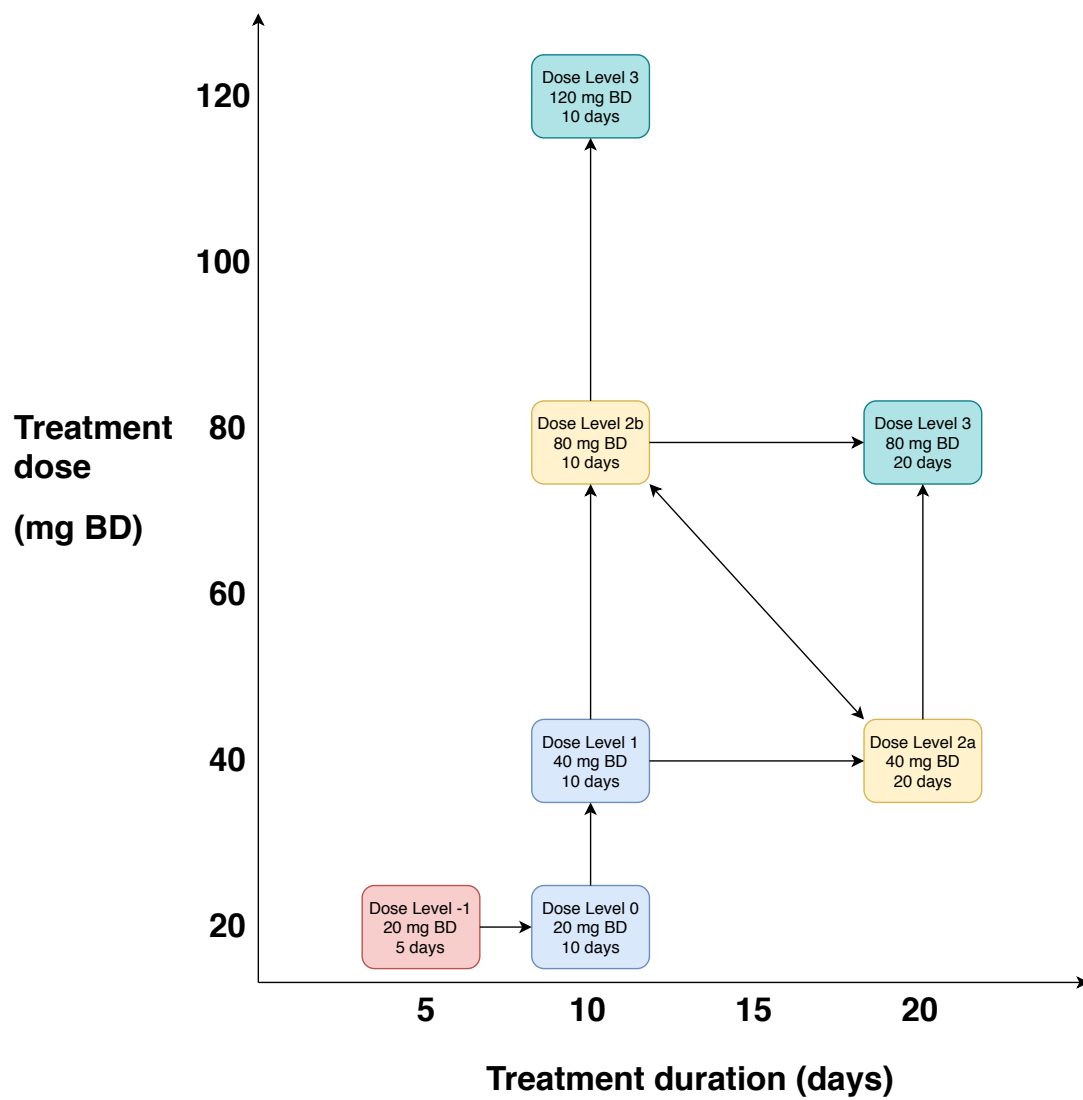
de-escalation if the initial dose was found to be too toxic. Dose escalation/de-escalation for subsequent cohorts would be determined from the two-stage PO-TITE-CRM. A two-stage design allows for escalation according to a pre-defined escalation scheme similar to a '3+3' design. The first stage dictates that if no DLT's are observed in the current cohort the dose allocated to the next cohort is the following dose in the escalation scheme. Dose levels continue to be incremented in this fashion until the first DLT is observed. In stage two, dose levels are determined by the PO-TITE-CRM.

Typically CRM designs begin by testing the first patient, or cohort, at the prior guess of TD. However, clinicians may have safety concerns beginning the trial at higher dose levels as well as escalating to higher dose levels without testing lower ones. Investigators in ADePT-DDR expressed similar concerns as such a two-stage design was adopted. The escalation scheme used in stage one of ADePT-DDR will follow that of the first ordering ($d_{-1} \rightarrow d_0 \rightarrow d_1 \rightarrow d_{2a} \rightarrow d_{2b} \rightarrow d_3$). If patients in the first cohort (assigned to dose level 0) don't experience a DLT the next cohort will be allocated to dose level 1 and then if no DLTs are observed again the third cohort will be allocated to dose level 2a and so on and so forth. The dose escalation scheme was determined based on the prior probabilities of toxicity generated for each dose level.

Information elicited from the investigators helped generate prior probabilities of toxicity for each dose level. They believed that dose level 2b would be the TD25 with 2a being less toxic. This was used in conjunction with the getprior function from the dfcrm R package [22] which yielded priors of 0.012, 0.036, 0.084, 0.157, 0.25 and 0.355 for dose levels -1, 0, 1, 2a, 2b and 3 respectively. Prior probabilities are also required for the plausibility of each model and even though the clinicians think that 2b will be more toxic than 2a there is no clear evidence and still a lot of uncertainty. As such it is sensible to assume a plausibility probability of 0.5 for each ordering, implying both orders

are equally likely to be the true ordering of these dose levels.

2.3.2 The TITE component

The observation window for this trial will be longer as the combination of radiotherapy with AZD6738 is anticipated to cause late-onset toxicity. The Acute DLT observation period is 84 days (12 weeks) post radiotherapy end with a minimum of 56 days (8 weeks) for the last patient of each cohort. However, patients will continuously be monitored for occurrence of DLT for at least 84 days (12 weeks), i.e. at least 84 days (12 weeks) from the end of radiotherapy. Th full window will last for 365 days (52 weeks).

The TITE component incorporates a weighting contribution for each patient dependent on how long that patient has been evaluable in the study. This allows a patient to be evaluated once they have been observed for the minimum DLT period of 56 days (8 weeks). The weighting at this point is 60% rising to 80% at 84 days (12 weeks). A patient will not contribute fully to the model until they have completed 365 days (52 weeks) follow up (or have experienced a DLT at any stage in which case they will be weighted as a whole contribution). Linear weighting functions will be employed for any patient with a length of follow up between these three time points. One weighting function to calculate weight from 56-84 days (8-12 weeks) and another for weights from 84-365 days (12-52 weeks). For the weighting function $w(u; T)$ where u is the time-to-toxicity of patient j and T is the time period, let t_1, t_2, t_3 take values 56, 84 and 365 respectively such that $T \in (t_1, t_2, t_3)$. Then for $t_1 \leq u \leq t_3$

$$w(u; T) = 0.6 + 0.2\frac{min(u - t_1, t_2 - t_1)}{t_2 - t_1} + 0.2\frac{max(0, u - t_2)}{t_3 - t_2}. \qquad (2.9)$$

All patients will have a minimum weight of 60% as that is the prescribed weighting to the minimum follow up period before dose escalation/de-escalation

decisions can be made. For each additional week the patient is observed, without a DLT occurring, between weeks 8 and 12 their weighting increases by 5%. Similarly for each week between 12 and 52 weeks, without a DLT, weighting increases by 0.5%.

The TITE-CRM originally presented by Cheung and Chappel [10] did not incorporate a minimum follow-up period and their design allowed for the continual recruitment of patients whenever they became available. There are some practical considerations which make this infeasible. The model would need to be run each time a new patient entered the study which requires statistical input hence the introduction of cohorts. Clinicians may also have safety concerns if we see rapid recruitment at the start of the trial and the model keeps escalating so we impose a minimum follow-up period. Initially this was set at 12 weeks (at 80% weighting) however, statisticians AP and PG pointed out that dose escalation/de-escalation decisions would have to take place 19 weeks (7 weeks treatment and 12 weeks follow-up) after recruitment of the third patient in the cohort. Dependent on the recruitment rates this could extend the duration of the trial and negates the benefits of using a TITE design. The investigators aslo agreed this was too long and settled on lowering this period to 8 weeks (at 60% weighting) whilst also including the original 12 week weighting of 80%.

2.3.3 Stopping Rules

A practical modification was included to allow for early stopping of the trial if there is sufficient evidence that the TD has been reached. Sufficient evidence is achieved once 15 patients (five cohorts) have been treated at the same dose level and the model allocates that dose level again to a sixth cohort. This rule

evolved from the original designs of the trial which involved 30 patients with a dose expansion cohort to ensure at least 15 patients were treated at the TD.

Initial simulations highlighted the inadequacy of these design parameters as operating characteristics for various scenarios were poor, specifically in terms of correct TD selection. Clinicians explained the inclusion of the dose expansion cohort was to ensure the dose-finding aspect of the trial did not take a large amount of time whilst also allowing safety to be assessed at the TD. In order to ensure that a reasonable amount of patients would be treated at the TD, the trial wouldn't take longer than necessary and operating characteristics improved, the sample size was increase and this rule was introduced.

A rule was also implemented to allow for early termination of the trial in the case of excess toxicity at the lowest dose. If the posterior probability of DLT at the lowest dose is higher than 0.35 with a probability of 80% and has been tested the trials safety committee will be alerted and will recommend if the trial should be stopped. As the trial starts at dose level 0, which is not the lowest dose, it is possible for the trial to recommend terminating without ever allocating patients to the lowest dose level. As such it was decided early termination would only occur once at least 3 patients (1 cohort) have been allocated dose level -1.

2.3.4 Operating Characteristics

Simulations were continually utilised during the design process of the trial to assess how various changes impact the overall performance. These changes to design features such as the sample size, weight function and stopping rules helped inform decisions which led to the final design.

Functions from pocrm package in R [18], [19] were modified in order to perform simulations and conduct the trial. The majority of work involved

integrating the TITE component and the stopping rules into the code. In standard CRM designs a binary outcome for toxicity is generated for each patient based on a pre-specified true DLT rates for the dose they are assigned. Adding the TITE component means the time the toxicity occurs also has to be generated, the simulation must also track this time and incorporate this information into the PO-TITE-CRM model when it needs to make dose allocation decisions for the next cohort. We defined multiple scenarios to reflect various real life possibilities in order to assess the designs performance.

Standard scenarios to run involve adjusting the true DLT rates to reflect each dose being the TD25. For each of these we calculate the probability of selecting each dose as the TD25. It would be expected the dose with the highest probability of being selected has its true DLT rate set at 25% to match the target rate. A high probability of selection for the correct dose implies the design works well in the specified scenario. Additional characteristics such as the average number of patients at each dose level are also investigated. This can be used to look at how many patients may potentially be allocated to a toxic dose. It is also necessary to consider performance when all doses are too toxic, here we would want the design to recommend stopping early. Usually the true DLT rates used to define these scenarios abide by the monotonicity assumption. Due to the partial ordering we consider scenarios in which the true DLT rates follow both orders. For trials with a large amount of orders it may be unfeasible to run so many simulations. However, as ADePT-DDR only has two orders we explored all scenarios for each ordering.

We simulated 2000 trials for each scenario using the finalised design detailed in section 2.3. A fixed recruitment rate of one patient per month was utilised. The occurrence of DLT's were randomly generated for patients in each cohort using a Bernoulli distribution with the probability set at the true

DLT rate for the cohorts assigned dose level in the specific scenario. For patients who had a DLT occur, the time at which the DLT occurred was randomly generated using a uniform distribution which spanned the start of treatment to the end of follow-up.

Table 2.3 details simulations for eight scenarios to test the performance of the PO-TITE-CRM design using true DLT rates which reflect the first ordering. We analyse scenarios where each dose is the TD25 (scenarios 1-6) and when all doses are too toxic (scenario 8). Additionally, we also investigate performance under conditions where the probability of DLT is fairly similar between doses (scenario 7). This is a notoriously difficult circumstance for CRM designs to deal with as the limited number of patients and events at each dose make it hard to accurately estimate toxicity probabilities if they are similar. Simulation results for ordering 2 are shown in Table 2.4 where dose level 2a is considered more toxic than 2a. This is achieved by altering the true DLT rates so 2b has a lower probability of DLT compared to 2a.

Ideally we want the probability of selection for the dose allocated at TD25 to be as high as possible and greater than other dose levels. For scenarios 1-7 the TD25 is highlighted in bold along with results from the simulations. However, for scenario 8 where all doses are too toxic we expect the trial to terminate early, here 'stop' should be selected and its associated probability of stopping is shown in bold.

In scenarios 1 - 6 (Table 2.3), this design correctly selects the TD25 with probabilities between 43% and 78%, under the assumption 2b is more toxic than 2a. Likewise, for the ordering where 2a is more toxic than 2b, scenarios 9-14 (Table 2.4) have probabilities between 42% and 78% of correctly selecting the TD25. Correct selection probabilities are generally higher when the TD25 is at the first and last dose levels compared to dose levels 2a and 2b. However, these dose levels are still chosen with the highest probability as the TD25 in

their given scenarios. For scenarios 7 and 15, the probabilities of toxicity are equally spaced, approximately 5% apart. This is a relatively diffcult scenario for dose-finding studies to handle. The probability of selecting the TD25 is 28% and 30% for orderings 1 and 2 respectively. In scenarios 8 and 16, where all the doses are too toxic, the design very seldom allocates patients higher than the first three doses and there is a high chance (74%) that the trial will recommend early stopping.

TABLE 2.3: Operating Characteristics of the two-stage PO-TITE-CRM (with true DLT rates that imply 2b is more toxic than 2a) based on 2000 simulated trials. Definitions: DLT: Dose-limiting toxicity. P(select): Probability of selecting a dose as the TD25.

		Dose Levels						
		-1	0	1	2a	2b	3	Stop
Scenario	Prior DLT	0.01	0.04	0.08	0.16	0.25	0.35	
	True DLT rate	0.25	0.4	0.45	0.5	0.55	0.6	
1: MTD @-1	P(select)	0.63	0.18	0.05	0	0	0	0.13
	% of patients	39	32	20	6	3	0	
	Mean number of patients	10.08	8.32	5.3	1.56	0.67	0.08	
	True DLT rate	0.12	0.25	0.4	0.45	0.5	0.55	
2: MTD @0	P(select)	0.21	0.53	0.19	0.02	0.02	0	0.03
	% of patients	18	36	30	11	5	1	
	Mean number of patients	5.33	10.52	8.74	3.22	1.57	0.23	
	True DLT rate	0.09	0.12	0.25	0.4	0.45	0.5	
3: MTD @1	P(select)	0.01	0.19	0.56	0.13	0.1	0.01	0
	% of patients	4	20	35	22	16	3	
	Mean number of patients	1.12	6.51	11.15	6.99	5.05	0.99	
	True DLT rate	0.06	0.09	0.12	0.25	0.4	0.45	
4: MTD @2a	P(select)	0	0.02	0.23	0.47	0.24	0.04	0
	% of patients	1	12	20	31	25	10	
	Mean number of patients	0.48	3.9	6.84	10.43	8.27	3.46	
	True DLT rate	0.03	0.06	0.09	0.12	0.25	0.4	
5: MTD @2b	P(select)	0	0	0.03	0.28	0.44	0.25	0
	% of patients	1	10	13	25	28	24	
	Mean number of patients	0.25	3.37	4.29	8.24	9.36	8.14	
	True DLT rate	0.01	0.03	0.06	0.09	0.12	0.25	
6: MTD @3	P(select)	0	0	0	0.08	0.13	0.79	0
	% of patients	0	10	11	17	18	42	
	Mean number of patients	0.12	3.16	3.51	5.44	5.68	13.21	
	True DLT rate	0.05	0.1	0.15	0.2	0.25	0.3	
7: Equal steps in DLT rate	P(select)	0	0.02	0.12	0.31	0.27	0.27	0
	% of patients	2	13	19	26	22	19	
	Mean number of patients	0.51	4.04	5.88	8.35	6.88	5.97	
	True DLT rate	0.5	0.6	0.65	0.7	0.75	0.8	
8: All toxic	P(select)	0.29	0	0	0	0	0	0.71
	% of patients	56	26	15	2	0	0	
	Mean number of patients	9.05	4.2	2.42	0.32	0.03	0	

TABLE 2.4: Operating Characteristics of the two-stage PO-TITE-CRM (with true DLT rates that imply 2a is more toxic than 2b) based on 2000 simulated trials. Definitions: DLT: Dose-limiting toxicity. P(select): Probability of selecting a dose as the TD25.

		Dose Levels						Stop
		-1	0	1	2a	2b	3	
Scenario	Prior DLT	0.01	0.04	0.08	0.16	0.25	0.35	
9: MTD @-1	True DLT rate	0.25	0.4	0.45	0.55	0.5	0.6	
	P(select)	0.63	0.18	0.05	0	0	0	0.12
	% of patients	40	32	20	5	2	0	
	Mean number of patients	10.3	8.26	5.14	1.38	0.58	0.04	
10: MTD @0	True DLT rate	0.12	0.25	0.4	0.5	0.45	0.55	
	P(select)	0.22	0.52	0.21	0.02	0.02	0	0.02
	% of patients	18	36	29	10	6	1	
	Mean number of patients	5.28	10.67	8.76	3.1	1.73	0.2	
11: MTD @1	True DLT rate	0.09	0.12	0.25	0.45	0.4	0.5	
	P(select)	0.02	0.21	0.54	0.09	0.14	0.01	0
	% of patients	4	21	35	20	17	3	
	Mean number of patients	1.23	6.55	11.15	6.41	5.37	1.03	
12: MTD @2a	True DLT rate	0.06	0.09	0.12	0.25	0.15	0.45	
	P(select)	0	0.02	0.08	0.44	0.31	0.14	0
	% of patients	2	12	16	31	22	18	
	Mean number of patients	0.51	3.94	5.36	10.24	7.51	5.95	
13: MTD @2b	True DLT rate	0.03	0.06	0.09	0.35	0.25	0.4	
	P(select)	0	0	0.15	0.31	0.43	0.11	0
	% of patients	1	11	19	30	28	13	
	Mean number of patients	0.25	3.51	6.15	9.84	9.1	4.2	
14: MTD @3	True DLT rate	0.01	0.03	0.06	0.12	0.09	0.25	
	P(select)	0	0	0	0.13	0.09	0.78	0
	% of patients	0	10	11	19	17	42	
	Mean number of patients	0.13	3.15	3.54	5.95	5.17	13.11	
15: Equal steps in DLT rate	True DLT rate	0.05	0.1	0.15	0.25	0.2	0.3	
	P(select)	0	0.02	0.13	0.3	0.27	0.28	0
	% of patients	2	13	19	27	21	19	
	Mean number of patients	0.52	4.09	5.89	8.48	6.64	5.91	
16: All toxic	True DLT rate	0.5	0.6	0.65	0.75	0.7	0.8	
	P(select)	0.27	0	0	0	0	0	0.73
	% of patients	56	27	15	2	0	0	
	Mean number of patients	8.66	4.15	2.23	0.27	0.04	0	

Additionally, we assess designs based on how doses are allocated to patients. Designs may correctly select the TD however, this could be undesirable and unethical if the majority of patients are over dosed at the more toxic dose levels. The average number and the percentage of patients at each dose level, for each scenario, is recorded in Tables 2.3 and 2.4.

The percentage of patients treated at the TD25 ranges between 23% and 43% for each scenario under both orderings. The design also allocates the most

patients on average to the TD25 apart from in scenario 7. In this case more patients were allocated to the next lowest dose, we have already discussed the difficulties of this scenario so this characteristic is not too concerning. The mean number of patients recruited for scenarios 1-6 is 26, 30, 32, 33, 34 and 31 respectively. Similarly for scenarios 9-14 its 26, 30, 32, 34, 33 and 31. Even though we allow for up to 60 patients the majority of trials terminate early based on the pre-defined rules for selecting the TD25.

Overall, the simulation results show the specification of this design performs relatively well in a number of scenarios. We have shown there is a high probability of the trial stopping early if all dose-levels are too toxic. We have also shown the design behaves in an appropriate manner when there is a lack of disparity between dose-levels in terms of toxicity. Finally, we have demonstrated that regardless of the ordering we observe the PO-TITE-CRM has a high probability of selecting the correct dose. There are a number of limitations to the operating characteristics presented here which are due to the specification of the simulations and trial design. Section 2.4 explores and discusses these limitations in more detail.

2.4 Inspection of the operating characteristics

- Limitations of the simulation study

  - Selection of True DLT rates

  - Generation of Time-to-event data

  - Specification of the accrual rate (time of recruitment)

  - Assumes analysis is done instantaneously

  - Mention issues with each point the impact and other alternatives for simulation.

- Limitations of the trial design

    - Conflict of the sample size and the stopping rule

    - Using a time-to-event component and using cohorts

    - Simulations without cohorts, different N, no stopping rules etc.

    - General limitations with looking at operating characteristics - compare to a TITE-CRM. Can look at the difference as the price to pay for the uncertainty of not knowing order of 2a and 2b

2.5 Dose Transition Pathways in a TITE Setting

- Explain dose transition pathways

- Issues with why they are difficult to utilise in a TITE-CRM

- Show examples of TITE-CRMs ???

- Dose transition pathways for ADePT-DDR

2.6 Discussion

2.7 Conclusion

Bibliography

[1] N. Stransky, A. M. Egloff, A. D. Tward, *et al.*, "The Mutational Landscape of Head and Neck Squamous Cell Carcinoma," *Science (New York, N.Y.)*, vol. 333, no. 6046, pp. 1157–1160, Aug. 26, 2011, ISSN: 0036-8075. DOI: `10.1126/science.1208130`. pmid: `21798893`. [Online]. Available: `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3415217/`.

[2] Cancer Reaserch UK. (). "Head and neck cancers statistics," [Online]. Available: `https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/head-and-neck-cancers`.

[3] C. E. Round, M. V. Williams, T. Mee, *et al.*, "Radiotherapy Demand and Activity in England 2006–2020," *Clinical Oncology*, vol. 25, no. 9, pp. 522–530, Sep. 1, 2013, ISSN: 0936-6555. DOI: `10.1016/j.clon.2013.05.005`. [Online]. Available: `http://www.sciencedirect.com/science/article/pii/S0936655513002343`.

[4] D. M. Cognetti, R. S. Weber, and S. Y. Lai, "Head and Neck Cancer: An Evolving Treatment Paradigm," *Cancer*, vol. 113, pp. 1911–1932, 7 0 Oct. 1, 2008, ISSN: 0008-543X. DOI: `10.1002/cncr.23654`. pmid: `18798532`. [Online]. Available: `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2751600/`.

[5] J. Pignon, J Bourhis, C Domenge, *et al.*, "Chemotherapy added to locoregional treatment for head and neck squamous-cell carcinoma: Three

meta-analyses of updated individual data," *The Lancet*, vol. 355, no. 9208, pp. 949–955, Mar. 18, 2000, ISSN: 0140-6736. DOI: 10.1016/S0140-6736(00) 90011-4. [Online]. Available: http://www.sciencedirect.com/science/ article/pii/S0140673600900114.

[6]   A. Chalmers, "Science in Focus: Combining Radiotherapy with Inhibitors of the DNA Damage Response," *Clinical Oncology*, vol. 28, no. 5, pp. 279–282, May 2016, ISSN: 09366555. DOI: 10.1016/j.clon.2016.01.035. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/ S0936655516000649.

[7]   L. Mei, J. Zhang, K. He, *et al.*, "Ataxia telangiectasia and Rad3-related inhibitors and cancer therapy: Where we stand," *Journal of Hematology & Oncology*, vol. 12, no. 1, p. 43, Apr. 24, 2019, ISSN: 1756-8722. DOI: 10. 1186/s13045-019-0733-6. [Online]. Available: https://doi.org/10. 1186/s13045-019-0733-6.

[8]   C. Le Tourneau, J. J. Lee, and L. L. Siu, "Dose Escalation Methods in Phase I Cancer Clinical Trials," *JNCI: Journal of the National Cancer Institute*, vol. 101, no. 10, pp. 708–720, May 20, 2009, ISSN: 0027-8874. DOI: 10.1093/jnci/djp079. [Online]. Available: https://academic.oup. com/jnci/article/101/10/708/969691.

[9]   J. O'Quigley, M. Pepe, and L. Fisher, "Continual Reassessment Method: A Practical Design for Phase 1 Clinical Trials in Cancer," *Biometrics*, vol. 46, no. 1, pp. 33–48, 1990, ISSN: 0006-341X. DOI: 10.2307/2531628. JSTOR: 2531628.

[10]   Y. K. Cheung and R. Chappell, "Sequential Designs for Phase I Clinical Trials with Late-Onset Toxicities," *Biometrics*, vol. 56, no. 4, pp. 1177–1182, 2000, ISSN: 1541-0420. DOI: 10.1111/j.0006-341X.2000.01177.x.

[Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.0006-341X.2000.01177.x.

[11] P. F. Thall, R. E. Millikan, P. Mueller, *et al.*, "Dose-Finding with Two Agents in Phase I Oncology Trials," *Biometrics*, vol. 59, no. 3, pp. 487–496, 2003, ISSN: 1541-0420. DOI: 10.1111/1541-0420.00058. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/1541-0420.00058.

[12] G. Yin and Y. Yuan, "Bayesian dose finding in oncology for drug combinations by copula regression," *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 58, no. 2, pp. 211–224, 2009, ISSN: 1467-9876. DOI: 10.1111/j.1467-9876.2009.00649.x. [Online]. Available: https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9876.2009.00649.x.

[13] N. A. Wages, M. R. Conaway, and J. O'Quigley, "Continual reassessment method for partial ordering," *Biometrics*, vol. 67, no. 4, pp. 1555–1563, Dec. 2011, ISSN: 1541-0420. DOI: 10.1111/j.1541-0420.2011.01560.x. pmid: 21361888.

[14] ——, "Using the time-to-event continual reassessment method in the presence of partial orders," *Statistics in Medicine*, vol. 32, no. 1, pp. 131–141, Jan. 15, 2013, ISSN: 1097-0258. DOI: 10.1002/sim.5491. pmid: 22806898.

[15] S. Liu, G. Yin, and Y. Yuan, "BAYESIAN DATA AUGMENTATION DOSE FINDING WITH CONTINUAL REASSESSMENT METHOD AND DELAYED TOXICITY," *The annals of applied statistics*, vol. 7, no. 4, pp. 1837–2457, Dec. 1, 2013, ISSN: 1932-6157. DOI: 10.1214/13-AOAS661. pmid: 24707327. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3972824/.

[16]   G. M. Wheeler, M. J. Sweeting, and A. P. Mander, "A Bayesian model-free approach to combination therapy phase I trials using censored time-to-toxicity data," *Journal of the Royal Statistical Society. Series C, Applied statistics*, vol. 68, no. 2, pp. 309–329, Feb. 2019, ISSN: 0035-9254. DOI: `10.1111/rssc.12323`. pmid: `30880843`. [Online]. Available: `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6420054/`.

[17]   N. A. Wages and M. R. Conaway, "Specifications of a continual reassessment method design for phase I trials of combined drugs," *Pharmaceutical statistics*, vol. 12, no. 4, 2013, ISSN: 1539-1604. DOI: `10.1002/pst.1575`. pmid: `23729323`. [Online]. Available: `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3771354/`.

[18]   N. A. Wages and N. Varhegyi, "Pocrm: An R-package for phase I trials of combinations of agents," *Computer Methods and Programs in Biomedicine*, vol. 112, no. 1, pp. 211–218, Oct. 2013, ISSN: 1872-7565. DOI: `10.1016/j.cmpb.2013.05.020`. pmid: `23871691`.

[19]   N. A. Wages, *Pocrm: Dose Finding in Drug Combination Phase I Trials Using PO-CRM*, version 0.12, Mar. 15, 2019. [Online]. Available: `https://CRAN.R-project.org/package=pocrm`.

[20]   N. A. Wages, A. Ivanova, and O. Marchenko, "Practical designs for phase I combination studies in oncology," *Journal of biopharmaceutical statistics*, vol. 26, no. 1, pp. 150–166, 2016, ISSN: 1054-3406. DOI: `10.1080/10543406.2015.1092029`. pmid: `26379085`. [Online]. Available: `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4720553/`.

[21]   P. Lesueur, J. Lequesne, J.-M. Grellard, *et al.*, "Phase I/IIa study of concomitant radiotherapy with olaparib and temozolomide in unresectable or partially resectable glioblastoma: OLA-TMZ-RTE-01 trial protocol," *BMC Cancer*, vol. 19, Mar. 4, 2019, ISSN: 1471-2407. DOI: `10.1186/s12885-`

`019-5413-y`. pmid: `30832617`. [Online]. Available: `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6399862/`.

[22] K. Cheung, *Dfcrm: Dose-Finding by the Continual Reassessment Method*, version 0.2-2.1, Jan. 26, 2019. [Online]. Available: `https://CRAN.R-project.org/package=dfcrm`.