Mathematical modelling for the selection of optimal vaccine dose

John Helier Benest

Thesis submitted in accordance with the requirements for the degree of

**Doctor of Philosophy
of the**
University of London

August 2022

**Department of Infectious Disease Epidemiology**

Faculty of Epidemiology and Population Health

London School of Hygiene and Tropical Medicine

I, John Benest, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, this has been indicated in the thesis. I have read and understood the school's definition of plagiarism and cheating given in the Research Degrees Handbook.

John Benest
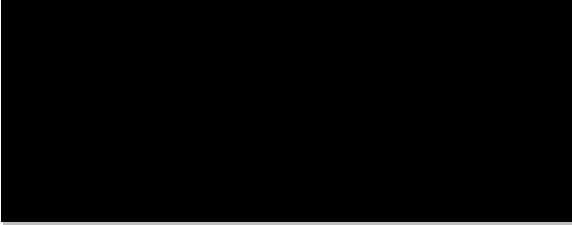
August 2022

Abstract

Background

Vaccines are an important tool in global disease burden reduction, with vaccine dose amount (hereafter 'dose') being a key decision during vaccine development. Vaccine dose selection is often conducted through empirical comparison of a small number of potential doses, which is likely to fail to find the optimal dose if none of the doses are optimal. Mathematical modelling has been suggested as a method for identifying optimal vaccine dose and has been historically important in determining optimal dose of, and informing clinical trials for, drug development. Mathematical modelling is however not commonly used in either the design of vaccine dose ranging trials nor in the selection of optimal vaccine dose based on the resulting clinical trial data.

To address this gap, recently 'Immunostimulation/Immunodynamic' (IS/ID) modelling has been proposed to encompass quantitative modelling for vaccine dose optimisation. Initial IS/ID work has been used to find the maximally immunogenic dose for tuberculosis and influenza vaccines, and has shown that, contrary to widespread belief, vaccine dose-efficacy response may be peaking rather than saturating. However, as the field is new, there are many gaps including: uncertainty in the prevalence of such peaking dose-response curve shape, primarily only efficacy-maximisation has been considered, the impact of incorrectly assuming a peaking/saturating dose response has not been assessed, and mathematical modelling has been performed retrospectively of clinical trials rather than informing them during the trial itself (e.g. in an adaptive trial design). Further, there has also been little research into multi-dimensional vaccine dose-optimisation, where there is a need to choose prime doses, boost doses, adjuvant doses, and/or time between doses, which may complicate the dose-response relationship through potential for synergism/antagonism.

3

My aim for this thesis was to explore and expand the field of IS/ID and mathematical modelling for vaccine dose optimisation, addressing the gaps described above. My objectives were:

1. To gather dose-response data through a systematic review of dose-ranging studies for a specific class of vaccine (adenoviral vector), and to find the distribution of the number of doses typically investigated in these studies.

2. Using dose-response data from objective one and mathematical models, determine the prevalence of predicted saturating versus peaking dose-response curves.

3. To extend IS/ID beyond efficacy-maximisation into multi-factor dose optimisation by proposing alternative utility functions and investigate the impact of the choice of utility function on the selection of 'optimal' dose.

4. To evaluate the potential impact of correctly or incorrectly assuming a peaking/saturating dose-efficacy response, along with the impact of adaptive trial design, on optimal vaccine dose selection.

5. To evaluate the use of a non-parametric dose-response model for the purpose of optimal vaccine dose selection in the adaptive trial design setting, with emphasis on multi-dimensional vaccine dose-optimisation.

Methods

For objective one, a class of vaccine (adenoviral vector) was selected, and dose-response data were extracted from a systematic review of single-dose dose-ranging studies. I conducted a descriptive analysis of these studies to investigate the number of doses that were typically investigated.

For objective two, representative peaking and saturating dose-response models were calibrated to the data from objective one. I assessed which of the two mathematical models best described the data through the use of Akaike Information Criterion. I determined the prevalence of dose-response data which was peaking or saturating and investigated potential covariates that may impact dose-response shape.

4

For objective three, I calibrated dose-response models to efficacy and toxicity data from a phase I dose-ranging study of a recombinant adenovirus type-5 COVID-19 single-dose vaccine (Ad5-nCoV). Using these mathematical models, I predicted optimal dose for three potential dose selection criteria, namely i) achieving herd immunity, ii) balancing efficacy and toxicity, and iii) balancing efficacy, toxicity, and cost.

For objective four, I used a simulation-based study to assess the impacts of different assumed efficacy models and trial dose selection methods on optimal dose selection and ethical trial design. Comparison was done using simulated clinical trials, using a utility function that involved both efficacy and ordinal toxicity and both peaking and saturating dose-efficacy curves across 14 dose-optimisation scenarios.

For objective five, I conducted a second simulation-based study to assess a novel non-parametric dose-response model, the 'Continuously Correlated Beta Process' model, in identifying optimal dose. This was compared to other mathematical model-based and mathematical model-free methods of vaccine dose optimisation. The simulation study included both single-dose and multi-dimensional dose-optimisation scenarios.

Results

For objective one, data from 35 studies were extracted and I found that adenoviral vector vaccine dose ranging trials were designed around selecting between a small number of doses (94% of studies investigated < six doses).

For objective two, I found that the data from the available dose-ranging trials were often insufficient to provide evidence for either a peaking or saturating dose-efficacy curve, with the peaking model best describing 22% of the data, the saturating model best describing 4.7% of the data, and there being no significant difference for 73.3% of the data. Further, the species being vaccinated and the response type of interest may be more predictive of dose-response curve shape than the adenoviral serotype or route of administration of a vaccine.

For objective three, I found that vaccine optimal dose depends on the utility function that is being maximised, with the optimal doses for the Ad5-nCoV vaccine being i) 1.3 x 10^{11}, ii) 1.5 x 10^{11} or iii) 1.1 x 10^{11} viral particles.

For objective four, I showed that assuming a peaking dose-efficacy curve or using weighted model averaging was typically preferable to assuming a saturating dose-efficacy curve for the purpose of selecting optimal dose. Adaptive trial design may not typically improve dose selection relative to a 'sufficiently explorative' trial design but may lead to trial participants receiving more optimal doses.

For objective five, I found that the non-parametric model was able to optimise dose well despite using a simpler set of assumptions than the parametric models, in particular for multi-dimensional dose-optimisation problems. Using mathematical model-based and/or adaptive design-based approaches of vaccine dose optimisation consistently selected a more optimal dose using less trial participants than using neither mathematical modelling nor adaptive design.

Discussion

In this work I explored and expanded the field of IS/ID and mathematical modelling for vaccine dose optimisation.

I collated and summarised adenoviral vector vaccine data that can now be used for quantitative adenoviral vectored vaccine dose optimisation analysis. There is evidence from these dose-ranging trial data to support the hypothesis that for some adenoviral vector vaccines the dose-efficacy response was peaking, so vaccine adenoviral vector developers should not assume that increasing dose always leads to more efficacious vaccine response.

I have shown that the 'optimal dose' predicted by modelling is likely to depend on which utility function is used to define 'optimal', so vaccine developers should have a clear definition of optimal dose prior to conducting clinical trials. I have also shown that adaptive trial design informed by mathematical modelling can be used both to

6

improve vaccine dose selection and benefit clinical trial participants, suggesting that vaccine developers could consider adaptive trial design.

Additionally, I showed that using a weighted average of peaking and saturating models to describe the dose-efficacy response may be beneficial, suggesting that vaccine developers should consider using a weighted average of peaking and saturating models to describe dose-efficacy response. However, my results also suggested that a non-parametric model was able to optimise dose at least well as using parametric models, despite using a simpler set of assumptions, in particular for multi-dimensional dose-optimisation problems, suggesting that vaccine developers could consider using non-parametric models as an alternative.

Finally, my results showed that using mathematical modelling and/or adaptive trial design reduced the number of trial participants required to find an optimal vaccine dose when compared to using neither modelling nor adaptive design, and so vaccine developers should consider using modelling and/or adaptive design in vaccine dose-finding trials to increase efficiency.

I believe there is merit to continued development and validation of IS/ID methods in order to provide tools for identifying optimal vaccine dose, ultimately saving lives.

Contents

Acknowledgements

Fifthly, I thank my family members: my mother, my father, my sister Sophie, my grandmother, and especially my partner Charlotte. You have all have shown a willingness to read and discuss vaccine dose-response modelling beyond any level of reasonable expectation. Your support has been there when it has been needed most over these years.

Abbreviations

%CI	**Percentage Confidence Interval**
ΔAIC	**Difference in AIC**
Ad5	**Adenovirus Serotype 5**
AIC	**Akaike Information Criterion**
AICc	**corrected Akaike Information Criterion**
anti-HA	**anti-hyaluronic acid**
AR	**Average Regret**
BBSRC	**Biotechnology and Biological Sciences Research Council**
BCG	**Bacillus Calmette–Guérin vaccine**
BIC	**Bayes Information Criterion**
BOD	**Bayesian Optimal Design (Dose Optimisation Approach)**
BOD	**Biologically Optimal Dose**
CCBP	**Continuous Correlated Beta Process**
CD4	**Cluster of Differentiation 4**
CD8	**Cluster of Differentiation 8**
ChAd	**Chimp Adenovirus**
ChAdOx1/ChAdOx2	**Chimp Adenovirus Oxford 1/2**
CoBe	**Correlated Beta (Dose Optimisation Approach)**
COVID	**Coronavirus Disease**
CRM	**Continual Reassessment Method**
DALY	**Disability Adjusted Life Years**
DLT	**Dose Limiting Toxicity**
DOA	**Dose Optimisation Approach**
ELISPOT	**Enzyme-linked immunosorbent spot**
EMA	**European Medicines Agency**

GBP	**British pound sterling**
HPIV	**Human Parainfluenza Virus**
IA	**Inaccuracy**
IAV	**Influenza A virus**
IFN-y	**Interferon gamma**
IL 17	**Interleukin 17**
IL 2	**Interleukin 2**
IM	**Intramuscular**
IS/ID	**Immunostimulation/Immunodynamic**
LHS	**Left Hand Side**
LIDo	**London Interdisciplinary Doctoral programme**
LSHTM	**London School of Hygiene and Tropical Medicine**
MBDD	**Model-based drug development**
ml	**millilitre**
MOP	**Multi-objective Optimisation Problem**
MRNA	**Messenger RNA (Ribonucleic acid)**
MSE	**Mean squared error**
MTD	**Maximally Tolerated Dose**
n.a	**Not Applicable**
OP	**Optimisation Problem**
PAR	**Percentage Average Regret**
PD	**Pharmacodynamics**
PDF	**Probability Density Function**
PFU	**Plaque Forming Unit**
PK	**Pharmacokinetics**
PK/PD	**Pharmacokinetics/Pharmacodynamics**
PSR	**Percentage Simple Regret**
PU	**Particle Unit**

14

QALY	**Quality Adjusted Life Years**
RHS	**Right Hand Side**
RMSE	**Root Mean Squared Error**
RoA	**Route of Administration**
SARS-CoV-2	**Severe Adverse Respiratory Syndrome Coronavirus 2**
SBC	**Smooth Binary Correlation**
s.d	**Standard Deviation**
SQ	**Subcutaneous**
SR	**Simple regret**
SS	**Sum of Square**
SSE	**Sum of Squared Error**
STEIN	**Simple Toxicity and Efficacy Interval Design**
TB	**Tuberculosis**
TNF	**Tumour-Necrosis Factor**
UK	**United Kingdom**
US (or USA)	**United States of America**
USD	**United State Dollars**
VP	**Viral Particle**
w.r.t	**with respect to**
YLL	**Years of Lost Life**

List of Tables

The tables in this thesis follow three numbering systems, those written into the thesis body (which are numbered following the chapter they feature in), and tables in the main papers and supplementary material (labelled with S preceding the number) referenced in the papers.

List of Figures

The figures in this thesis follow three numbering systems, those written into the thesis body (which are numbered following the chapter they feature in), and figures in the main papers and appendix/supplementary material (labelled with an A or S preceding the number) referenced in the papers.

23

S7.1	Mean true efficacy at the predicted optimal dose and mean cumulative sum of efficacy against trial size for all seven objective S7.1 scenarios.	299
S7.2	Mean true efficacy at the predicted optimal dose and mean cumulative sum of efficacy against trial size for all two objective S7.2 scenarios.	304
S7.3	Mean true efficacy at the predicted optimal dose and mean cumulative sum of efficacy against trial size for all seven objective S7.3 scenarios.	306
S7.4	Mean true efficacy at the predicted optimal dose and mean cumulative sum of efficacy against trial size for all seven objective S7.4 scenarios.	307
S7.5	Mean true efficacy at the predicted optimal dose and mean cumulative sum of efficacy against trial size for all seven objective S7.5 scenarios.	310
S7.6	Mean true efficacy at the predicted optimal dose and mean cumulative sum of efficacy against trial size for all seven objective S7.6 scenarios.	314

Chapter 1. Background and thesis overview

Immunostimulation/Immunodynamic (IS/ID) modelling is a new field that aims to allow vaccine developers to better select optimal vaccine doses through the utilisation of mathematical modelling. Towards this aim, IS/ID modellers must consider not only mathematical modelling techniques, but also how these modelling techniques can best be applied to find optimal vaccine dose, and indeed what is meant by 'optimal' vaccine dose.

In this chapter, I will describe vaccines and vaccination, along with the importance of selecting optimal vaccine doses. I will then provide a narrative review of mathematical modelling ideas that may be relevant to vaccine dose selection, drawing inspiration from the field of model-based drug development. This will be followed by the thesis rationale, thesis aims and objectives, and then the thesis overview. Not all of the topics of this chapter are explicitly investigated within the research papers of this work. Similarly, this does not represent an exhaustive review of all possible considerations in modelling dose-response or dose-optimisation. However, these topics were included to provide a holistic overview of the concepts that are relevant to mathematical dose-response modelling and vaccine dose selection. In this chapter I hope to provide a background that highlights the research gaps I aimed to fill, the wealth of literature that exists outside of the field of IS/ID modelling, and the relevance of these gaps within the continued development of the field of IS/ID modelling.

Vaccination and difficulty in finding optimal vaccine dose

In this section, I will discuss vaccination and difficulty in finding optimal vaccine dose. I will discuss the history and present state of vaccination globally, then highlight the vaccine development pathway. In particular, I will focus on the selection of vaccine dose as part of this pathway, including a discussion of factors that developers may need to consider when selecting optimal vaccine dose. These include: prime/boost administration vaccines, antigen/adjuvant administration vaccines, immunological surrogates and correlates of protection, and vaccine toxicological profile.

A brief history of vaccination

Vaccination is an important and highly relevant tool in global disease control and eradication. Whilst variolation was reportedly practised prior, the work of Edward Jenner in 1796 is often considered the foundation of modern vaccination[1]. This early work involved gathering pus from the wounds of individuals recovering from cowpox and then administering this matter subcutaneously to healthy individuals. The individuals that were inoculated in this way were bestowed a degree of protection against fatal smallpox infection.

This early work has been expanded upon and has led to vaccination becoming a widespread public health measure globally, with the United Kingdom's recommended childhood immunisation schedule including vaccination against eight diseases prior to the first birthday [2]. There now exist many vaccine platforms that have been investigated or seen clinical use [3–5]. Classically these platforms used either inactivated or attenuated whole virus or bacteria, or protein/virus-like subunits that are co-administered with an adjuvant. Next-generation vaccine development has also considered viral vector and nucleic acid platform vaccines, which are hoped to better induce cellular immunity in vaccinated individuals[6]. Vaccines can also be used in both prophylactic and therapeutic settings, which respectively aim to prevent infection or disease in presently healthy individuals and alleviate symptoms in presently diseased individuals [5].

Such implementation of vaccination programmes has led to the eradication or near eradication of smallpox, rinderpest, and polio in the modern world [7]. Despite this, due to public hesitance around vaccines and vaccination and the burden of care, there is still a great need to ensure that vaccines are developed in such a way that the vaccines that are selected for wide-scale implementation are not only effective in offering protection but also safe and affordable [8].

Vaccine development and dosing

Vaccine development

The vaccine development pathway is the standardised route for vaccines to go from initial research to clinical use. This can be divided into the exploratory, pre-clinical, clinical, regulatory, manufacturing, and quality control stages, after which the vaccine can be marketed/distributed [9–12]. The clinical stage is divided into phases, phase I, phase II and phase III, though phase II is often divided into phase IIa and IIb and sometimes clinical trials are hybrids of these phases [figure 1]. Pre-clinical trials are often animal models, where efficacy and safety in some animal species such as mice or non-human primates is assumed to be indicative that the vaccine may be safe and effective in humans. These may also inform the doses that will be used in phase I trials. Phase I-III trials are conducted on human participants and increase in size as the vaccine development pathway is furthered, from n=20-80 in phase I trials [13] to potentially n>1000 in phase III trials [14].

	Exploratory	Pre-clinical	Phase I	Phase II	Phase III	Regulatory/ Manufacturing/ Quality Control	Marketing/ Distribution	Phase IV
			Clinical Trials					
Relevance to dosing		Animal dose-response	Human dose-response and selection of human vaccine dose					Selection of dose for special populations
Purposes and notes	Basic research and identification of substances that may have prophylactic or therapeutic use	Show that the potential vaccine is immunogenic or efficacious and is safe in an animal model.	Show that the potential vaccine may induce immunogenicity and is safe response in human subjects. Identify promising doses/route of delivery. Often unblinded.	Identify vaccine immunogenicity and safety to a statistically relevant degree of accuracy. Identify promising doses. Often blinded	Provide evidence that the potential vaccine is efficacious and safe at a population level.	Get approval from regulatory bodies for vaccine marketing and rollout. Upscale vaccine production to mass scales. Prove that mass produced vaccine is equivalent to small-scale produced vaccine	Distribution of vaccine for the purpose of disease burden reduction. Surveillance or vaccine efficacy and adverse events	Evaluate for side-effects that were not observed in phase I-III trials. Show that the vaccine is safe and effective and selecting optimal dose for special populations and
Typical duration and total development time	2-5 years 2-5 years	2 years 4-7 years	2 years 6-9 years	2-3 years 8-12 years	5-10 years 13-22 years	2 years 15-22 years	Continuous for vaccine life-span	

Figure 1.1. Overview of the stages of the clinical pathway for vaccine development.
Market/distribution and phase IV stages are separated as these may not always be considered as part of the vaccine development pathway.

Phase I trials aim to determine whether a vaccine can cause immunogenicity and is safe at some dose level. These trials may be used to make some decisions regarding dose and mode/schedule of administration. Phase II trials are expected to have sufficient statistical power to provide clinically meaningful estimates of the immunogenicity and toxicity. Phase II may also be used to choose the dose that will be continued on into phase III trials and potentially clinical use. Phase III vaccine

27

trials aim to show that a vaccine at the chosen dose is safe, immunogenic, and protective at a population level. Phase IV trials are also sometimes conducted after a vaccine has already been made available to market and are either used to investigate vaccine effectiveness in special populations or to evaluate the vaccine for side-effects that were not observed in phase I-III trials[1] [15]. Preclinical and phase I-II trials are typically the phases that aim to determine a dose that is effective and safe and will be of most relevance to the scope of this thesis.

Whilst these steps are important to ensure that undue resources are not wasted on investigating vaccines that are unlikely to be suitable for wide scale implementations, this development pathway leads to vaccine development being expensive and time consuming, often costing over $300 million and taking over 10 years to develop a vaccine from preclinical development to clinical use [16,17]. Whilst there has been great effort made to accelerate vaccine development in response to the COVID-19 global health crisis, there is still much work needed to ensure safe and effective vaccination is achieved at shorter time scales [18].

Dose selection in clinical vaccine development

One of the key questions in preclinical and phase I-II clinical trials is the amount of vaccine that should be administered. This can be measured in volume, viral particles, infectious units, or various other units, but will be called 'dose' in this work. Many considerations exist regarding what is meant by optimal vaccine dose, and I will discuss some of these throughout this work. However, there are in brief three factors that I would consider to be most important: efficacy, safety/toxicity, and cost/dose-sparing.

Firstly, there is of course the need for the vaccine to be dosed in such a way that it is efficacious. This may mean protection in the case of prophylactic vaccines, or it may mean that the vaccine is capable of treating some disease in the case of therapeutic

---

[1] Phase IV trials to determine vaccine efficacy or toxicology in special populations may be relevant for the future development of IS/ID modelling to refine optimal vaccine dose for older or immunocompromised populations but will not be discussed in this work.

vaccines. There would likely be no reason to administer a vaccine that is not efficacious for the purpose of preventing or treating some disease.

Secondly, vaccine dose must be selected such that the vaccine is 'safe'. This is to say that the risk of adverse events is justified by the vaccine's effectiveness or that the vaccine induces toxicity below some regulatory defined threshold [19]. Given that vaccines may be prophylactic, it is often preferable that only a small proportion of vaccinated individuals should experience adverse events that prohibit their daily life. Toxicity in vaccine dose-ranging trials has historically been low for this reason [20], as developers may be unlikely to test doses with a high risk of high-grade toxicity.

Finally, vaccine dose must be selected with dose sparing in mind. This may be to reduce the cost of a population wide vaccine rollout, or to allow for a wider vaccine rollout in the case that there be limited capacity for vaccine manufacture. In particular, for the 17DD yellow fever vaccine it was found that a tenfold in dose decrease relative to the approved dose was similarly protective and would have allowed to a greater number of vaccinations and saved lives if that reduced dose had been used [21]. These and a myriad of other factors may be relevant in vaccine dose selection.

The actual selection of vaccine dose in preclinical and phase I-II is typically empirical. A number of different dosing groups are suggested, trial individuals are divided among these groups, and these individuals receive the dose assigned to their group. Various immunological readouts and incidence of adverse events are recorded for each group. Statistical analysis for significant differences between dosing groups is conducted, and then clinicians use the resulting data and statistical analysis to select which dosing group(s) had the response that they believe is most likely to be optimal. Optimal may be defined in this case as the dose that maximises/minimises the observed immunological/toxicological profile, or as any dose for which the lower confidence bound for the posterior prediction of efficacy is above some previously defined threshold [22]. This dose or doses are then carried forward for investigation further down the vaccine development pathway, or development of the vaccine may be discontinued if none of the doses are considered to have a favourable efficacy and toxicity profile.

It has been suggested that this empirical dose selection methods may have historically caused suboptimal doses to be selected for clinical use, for example the 17DD yellow fever vaccine mentioned above. It has also been suggested that this empirical procedure (sometimes called 'data-analytic' or 'hypothesis testing' approaches) requires more individuals to identify optimal dose relative to modern dose-selection methods that have been developed to select optimal dose in drug development [23–25].

Prime and boost dosing

When conducting pre-clinical and clinical trials and selecting optimal dose for a vaccine, developers may need to consider vaccines that are administered using the 'prime-boost' paradigm. Whereas the 'single-administration' paradigm involves giving a single dose of a vaccine, the 'prime-boost' paradigm involves giving two or more vaccine doses, with the first called the 'prime' and latter doses being called 'boost' doses[26,27]. These may be administrations of the same vaccine, called 'homologous' prime-boosting, or of different vaccines, called 'heterologous' prime boosting. For both homologous and heterologous prime-boost vaccines, vaccine developers must choose a dose for each administration.

There has been discussion about how best to dose vaccines using a prime-boost paradigm. Clearly, a dose for each prime/boost must be selected, but this may not be the dose that would be safe/effective if that prime/boost dose were given as a single administration. It has been suggested that a decreased 'prime' dose may allow for a greater immunological response following the boost dose[2], meaning that these dose choices cannot be made independently of each other. Further, there is also the question of the time between administrations. It is possible that increasing or decreasing the time between administrations may lead to changes in the immunogenicity/toxicity that is observed [28,29].

---

[2] One theorised mechanism behind this effect is that the smaller prime dose reduces antigen availability in germinal centres, which in turn means that only naive immune cells with high affinity to the antigen would be activated. When boosted, these high affinity activated immune cells proliferate to a high degree, resulting in a large cohort of high affinity immune cells and improved protective efficacy.

Vaccine and adjuvant dosing

As noted above, there are a number of vaccine platforms that have been investigated and seen clinical use. One of the classical platforms are 'antigen/adjuvant' vaccines. Here the vaccine consists of some antigenic material (typically a protein or virus-like subunit) and an adjuvant.

The antigen contains the genetic material that allows the adaptive immune system to develop specific responses to the pathogen of interest but is not typically capable of stimulating an immune response [30]. An adjuvant is therefore given at the same time as the antigen, functioning as the necessary delivery or immunostimulation molecules to allow shuttling of antigen to the primary lymphatic organs and hence initiate the proliferation of specific adaptive immune response. Historical adjuvants include aluminium salts, oil-water emulsions, and lipid vesicles adjuvants. Therefore, like for prime-boost paradigm vaccines, there is the need to choose multiple different doses, one each for adjuvant and antigen. These are co-administered, and so the time between administrations does not need to be chosen. Both antigen and adjuvant dose may be important for ensuring that a vaccine is both safe and effective [31].

Surrogates and correlates of protection in vaccine development

I noted above that a chosen vaccine dose should be effective. In the case of therapeutic vaccines, this may be obviously measurable as reduction in symptoms or adverse disease outcomes. However, in the case of prophylactic vaccines, effectiveness may be defined by prevention of future infection, symptoms, hospitalisation and/or death in individuals who are presently healthy. In order to determine whether prophylactic vaccines are effective, typically either challenge studies or large phase III population level trials must be conducted [32]. Challenge studies involve exposing individuals to potential infection, which may not be ethically justifiable [33], and large population level trials are expensive and time consuming [34]. Therefore, in order to select dose it is common to use correlates or surrogates of protection [35].

A correlate of protection is some immunological response that can be measured after vaccine administration, and that it is assumed an increase in that response reflects an increase in the protection that the vaccine offers. It has previously been suggested that it may be possible to directly predict probability of protection using these immunological correlates of protection, however data to allow for such prediction typically requires challenge or population studies [36], which I have previously stated may not be feasible. Instead, sometimes a 'surrogate of protection' may be used. This is some threshold for which, if a vaccinated individual had an immune response in excess of this threshold, they would be considered protected, and otherwise they would be considered unprotected. An example surrogate of protection is used in considering influenza vaccines, with individuals with a post-vaccination anti-HA antibody titre greater than 40 being considered 'seroprotected' [37]. The dose which best maximises these correlates/surrogates of protection may be predicted to be considered to be most effective/protective.

There have been conflicting results regarding whether common correlates and surrogates of protection are effective targets when attempting to maximise vaccine protection. For example, the works of Khoury [38] and Gilbert [39] found that fold-increase in antibody titre was effective as both a correlate and surrogate of vaccine efficacy/protection. Later work by Gilbert [40] however claimed that fold increase in antibody titre may not be a 'mechanistic' correlate of protection, which is to say that an increased antibody titre was not causal of greater protection, and that there may be some unmeasured variable that causes both an increase in antibody titre and increase in protection. Regardless of this, correlates and surrogates of protection are still commonly used is practice when selecting vaccine dose.

Toxicity and vaccine safety

It is possible that a vaccinated individual may experience adverse events that they may not have experienced if they had not been vaccinated. This is referred to as vaccine associated toxicity or just 'toxicity', and vaccine safety involves reducing the

risk of vaccine toxicity[3]. Typically, vaccine related adverse events are acute and non-life threatening, requiring at most mild pain-relief and disruption of daily activity that lasts less than a week [41]. Whilst serious toxicity has been associated with some vaccines [42], these are typically rare and may be unlikely to occur during smaller vaccine clinical trials. A serious adverse event that involves hospitalisation or death during a vaccine clinical trial may require the stopping or pausing of that clinical trial until evidence to regarding whether this was associated with the vaccination can be determined [43]. When serious adverse events have previously occurred during a clinical vaccine trial, it has often been found that the serious adverse event was not related to the vaccine [20]. Regardless of the minimal risk of serious vaccine toxicity, the reduction of vaccine toxicity is still considered important, both to meet regulatory requirements set in place by governing bodies or to reduce vaccine hesitancy and increase vaccine uptake. Therefore, it is important to choose a dose with vaccine safety in mind.

A narrative review of mathematical dose decision making in vaccines and drugs

In the previous section we highlighted that vaccine dose is a crucial decision in accelerating vaccine development whilst ensuring that effective, safe, and affordable dosing is achieved. It may be that the empirical method of dose selection that is presently used for vaccine dose selection could be improved. In this section we discuss other approaches that could be used to select optimal vaccine dose, with a focus on introducing mathematical modelling techniques used in the drug development process.

Applying techniques from drug dose decision making to vaccines

Mathematical model-based drug development (MBDD) and pharmacokinetic/pharmacodynamic modelling (PK/PD) have been well established for making decisions relating to drug dosing. There exists a vast literature and a multitude of methodologies that have previously been suggested, evaluated, and

---

[3] Risk may be a factor of probability and/or severity of these adverse events.

used to accelerate and improve the decisions processes in drug development [24,44–47]. Recently the field of vaccine IS/ID has been suggested [23], which aims to be an equivalent to MBDD and PK/PD methodologies in the vaccine development setting through adapting techniques already in the MBDD and PK/PD literature. However, whilst there have been a number of works that explore these methodologies [48–52], IS/ID is still novel and not commonly used as part of practical vaccine development.

Whilst MBDD and PK/PD have seen effective practical use in drug dose decision making, these techniques may need to be adapted for the vaccine setting. For example, many PK/PD mathematical models are derived through the assumption that drug kinetics and response follow 'mass action laws' [53], which may not be a reasonable assumption for vaccines. Additionally, in drugs it is typically reasonable to assume that an increase in dose leads to either an increased or equivalent efficacy and toxicity, the so-called 'saturating' dose-response curve [54]. This may not be reasonable in vaccines. Recent work showed that for both the TB vaccine H56+IC31 and IAV/HPIV influenza inoculation the relationship between dose and immunogenicity/protection was 'peaking', meaning there was some dose that maximised immunogenicity/protection and that increasing the dose beyond this value only decreased effectiveness [49,50]. Thus, many mathematical models and techniques that have been effective in PK/PD may not be reasonable in vaccine development.

Despite this, the potential for accelerated, quantitatively informed vaccine development or improvements in vaccine efficacy, safety and affordability means that IS/ID and mathematical modelling for vaccine dose selection warrant further investigation. The remainder of this section will give description of prior techniques that have been used in drug and vaccine dose selection, drawing from drug literature where there has been little previous investigation in vaccines.

Mathematical models

In this work 'mathematical models' refer to equations or systems of equations that can be used to describe the relationship between dose and a response variable of

interest. Typically, this will be either a correlate of protection, probability of a surrogate of protection, probability of an adverse event occurring, or cost. These are typically functions of dose with some parameters that are either known and fixed to a value using pre-existing literature or must be estimated using clinical trial data. Broadly speaking there are two types of mathematical model that are used in MBDD and IS/ID modelling; mechanistic and statistical [55]. These are also sometimes referred to respectively as 'compartmental' and 'noncompartmental' models in MBDD. Statistical models have also been called 'empirical models', but this name is not used in this thesis due to potential confusion with traditional empirical vaccine dose selection.

Statistical versus mechanistic models

Mechanistic models leverage knowledge of the underlying biology to describe vaccine dose-response. Such models have been used for modelling of both vaccine [49,50,56] and drug [57–59] dose-response and are sometimes called 'Quantitative Systems Pharmacology' models. These typically are composed of either systems of nonlinear differential equations [49] or use agent-based models [60–62], and these models are typically used to describe longitudinal data. These models may benefit from leveraging biological knowledge to allow extrapolation [63], and may benefit from values for some model parameters already being known from prior studies. These two factors combine to make mechanistic modelling very useful for including sub-population analysis [64–66]. The downside of mechanistic modelling is that it requires knowledge of the underlying biological processes that determine dose-response. It may not always be the case that such knowledge is available, for example many different mechanistic models of CD8+ T Cell proliferation have been suggested [67–70], and which may lead to different dose decisions depending on which model is assumed to be true. Additionally, underlying biological processes may vary for different immunological responses (CD4+ T cell, CD8+ T cell, B Cell, or antibodies as examples), or for different vaccine platforms (MRNA, Adjuvant, Viral Vector). This may complicate the modelling process. An additional weakness of these models is that they require large amounts of data at multiple time points in

order to allow for parameter estimation. For example, Tam et al. found that for a relatively simple[4] mechanistic model of rifampicin PK in mice, data would have to be collected at eight time points for dosing groups defined by 11 different dosing strategies in order to estimate the model parameters with sufficient accuracy [71].

The other type of modelling that has commonly been used is statistical modelling. Here modellers do not aim to explicitly describe the underlying biology, and instead a simpler set of assumptions about the general shape of the dose-response curve are used. These statistical models typically have fewer parameters than are used for mechanistic models, and thus may require less data to estimate model parameters than mechanistic models and also have less risk of overfitting to small data sets [72]. As these models only assume a certain shape and make minimal assumptions about the underlying biology, they are often able to be applied quite generally across different classes of drugs. However, extrapolation may be less valid than it would be if an appropriate mechanistic model was used [55,73,74] as these models do not consider the casual immunological pathways behind vaccine response. Hence, they are not limited to predicting biologically feasible dose-response curves, therefore extrapolated predictions are neither justified by data nor biological theory. Statistical models are less commonly used for describing longitudinal data but are commonly used in MBDD for describing the relationship between dose and toxicity.

Recently a third class of models has been suggested as potentially useful for selecting optimal dose. These are called 'non-parametric' or 'curve-free' models [75–78]. Here neither assumptions regarding underlying biology nor dose-response curve shape are made. Instead, simpler assumptions are made, such as that a small change in dose should result in a small change in response. Non-parametric modelling has been suggested as beneficial for model dose-response when a saturating dose-response curve cannot be assumed, as may be the case for vaccines, and so these may merit investigation for potential use in the future development of IS/ID modelling.

---

[4] A two-compartmental model with 9 parameters describing clearance and absorption.

For all types of models, these can be 'fit' or 'calibrated' to observed clinical data from dose-ranging trials, which allows modellers to predict dose-response mathematically.

In chapters 2 through 6 I considered only statistical and non-parametric models, as these may be more widely applicable in cases where the underlying immunodynamics are less well known and may require less data. For simplicity, when I refer to statistical modelling, I also include non-parametric modelling unless otherwise specified. A description of many potential statistical dose-response models is given in Appendix A.A.

Modelling dose-efficacy

Vaccines must be effective at protection in the case of prophylactic vaccines or treatment in the case of therapeutic vaccines, and that this efficacy may be defined by many clinical endpoints. Reduction in transmission, symptoms, hospitalisation, and death are all clinical endpoints that may be of interest. Broadly we may consider these to be binary outcomes, either a vaccine reduces the risk of transmission, symptoms, hospitalisation, or death, or it does not. However, as stated before, clinical trials measuring efficacy directly may not be possible. Instead, we can use surrogates or correlates of protection. In this section I consider how dose-efficacy may be modelled either directly through surrogates of protections or indirectly through a combination of dose-immunogenicity and immunogenicity-efficacy models.

Modelling of vaccine dose-efficacy

If a surrogate of protection is available, then it is possible to approximate the modelling of dose-efficacy by modelling the relationship between dose and the proportion of individuals for which the immunological surrogate of protection is achieved. Such a binary surrogate of protection may involve humoral immune responses, cellular immune responses, or some combination of immune responses. Historically, modelling of dose-efficacy has been conducted using statistical modelling techniques in MBDD, where efficacy is often directly measurable [79]. Statistical models such as the sigmoid saturating function have commonly been used to model dose-efficacy in drugs [80].

Modelling of vaccine dose-immunogenicity

If a surrogate of protection is not available, then instead dose-immunogenicity may be modelled to predict what the expected value for a correlate of protection would be for a given dose. Either statistical or mechanistic models can be used for this purpose. For mechanistic modelling of dose-immunogenicity, we can see examples in the works of Handel et al., Moore et al., Mayer et al., Zarnitsyna et al. and Farhang-Sardroodi et al. [49,67–69,81]. Handel used a mechanistic model to describe the impact of dose on generation of humoral immune response IAV/HPIV influenza inoculation. Moore et al., Mayer et al., and Zarnitsyna al. considered mechanistic models of CD8+ T Cell response to yellow fever vaccine, listeria monocytogene infection and influenza A infection respectively, though only Mayer et al. discuss using their model to predict optimal vaccine dose. Farhang-Sardroodi considered mechanistic models of cellular, humoral and cytokine responses to the COVID-19 ChAdOx1-S vaccine, though they do not explicitly discuss the relationship between dose and immunogenicity. Rhodes used statistical modelling techniques to describe the dose-immunogenicity of IFN-y T cells in mice when administered the H56+IC31 tuberculosis vaccine [50].

Modelling of immunogenicity-efficacy

Modelling of vaccine dose-immunogenicity on its own may be reasonable if the definition of optimal dose is defined strictly as the dose that maximises some immunological correlate of protection. However, it may be possible to also model the relationship between immunogenicity and efficacy. This has been suggested by Dudášová et al., who used a statistical modelling approach to describe the relationship between specific antibody titre and protection from influenza, dengue fever, and herpes zoster, and the implications that this could have for vaccine decisions making [36]. Again, building such a model requires data from challenge or population level trials that gather both immunological and clinical outcome data, which may not be practical or ethical, as was discussed in the section 'Correlates and surrogates of protection'.

Such models may be beneficial as they can be used to predict the potential clinical outcome benefit that results from some dose-dependent increase in immunogenic

response. If both dose-immunogenicity and immunogenicity-efficacy models are available, then these could be combined to form a two-stage dose-efficacy model. This two-stage modelling approach was used by Handel et al.[49], but due to the aforementioned difficulty in parameterising immunogenicity-efficacy models they were forced to choose an arbitrary parameterisation for that model. The work of Kumbhari et al. [82] also combined both dose-immunogenicity and immunogenicity-efficacy modelling to predict a dose-efficacy curve using a complex mechanistic model of peptides, dendritic cells, Helper T cells and Cytotoxic T cells. This was however for a therapeutic cancer vaccine, so an efficacy outcome 'percentage of cancer cells killed' could be measured directly. As stated, this may not be feasible for prophylactic vaccines.

Modelling dose-toxicity

Modelling of dose-toxicity has been relevant in MBDD, where statistical models of dose-toxicity are commonly used. Toxicity is typically defined by either binary or ordinal outcome variables.

Binary toxicity response

A common method of describing toxicity in vaccines of drugs and then modelling the dose-toxicity relationship is to consider toxicity as a binary outcome measure. Different toxicity outcomes may be of interest, for example pain, nausea, or inflammation, but toxicity is often considered to have occurred for a vaccine recipient if any of these potential outcomes exceeded some level. If any of these toxicity outcomes exceeds some thresholds set by clinicians or regulators, then that individual can be said to have experienced a toxicity response. It is typically assumed that the vaccine dose-toxicity curve is saturating, and that therefore a sigmoid saturating model is often appropriate [83]. Mandreker et al. [79] discussed application of a trivariate model of binary outcomes for both efficacy and toxicity with potential for application for both vaccines and drug dose selection.

Ordinal toxicity response

Ordinal grading systems of toxicity have been suggested for use in both drugs and vaccines. These ordinal grading systems typically classify toxicity outcomes in

increasing order of severity. For example, these gradings might be; None, Mild, Moderate, Severe, Serious [84]. A numeric rating of toxicity may also be given, with grades 0 through 4 corresponding respectively to grades None through serious. Grade 5 (death) may also be considered, though this would likely be an unacceptable vaccine outcome. For each of these gradings, different prevalence may be acceptable. Models have been suggested for this for the purpose of selecting optimal drug dose when attempting to minimise ordinal measures of toxicity [85]. Two such models are the probit model and proportional odds model.

Maximum tolerated dose

In cases where dose-efficacy can typically be assumed to be saturating, a common technique in MBDD is to use the so-called 'Maximum tolerated dose' [86]. Here, a threshold on the prevalence of some adverse event is set, then the purpose of dose-finding studies is to find the maximum dose for which the prevalence of toxicity is less than this threshold [87]. This is the maximum tolerated dose (MTD) and has typically been suggested as the dose that should then be continued onto further research and clinical use. It has been shown repeatedly through simulation studies that using statistical models of dose-toxicity can reduce the number of trial participants that are required to find the MTD [88]. The use of an MTD may not be appropriate when dose-efficacy cannot be assumed to be saturating, or if efficacy can be assumed to have saturated for a dose that is less than the maximally tolerated dose. In that case, a multi-objective optimisation approach is preferred [89,90], aiming to find the dose that best balances efficacy against toxicity (in some literature called the 'optimum biological dose'). Multi-objective optimisation will be discussed in detail later in this chapter.

Modelling dose-cost

Previous work related to optimal vaccine dose selection has highlighted economic considerations that arise in selecting optimal vaccine dose. Typically, the models and parameters associated with dose-cost relationships can be calculated explicitly, and do not rely on clinical trials.

I consider that the economics involved in vaccine dose selection can be broadly described in two major economic settings for public rollout of vaccine, the limited supply, and the unlimited supply setting.

In the limited supply setting, there is only some limited supply of vaccine that is available as part of a vaccination campaign, and this must be used efficiently to maximise disease burden reduction. For example, say that there is only sufficient vaccine supply to vaccinate 50% of a population with a standard dose. If a standard dose is 90% protective against symptomatic disease, and a half dose is 80% protective, it may be preferable at a population level to roll out the vaccine using the half-dose [91]. Epidemiological modelling has shown that fractioning of dose may be beneficial when vaccine supply is limited [92].

In the unlimited supply setting, there is capacity to produce sufficient vaccine supply for an entire population of interest. Reducing or increasing vaccine dose will not alter the proportion of the population that is vaccinated, however increasing the dose would lead to a vaccination campaign being more expensive. Where a vaccine campaign is government funded, taxpayers may only have a certain 'willingness-to-pay for such a campaign, and thus a dose must be selected that is cognizant of the willingness-to-pay' per year of lost life or disability adjusted life year that is averted by the vaccine campaign [93]. Where vaccination is privately funded, increased cost may reduce the number of individuals that are willing to pay for a vaccine, therefore reducing the potential benefit that the vaccine will have [94].

In both the limited supply and unlimited supply settings dose has economic implications that may impact public opinion towards, uptake of vaccines, and overall public health benefit.

When considering economic factors of dose, it is important to note that there will be costs associated with a vaccination campaign that are not dose-dependent, so a halving of dose would not cause a halving of the costs associated with the vaccine campaign. Du et al. estimated that a 50% reduction in standard dose of the COVID-19 ChAdOx1-S vaccine would reduce the cost of vaccine administration per person from $12.00 to $10.50, which would only be a 12.5% reduction in cost [93].

41

Multi-objective optimisation and utility functions

I have argued that dose is a crucial decision in vaccine development, with dose impacting efficacy, toxicity, and cost. Ideally a dose should be chosen that maximises efficacy/immunogenicity, minimises risk of vaccine toxicity, and minimises cost. There may also exist other potential factors that must be minimised/maximised when choosing an optimal vaccine dose. It may be possible that there does not exist a single dose that is optimal for all of these objectives, and hence we would have to consider vaccine dosing as a multi-objective optimisation[5] problem. Using modelling to predict optimal solutions to multi-objective optimisation problems is commonly used in many fields [95].

Multi-objective optimisation is a problem that occurs whenever there are multiple objectives that need to be optimised. Two potential methods of solving such problems are to consider either 'Pareto fronts' or 'utility functions'. When these problems are non-trivial (there does not exist a single solution/dose that is optimal for all objectives), the concept of Pareto optimality must be considered [96]. A dose would be considered Pareto optimal if, for all other solutions/doses, an improvement with regards to each objective function could only be achieved by deteriorating performance with respect to at least one other objective function. The set of all such solutions is called the Pareto front, and any solution along this front could be considered optimal.

With regards to drug or vaccine multi-objective single-administration dose-optimisation where efficacy/immunogenicity must be maximised, toxicity must be minimised, and efficacy and toxicity are assumed to be saturating with respect to dose, all doses would lie on this Pareto front [appendix A.B.]. This means that all doses could be considered optimal, so consideration of the Pareto front may not bring useful insight with regards to vaccine dose selection.

---

[5] 'Multi-objective optimisation', 'multi-factorial optimisation', and 'multi-factor optimisation' are all used interchangeably, both within this work and other literature.

The utility function method is to use some function that maps the multiple objective function values (probability of efficacy/immunogenicity, probability of toxicity, cost, etc.) onto a single numerical value. The dose or doses that maximise this value are considered optimal[6]. Utility functions have been commonly used in drug dose-optimisation and other multi-objective optimisation problems outside of dosing [97–99]. Model predictions in combination with utility functions can be used to select an optimal dose.

Choosing which utility function will be used is an important question, as the utility function should include any factors that vaccine developers consider relevant when deciding dose (efficacy, toxicity, etc,) and should align closely with vaccine developers' intuitions regarding the relative importance of these objectives. Choosing the form and parameters of utility functions is suggested to involve modellers eliciting information from vaccine developers, clinicians, and stakeholders, and building a utility function that reflects this information. Thall and Cook [100] and Branke [101] both suggest methods of calculating utility function parameters through stakeholder elicitation, though Peterson [102] highlights that such methods may have faults, namely assuming that stakeholders have preferences that are self-consistent and obey the laws of utility-expectation maximisation. In the next two sections I highlight some utility functions that have historically been used for selection of optimal drug dose, and then discuss some potential objectives that could be considered when designing a utility function that would be used for defining optimal vaccine dose.

Historical utility Functions

Within vaccine and drug dose-optimisation there have been a number of utility functions that have been considered. Thall and Russel [103,104] considered a trivariate approach to selecting optimal dose, where three possible outcomes are possible with probabilities depending on dose; neither efficacy nor toxicity, efficacy but no toxicity, or toxicity (regardless of efficacy). The optimal dose was that which

---

[6] Whether the optimal solution is that for which the utility function is maximised or minimised depends on the context and are equivalent. Any utility function f() which is to be minimised can be equivalently described by a utility function g() to be maximised by setting g() = -f(). I will therefore refer to optimal doses as those that maximise the relevant utility function.

maximises the probability of efficacy without toxicity in vaccinated individuals. This is technically a single-objective optimisation problem but represents how dose-response modellers have attempted to address multi-objective dose-optimisation without the need to consider Pareto fronts or utility functions.

Simple threshold utility functions have been used for drug dose decision making [105,106]. Here, for each objective there is some set of thresholds. If modelling or data predict that a dose would be within bounds for all objectives, then it is considered acceptable/optimal, but there is no distinction between doses otherwise. A more complicated utility function was suggested by Brocke et al. [107]. for defining optimal dose of the drug ponatinib, with the utility function considering both efficacy and toxicity probabilities and being called the utility contour utility function. Handel et al. [49] proposed a utility function that aimed to maximise protection and minimise morbidity as measured by a transform of the innate immune response. See [Appendix A.C.] for a more detailed discussion of these and other utility functions that have been used in dose optimisation in drugs and vaccines.

Potential objectives for vaccine dose optimisation utility functions

Efficacy, toxicity, and cost/dose-sparing are all important objectives to consider when choosing optimal vaccine dose as stated above. However, when selecting optimal vaccine doses, these may not be the only objectives that may be influential. Here I discuss some objectives that may need to be considered in addition to or in replacement of efficacy, toxicity, or cost. I also discuss how the topics of disability weighting, epidemiological modelling and game theoretic modelling may be useful when designing utility functions and selecting optimal vaccine dose.

Practical limitations on maximum and minimum vaccine doses

There may be practical limitations which may affect which dose should be selected as optimal. For example, vaccine dose may be limited by administration volumes. For example, the maximum bolus of vaccine that can be delivered intramuscularly to the deltoid region is 1 millilitre [108]. A standard industrial process for production of an adenoviral vector vaccine was estimated to produce an average concentration of 6.8×10^{14} [109] viral particles per litre, equivalent to 6.8×10^{11} viral particles per

millilitre. Therefore, even if modelling predicted that a dose of 10^{12} viral particles would be optimal in terms of maximising efficacy, such a dose could not be practically given. Whilst dilution may be used to reduce the concentration of a vaccine and allow for precise delivery of smaller doses, there may be practical limitations that provide a minimum possible dose. For a trivial example, given that viral particles are discrete, any dose that was less than $10^0 (=1)$ viral particles would be equivalent to the placebo dose. Practical limitations on dose have been noted by the Food and Drug Administration, who suggest that it is common for toxicity in therapeutic cancer vaccines to be so low that dose is limited by anatomic or manufacturing issues rather than toxicity [110].

Immunogenicity

I have noted that correlates and surrogates of protection may be used to predict vaccine efficacy, and noted that these predictions of vaccine efficacy or protection may be used as an objective to be maximised as part of a utility function which defines optimal vaccine dose as suggested by Handel et al [49]. For vaccines where predictive immunogenicity-efficacy models are not available, such approaches may be limited. In these cases, immunogenic responses could be used as objectives which the utility function aims to maximise. For example, Rhodes et al. [50] defined optimal dose as the dose that maximised interferon-gamma spot forming units per million splenocytes. They did not consider in this a prediction of efficacy or protection that such immunogenicity would cause. This was likely because there were no reliable correlates of protection for tuberculosis [111]. If only one immunogenic response would be used to predict efficacy, the utility function includes only one objective (maximising efficacy), and the relationship between that immunogenic response and efficacy/protection can be assumed to be monotonically increasing, then such a simplification is valid.

However, for multi-objective vaccine dose optimisation, using efficacy/protection may be preferable to using immunogenicity where this is possible. This is because it may be simpler to elicit stakeholder preference for trade-offs between efficacy and toxicity/cost than it would be to elicit stakeholder preference for trade-offs between immunogenicity and toxicity/cost. For example, whether a 10% increase in protection

would justify a 5% increase in prevalence of severe (grade 3) adverse events may be easier to answer than whether an increase of 100 interferon-gamma spot forming units per million splenocytes would justify a 5% increase in prevalence of severe adverse reactions.

Disability metrics

For decisions regarding vaccine rollout and policy, it is common to use disability metrics to compare different vaccination policies. Examples of such metrics are disability adjusted life years (DALYs), quality adjusted life years (QALYs), or years of lost life (YLLs) [112]. Salomon et al. suggested that different negative clinical outcomes could be described using 'disability weights', which can be used to calculate DALYS averted, QALYS averted, or YLLs averted [113]. The vaccination policy that averts the most disability, as measured by any of these metrics, is then considered as optimal. For example, Drolet et al. used 'DALYs averted' as the metric to suggest that the optimal vaccination policy for a human papillomavirus vaccine was to vaccinate 14-year-old girls only rather than to also vaccinate boys [114]. In that sense, disability metrics were used as a utility function to predict optimal vaccination policy, which might suggest that it would be reasonable to consider disability metrics when choosing vaccine dose.

As noted, vaccine dose may affect both efficacy and toxicity. An increase in efficacy could lead to a decrease in risk of infection and disease, leading to a reduction in expected disability for vaccinated individuals. An increase in vaccine toxicity could instead increase expected disability in vaccinated individuals. Therefore, utility functions could be designed to consider the average disability that would be expected for vaccinated individuals and use minimisation of this relative to unvaccinated individuals as an objective function.

A similar comparison was conducted to predict whether benefits (efficacy) outweigh the risks (adverse events) for vaccines. The European Medicines Agency considered reduction in COVID-19 hospitalisation, intensive care unit admission and death relative to risk of thrombosis with thrombocytopenia syndrome for individuals who were administered the Vaxzervria vaccine [115]. Funk et al. conducted a similar

analysis for the Pfizer-BioNTech COVID-19 vaccine [116]. In these cases, the comparison was only between vaccination and no vaccination, but this could easily be adapted for consideration of different potential vaccine doses. Such metrics therefore may be reasonable to consider, and method for computing immunisation related DALYs has been discussed by McDonald et al [117].

Epidemiology

For dose optimisation of typical drugs, only the individuals treated by a drug will have been directly impacted by the drug and hence only they are directly impacted by any dosing decisions. This means that utility functions might only require considering the expected benefits (efficacy) and deficits (toxicity, cost) for each individual that receives the drug. Whilst such utility functions might also be practical for vaccines, and have been used in the works described above, the concept of 'herd immunity' means that population level effects of prophylactic vaccine efficacy might also be considered. Herd immunity is the epidemiological concept that as a larger proportion of a population develops immunity to infection, there are fewer individuals available to continue proliferation of the disease throughout a population, which will indirectly protect individuals who have not developed immunity [118]. This can lead to disease eradication if a sufficient proportion of a population is immune to infection, with that proportion not needing to be 100%.

One method for determining the population effects of increased vaccine efficacy might be to consider epidemiological models of infectious diseases. These are models which attempt to describe and predict the transmission and burden of disease at a population level. These can be used to predict the reduction in incidence, cases, hospitalisation, or mortality that might be expected given a certain vaccine efficacy when a vaccination programme is used in a population. For example, modelling of a vaccination programme against COVID-19 in New South Wales predicted that a 60% efficacious vaccine would prevent ~10400 deaths relative to a 50% efficacious vaccine, but that a similar 10% efficacy increase from 80% to 90% would only prevent an excess 300 deaths [119].

These epidemiological predictions could then be used to predict DALY metrics associated with a vaccine programme for specific efficacy and toxicity probabilities, with these then being used as part of the utility function used to define optimal dose.

Vaccine uptake and vaccination game theory

A final consideration regarding vaccine dose regards the impact that toxicity may have on the rollout of a vaccine at the population level. As stated above, the efficacy of a vaccine may impact vaccine benefit to the population through herd immunity. Whilst there is no exact parallel for toxicity (there is no 'herd toxicity'), a review of vaccine related game theoretic analysis by Chang et al. has suggested that increased vaccine toxicity may lead to decreased vaccine uptake at a population level [120].

This review also discusses that as a population tends towards a herd immunity threshold, a rational population may begin to consider the small risks associated with vaccines to outweigh the benefits they would receive through vaccine efficacy. This may lead to so-called 'free-riders', individuals who receive protection against disease through herd immunity but who themselves are not willing to be vaccinated [121]. Indeed, analysis has suggested that a completely rational and self-interested population with perfect information would never have enough individuals willingly vaccinated to achieve complete herd immunity and disease eradication [122].

Whilst this might be the case theoretically, given that individuals in a population do not have access to perfect information, they may make decisions based on their perceived notion of vaccine safety within their social network. A further game theoretical analysis discussed the possibility that the occurrence of rare but severe vaccine related adverse events may lead to a reduction in vaccine uptake that is not proportional to the actual risks associated with such events and are hence not 'rational' [123].

'Free-riders', vaccine hesitancy due to rare but severe adverse events, and the emergence of 'conspiratorial thinking' [121] are all game theoretic factors that have been suggested to cause a reduction in vaccine uptake. It has been suggested that mandatory safety practices such as vaccination are a necessity to ensure complete

48

disease control [122]. Where this is not deemed ethical, choosing a vaccine dose that minimises the risk of adverse events might lead to increased uptake [124]. Unfortunately, there may be a possibility of rare but serious adverse events being undetected during dose-ranging trials due to the small size of these trials and rarity of these events, which may limit ability to account for such occurrences when selecting optimal dose.

Multi-dimensional optimisation: Optimisation of prime/boost dose or interval, or antigen/adjuvant dose

Much work in MBDD for selection of optimal dose focuses on drugs for which only a single dose must be selected. For vaccine development, developers may wish to optimise not only single administration vaccine doses, but also to determine optimal prime/boost dosing or interval between doses or optimal doses of co-administered antigen/adjuvant vaccines. In the language of multi-objective optimisation, the number of factors that developers and modellers are trying to optimise would be called the dimension of the optimisation problem [125]. For example, if developers must predict optimal prime dose, optimal boost dose, and the optimal time between these two doses, then this would be a three-dimensional optimisation problem.

There has been work in drugs and vaccines that we should consider when considering multi-dimensional dose optimisation problems, but these are all presently limited.

Mechanistic models have been used in vaccines for optimisation of both prime/boost dose and time between doses. De Boer and Perelson used a mechanistic model to describe proliferation and expansion of naive, activated, and memory T cells in response to a prime/boost administration of lymphocytic choriomeningitis virus [126,127]. Whilst this model was able to effectively describe the longitudinal time-immunogenicity relationship well, the data of this work was not from a dose-ranging study, the model predictions were not validated, and dose optimisation was not discussed.

Mayer et al. presented a simple mechanistic model of peptide-affinity dependent T Cell expansion, and predicted that T Cell fold-expansion is inversely proportional to

the magnitude of initial T cell population [69]. They calibrated this model to data and used this model to predict that administration of pigeon cytochrome C and lipopolysaccharide over three prime/boost/second-boost doses would be more immunogenic than giving the same volume as a single bolus. However, the data of this work was not from a dose-ranging study [128], the model predictions were not validated, and this mechanistic model would be unable to explain peaking dose-immunogenicity response. They also do not include the dose of the adjuvant lipopolysaccharide as part of the model, reducing the dimension of the optimisation problem.

Kumbhari, Kim, and Lee [82] used mechanistic modelling to predict optimal dose and interval between doses for an anti-cancer vaccine (melanoma antigen glycoprotein 100). They first built a complicated mechanistic immunological model that included more than 16 compartments[7] for peptides and 39 parameters. 35 of these parameters were literature derived, and 4 were found through model calibration to data [129]. In a second work they then used this 'full' model to generate data that could be used to fit a simpler 'minimal' model [130]. With each model they used in-silico simulation to predict optimal dose and dosing interval, as defined by reduction in cancer concentration at 60 days or mean avidity difference. Hence, whilst this was not described as IS/ID modelling, it fits well within the scope of the field. The authors attempted to validate the model using data from other studies. Whilst the authors claim that the model was validated, their model predicted that T cell expansion peaked at approximately 2 days post-vaccination, where-as the experimental data suggested the peak was at ten days. Additionally, the predictions of the 'minimal' model were qualitatively different to those of the 'full' model for high avidity killer T cell compartment. Thus, I believe further validation of their model and findings would be needed.

---

[7] The models divided some of its compartments into a number of subscripted compartments reflecting the amount of vaccine associated peptide-major histocompatibility complexes presented on the surface of maturing dendritic cells, and also reflecting different avidities of naive and effector T cells. The number of these sub-compartments is not detailed in this work, so the total number of compartments was not specified.

Statistical models have also been used for such multi-dimensional problems. For optimisation of prime/boost dose within vaccines, Rhodes et al used IS/ID modelling to predict the prime/boost dose of the antigen/adjuvant paradigm vaccine TB10.4/Ag85B which was maximally immunogenic [50]. However, this work assumed that the prime and boost dose should be equal at both the prime and boost time points. They also did not attempt to optimise time between doses or adjuvant dose in this work, as such analysis would have required a more complex and expensive 'checkerboard' experimental design. This assumption meant that the optimisation was still single-dimensional.

Within MBDD, statistical modelling has been used to predict effectiveness or toxicity for treatments that require multiple drug administrations (either at the same time or at different time points). This is typically done through the use of Bliss/Loewe independence and additivity models [131]. For each 'monotherapy', a marginal dose-response curve is found. The effect of the joint therapy for each dose combination is predicted by assuming that the treatment effects are independent. A joint therapy is then tested, and modellers compare the predicted joint treatment effect to the empirical joint treatment effect in order to build a multi-dimensional drug dose-response model that can account for treatment being synergistic or antagonistic [132].

Bliss/Loewe model may not be appropriate for multi-dimensional vaccine dose-response modelling and optimisation. Firstly, the best-known Bliss independence models require the assumption that an increase in either of the doses cannot lead to a decrease in response [133]. As noted, IS/ID modelling has already suggested that this assumption may not be reasonable in vaccine dose-response. Secondly, these models typically rely on 'copula' methods [131], which may not be appropriate for vaccines. For example, if the respective seroprotective rate for given monotherapy doses of drug A and B were 5% and 7% of patients, then the maximum seroprotective rate that could be predicted by a copula method for the joint therapy of these doses would be 12% (=5% + 7%) and the minimum would be 7%(=minimum of 7% and 5%)[134]. This limitation may render such methods inappropriate for modelling antigen/adjuvant dose-response, as it would be common that monotherapies of antigen or adjuvant are not efficacious, but the joint treatment is.

51

Thirdly, Bliss/Loewe models have not been used for optimising time between doses to my knowledge.

I therefore believe that previous findings in drugs and vaccines using either mechanistic or statistical modelling are presently limited in their implications for multi-dimensional IS/ID modelling, and that more research is needed into building models that can be used for multi-dimensional vaccine dose optimisation.

Model uncertainty: selection and averaging

As stated, IS/ID modelling has suggested that for some vaccines dose-immunogenicity and dose-efficacy may be better described by a peaking dose-response curve as opposed to a saturating dose-response curve. Hence there may be uncertainty in the correct choice of assumed dose-response curve for IS/ID modelling of vaccine dose-efficacy. The topic of 'model uncertainty' is common in many mathematical modelling fields, and methods of accounting for model-uncertainty in drug dose-response modelling have been previously developed [135–137]. These have been used even when efficacy could be assumed to always increase with dose, as there may be uncertainty between multiple 'saturating' models [138].

There are two common methodologies used for accounting for model uncertainty: selection and averaging. For both selection and averaging, for each candidate model a 'goodness-of-fit' metric is calculated. The 'goodness-of-fit' metrics may be an information theoretic criterion such as the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), or corrected AIC (AICc) [139]. Alternatively, squared-error metrics such as mean squared error (MSE) or root mean squared error (RMSE) have also been suggested [140]. Models which minimise these metrics for a given dataset are said to be a better description of the data. The information theoretic criteria normally include penalty terms which aim to prevent 'overfitting'. For example, the formula for the AICc for a model that has been calibrated to a given [141] dataset has been given as

$$AICc = n \times \ln\left(\frac{SS}{n}\right) + 2K + \frac{2K(K+1)}{N-K-1}$$

With N as the number of data points, K as the number of fitted model parameters and SS is the sum of the squared residuals of the calibrated model. The $\ln\left(\frac{SS}{n}\right)$ term is decreased for models that better approximate the observed data, and the $2K + \frac{2K(K+1)}{N-K-1}$ term is the penalty term that increases with the number of model parameters.

With the 'selection' method of accounting for model uncertainty, the model which minimises the goodness-of-fit metric is selected as the ''best' model and used to make any predictions or decisions. Some modellers suggest using only this best model. Others suggest that only models which have a significantly worse goodness-of-fit should be rejected as candidate models, and that models which are not significantly worse than the best model should not be rejected and should still be considered as potential alternative models when making predictions/decisions [142].

With the 'averaging' method, all candidate models contribute to an overall prediction, with models contributing to the overall prediction with a weighting that is dependent on their goodness-of-fit metric. Models which have a poor goodness-of-fit metric compared to the 'best' model contribute little to the overall prediction, whereas if two models have equal goodness-of-fit metrics they will contribute equally to the overall prediction.

There is not a consensus that either of selection or averaging is better than the other for accounting for model uncertainty. The selection method might be more reasonable if the modeller is primarily interested in determining which model best describes a dose-response relationship. For the purposes of response prediction, model averaging has been found to outperform model selection in simulation studies [143].

Clinical trial design: gathering the data that are used to choose optimal dose

An important consideration for vaccine dose-selection, regardless of whether mathematical modelling will be used, is in how dose-response data will be generated. Typically, these data are collected through the use of dose-ranging studies. Therefore, vaccine developers and modellers must consider the clinical trial

design of such dose-ranging trials, and any impact this may have on the selection of optimal vaccine dose.

Clinical trial design includes many factors, for example, the total number of trial participants, the number of dosing groups that should be investigated, how trial participants are to be distributed between those groups, which responses are to be measured, whether any interim analysis will be conducted, and any safety considerations that may result in early trial termination. One key consideration in trial design is the number of dosing groups. For a given total trial size N, if there are K equally sized dosing groups, then each dosing group contains n=N/K individuals. With fixed N, decreasing K increases n, increasing statistical power to distinguish potential differences in response between dosing groups. For traditional empirical dose selection, increased statistical power might be desirable, however it has been suggested that using too few dosing groups (small K) may inhibit selection of optimal dose [144].

There are many clinical trial designs that have been used or suggested for use in conducting dose-finding trials in vaccines and drugs. One method of categorising these trial designs would be in whether they include elements of 'adaptive' trial design[145]. Adaptive trial design involves conducting analysis at interim time points during a dose-ranging study, and preferentially assigning trial participants to certain dosing groups depending on the interim analysis. In this section I provide examples of potential vaccine dose-ranging trial designs, both adaptive and non-adaptive.

A common, non-adaptive design for vaccine dose-ranging trials is to randomise all trial participants approximately evenly between different dosing groups. Such trial designs can include or exclude dose-response modelling when predicting optimal dose. This was the trial design used by Zhu et al. [146], Pollock et al. [147], and Sadoff et al. [35], who did not discuss using modelling to select optimal dose. Rhodes et al. used this trial design to inform dose-response modelling for the selection of optimal vaccine dose [50]. Despite its frequency of use in vaccine dose-ranging trials, I have not found a naming convention for this trial design within dose-ranging literature, other than potentially '1:1 randomisation'. Consideration of vaccine

dose-optimisation as a so-called 'multi-armed bandit' problem[8] would label this as 'uniform exploration' [148].

Adaptive designs without the use of mathematical modelling have been suggested for use in clinical trials since first described by Thompson in 1933 [149]. Whilst initially discussed as a method of ethical trial design when comparing multiple different treatments, adaptive design has been applied for locating optimal drug and vaccine dose. Aziz et al. provide a description of these adaptive designs for dose-ranging trials [150], and Lui et al. recommend further adoption of such methods in vaccine dose-ranging trials in order to accelerate vaccine development and better identify optimal vaccine dose [151]. Chen at al. used an adaptive trial design during a phase IIb/III clinical vaccine trial to determine an optimal treatment arm between a four-valent human papillomavirus vaccine and three different dosing groups (low, medium, high) of a nine-valent human papillomavirus vaccine [152]. 1240 human trial participants were uniformly divided between the four treatment arms. Interim analysis was conducted, and only the four-valent human papillomavirus vaccine and middle dosing groups of a nine-valent human papillomavirus vaccine were continued into the second phase of the trial. Again, this form of adaptive design has a canonical name within multi-armed bandit literature, 'sequential halving' [148].

The predominant form of adaptive design in vaccine dose-ranging trials are so called 'rule-based' trial designs, which are specifically prevalent in conducting dose-ranging trials for therapeutic cancer vaccines. The typical example given for 'rule-based' trial design is the '3+3' design, which is used to locate the minimum dose at which the probability of 'dose-limiting toxicity' (DLT) is no longer acceptable [153]. DLT is

---

[8] Multi-armed bandit problems are problems in which an investigator is faced with multiple potential actions that can be taken, with each action having a stochastic reward associated with it. The investigator does not know a-priori the distribution of rewards for each action and must find a strategy to locate optimal actions. These may aim to maximise the total reward achieved over a finite number of actions taken (the 'average regret' setting), or to predict an 'optimal' action after a finite number of actions taken (the 'simple regret' setting).

Considering 'actions' as 'doses' and 'rewards' as efficacy, toxicity, etc., we can see that vaccine dose-ranging trials fall into this class of problem. Multi-armed bandit problems and their solutions also were initially conceived as a method of conducting ethical clinical trials, which is why I consider this topic and its language relevant to vaccine dose-optimisation.

typically defined as the occurrence of a grade 3/severe adverse event. With this design, three trial participants initially receive the lowest possible dose. If none of these individuals experience DLT, then the next cohort of three trial participants are enrolled at an increased dose. This is repeated until at least one individual experiences DLT. If only one individual experiences DLT, another cohort of three individuals receives the same dose. If two or more out of the six individuals for a dose experience DLT, then dose de-escalation occurs, and the prior dose level is expanded to six trial participants. If after six individuals have received a given dose no more than one of these experienced DLT, the trial concludes, and that dose is chosen as the maximum tolerated dose (MTD). The original mathematical theory behind this design suggested that the probability for DLT at that dose is selected by such a design is ~17%, which the Food and Drug Administration (FDA) defined as the threshold above which DLT prevalence is not acceptable for therapeutic cancer vaccines [19].

Whilst rule-based designs have seen common use in drug dose selection and dose selection for therapeutic cancer vaccines, they have also been criticised. Theoretical analysis has repeatedly shown that these designs are not reliable in their prediction of the MTD and may require more trial participants than other trial designs [153]. In particular, if the dose-efficacy relationship cannot be assumed to be saturating, or can be assumed to saturate for a dose that is less than the MTD, then this design is not recommended [154]. The Medical Research Council Network of Hubs for Trials Methodology Research's Adaptive Designs Working Group strongly recommends that these designs should not be used for phase I clinical trials [155], and state that such dosing-ranging trial designs require more trial participants due to being 'memory-less'. If adaptive design is to be used, then it has therefore been suggested that trial designs that are model-based are superior to rule-based designs [156].

Model-based adaptive trial design designs aim to leverage mathematical modelling to improve dose-selection, increase benefit to trial participants, and reduce dose-ranging trial size relative to the designs described above. In 1990, O'Quigley, Pepe and Fisher published a description of the 'Continual Reassessment Method' (CRM) of clinical trial design for locating the MTD [157]. This is conducted by performing mathematical modelling at interim time points throughout a clinical trial. After each

cohort, any dose-response models are updated using all available data, and then individuals in the next cohort are given doses that are predicted to be optimal whilst following potential dose-escalation rules. This has since come to be accepted as a highly effective method of locating optimal doses [158], with O'Quigley suggesting that the CRM may approach the theoretical upper bound for optimal trial design for locating MTD [159]. The CRM has since been extended broadly [85]. In particular, an early extension by Thall and Cook was to include modelling of dose-efficacy as well as dose-toxicity [100]. This 'EffTox' variant of the CRM has been used in locating optimal dose of the drug Ponatinib, which similarly to vaccines has been suggested to have a potentially peaking dose-efficacy curve shape [107]. Wheeler et al. provide a description of how dose-ranging clinical trials can be conducted using the CRM [160].

If mathematical modelling is to be used for the selection of optimal dose, then optimal design theory can also be considered [161,162]. Trials designed using optimal design theory aim to support mathematical modelling by distributing trial participants between dosing groups in a way that maximises expected gain in information regarding model parameters or predictions. Here modelling may be used before any data are gathered if modellers have access to a-priori knowledge regarding likely dose-response models and parameters. Trial designs that use optimal design in an adaptive trial design manner have also been proposed [163].

There have been a number of other designs that have been suggested for selection of optimal drug dose, but the above represent the most common methods of trial design. For example, the 'Simple Toxicity and Efficacy Interval Design' (STEIN) is a more recent rule-based trial design that was suggested as a method of trial design for dose ranging trials when efficacy cannot be assumed to be increasing with dose, and the authors of that recommend its use over CRM style trial design [77]. However, Takahashi et al. found that STEIN was outperformed by the EffTox variant of the CRM, which used a parametric model, and by their Bayesian Optimal Design (BOD) trial design, which used a non-parametric model [75]. Zang et al. suggested another model-based adaptive trial design that was not considered as a CRM-style trial design [164]. This leveraged a double-isotonic modelling to describe dose-

efficacy and was found to have satisfactory performance for the selection of optimal dose in simulation studies.

A remark that is important to make is that dose-ranging trials may include interim analysis to reduce the risk that any trial dose has an unacceptable risk of harm, without being 'adaptive' in the sense that is typically meant. For example, Pollock et al. conducted a dose-escalation study of the COVAC1 vaccine [147]. A sub-study of this investigated immunogenicity and toxicity of three potential doses, 0.1, 0.3, 1.0 µg. They conducted an initial dose-escalation phase, where one individual received 0.1 µg of COVAC1 vaccine. If this was 'safe' (which was defined by adverse events being mild or short lived if they were severe), three more individuals were to be vaccinated with the same dose. If this was 'safe', then this process was to be repeated with 0.3 µg, and then twice with 1.0 µg. If safety for all doses was acceptable (which it was), the remaining trial participants were randomised uniformly between those three doses. I do not believe that, given the final randomisation was uniform between the three doses, that this should be considered adaptive design. However, it is representative of how interim analysis can presently be used to inform vaccine dose-ranging trial design.

An aspect of Pollock et al. that is interesting from the perspective of developing mathematical modelling methods for vaccine dose selection is that these researchers considered a linear model of dose-immunogenicity when planning this study. This was not used to guide selection of clinical vaccine dose, as is recommended in IS/ID modelling, but instead was used to aid in power calculations for determining sample size. Furthermore, in the supplementary material the authors suggested a non-linear dose-immunogenicity relationship could prohibit such sample size calculations, suggesting that uncertainty in the curve shape of vaccine dose-immunogenicity and dose-efficacy may already be a limiting factor in current vaccine dose-finding studies.

Leveraging expert knowledge and existing data for trial design/dose selection

When selecting an optimal vaccine dose for a vaccine with minimal (or no) existing human data it may be beneficial to consider relevant expert knowledge and existing

data regarding 'similar' vaccine dose-response and preclinical dose-response data. These could also be beneficial, and are often used, in guiding design of early phase clinical trials. Here I will describe previous works in vaccines and drugs that aimed to improve modelling for trial design/dose selection through leveraging such existing knowledge or data.

Leveraging expert knowledge and opinion for trial design/dose selection

Expert opinion is an important consideration in designing vaccine dose-ranging trials. The selection of the doses to be investigated in dose-ranging trials is currently primarily decided by consideration of the set of doses that clinicians and experts believe are most likely to be acceptable or optimal (personal communications, T Evans). Therefore, it may be important to leverage this expert knowledge to inform modelling. Various methods have been proposed in order to accomplish this, though presently these all have noted limitations.

A common type of model that has been used for dose-finding trials in the style of the CRM is the so-called 'power skeleton', [Appendix A.A.3.Power Skeletons]. This model easily incorporates expert opinion, as rather than directly model the relationship between dose and some response, modellers elicit expected response probabilities for all potential doses in the clinical trial to form a model 'skeleton' [165]. The calibration of the model in response determines whether the expert opinions were under-estimations or over-estimations. The limitation of this approach is that the default power-skeleton model assumes that the elicited expert predictions are accurate up to linear translation. This is to say that if experts predict that a certain dose is most efficacious, then the model will always predict that as the maximally efficacious dose regardless of any observed dose-efficacy data. This limitation can be minimised by using model-averaging of multiple skeletons but doing so is computationally expensive and may be less intuitive with regards to stakeholder elicitation [166].

Alternative methods for incorporating expert opinion into mathematical modelling have been suggested for when statistical mathematical models are used; notably either sigmoid saturating or latent quadratic models [appendix A.A.]. These are

parametric models which have used between two and four parameters and that are common in CRM-style clinical trials. If these models are to be calibrated using either maximum likelihood or least-squares methodologies, then the use of pseudo-data has been suggested [167]. Pseudo-data are simulated trial participants that are simulated to have observed responses that are dependent on dose-response predictions elicited from experts and are down-weighted relative to real data during model calibration. However, in practice pseudo-data are more commonly recommended to be minimally informative, functioning to stabilise the model calibration rather than capturing expert opinion [85].

Alternatively, statistical models can be calibrated using Bayesian inference, with Bayesian priors being placed over the parameters. Lee and Cheung suggested that these priors should be calculated by simulating CRM-style clinical trials over a so-called 'calibration set' of potential dose-response curves that experts have confirmed could be possible for the drug for a number of different parameter-priors [168]. The prior that most often led to selection of the 'true' optimal dose in these simulations should then be used. This is intuitive but is computationally expensive and reliant on the scenarios that experts suggest being reflective of the true dose-response curve for the drug or vaccine. Thall proposed an algorithm involving elicitation of prior probability intervals for expected predicted efficacy at each dose, calibration of priors to jointly minimise least squares error between expert predictions and prior model prediction whilst targeting an 'effective sample size', which can be considered as the number of individuals worth of data that the parameter priors will contribute to the model prediction [169]. Thall suggested that this was effective but would be expensive in terms of computation and labour, and also suggested that the algorithm he proposed was sensitive to certain hyper-parameters that the modeller must choose which may limit practical use.

Leveraging historical data for trial design/dose selection

Similar to the inclusion of prior expert knowledge, modelling methods for conducting vaccine dose-ranging trials and selecting optimal vaccine dose may be improved if historical dose-response data can be leveraged to inform mathematical models. Specifically Rhodes et al. suggested that open access to historical dose-response

data may be needed to maximise effective IS/ID modelling [23]. When planning dose-ranging trials or selecting optimal dose for a novel vaccine, modellers and developers could consider historical data from trials of vaccines that are 'similar' to the novel vaccine and assume that 'similar' vaccines have 'similar' dose-response curves. In theory, this historical dose-response data could be incorporated into modelling and trial design in the same ways I mentioned for including prior expert knowledge, for example as down-weighted pseudodata. I have however not been able to find any instance of such a methodology being used for vaccine dose-response modelling. There is also likely to be a question of what would be meant by 'similar' vaccines which would have to be answered as part of the modelling process. Modellers have also used immunological parameter values derived from literature as part of mechanistic modelling [130].

Historical data may already be being used to inform the design of vaccine ranging trials already, though without mathematical modelling. As an example, dose-ranging trials for two novel ChAdOx1 adenoviral vectored vaccines against middle eastern respiratory syndrome and COVID19 both investigated the same three potential doses, 5.0×10^9, 2.5×10^{10}, and 5.0×10^{10} viral particles [170,171]. This may suggest that the vaccine developers considered similar historical data when determining the doses to be investigated.

Leveraging pre-clinical data to inform clinical trials: Interspecies allometric scaling

The field of MBDD has been benefited by the possibility of 'interspecies allometric scaling' [172]. This comprises methods which allow developers and modellers to predict human dose-response using dose-response data from pre-clinical dose-response experiments in animals. Typically, a prediction of 'optimal' drug dose in an animal species can be multiplied by a scalar to predict a reasonable estimate of 'optimal' human dose [173]. As with expert opinion and historical data, vaccine modellers and developers may hope to leverage this preclinical data to better inform dosing decisions. Rhodes et al. suggested that allometric scaling is another concept from MBDD that could benefit IS/ID modelling if it is possible to replicate in vaccines [23], and provided a proof-of-concept study for allometric scaling to predict human interferon-gamma response using macaque time-response data [51].

61

There are however a number of potential limitations in interspecies allometric scaling for vaccine development. Firstly, the biological assumptions that justify allometric scaling of drugs is that the dominant predictor of drug clearance and metabolism is body mass [174], and that metabolic rate is highly predictive of drug effect given that concentration is predictive of drug effect. The so-called '¾ rule' for allometric scaling, which states that the parameter value of the exponent of the power relation between body mass and metabolic rate equals 0.75, can be derived theoretically [174] and has been validated empirically [172]. There may not be as much theoretical or empirical justification to assume that there is a causative path between metabolic rate and vaccine effect.

Secondly, the parameters used to define allometric scaling factors for drugs are typically consistent. For example, whilst there have been detractors of the ¾ rule that claim that ⅔ would be a better value for the value of the exponent parameter, the majority of data and studies of metabolic rate are consistent with the value being 0.75, and most species' empirical metabolic clearance lies near the prediction that would be expected based on their mass using this rule [172]. This contrasts with the observation that estimates of potential allometric scaling parameters for prediction between murine and human dose for tuberculosis vaccines have varied between 0.5 and 100 [23].

Thirdly, we must be cautious to define what allometric scaling is being used to predict. If 'optimal' dose is being predicted, then as I noted before this may be dependent on the utility function used to define optimality. If model parameters or biological quantities such as immunological response are being predicted, then it may be necessary to assume that the same or similar models can be used to describe dose-response in both species. Whilst this assumption did not limit the work of Rhodes et al. [51], validation of consistency in dose-response curve shape between host species may be needed.

Finally, though linked to the third limitation, it may be questioned whether there exists a biological reason to assume that animal experiments are predictive of human vaccine response. It may be necessary to consider for which animal species immunological and toxicological response will most reflect human immunological and

toxicological vaccine response. Rhodes et al. cautiously suggested that data from Indonesian cynomolgus macaques best predicted longitudinal human interferon-gamma response to BCG vaccination, and that Chinese cynomolgus macaque were a comparatively less predictive animal model of human dose-response for that vaccine. More generally, extrapolation to predict that a vaccine is immunogenic/efficacious in humans from the observation that it was immunogenic/efficacious in an animal model is not always justified [175]. This is due to inherent differences in immune mechanisms between species. If intraspecies prediction of whether a vaccine is immunogenic/efficacious can only be done cautiously, there may need to be caution in predicting dose-response curves or optimal dose through interspecies allometric scaling.

Incorporating individualised dosing

For certain classes of drugs, particularly those that may be highly toxic, mathematical modelling can be effective for creating 'individualised dosing' schedules [176]. This 'individualised dosing' allows modellers to predict doses that are not only optimal on average for a population but are instead predicted to be optimal for the specific individual that will receive the drug. This has been historically done using 'covariate' models. Example covariates that have been used in MBDD include weight, age, co-medication status, or genotype [177]. Separate models can be calibrated based on data stratified on these covariates, or alternatively a hierarchical or mixed-effects modelling approach to parameterisation can be used, where model parameters may potentially vary between individuals in a population [178]. The parameters for each individual cannot be directly measured but can be assumed to follow some distribution that will be determined through model calibration and that may depend on covariates. This can allow modellers to determine which covariates may affect individual response, and hence which covariates may need to be accounted for when predicting individualised dosing.

Once these covariate models are parameterised through a combination of literature review or calibration to clinical trial data, these can be used to predict individually optimal doses. For example, modellers built a mechanistic covariate model to predict optimal dose for the oncology drug Paclitaxel [179]. One of the parameters of this

model, 'maximum elimination capacity', was found to be typically higher in men. Therefore, modellers predicted that the optimal doses per square metre of height were respectively 200 and 175 milligrams for men and women. When clinicians determine that a patient should receive this treatment, they can individualise the dose depending on the patient's sex.

A further extension of individualised dosing that can be used for patients receiving multiple administrations of a drug is 'therapeutic drug monitoring'/'model informed precision dosing' [180]. A covariate model using a hierarchical/mixed effects interpretation of model parameters is used to determine the initial dose(s) that a patient will receive [181]. Observations after these initial doses administrations allow modellers to update the model to better reflect that individual's data, and then better predict individualised optimal dose for later administrations.

In consideration of individualised vaccine dosing using IS/ID modelling, Rhodes et al.'s modelling of longitudinal interferon-gamma response to BCG vaccination in humans used a covariate model [51]. They found that prior BCG vaccination status was an important covariate to include in the model. This was done to account for variability in response between individuals, not to predict an optimal dose dependent on the covariate of prior BCG vaccination.

Individualised dosing has been effective in the field of MBDD, and the potential for covariate models to be used in IS/ID modelling has been shown. Research into mathematical modelling for individualised vaccine dosing may be justified if it is believed that covariates may affect optimal dose and that it is clinically practical to individualise vaccine dose for different individuals.

Further, model informed precision dosing is primarily recommended for highly toxic drugs which will be administered many times and for which monitoring of biomarkers at timepoints between administrations will be conducted, for example hospital patients receiving a course of rifampicin [182]. This is very dissimilar to the typical way in which prophylactic vaccines are administered, and so it may be unlikely that model informed precision dosing would be useful in IS/ID modelling to select optimal vaccine dose.

Simulation studies: how to evaluate approaches for selecting optimal dose

Many approaches have been suggested for locating optimal dose ('dose-optimisation approaches'). Developers must determine which approaches to use, including the choices of whether to use modelling or empirical comparison, which modelling techniques to use, which trial design to follow, how to include potential historical/pre-clinical/expert knowledge into dose selection and trial design, and many other factors that I have discussed throughout this background.

Vaccine developers may want to know which dose-optimisation approaches are effective, safe, and efficient. For example, they may want to know whether conducting a CRM style dose-ranging trial improves dose selection relative to more traditional dose-finding trial design. To assess this empirically, multiple dose-ranging trials could be conducted for a given vaccine, half using a CRM-style dose-optimisation approach and the other half conducted under a more traditional vaccine trial dose-optimisation approach. Due to dose-ranging trials being expensive, and caution over whether novel trial designs and dose-selection methods will be ethical, such studies are not likely to be feasible. Additionally, due to the stochastic nature of vaccine dose-ranging trials, it is possible that even if one of these approaches is "better" for locating optimal dose 'on average' than the other, the "better" approach may appear less effective or less safe than the "worse" approach if only a small number of empirical trials were conducted. This stochastic nature also means that it would not be possible to know whether a clinical trial found the 'true' optimal dose, and so we would likely still be unable to say which approach was truly most effective for locating optimal vaccine dose safely and efficiently.

These problems can be addressed by conducting 'simulation studies', which are important methods for evaluating statistical and mathematical modelling tools when empirical evaluation is limited by expense, ethical requirements, or stochastic effects [183–185]. Simulation studies require modellers to specify 'true' dose-response curve called 'scenarios'. Modellers then simulate a large number of clinical trials; each being conducted using the different dose-optimisation approaches of interest for the different scenarios. This addresses the three limitations of empirical investigation mentioned above, as computation simulation is considerably less

expensive than empirical dose-ranging trials, can be repeated many times, and as the 'true' dose-response is known modellers can determine how well an approach optimises dose.

The weakness of these methods is that, given these scenarios are defined by modellers, they may not reflect real life vaccine dose-response curves. This represents a potential decrease in ecological validity, but the benefits discussed have led to simulation studies being very prevalent in assessing MBDD dose-optimisation approaches. Ecological validity can be improved by increasing the number of scenarios used to evaluate dose-optimisation approaches, or by ensuring that there are scenarios with dose-response curves that reflect the real-life vaccine dose-response expected by relevant vaccine experts and stakeholders.

Previous simulation studies commonly find that we should not expect that any dose-optimisation approach will always be able to locate optimal dose, particularly with the small number of trial participants in typical vaccine dose-finding trials [75,159]. Based on results in multi-armed bandit literature, we should also not expect that a single dose-optimisation approach exists which would be most effective or safe for all potential vaccines [186]. However, simulation studies can be used to assess which dose-optimisation approaches are typically most effective for some expected real-life vaccine dose-response by carefully choosing the scenarios.

Thesis Rationale: Mathematical modelling for optimal selection of vaccine dose.

Vaccine dose is an important factor in how efficacious, safe, and cost-effective vaccines are, and sub-optimal dosing may lead to the public health benefit from life-saving vaccines being reduced. There is reason to believe that traditional vaccine dose-optimisation approaches may not be maximally effective, fast, cost-efficient, or beneficial to trial participants. Mathematical Immunostimulation/Immunodynamic (IS/ID) modelling has been suggested to address this, and it is hoped that developing this field may potentially improve the process of vaccine dose-selection and the quality of selected vaccine doses. This suggestion is supported by preliminary research in vaccine dose-response modelling; however, this field is presently small. This suggestion is also supported by findings in model-based drug development (MBDD), which are numerous but may not be applicable for vaccine dose-optimisation. It would be beneficial to consider or augment techniques from MBDD in a vaccine dose-optimisation setting, aiming to validate and apply techniques that can be used generally across the majority of future vaccines.

This thesis will aim to expand the field of IS/ID mathematical modelling. The broad terminal goal of this field is to use mathematical modelling to improve selection of optimal vaccine dose. In order to achieve this goal, we need to establish:

- 'What mathematical model(s) should be used to model vaccine dose-response'?
- 'What is meant by 'optimal' vaccine dose?'
- 'How should we use these models and incorporate them into vaccine-dose selection?'

Consideration of these questions and the narrative review of the relevant literature led to the objectives that I considered in this thesis.

Which model(s) should be used to model vaccine-dose response to improve vaccine dosing?

Many different mathematical models could potentially be used to model vaccine dose-response. I intend to investigate statistical mathematical models rather than

mechanistic models, due to their comparative simplicity and ease of use, potential to be applied to vaccines more generally, and to their prevalence within MBDD.

To reflect the previous findings that an increased vaccine dose may not always increase vaccine immunogenicity and efficacy, IS/ID modellers may need to consider 'peaking' dose-response models. Research is needed into the prevalence of vaccine data for which a 'peaking' curve better describes vaccine dose-efficacy, and into whether incorrectly assuming a peaking or saturating curve for dose efficacy would actually decrease the effectiveness of modelling-based dose optimisation. Further, both 'selection' and 'averaging' methods for accounting for model uncertainty have been suggested in other fields and warrant investigation in vaccine dose-response modelling.

For many vaccines there may not be pre-existing data available to define a known parsimonious parametric model of dose response, meaning a model which is capable of describing dose-response accurately and with minimal complexity or number of parameters. In the absence of a parsimonious model, non-parametric modelling could be useful. In particular, investigation is warranted for non-parametric models of prime/boost paradigm vaccine dose-efficacy, as there does not presently exist well validated IS/ID modelling which can account for potential response interactions between the prime and boost doses.

I will investigate the prevalence of dose-immunogenicity data for a class of vaccines (adenoviral vector vaccines) for which dose-response is best described by peaking or saturating curves. This will be done by conducting a systematic review and meta-analysis using model-selection methods. I will use simulation study methods to investigate the potential impact of model-misspecification on optimal vaccine dose selection, and whether model-averaging methods or non-parametric dose-response models can address this. I will also use simulation studies to compare these model informed adaptive designs to traditional vaccine dose-optimisation approaches to evaluate whether there is potential for model informed adaptive designs to improve vaccine dose selection and benefit to vaccine clinical trial participants.

What is meant by 'optimal' vaccine dose?

To investigate what is meant by 'optimal dose', I will follow on from the work of Handel and formally describe vaccine dose-optimisation under the perspective of multi-objective optimisation. Within the narrative review I highlighted that there are a number of potential objectives that may need to be considered when optimising vaccine dose, regardless of whether that is done by mathematical modelling methodologies or not. I also described the concept of 'utility functions', how these can be used to quantitatively define stakeholder/developer preference, and examples of these being used in MBDD. Utility functions may need to be specified to quantitatively define what is meant by 'optimal' vaccine dose.

To demonstrate that multiple definitions of 'optimal' can be used I will consider multiple such utility functions and how they may be used to define 'optimal' dose in IS/ID modelling. These will include both single and multi-factor objective utility functions and span a number of objectives that may be highly relevant to vaccine developers. These objectives will include various cellular and humoral immunogenicity responses, efficacy, any-grade toxicity, high-grade toxicity, ordinal toxicity, and cost. I will also describe that qualitatively different vaccine doses may be selected depending not only on modelling choices, but also on this choice of utility function.

How should we use these models and incorporate them into vaccine dose selection?

I intend not only to investigate the statistical modelling of vaccine dose-response, but also to consider the use of these for vaccine dose-optimisation and whether such modelling can actually be used to improve vaccine dosing and clinical trial design. I intend to do this by considering how modelling has been presently applied within the context of IS/ID modelling, the data that may be available to modellers from current vaccine dose-ranging trials, and clinical trial designs from MBDD that may benefit and benefit from IS/ID modelling.

The application of IS/ID modelling has been retrospective thus far. Data from dose-ranging trials were used to calibrate mathematical models, which were used to predict optimal dose, but these predictions were not used to guide future dose-

response studies. If retrospective modelling is to continue to be prevalent in IS/ID modelling, modellers may need to consider what data are typically available for this purpose. For example, modelling may not be feasible if the number of parameters of dose-response models exceeds the number of dosing groups. Therefore, modellers may need to be aware that the complexity of retrospective modelling may be limited by trial design and should be aware of the number of dosing groups that are typically used.

Rhodes et al. [23] recommend that models could be used to guide not only dose selection but also future trial design. Given the effectiveness of model-informed adaptive design and CRM-style trial design in MBDD, investigation into the application of such methods in vaccines was warranted.

Additionally, traditional trial designs are optimised to determine whether there exists a statistical difference between dosing groups, and hence select optimal dose using that analysis. If modelling is to be used to select optimal dose retrospectively, then it may be beneficial to design trials to support this modelling rather than to support traditional hypothesis testing. For example, this might mean investigating many small dosing groups rather than few large groups.

I therefore consider modelling published vaccine dose-response data to provide a case study of retrospective modelling in selecting optimal vaccine dose. I also investigate through simulation study methods the capacity for model-informed and model-free adaptive trial design to improve vaccine dose selection. I also use simulation study methods to determine whether a large number of small dosing groups can be effective for locating optimal vaccine doses using mathematical modelling.

Summary of thesis data, models, and model calibration

Full details of data, models and calibration are outlined in the relevant chapters, but an overview is given here.

For chapters 2 and 3 I used dose-immunogenicity data that was gathered from published dose-response studies of single-administration paradigm, replication-

deficient adenoviral vector vaccines. These were gathered using a systematic literature review. Full details of inclusion and exclusion criteria for that review were included in chapter 2. For chapter 4, I used publicly available data from a recombinant adenovirus type-5 COVID-19 single-dose vaccine (Ad5-nCoV) dose-ranging study. In that work I assumed that a four-fold increase in neutralising antibody titre ('seroconversion') was a surrogate of protective immunity ('efficacy'). For chapters 5 and 6, all data were generated through simulation study methodologies. The code used to generate these simulation studies, along with the data that were generated, were made publicly available to aid reproducibility and auditing of my findings.

All models in this work were statistical dose-response models. For all chapters, a three-parameter sigmoid saturating model was used as the representative dose-response model for 'saturating' vaccine dose-immunogenicity/efficacy for single administration vaccines. For chapters 2, 3, and 4, the three-parameter gamma PDF model was used as the representative dose-response model for 'peaking' vaccine dose-immunogenicity/efficacy for single administration vaccines. These followed the suggestions of Rhodes et al. [50]. For chapters 5 and 6, the three-parameter latent-quadratic model was used as the representative dose-response model for 'peaking' vaccine dose-efficacy for single administration vaccines. This followed the suggestions of Brocke and O'Quigley [45,107]. For chapter 6, five and seven parameter latent-quadratic models were used as the representative dose-response model for vaccine dose-efficacy for prime/boost and prime/boost/second-boost administration vaccines respectively. These were my own extensions of the three-parameter latent-quadratic model, which was found to be useful for dose-optimisation of vaccine dose in chapter 5.

For chapters 4 and 6, a two-parameter sigmoid saturating model was used as the representative dose-response model for vaccine dose-toxicity for single-administration vaccines. For chapter 5, a four-parameter probit model was used as the representative dose-response model for ordinal vaccine dose-toxicity for single-administration vaccines. For chapter 6, three and four parameter sigmoid-saturating models were used as the representative dose-response model for vaccine dose-

toxicity for prime/boost and prime/boost/second-boost administration vaccines respectively.

In chapter 6 I also used the continuous correlated beta process (CCBP) model to model dose-efficacy and dose-toxicity for single-administration, prime/boost, and prime/boost/second-boost administration vaccines.

For chapters 2, 3 and 4 I conducted model calibration by minimising the sum of squared error between data and models. Calibration was conducted using the R statistical programming language, using the in-built 'nls' function [187]. For chapters 5 and 6 I conducted model calibration by maximising likelihood. Calibration was conducted using the python programming language, using the 'optimize.minimize' function of the SciPy package [188]. I implemented the CCBP model in python myself, using the 'stats.beta' function of the SciPy package, with the CCBP models being updated in response to data using the algorithms detailed in that chapter.

Thesis Aims and Objectives

The overall purpose of this thesis was to explore and expand the field of IS/ID and mathematical modelling for vaccine dose optimisation, addressing gaps that existed in the field.

This aim was achieved using the following objectives:

Objective one: Preliminary gathering of dose-ranging data through systematic review of dose-ranging studies. These were limited to dose-ranging studies for a specific class of vaccine (replication-incompetent adenoviral vector vaccines).

Objective two: Determination through mathematical modelling of the data from objective one the prevalence of predicted saturating versus peaking dose-response curves.

Objective three: Extension of IS/ID to multi-factor dose optimisation using 'utility functions' to balance efficacy, toxicity, and cost and consideration of the impact of the choice of utility function on the selection of 'optimal' dose.

Objective four: Evaluation of the potential for misspecification of dose-efficacy models to impact optimality of model-predicted optimal dose, along with the potential impact of model informed adaptive trial design.

Objective five: Design and evaluation of a novel vaccine dose-optimisation approach combining non-parametric modelling and adaptive trial design, with consideration of vaccine dose optimisation as a potentially multi-dimensional, multi-objective optimisation problem.

Thesis Overview

Figure 1.2 outlines the aim of this thesis, the objectives to achieving this aim, how the objectives align with the thesis chapters, research papers, and the data and methods required for completing the objectives.

Aim

Aim: explore and expand the field of IS/ID and mathematical modelling for vaccine dose optimisation, addressing gaps that existed in the field.

Objectives

Preliminary gathering and descriptive analysis of dose-ranging data through systematic review of dose-ranging studies for adenoviral vector vaccines.	Determination of the prevalence of predicted saturating versus peaking dose-response curves in adenoviral vector vaccines.	Extension to multi-factoral dose optimisation and consideration of the impact of the choice of utility function on the selection of 'optimal' dose.	Evaluation of the potential for mispecification of dose-efficacy models to impact optimality of model-predicted optimal dose, along with the potential impact of model informed adaptive trial design.	Design and evaluation of a novel vaccine dose-optimisation approach combining non-paramteric modelling and adaptive trial design, considering multi-dimensional, multi-objective vaccine dose optimisation.

Chapters

Chapter 2 Exploration into design of adenoviral vector vaccine dose-ranging studies	**Chapter 3** Analysis of adenoviral vector vaccine dose-response curves in prior studies	**Chapter 4** A case study in multi-factor optimisation: a modelling study to maximise vaccine safety, efficacy and affordability	**Chapter 5** Theoretical analysis of mathematical modelling for vaccine dose optimisation: efficacy curve shape, trial size, and trial dose selection	**Chapter 6** Evaluation of a novel non-parametric modelling approach to optimisation of vaccine dose

Papers

Paper 1 Immunological Dose-Response to Adenovirus-Vectored Vaccines in Animals and Humans: A Systematic Review of Dose-Response Studies of Replication Incompetent Adenoviral Vectors when Given via an Intramuscular of Subcutaneous Route	**Paper 2** Response Type and Host Species may be Sufficient to Predict Dose-Response Curve Shape for Adenoviral Vector Vaccines	**Paper 3** Optimising Vaccine Dose in Inoculation against SARS-CoV-2, a Multi-Factor Optimisation Modelling Study to Maximise Vaccine Safety and Efficacy	**Paper 4** Mathematical Modelling for Optimal Vaccine Dose Finding: Maximising Efficacy and Minimising Toxicity	**Paper 5** The Correlated Beta Dose Optimisation Approach: Optimal Vaccine Dose Selection using Mathematical Modelling and Adaptive Trial Design

Data

Immunological dose-response data for replication incompetent adenoviral vector vaccines from the systematic literature review.	Dose-response data for percentage seroconversion, percentage grade 1+ adverse event and percentage grade 3+ adverse event from a published phase Ib dose-ranging study.	No additional data required. Data was generated through simulation of clinical trials.

Methods

Systematic review Descriptive analysis	Least squares regression of statistical dose-response models to published data Model selection using Akaike Information Criteria	Simulation study methods (simulation of clinical trials) Maximum likelihood regression of statistical dose-response models

Figure 1.2. Aim of the thesis with corresponding objectives thesis chapters, papers, data requirements and methods

74

There are seven chapters in this thesis. Chapter one provides a detailed background to the thesis. I highlight the vaccine development pathway with emphasis on the importance of selecting optimal vaccine dose and provide a narrative review of topics related to mathematical modelling for vaccine dose optimisation and 'Immunostimulation/Immunodynamic' (IS/ID) modelling. This also includes topics that have been relevant in mathematical modelling for the optimisation of drug dose. In the concluding section of this chapter, I identify the research gaps highlighted by the narrative review that limited the field of IS/ID, present my aim and objectives to address these gaps, and present this thesis overview.

Chapters 2 through 6 are research papers, four of which are published and one is submitted at the time of writing. Research paper chapters are structured to include an introduction, then the paper. Any section of supplementary material considered important in the context of this overall thesis are then included. Due to the length of some of the supplementary documents, some sections of the supplementary documents for the papers were included in the appendix [Appendix A.D.]. The closing chapter of the thesis includes discussion of the findings of the thesis and implications for future research and development of the field of IS/ID mathematical modelling for the selection of optimal vaccine dose.

I note here that these research papers were written as standalone articles. Therefore, there was some repetition of information. Effort was made to use consistent terminology; however, this was not always possible. Partially this was due to some of the topics discussed in this thesis having been inconsistently named across available literature. Partially this was done to maximise effective parsing of each work individually. For example, in chapter 5/paper 4 I define 'dose-optimisation approaches' as a combination of the mathematical dose-efficacy model that was used, the method through which clinical trial doses were selected, and the size of simulated clinical trials. In the broader scope of the thesis, I would consider a 'dose-optimisation approach' to comprise any choice made by modellers or developers that may influence how vaccine dose is selected, but I believed that the simpler definition used in paper 4 would aid readers of that work.

References for the thesis body are at the end of the thesis main body. Each paper has its own set of references, as does the appendix.

Chapter 2 is a research paper (paper 1) gathering dose-immunogenicity data for single-administration replication-incompetent adenoviral vector vaccines across multiple host species, serotypes, routes of administration and immunogenicity responses.

Main objectives of this paper:

1. Assess the number of available papers, including adenoviral dose-response studies, and the distribution of host species and adenoviral serotypes within these papers.
2. Assess which immunological responses dose-response data were available.
3. Assess the dosing strategies used in adenoviral dose-ranging studies, including number and magnitude of dose levels.

This chapter corresponds to objective 1 of this thesis.

Citation: Afrough S, Rhodes S, Evans T, White R, Benest J. Immunologic Dose-Response to Adenovirus-Vectored Vaccines in Animals and Humans: A Systematic Review of Dose-Response Studies of Replication Incompetent Adenoviral Vaccine Vectors when Given via an Intramuscular or Subcutaneous Route. Vaccines (Basel). 2020 Mar 17;8(1):131. doi: 10.3390/vaccines8010131.

Chapter 3 is a research paper (paper 2) using the dose-immunogenicity data from chapter 2/paper 1 to assess the prevalence of adenoviral vector vaccine data that is best described by a peaking or saturating dose-response curve.

Main objectives of this paper:

1. Assess the prevalence of peaking/saturating dose-response curve shapes in published adenoviral vector vaccine studies.
2. Assess whether dose-response curve shape may be predicted by response type, host species, adenoviral species, and route of administration (RoA).
3. Assess which of host species, adenoviral species and RoA are the most likely predictors of dose-response curve shape.

This chapter corresponds to objective 2 of this thesis.

Citation: Benest J, Rhodes S, Afrough S, Evans T, White R. Response Type and Host Species may be Sufficient to Predict Dose-Response Curve Shape for Adenoviral Vector Vaccines. Vaccines (Basel). 2020 Mar 30;8(2):155. doi: 10.3390/vaccines8020155.

Chapter 4 is a research paper (paper 3) using published dose-efficacy and dose-toxicity data from a phase I dose-ranging study of a recombinant adenovirus type-5 COVID-19 single-dose vaccine to perform a case study of multi-factor model-based dose-optimisation.

Main objectives of this paper:

1. Using published data, calibrate mathematical models to the relationship between dose and seroconversion, safety, and cost of a single inoculation.
2. Identify the minimum dose that is predicted to theoretically induce herd immunity.
3. Identify the dose that maximises immunogenicity and safety.
4. Identify the dose that maximises immunogenicity and safety whilst minimising cost.

This chapter corresponds to objective 3 of this thesis.

Citation: Benest J, Rhodes S, Quaife M, Evans TG, White RG. Optimising Vaccine Dose in Inoculation against SARS-CoV-2, a Multi-Factor Optimisation Modelling

Study to Maximise Vaccine Safety and Efficacy. Vaccines (Basel). 2021 Jan 22;9(2):78. doi: 10.3390/vaccines9020078.

Chapter 5 is a research paper (paper 4) using simulation study methodologies to evaluate the impact that assumed statistical dose-efficacy model, trial size, and clinical trial design (method of trial dose selection) may have on mathematical modelling-based vaccine dose optimisation.

Main objectives of this paper were to investigate:

1. When the method of trial dose selection is fixed, how dose-optimisation approaches are affected by the assumed statistical efficacy model and trial size.
2. When trial size is fixed, how dose-optimisation approaches are affected by the assumed statistical efficacy model and method of trial dose selection.

This chapter corresponds to objective 4 of this thesis.

Citation: Benest J, Rhodes S, Evans TG, White RG. Mathematical Modelling for Optimal Vaccine Dose Finding: Maximising Efficacy and Minimising Toxicity. Vaccines (Basel). 2022 May 11;10(5):756. doi: 10.3390/vaccines10050756.

Chapter 6 is a research paper (paper 5) on the design and evaluation through simulation study methodologies of the novel correlated beta approach for optimisation of vaccine dose, which combines non-parametric Continuous Correlated Beta Process models with adaptive trial design. This evaluation considered vaccine dose optimisation as a potentially multi-factorial and multi-dimensional optimisation problem.

Main objectives of this paper:

1. Evaluate the Correlated Beta Dose Optimisation Approach for optimising vaccine efficacy for a single dose administration.
2. Evaluate the Correlated Beta Dose Optimisation Approach for optimising vaccine efficacy for a prime-dose/boost-dose administration.
3. Evaluate the Correlated Beta Dose Optimisation Approach for optimising vaccine utility, maximising efficacy, and minimising toxicity.
4. Evaluate the use of expert knowledge informed Continuously Correlated Beta Process priors for vaccine dose-optimisation.

This chapter corresponds to objective 5 of this thesis.

Citation: Benest J, Rhodes S, Evans TG, White RG. The Correlated Beta Dose Optimisation Approach: Optimal vaccine dose selection using mathematical modelling and adaptive trial design (Submitted)

Author Contributions

The overall idea for this PhD project to further extend and improve the field of IS/ID modelling for optimising vaccine dose was generated by Dr Sophie Rhodes, Dr Thomas Evans, and Professor Richard White. Author contributions for papers 1-5 are outlined in the associated chapters.

Funding

Chapter 2. Collation of dose-response data and exploration into trial design for adenoviral vector vaccine dose-ranging studies:

Chapter 2 Introduction

The objective of paper two was to collate and describe dose-response data that are available within the published literature on replication-deficient adenoviral vector vaccines. This data could then be used to conduct the analysis of dose-response curve shape for chapter 3/paper 2. This work addresses objective 1 of this thesis.

Another objective of this paper was to assess the dosing strategies used in adenoviral dose-ranging studies, including number and magnitude of dose levels. Providing the distribution for the number of dosing groups that are typically considered within adenoviral vector vaccine dose-ranging trials may be informative to future modellers when choosing the complexity of vaccine dose-response models. Similarly, the distribution of the magnitude of the dosing levels is important, to approximate the range of dosing magnitudes over which adenoviral vector vaccines may currently be assumed to have 'optimal' dose.

This work was intended to be descriptive and enable future modelling. The search strategy and inclusion/exclusion criteria for this systematic review had been decided prior to my period of study.

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed <u>for each</u> research paper included within a thesis.

SECTION A – Student Details

Student ID Number	lsh1804914	Title	Mr
First Name(s)	John Helier		
Surname/Family Name	Benest		
Thesis Title	Immunologic Dose-Response to Adenovirus-Vectored Vaccines in Animals and Humans: A Systematic Review of Dose-Response Studies of Replication Incompetent Adenoviral Vaccine Vectors when Given via an Intramuscular or Subcutaneous Route		
Primary Supervisor	Richard G. White		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?	MDPI Vaccines		
When was the work published?	17 March 2020		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	N.A		
Have you retained the copyright for the work?*	Yes	Was the work subject to academic peer review?	Yes

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	

Please list the paper's authors in the intended authorship order:	
Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I collaborated with Afrough on the extraction and validation of data from these papers. I performed the statistical analysis and visualisation for this work. I collaborated with Afrough on writing this paper, in particular the methods and introduction, however I carried out the majority of the writing on the abstract, objectives, results, and discussion. All authors reviewed the paper. The interpretation of the results was my own work.

SECTION E

Student Signature	████████████
Date	11 August 2022

Supervisor Signature	████████████
Date	12 August 2022

82

Paper 1 Title: Immunologic Dose-Response to Adenovirus-Vectored Vaccines in Animals and Humans: A Systematic Review of Dose-Response Studies of Replication Incompetent Adenoviral Vaccine Vectors when Given via an Intramuscular or Subcutaneous Route

Authors: Sara Afrough, Sophie Rhodes, Thomas G. Evans, Richard G. White, John Benest

Permission from copyright owner to include this work:

Immunologic Dose-Response to Adenovirus-Vectored Vaccines in Animals and Humans: A Systematic Review of Dose-Response Studies of Replication Incompetent Adenoviral Vaccine Vectors when Given via an Intramuscular or Subcutaneous Route

Sara Afrough [1],*, Sophie Rhodes [2], Thomas Evans [1], Richard White [2] and John Benest [2]

[1] Vaccitech Ltd., The Schrodinger Building, Heatley Road, The Oxford Science Park, Oxford OX4 4GE, UK; tom.evans@vaccitech.co.uk

[2] Department of Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK; Sophie.Rhodes@lshtm.ac.uk (S.R.); Richard.White@lshtm.ac.uk (R.W.); John.Benest@lshtm.ac.uk (J.B.)

* Correspondence: sara.afrough@UCB.com

Abstract: Optimal vaccine dosing is important to ensure the greatest protection and safety. Analysis of dose-response data, from previous studies, may inform future studies to determine the optimal dose. Implementing more quantitative modelling approaches in vaccine dose finding have been recently suggested to accelerate vaccine development. Adenoviral vectored vaccines are in advanced stage of development for a variety of prophylactic and therapeutic indications, however dose-response has not yet been systematically determined. To further inform adenoviral vectored vaccines dose identification, historical dose-response data should be systematically reviewed. A systematic literature review was conducted to collate and describe the available dose-response studies for adenovirus vectored vaccines. Of 2787 papers identified by Medline search strategy, 35 were found to conform to pre-defined criteria. The majority of studies were in mice or humans and studied adenovirus serotype 5. Dose-response data were available for 12 different immunological responses. The majority of papers evaluated three dose levels, only two evaluated more than five dose levels. The most common dosing range was 10^7–10^{10} viral particles in mouse studies and 10^8–10^{11} viral particles in human studies. Data were available on adenovirus vaccine dose-response, primarily on adenovirus serotype 5 backbones and in mice and humans. These data could be used for quantitative adenoviral vectored vaccine dose optimisation analysis.

Keywords: dosing; dose-response; adenovirus-vectored vaccines; immunogenicity; clinical; pre-clinical

1. Introduction

The methods of finding doses for optimal vaccine delivery in humans is an empirical science. Frequently, vaccine developers have relied on historic information to conduct small dose-ranging studies in animal models, and then used these data to design further studies in humans, despite the relationship between animal and human dose being unproven [1]. Unlike allometric analysis used in pharmacokinetic/pharmacodynamic assessments in small molecule drug development [2], there are no published or widely accepted allometric scaling factors to easily translate animal vaccine dose-responses to human vaccination. Thus, each vaccine development group collates the relevant literature, and their own data, to determine how to design initial animal or human dose-response

studies. Unfortunately, recent evidence suggests that this empirical method of dose selection has, in part, led to suboptimal dose identification in humans for diseases, such as yellow fever, meningitis and malaria [3–6].

Recently developed mathematical modelling methods, referred to as immunostimulation/immunodynamic (IS/ID) modelling, attempts to address these issues [7–10]. IS/ID modelling was developed to address the lack of quantitative methods in vaccine development [11]. The aim of IS/ID is to translate pharmacokinetic/pharmacodynamic (PK/PD) methodology to vaccine development, and preliminary IS/ID modelling has shown promise in accelerating vaccine dosing decisions. Modelling of the dose-response curve and cross-species translation of tuberculosis vaccine dosing data have predicted a lower human dose than previously tested [7,9], and showed that antibody response against human parainfluenza virus may be maximised by an intermediate dose [12]. To inform future IS/ID modelling, dose-response data must be collated. However, these data can also provide valuable insight into study designs that are currently used to explore vaccine dose-response, as understanding the scope of previous dose-ranging trials may be of use in determining the cause of suboptimal dosing.

Adenovirus vectored vaccines have been widely investigated for their ability to induce antibody and T cell responses against infectious diseases and cancers [13]. However, the dose-response for adenoviral vectored vaccines has not yet been systematically investigated. In this systematic review, we aim to explore and collate available adenoviral dose-response data for the purpose of informing adenoviral dosing towards safer and more effective vaccination. Our objectives were to:

1. Assess the number of available papers, including adenoviral dose-response studies, and the distribution of host species and adenoviral serotypes within these papers.
2. Assess which immunological responses dose-response data were available.
3. Assess the dosing strategies used in adenoviral dose-ranging studies, including number and magnitude of dose levels.

This systematic review should help inform adenoviral vaccine developers in choosing dose amounts for first-in-human trials. The collated data on dose-response, and replicating incompetent adenovirus-based vaccines, will also be used to inform IS/ID modelling studies for vaccine dose optimisation.

2. Materials and Methods

The study protocol was registered in PROSPERO (CRD42017080183).

2.1. Study Types, Study Design, Population, Intervention and Outcome Measures

Papers on clinical trials and in-vivo pre-clinical studies, that presented data from adenovirus vector-induced immunogenicity, were included in the review. These could include data from humans and animals of any age, sex and genetic background who received adenoviral vectored vaccines administered intramuscularly or subcutaneously. We did not assess study design aspects, such as methods of randomisation or use of control groups. The primary outcome measures were humoral and cellular immunity.

2.2. Search Strategy

The MEDLINE (PubMed) database was searched from inception to 27 November 2018. The search was limited to papers published in English and included terms relating to the following concepts: Adenovirus-vectored vaccines, immunogenicity, and dose-response (Appendix A, Criteria A1).

2.3. Paper Selection (Inclusion/Exclusion Criteria)

A three-stage screening process was used to systematically screen retrieved references and assess whether they met the inclusion criteria. Papers were first screened by title then by abstract before a full-text screen was conducted (Appendix A, Figure A1).

We included papers that presented data from studies with immunological response at three or more dose levels of an adenoviral vectored vaccine. We included papers that captured CD4+/CD8+ T-cell response, as measured by cytokine release using either ELISPOT or multiparameter flow cytometry and/or humoral responses, including binding and neutralising antibody titres against the vector and antigen. Exclusion criteria were chosen to minimize the probability of response being altered by a non-dosal effect, for example excluding cancer models and prime-boost paradigm vaccination (Appendix A, Figure A2).

2.4. Data Extraction

Using a pre-designed data extraction spreadsheet, information relating to study characteristics were extracted from studies that met the inclusion criteria. Numerical data from figures were extracted using GraphClick version 2.9.2 (Arizona Software, Los Angeles, CA, USA). Papers could contain data from multiple dose-response studies, and these studies may vary in adenoviral serotype, route of administration, host species, or disease.

2.5. Assessment of Methodological Quality

Bias was controlled for by having two individuals participate in the original search, and on abstract review. A review of 10 articles known to be relevant was conducted, to evaluate the degree of completeness. No statistical methods were performed to assess publication bias.

2.6. Comparing Doses

Three different units of measurement of dose were used in the extracted studies; viral particles (VPs), particle units (PUs), or plaque forming units (PFUs). Doses measured in VPs and PUs were considered equivalent as they both measure the number of physical viral particles [14]. PFUs were considered a separate outcome, as the ratio of VPs/PUs to PFUs were not constant in adenoviral vaccines studies [15].

3. Results

3.1. Objective 1: Assess the Number of Available Papers Including Adenoviral Dose-Response Studies, and the Distribution of Host Species and Adenoviral Serotypes within These Papers

Following removal of duplicate entries, 2787 references remained and were screened by title. 581 references were screened by abstract and 300 were screened by full text. After evaluation of the full text, 265 of the papers were excluded. Therefore, 35 papers were included in this review [16–50]. The majority of papers contained studies conducted in mice (60%), followed by humans (26%) (Table 1). Although, it is likely that many studies may have been carried out by industry using the same construct in mice and humans, the number of published studies using the same adenoviral strain, route and antigen insert across different species was limited.

Table 1. The number of papers that included dose-response studies for each host species identified in the review.

Number of Papers (%)	Host	Paper References
21 (60%)	Mouse	[18,22,24–26,29,32–36,39–41,43–45,47–50]
9 (26%)	Human	[16,17,19–21,23,27,28,30]
2 (6%)	Monkey	[38,42]
2 (6%)	Rat	[37,45]
1 (3%)	Rabbit	[31]
1 (3%)	Cattle	[46]

Out of all the adenoviral serotypes, the most common was human adenovirus 5 (46%), followed by human adenovirus 35 (26%) (Table 2). The route of administration was more frequently intramuscular (84%) than subcutaneous (16%).

Table 2. The number and percentage of papers that included dose-response data for each adenovirus serotype.

Human	Non-Human Primates
Ad5 (16, 46%) [22,24,25,29,31,33,34,38,39,42–46,49,50]	ChAd63 (3, 9%) [24,30,43]
Ad35 (9, 26%) [19,21,23,24,26,28,35,36,39]	ChAd3 (3, 9%) [16,17,24]
Ad26 (4, 11%) [26,27,32,37]	AdC6 (1, 3%) [41]
Ad6 (2, 6%) [38,47]	AdC7 (1, 3%) [40]
Ad28 (1, 3%) [24]	sAd11 (1, 3%) [24]
	sAd16 (1, 3%) [24]
	sAdv-36 (1, 3%) [48]
	ChAdOx1 (1, 3%) [20]

3.2. Objective 2: Assess for Which Immunological Responses Dose-Response Data Were Available

The immunogenicity data recorded also varied widely among the published studies, including antibody responses (both binding and neutralizing), T cell ELISpot data, and CD4+ and CD8+ T cell responses by intracellular cytokine staining. The majority of papers (51%) included studies of antibody dose-response. (Table 3).

Table 3. The number and percentage of papers that included dose-response data for each immunological response type.

Number of Papers (%)	Response Type	Paper References
18 (51%)	Antibody	[16,17,20–23,25–28,31,33,36,39,40,42,45,48]
12 (34%)	T cell count	[16,20,21,26–28,30,32,36,38,42,47]
12 (34%)	CD8+ T cell count	[19,22,24,32,34–36,38,39,48–50]
11 (31%)	Virus Neutralisation Titre	[22,25,27,29,30,34,36,37,43,44,46]
4 (12%)	CD4+ T cell count	[19,32,35,38]
3 (9%)	CD8+ T Cell, IFN-y+ Percentage	[19,21,41]
3 (9%)	CD4+ T Cell, IFN-y+ Percentage	[19,21]
2 (6%)	CD4+ T Cell, TNF-a+ Percentage	[19,21]
2 (6%)	CD8+ T Cell, TNF-a+ Percentage	[21]
2 (6%)	CD4+ T Cell, IL-2+ Percentage	[19,21]
2 (6%)	CD8+ T Cell, Il-2+ Percentage	[21]
1 (3%)	CD4+ T Cell, Il-17+ Percentage	[19]

3.3. Objective 3: Assess the Dosing Strategies Used in Adenoviral Dose-Ranging Studies, Including Number and Magnitude of Dose Levels

3.3.1. Number of Dose Levels

The majority of papers (60%) included studies with three dose levels, which was the minimum number of dose levels for a study to be included; 23% included four dose levels, and 20% included five or more levels (Table 4).

Table 4. The number and percentage of papers containing studies at each number of dosing levels.

Number of Papers (%)	Number of Dose Levels	Paper References
21 (60%)	3	[17,19,22,24,25,27–30,32,34,35,37,38,40,42,44,46,48,49]
8 (23%)	4	[16,20,21,23,26,36,40,43]
5 (14%)	5	[18,33,39,41,45]
1 (3%)	6	[31]
1 (3%)	7	[47]

3.3.2. Magnitude of Dose Levels

For VP/PU, the geometric mean human dose was 1.6×10^{10} (range 5×10^8–2×10^{11}) (Figure 1a). No human dose-response studies were measured in PFU. In VP, the geometric mean mouse dose was 4.9×10^7, (range 5×10^1–5×10^{11}) (Figure 1b). The mean human dose was therefore approximately 3.2×10^2 times larger than the mean mouse dose. Four mouse dose-response studies measured dose in PFU, with doses ranging between 1×10^4 and 1×10^9 PFU. Details on the magnitude of dose levels are found in Appendix A, Figure A2.
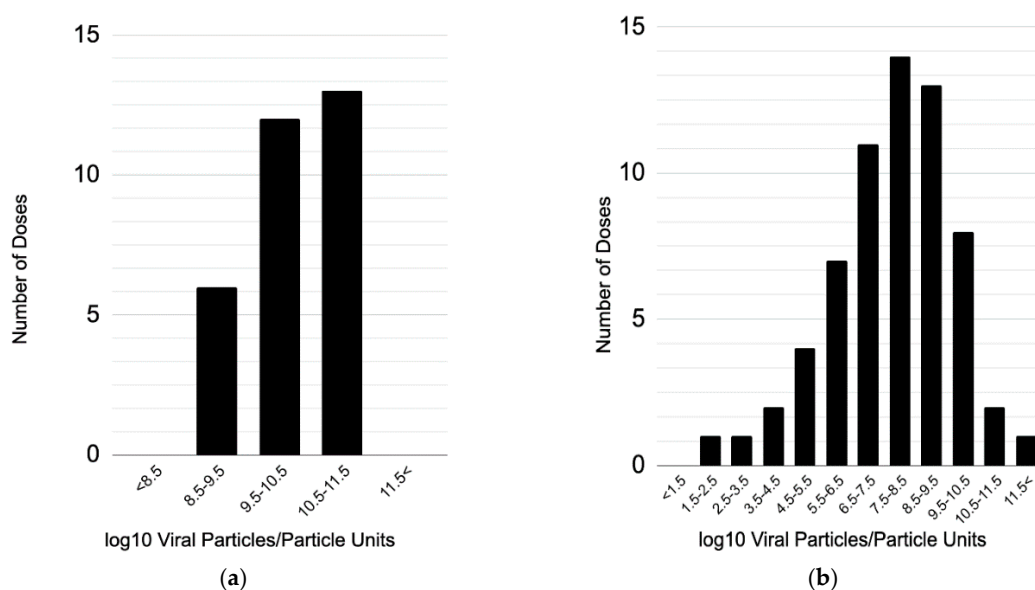
Figure 1. Frequency of dose magnitudes for all adenoviral vector vaccine dose-response studies with doses measured in VP/PFU in (**a**) humans and (**b**) mice.

4. Discussion

In this review we aimed to collate data on adenovirus-based vaccines in preparation for mathematical modelling to characterize the dose-response curve by host species, serogroup of the adenovirus or route of administration. After screening, 35 papers were extracted that provided dose-response immunogenicity data following intramuscularly or subcutaneously administered adenovirus vectors in animals and humans. Data were primarily from mouse and human studies, and included multiple different response types. From the adenoviral dose-response papers considered, studies typically used three dose levels, with the average human dose being two orders of magnitude larger than the average mouse dose. There were unfortunately very few comparator trials in which the same vaccine was used in human and animal models, and much of the pre-clinical data from larger industry companies are unlikely to have been published.

This review represents the first attempt to collate vaccine dose-response data, which has not yet been done for adenoviral vectored or non-adenovirus vectored vaccines. The review found that dose-response data existed for a wide range of immunological responses, both humoral and cellular. This suggests that published dose-response data may exist for many important correlates of protection. The broad spectrum of available data will be used to inform an IS/ID modelling study on adenoviral dose-response curve shape. However, the majority of studies used too few doses to allow for true dose response relationships to be clearly established, and thus, the majority of studies conducted are not sufficient to allow the authors to clearly justify their dose selection. To establish a true sigmoidal curve fit, at least five data points are needed to accurately model the response.

Whilst this review was able to identify 35 papers that may be useful in understanding adenoviral vectored vaccine dose-response behaviour, there are factors that may limit the utility of the collated

data. Firstly, it is likely that there exist vaccine dose-response studies that have not been published [51]. In order to predict dose-response in humans from animals for future vaccines, dose-response data from an existing vaccine in multiple species is required. The unavailability of these data may hinder attempts to develop an allometric scaling approach, therefore publishing of both clinical and pre-clinical dose-response studies is of great importance. Secondly, most of the doses were measured in viral particles, which may be a sub-optimal measure [52] as the infectious ratio, the number of viral particles per infectious unit, can vary between vaccines [15]. Therefore, the use of VP in measuring vaccine dose limits the comparisons in dose between different vaccines. Finally, when applying IS/ID modelling to define the dose-response curve, it is also possible that, whilst three dosing levels may be sufficient to theoretically define simple curves like a sigmoid function, this may not be a large enough number of doses to determine dose-response behaviour with an appropriate degree of certainty.

The strategies used to optimise vaccine dosing are likely to be suboptimal. There might exist mathematical descriptions of dose-response that are informative when choosing the various doses to use for a given construct in a given species which have not yet been identified. Indeed, both Darrah and Belovsky have shown that the highest dose was not the most effective dose for adenoviral vectored vaccination against *Leishmania* and non-adenovirus vectored vaccination against HIV [53,54], and yet the bias to choose the maximum safe dose remains among most vaccine developers.

New methods of vaccine dose optimisation need to be developed. Understanding adenoviral vaccine dose-response may be able to be achieved through reviewing and comparing historical dose-response data and combining these with mathematical modelling methods. This may aid in ensuring that the optimal dose for protection and safety is identified, while minimising the number of human and animal participants required to decide that dose.

Appendix A

Criteria A1. Search Terms, PubMed

Strategy	PUBMED Search	#
Concept 1 Adenovirus	adenovirus OR adenoviral OR adenovector OR adenovectors OR adenoviridae	1
Concept 2 Dose	dose OR doses OR dosage OR dosages OR dosing OR dosed OR dose response OR dose-response OR dose responses OR dose-responses OR dose response relationship OR dose-response relationship	2
Concept 3 Immune response	immunity OR immune OR immune-response OR immune response OR immune responses OR immune-responses OR immunostimulation OR immunodynamic OR immunodynamics OR immunisation OR immunisations OR immunization OR immunizations OR immunise OR immunises OR immunize OR immunizes OR immunised OR immunized OR immunising OR immunizing OR immunogenecity OR immunogenic OR immunology	3
Combine with AND	#1 AND #2 AND #3	
Add filter:	Humans, Other Animals, English	

Criteria A2. Exclusion Criteria

A study was excluded if it; (a) was a gene transfer study, (b) was conducted in cancer models, (c) presented no immunological readouts, (d) used replication-competent adenovirus, (e) did not administer adenovirus via an intramuscular or subcutaneous route, (f) used adenovirus as a boost in a heterologous prime-boost vaccination regimen, (g) used adenovirus as a prime, in a heterologous prime-boost regimen, and did not report on immunological parameters post-prime and pre-boost, (h) presented a homologous dosing regimen with no reported immunological parameters after the initial dose, (i) only reported on immune parameters following a disease challenge, (j) co-administered an adjuvant, administered an adjuvant prior to adenovirus delivery or used an ad-juvant-encoded adenovirus vector, (k) only reported on pulmonary immunity to the adenovirus (l) presented only data on gene expression, (m) used a sample size of less than five mice per group, or less than three for non-human primates, (n) presented in-vitro derived data, or (o) was a systematic review or meta-analysis.
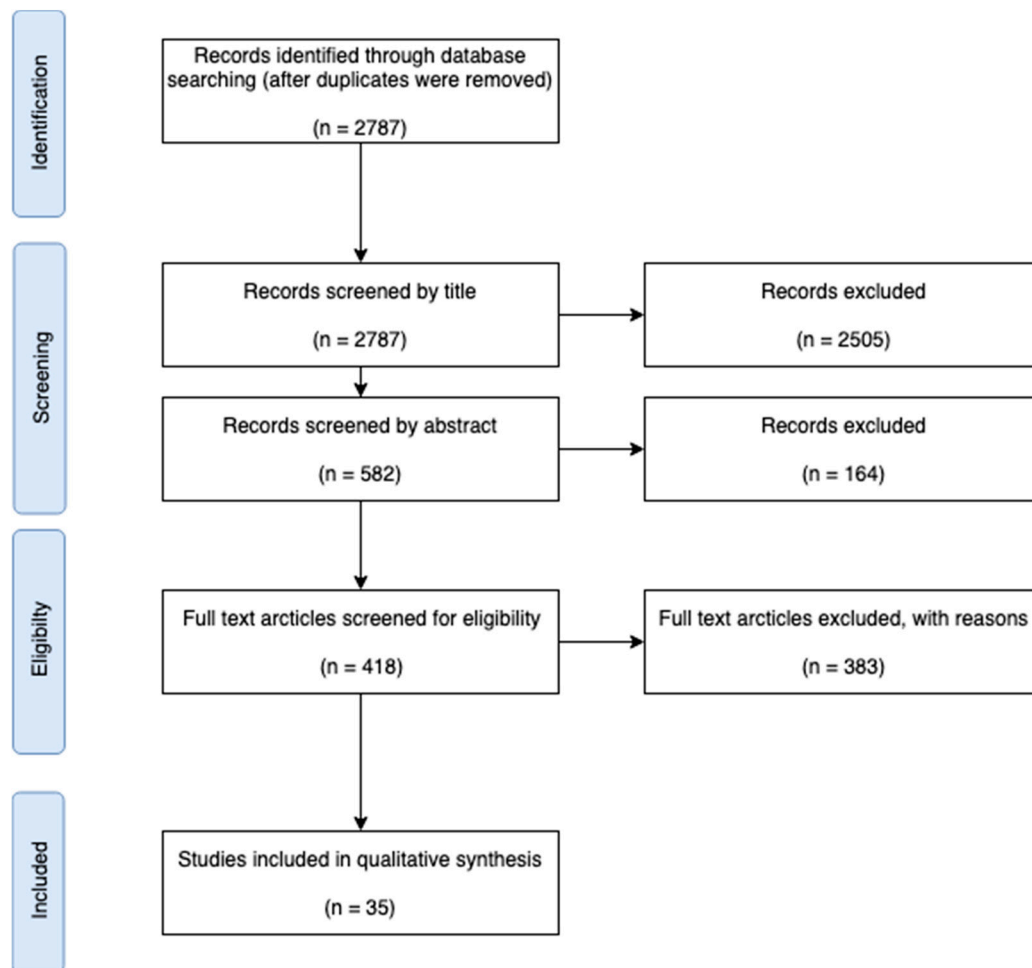
Figure A1. Flow chart diagram of the search and screening strategy.

Paper	Ref	Species	Dose 1	Dose 2	Dose 3	Dose 4	Dose 5	Dose 6	Dose 7	Units
Antrobus, 2014	20	Human	5.00E+08	5.00E+09	2.50E+10	5.00E+10				VP
Baden, 2013	27	Human	1.00E+09	1.00E+10	1.00E+11					VP
Capone, 2006	38	Monkey	1.00E+08	1.00E+10	1.00E+11					VP
Casimiro, 2003	42	Monkey	1.00E+07	1.00E+09	1.00E+11					VP
Creech, 2013	23	Human	1.00E+09	1.00E+10	5.00E+10	1.00E+11				VP
Eloit, 1995	45	Mouse	1.00E+05	1.00E+06	1.00E+07	1.00E+08	1.00E+09			VP
Eloit, 1995	45	Rat	1.00E+05	1.00E+06	1.00E+07	1.00E+08	1.00E+09			VP
Ewer, 2016	17	Human	1.00E+10	2.50E+10	5.00E+10					VP
Fonseca, 2017	48	Mouse	1.00E+06	1.00E+07	1.00E+10					VP
Hashimoto, 2005	40	Mouse	1.00E+08	1.00E+09	1.00E+10	1.00E+11				PU
Kagina, 2014	19	Human	1.50E+09	1.50E+10	1.00E+11					VP
Keefer, 2012	28	Human	2.00E+09	2.00E+10	2.00E+11					VP
Nazerai, 2016	49	Mouse	2.00E+05	2.00E+06	2.00E+07					PFU
Neilan, 2018	46	Cattle	3.13E+09	1.25E+10	5.00E+10					PU
O. J. a. E. Ophorst, 2006	39	Mouse	1.00E+06	1.00E+07	1.00E+08	1.00E+09	1.00E+10			VP
O. J. A. E. Ophorst, 2007	36	Mouse	1.00E+07	1.00E+08	1.00E+09	1.00E+10				VP
O'Hara, 2012	30	Human	1.00E+10	5.00E+10	2.00E+11					VP
Ondondo, 2014	18	Mouse	1.00E+05	1.00E+06	1.00E+07	1.00E+08	1.00E+09			VP
Ouédraogo, 2013	21	Human	1.00E+09	1.00E+10	5.00E+10	1.00E+11				VP
Pandey, 2012	29	Mouse	1.00E+07	1.00E+08	1.00E+09					PFU
Penaloza-MacMaster, 2016	50	Mouse	1.00E+08	1.00E+09	1.00E+10					VP
Pinto, 2003	41	Mouse	5.00E+07	5.00E+08	5.00E+09	5.00E+10	5.00E+11			VP
Quinn, 2013	24	Mouse	1.00E+07	1.00E+08	1.00E+09					PU
Rhee, 2010	32	Mouse	3.00E+07	3.00E+08	3.00E+09					VP
Richardson, 2009	34	Mouse	1.00E+04	1.00E+05	1.00E+06					VP
Rodríguez, 2008	35	Mouse	1.00E+08	1.00E+09	1.00E+10					VP
Steitz, 2010	33	Mouse	5.00E+01	5.00E+02	5.00E+03	5.00E+04	5.00E+05			VP
Sun, 2011	31	Rabbit	1.00E+04	1.50E+05	1.25E+05	2.50E+05	5.00E+05	1.00E+06		TCID50
Tang, 2017	47	Mouse	1.00E+06	4.00E+06	1.00E+07	4.00E+07	1.00E+08	4.00E+08	1.00E+09	VP
Tapia, 2016	16	Human	1.00E+10	2.50E+10	5.00E+10	1.00E+11				PU
Vemula, 2013	22	Mouse	2.00E+07	1.00E+08	5.00E+08					VP
Widjojoatmodjo, 2015	37	Rat	1.00E+06	1.00E+08	1.00E+09					VP
Z. Q. Xiang, 1996	44	Mouse	1.00E+02	1.00E+03	1.00E+04					PFU
Z. Q. Xiang, 1996	44	Mouse	2.00E+04	2.00E+05	2.00E+06					PFU
Z. Xiang, 2002	43	Mouse	1.00E+04	1.00E+05	1.00E+06	1.00E+07				PFU
Zahn, 2012	26	Mouse	1.00E+07	1.00E+08	1.00E+09	1.00E+10				VP
Zhou, 2013	25	Mouse	5.00E+06	5.00E+07	1.00E+08					PFU

Figure A2. Dose magnitudes for the collected studies, including the name of the paper, the host species, and the unit used to quantify dose.

References

1. Davis, H.L. Novel vaccines and adjuvant systems: The utility of animal models for predicting immunogenicity in humans. *Hum. Vaccines* **2008**, *4*, 246–250. [CrossRef] [PubMed]

2. Nair, A.B.; Jacob, S. A simple practice guide for dose conversion between animals and human. *J. Basic Clin. Pharm.* **2016**, *7*, 27–31. [CrossRef] [PubMed]

3. Campi-Azevedo, A.C.; de Almeida Estevam, P.; Coelho-dos-Reis, J.G.; Peruhype-Magalhães, V.; Villela-Rezende, G.; Quaresma, P.F.; de Lourdes Sousa Maia, M.; Farias, R.H.G.; Camacho, L.A.B.; da Silva Freire, M.; et al. Subdoses of 17DD yellow fever vaccine elicit equivalent virological/immunological kinetics timeline. *BMC Infect. Dis.* **2014**, *14*, 391. [CrossRef] [PubMed]

4. Guerin, P.J.; Naess, L.M.; Fogg, C.; Rosenqvist, E.; Pinoges, L.; Bajunirwe, F.; Nabasumba, C.; Borrow, R.; Frøholm, L.O.; Ghabri, S.; et al. Immunogenicity of fractional doses of tetravalent a/c/y/w135 meningococcal polysaccharide vaccine: Results from a randomized non-inferiority controlled trial in Uganda. *PLoS ONE Negl. Trop. Dis.* **2008**, *2*, e342. [CrossRef]

5.   Martins, R.M.; Maia, M.D.L.S.; Farias, R.H.G.; Camacho, L.A.B.; Freire, M.S.; Galler, R.; Yamamura, A.M.Y.; Almeida, L.F.C.; Lima, S.M.B.; Nogueira, R.M.R.; et al. 17DD yellow fever vaccine: A double blind, randomized clinical trial of immunogenicity and safety on a dose-response study. *Hum. Vaccines Immunother.* **2013**, *9*, 879–888. [CrossRef] [PubMed]

6.   Regules, J.A.; Cicatelli, S.B.; Bennett, J.W.; Paolino, K.M.; Twomey, P.S.; Moon, J.E.; Kathcart, A.K.; Hauns, K.D.; Komisar, J.L.; Qabar, A.N.; et al. Fractional Third and Fourth Dose of RTS,S/AS01 Malaria Candidate Vaccine: A Phase 2a Controlled Human Malaria Parasite Infection and Immunogenicity Study. *J. Infect. Dis.* **2016**, *214*, 762–771. [CrossRef]

7.   Rhodes, S.J.; Zelmer, A.; Knight, G.M.; Prabowo, S.A.; Stockdale, L.; Evans, T.G.; Lindenstrøm, T.; White, R.G.; Fletcher, H. The TB vaccine H56 + IC31 dose-response curve is peaked not saturating: Data generation for new mathematical modelling methods to inform vaccine dose decisions. *Vaccine* **2016**, *34*, 6285–6291. [CrossRef]

8.   Rhodes, S.J.; Knight, G.M.; Kirschner, D.E.; White, R.G.; Evans, T.G. Dose finding for new vaccines: The role for immunostimulation/immunodynamic modelling. *arXiv* **2018**, arXiv:1811.04024. [CrossRef]

9.   Rhodes, S.J.; Guedj, J.; Fletcher, H.A.; Lindenstrøm, T.; Scriba, T.J.; Evans, T.G.; Knight, G.M.; White, R.G. Using vaccine Immunostimulation/Immunodynamic modelling methods to inform vaccine dose decision-making. *Npj Vaccines* **2018**, *3*, 36. [CrossRef]

10.   Rhodes, S.J.; Sarfas, C.; Knight, G.M.; White, A.; Pathan, A.A.; McShane, H.; Evans, T.G.; Fletcher, H.; Sharpe, S.; White, R.G. Using Data from Macaques to Predict Gamma Interferon Responses after Mycobacterium bovis BCG Vaccination in Humans: A Proof-of-Concept Study of Immunostimulation/Immunodynamic Modeling Methods. *Clin. Vaccine Immunol.* **2017**, *24*. [CrossRef]

11.   Rhodes, S.J.; Knight, G.M.; Kirschner, D.E.; White, R.G.; Evans, T.G. Dose finding for new vaccines: The role for immunostimulation/immunodynamic modelling. *J. Theor. Biol.* **2019**, *465*, 51–55. [CrossRef] [PubMed]

12.   Handel, A.; Li, Y.; McKay, B.; Pawelek, K.A.; Zarnitsyna, V.; Antia, R. Exploring the impact of inoculum dose on host immunity and morbidity to inform model-based vaccine design. *PLoS ONE Comput. Biol.* **2018**, *14*, e1006505. [CrossRef] [PubMed]

13.   Zhang, C.; Zhou, D. Adenoviral vector-based strategies against infectious disease and cancer. *Hum. Vaccines Immunother.* **2016**, *12*, 2064–2074. [CrossRef] [PubMed]

14.   Sheets, R.L.; Stein, J.; Bailer, R.T.; Koup, R.A.; Andrews, C.; Nason, M.; He, B.; Koo, E.; Trotter, H.; Duffy, C.; et al. Biodistribution and Toxicological Safety of Adenovirus Type 5 and Type 35 Vectored Vaccines Against Human Immunodeficiency Virus-1 (HIV-1), Ebola, or Marburg Are Similar Despite Differing Adenovirus Serotype Vector, Manufacturer's Construct, or Gene Inserts. *J. Immunotoxicol.* **2008**, *5*, 315–335. [PubMed]

15.   Klasse, P.J. Molecular Determinants of the Ratio of Inert to Infectious Virus Particles. *Prog. Mol. Biol. Transl. Sci.* **2015**, *129*, 285–326. [PubMed]

16.   Tapia, M.D.; Sow, S.O.; Lyke, K.E.; Haidara, F.C.; Diallo, F.; Doumbia, M.; Traore, A.; Coulibaly, F.; Kodio, M.; Onwuchekwa, U.; et al. Use of ChAd3-EBO-Z Ebola virus vaccine in Malian and US adults, and boosting of Malian adults with MVA-BN-Filo: A phase 1, single-blind, randomised trial, a phase 1b, open-label and double-blind, dose-escalation trial, and a nested, randomised, double-blind, placebo-controlled trial. *Lancet Infect. Dis.* **2016**, *16*, 31–42.

17.   Ewer, K.; Rampling, T.; Venkatraman, N.; Bowyer, G.; Wright, D.; Lambe, T.; Imoukhuede, E.B.; Payne, R.; Fehling, S.K.; Strecker, T.; et al. A Monovalent Chimpanzee Adenovirus Ebola Vaccine Boosted with MVA. *N. Engl. J. Med.* **2016**, *374*, 1635–1646. [CrossRef]

18.   Ondondo, B.; Abdul-Jawad, S.; Bridgeman, A.; Hanke, T. Characterization of T-Cell Responses to Conserved Regions of the HIV-1 Proteome in BALB/c Mice. *Clin. Vaccine Immunol. CVI* **2014**, *21*, 1565–1572. [CrossRef]

19.   Kagina, B.M.N.; Tameris, M.D.; Geldenhuys, H.; Hatherill, M.; Abel, B.; Hussey, G.D.; Scriba, T.J.; Mahomed, H.; Sadoff, J.C.; Hanekom, W.A.; et al. The novel tuberculosis vaccine, AERAS-402, is safe in healthy infants previously vaccinated with BCG, and induces dose-dependent CD4 and CD8T cell responses. *Vaccine* **2014**, *32*, 5908–5917. [CrossRef]

20.   Antrobus, R.D.; Coughlan, L.; Berthoud, T.K.; Dicks, M.D.; Hill, A.V.; Lambe, T.; Gilbert, S.C. Clinical assessment of a novel recombinant simian adenovirus ChAdOx1 as a vectored vaccine expressing conserved Influenza A antigens. *Mol. Ther. J. Am. Soc. Gene Ther.* **2014**, *22*, 668–674. [CrossRef]

21. Ouédraogo, A.; Tiono, A.B.; Kargougou, D.; Yaro, J.B.; Ouédraogo, E.; Kaboré, Y.; Kangoye, D.; Bougouma, E.C.; Gansane, A.; Henri, N.; et al. A phase 1b randomized, controlled, double-blinded dosage-escalation trial to evaluate the safety, reactogenicity and immunogenicity of an adenovirus type 35 based circumsporozoite malaria vaccine in Burkinabe healthy adults 18 to 45 years of age. *PLoS ONE* **2013**, *8*, e78679. [CrossRef] [PubMed]

22. Vemula, S.V.; Amen, O.; Katz, J.M.; Donis, R.; Sambhara, S.; Mittal, S.K. Beta-defensin 2 enhances immunogenicity and protection of an adenovirus-based H5N1 influenza vaccine at an early time. *Virus Res.* **2013**, *178*, 398–403. [CrossRef] [PubMed]

23. Creech, C.B.; Dekker, C.L.; Ho, D.; Phillips, S.; Mackey, S.; Murray-Krezan, C.; Grazia Pau, M.; Hendriks, J.; Brown, V.; Dally, L.G.; et al. Randomized, placebo-controlled trial to assess the safety and immunogenicity of an adenovirus type 35-based circumsporozoite malaria vaccine in healthy adults. *Hum. Vaccines Immunother.* **2013**, *9*, 2548–2557. [CrossRef] [PubMed]

24. Quinn, K.M.; Da Costa, A.; Yamamoto, A.; Berry, D.; Lindsay, R.W.B.; Darrah, P.A.; Wang, L.; Cheng, C.; Kong, W.-P.; Gall, J.G.D.; et al. Comparative analysis of the magnitude, quality, phenotype, and protective capacity of simian immunodeficiency virus gag-specific CD8+ T cells following human-, simian-, and chimpanzee-derived recombinant adenoviral vector immunization. *J. Immunol. Baltim. Md 1950* **2013**, *190*, 2720–2735. [CrossRef] [PubMed]

25. Zhou, G.; Wang, H.; Wang, F.; Yu, L. Recombinant adenovirus expressing type Asia1 foot-and-mouth disease virus capsid proteins induces protective immunity against homologous virus challenge in mice. *Res. Vet. Sci.* **2013**, *94*, 796–802. [CrossRef]

26. Zahn, R.; Gillisen, G.; Roos, A.; Koning, M.; van der Helm, E.; Spek, D.; Weijtens, M.; Grazia Pau, M.; Radošević, K.; Weverling, G.J.; et al. Ad35 and ad26 vaccine vectors induce potent and cross-reactive antibody and T-cell responses to multiple filovirus species. *PLoS ONE* **2012**, *7*, e44115. [CrossRef]

27. Baden, L.R.; Walsh, S.R.; Seaman, M.S.; Tucker, R.P.; Krause, K.H.; Patel, A.; Johnson, J.A.; Kleinjan, J.; Yanosick, K.E.; Perry, J.; et al. First-in-human evaluation of the safety and immunogenicity of a recombinant adenovirus serotype 26 HIV-1 Env vaccine (IPCAVD 001). *J. Infect. Dis.* **2013**, *207*, 240–247. [CrossRef]

28. Keefer, M.C.; Gilmour, J.; Hayes, P.; Gill, D.; Kopycinski, J.; Cheeseman, H.; Cashin-Cox, M.; Naarding, M.; Clark, L.; Fernandez, N.; et al. A phase I double blind, placebo-controlled, randomized study of a multigenic HIV-1 adenovirus subtype 35 vector vaccine in healthy uninfected adults. *PLoS ONE* **2012**, *7*, e41936. [CrossRef]

29. Pandey, A.; Singh, N.; Vemula, S.V.; Couëtil, L.; Katz, J.M.; Donis, R.; Sambhara, S.; Mittal, S.K. Impact of preexisting adenovirus vector immunity on immunogenicity and protection conferred with an adenovirus-based H5N1 influenza vaccine. *PLoS ONE* **2012**, *7*, e33428. [CrossRef]

30. O'Hara, G.A.; Duncan, C.J.A.; Ewer, K.J.; Collins, K.A.; Elias, S.C.; Halstead, F.D.; Goodman, A.L.; Edwards, N.J.; Reyes-Sandoval, A.; Bird, P.; et al. Clinical assessment of a recombinant simian adenovirus ChAd63: A potent new vaccine vector. *J. Infect. Dis.* **2012**, *205*, 772–781. [CrossRef]

31. Sun, Y.; Li, H.-Y.; Tian, D.-Y.; Han, Q.-Y.; Zhang, X.; Li, N.; Qiu, H.-J. A novel alphavirus replicon-vectored vaccine delivered by adenovirus induces sterile immunity against classical swine fever. *Vaccine* **2011**, *29*, 8364–8372. [CrossRef] [PubMed]

32. Rhee, E.G.; Kelley, R.P.; Agarwal, I.; Lynch, D.M.; Porte, A.L.; Simmons, N.L.; Clark, S.L.; Barouch, D.H. TLR4 Ligands Augment Antigen-Specific CD8+ T Lymphocyte Responses Elicited by a Viral Vaccine Vector. *J. Virol.* **2010**, *84*, 10413–10419. [CrossRef] [PubMed]

33. Steitz, J.; Barlow, P.G.; Hossain, J.; Kim, E.; Okada, K.; Kenniston, T.; Rea, S.; Donis, R.O.; Gambotto, A. A Candidate H1N1 Pandemic Influenza Vaccine Elicits Protective Immunity in Mice. *PLoS ONE* **2010**, 5. [CrossRef] [PubMed]

34. Richardson, J.S.; Yao, M.K.; Tran, K.N.; Croyle, M.A.; Strong, J.E.; Feldmann, H.; Kobinger, G.P. Enhanced protection against Ebola virus mediated by an improved adenovirus-based vaccine. *PLoS ONE* **2009**, *4*, e5308. [CrossRef]

35. Rodrigues, E.G.; Zavala, F.; Eichinger, D.; Wilson, J.M.; Tsuji, M. Single immunizing dose of recombinant adenovirus efficiently induces CD8+ T cell-mediated protective immunity against malaria. *J. Immunol. Baltim. Md 1950* **1997**, *158*, 1268–1274.

36. Ophorst, O.J.A.E.; Radosević, K.; Klap, J.M.; Sijtsma, J.; Gillissen, G.; Mintardjo, R.; van Ooij, M.J.M.; Holterman, L.; Companjen, A.; Goudsmit, J.; et al. Increased immunogenicity of recombinant Ad35-based malaria vaccine through formulation with aluminium phosphate adjuvant. *Vaccine* **2007**, *25*, 6501–6510. [CrossRef]

37. Widjojoatmodjo, M.N.; Bogaert, L.; Meek, B.; Zahn, R.; Vellinga, J.; Custers, J.; Serroyen, J.; Radošević, K.; Schuitemaker, H. Recombinant low-seroprevalent adenoviral vectors Ad26 and Ad35 expressing the respiratory syncytial virus (RSV) fusion protein induce protective immunity against RSV infection in cotton rats. *Vaccine* **2015**, *33*, 5406–5414. [CrossRef]

38. Capone, S.; Meola, A.; Ercole, B.B.; Vitelli, A.; Pezzanera, M.; Ruggeri, L.; Davies, M.E.; Tafi, R.; Santini, C.; Luzzago, A.; et al. A novel adenovirus type 6 (Ad6)-based hepatitis C virus vector that overcomes preexisting anti-ad5 immunity and induces potent and broad cellular immune responses in rhesus macaques. *J. Virol.* **2006**, *80*, 1688–1699. [CrossRef]

39. Ophorst, O.J.A.E.; Radošević, K.; Havenga, M.J.E.; Pau, M.G.; Holterman, L.; Berkhout, B.; Goudsmit, J.; Tsuji, M. Immunogenicity and Protection of a Recombinant Human Adenovirus Serotype 35-Based Malaria Vaccine against Plasmodium yoelii in Mice. *Infect. Immun.* **2006**, *74*, 313–320. [CrossRef]

40. Hashimoto, M.; Boyer, J.L.; Hackett, N.R.; Wilson, J.M.; Crystal, R.G. Induction of Protective Immunity to Anthrax Lethal Toxin with a Nonhuman Primate Adenovirus-Based Vaccine in the Presence of Preexisting Anti-Human Adenovirus Immunity. *Infect. Immun.* **2005**, *73*, 6885–6891. [CrossRef]

41. Pinto, A.R.; Fitzgerald, J.C.; Giles-Davis, W.; Gao, G.P.; Wilson, J.M.; Ertl, H.C.J. Induction of CD8+ T cells to an HIV-1 antigen through a prime boost regimen with heterologous E1-deleted adenoviral vaccine carriers. *J. Immunol. Baltim. Md 1950* **2003**, *171*, 6774–6779. [CrossRef] [PubMed]

42. Casimiro, D.R.; Chen, L.; Fu, T.-M.; Evans, R.K.; Caulfield, M.J.; Davies, M.-E.; Tang, A.; Chen, M.; Huang, L.; Harris, V.; et al. Comparative immunogenicity in rhesus monkeys of DNA plasmid, recombinant vaccinia virus, and replication-defective adenovirus vectors expressing a human immunodeficiency virus type 1 gag gene. *J. Virol.* **2003**, *77*, 6305–6313. [CrossRef] [PubMed]

43. Xiang, Z.; Gao, G.; Reyes-Sandoval, A.; Cohen, C.J.; Li, Y.; Bergelson, J.M.; Wilson, J.M.; Ertl, H.C.J. Novel, Chimpanzee Serotype 68-Based Adenoviral Vaccine Carrier for Induction of Antibodies to a Transgene Product. *J. Virol.* **2002**, *76*, 2667–2675. [CrossRef] [PubMed]

44. Xiang, Z.Q.; Yang, Y.; Wilson, J.M.; Ertl, H.C. A replication-defective human adenovirus recombinant serves as a highly efficacious vaccine carrier. *Virology* **1996**, *219*, 220–227. [CrossRef]

45. Eloit, M.; Adam, M. Isogenic adenoviruses type 5 expressing or not expressing the E1A gene: Efficiency as virus vectors in the vaccination of permissive and non-permissive species. *J. Gen. Virol.* **1995**, *76*, 1583–1589. [CrossRef] [PubMed]

46. Neilan, J.G.; Schutta, C.; Barrera, J.; Pisano, M.; Zsak, L.; Hartwig, E.; Rasmussen, M.V.; Kamicker, B.J.; Ettyreddy, D.; Brough, D.E.; et al. Efficacy of an adenovirus-vectored foot-and-mouth disease virus serotype A subunit vaccine in cattle using a direct contact transmission model. *BMC Vet. Res.* **2018**, *14*. [CrossRef]

47. Tang, A.; Freed, D.C.; Li, F.; Meschino, S.; Prokop, M.; Bett, A.; Casimiro, D.; Wang, D.; Fu, T.-M. Functionally inactivated dominant viral antigens of human cytomegalovirus delivered in replication incompetent adenovirus type 6 vectors as vaccine candidates. *Hum. Vaccines Immunother.* **2017**, *13*, 2763–2771. [CrossRef]

48. Fonseca, J.A.; McCaffery, J.N.; Kashentseva, E.; Singh, B.; Dmitriev, I.; Curiel, D.T.; Moreno, A. A prime-boost immunization regimen based on a Simian Adenovirus 36 vectored multi-stage malaria vaccine induces protective immunity in mice. *Vaccine* **2017**, *35*, 3239–3248. [CrossRef]

49. Nazerai, L.; Bassi, M.R.; Uddback, I.E.M.; Holst, P.J.; Christensen, J.P.; Thomsen, A.R. Early life vaccination: Generation of adult-quality memory CD8+ T cells in infant mice using non-replicating adenoviral vectors. *Sci. Rep.* **2016**, *6*, 38666. [CrossRef]

50. Penaloza-MacMaster, P.; Alayo, Q.A.; Ra, J.; Provine, N.M.; Larocca, R.; Lee, B.; Barouch, D.H. Inhibitory receptor expression on memory CD8 T cells following Ad vector immunization. *Vaccine* **2016**, *34*, 4955–4963. [CrossRef]

51. Manzoli, L.; Flacco, M.E.; D'Addario, M.; Capasso, L.; Vito, C.D.; Marzuillo, C.; Villari, P.; Ioannidis, J.P.A. Non-publication and delayed publication of randomized trials on vaccines: Survey. *BMJ* **2014**, *348*. [CrossRef] [PubMed]

52. Gallaher, S.D.; Berk, A.J. A rapid Q-PCR titration protocol for adenovirus and helper-dependent adenovirus vectors that produces biologically relevant results. *J. Virol. Methods* **2013**, *192*, 28–38. [CrossRef] [PubMed]

53. Darrah, P.A.; Patel, D.T.; De Luca, P.M.; Lindsay, R.W.B.; Davey, D.F.; Flynn, B.J.; Hoff, S.T.; Andersen, P.; Reed, S.G.; Morris, S.L.; et al. Multifunctional T H 1 cells define a correlate of vaccine-mediated protection against Leishmania major. *Nat. Med.* **2007**, *13*, 843–850. [CrossRef] [PubMed]

54. Billeskov, R.; Wang, Y.; Solaymani-Mohammadi, S.; Frey, B.; Kulkarni, S.; Andersen, P.; Agger, E.M.; Sui, Y.; Berzofsky, J.A. Low Antigen Dose in Adjuvant-Based Vaccination Selectively Induces CD4 T Cells with Enhanced Functional Avidity and Protective Efficacy. *J. Immunol. Baltim. Md 1950* **2017**, *198*, 3494–3506. [CrossRef]

Chapter 3. Analysis of adenoviral vector vaccine dose-response curves in prior studies: Paper 2

Chapter 3 Introduction

For chapter 3, I used the data that was gathered through the systematic review described in chapter 2/paper 1 to evaluate the prevalence of single-administration non-replicating adenoviral vector vaccine dose-response that was best described by representative peaking and saturating curve shapes. This was a modelling meta-analysis, using a 'model-selection' approach where the representative curves were calibrated to the available dose-response data, and the model that best described the data determined by the Akaike Information Criterion. This chapter addresses objective 2 of this thesis.

As I have noted, previous work has observed the potential for vaccine dose-immunogenicity response to be better described by peaking that saturating curves. The relevance of this observation to clinical vaccine development depends on the prevalence of 'peaking' or 'saturating' dose-response. For example, if peaking dose-response is rarely observed then the negative impacts of a saturating dose-response assumption in practical vaccine development may be minimal. Additionally, if the true prevalence of peaking/saturating dose-response could be determined, then this information could be used to inform priors for model averaging methods in future vaccine development [189]. Whilst such application across all vaccines would require a meta-analysis of as many vaccines as possible, I considered only single-administration non-replicating adenoviral vector vaccines. This was done to limit scope and to leverage the domain expertise in adenoviral vector vaccines that was available through my supervisory team.

I hypothesised given previous IS/ID findings that there would be available both dose-response data that were best described as peaking and that were best described as saturating [49,50]. If this was true, it may be relevant to be able to predict which dose-response curve shape will best describe dose-response for a novel vaccine.

Discussion with my supervisor Thomas Evans suggested that the vaccine attributes of immunological response type, host species, adenoviral species and route of administration may be attributes that could affect dose-response curve shape. Therefore, I also considered whether these attributes may be likely predictors of vaccine dose-response curve shape. In objective 2 of this chapter, I considered whether curve shape is typically the same for vaccines that are alike in all four of these attributes. In objective 3, I considered whether curve shape is typically the same for vaccines that were alike in three of these attributes to investigate which of these attributes that may be less likely to affect curve shape.

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed <u>for each</u> research paper included within a thesis.

SECTION A – Student Details

Student ID Number	lsh1804914	**Title**	Mr
First Name(s)	John Helier		
Surname/Family Name	Benest		
Thesis Title	Response Type and Host Species may be Sufficient to Predict Dose-Response Curve Shape for Adenoviral Vector Vaccines. Vaccines		
Primary Supervisor	Richard G. White		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?	MDPI Vaccines		
When was the work published?	30 March 2020		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	N.A		
Have you retained the copyright for the work?*	Yes	Was the work subject to academic peer review?	Yes

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	

Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	The conceptualisation of this work was conducted by Sophie Rhodes, Richard White, and I. I conducted all modelling and statistical analysis. I wrote the first draft of this work and performed re-drafting in line with co-author comments. All authors reviewed the paper. The interpretation of the results was my own work

SECTION E

Student Signature	███████████
Date	11 August 2022

Supervisor Signature	███████████
Date	12 August 2022

99

Paper 2 Title: Response Type and Host Species may be Sufficient to Predict Dose-Response Curve Shape for Adenoviral Vector Vaccines

Authors: John Benest, Sophie Rhodes, Sara Afrough, Thomas G. Evans, Richard G. White

Permission from copyright owner to include this work:

Response Type and Host Species may be Sufficient to Predict Dose-Response Curve Shape for Adenoviral Vector Vaccines

John Benest [1],*, Sophie Rhodes [1], Sara Afrough [2], Thomas Evans [2] and Richard White [1]

[1] Department of Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK; sophie.rhodes@lshtm.ac.uk (S.R.); Richard.White@lshtm.ac.uk (R.W.)

[2] Vaccitech Ltd., The Schrodinger Building, Heatley Road, The Oxford Science Park, Oxford OX4 4GE, UK; sara.afrough@UCB.com (S.A.); tom.evans@vaccitech.co.uk (T.E.)

* Correspondence: john.benest@lshtm.ac.uk

Abstract: Vaccine dose-response curves can follow both saturating and peaking shapes. Dose-response curves for adenoviral vector vaccines have not been systematically described. In this paper, we explore the dose-response shape of published adenoviral animal and human studies. Where data were informative, dose-response was approximately five times more likely to be peaking than saturating. There was evidence that host species and response type may be sufficient for prediction of dose-response curve shape. Dose-response curve shape prediction could decrease clinical trial costs, accelerating the development of life-saving vaccines.

Keywords: dosing; dose-response; adenovirus-vectored vaccines; dose dynamics

1. Introduction

Vaccination is effective globally at preventing disease and reducing morbidity and disability [1]. Over recent decades, adenovirus, a vaccine vector used for prophylactic and therapeutic vaccination, has been widely applied, due to both the safety and induction of specific antibodies and T-cells by adenoviral-vectored vaccines [2]. However, adenoviral vaccine developers must still take care to avoid potentially severe adverse effects whilst ensuring that the developed vaccines are efficacious and affordable [3].

A key step in implementing a new vaccine is optimisation of the dosing quantity (hereafter 'dose') [4]. As the dose per individual is increased, the cost per individual vaccinated and vaccine toxicity may also increase [5]. We might also assume that the protective efficacy of a vaccine may vary with dose. Optimising a vaccine dose requires establishing a dose that is protective, or induces the highest desired immunological response, whilst avoiding dose-dependent toxicity and minimising cost. Optimising adenoviral vaccines should therefore be approached using multi-objective optimisation methods.

To effectively optimise dose, the relationship between magnitude of dose and immunological response must first be understood. Qualitatively, one might assume that as a dose increases, two types of dose-response relationships may occur, saturating or peaking. A saturating relationship, usually referred to as a sigmoidal response, implies the response is strictly increasing as dose increases, but plateauing so that an increase in dose gives a negligible increase in response beyond a certain threshold. A peaking relationship implies that there exists some dose for which the response is maximised, and that an increase in dose would lead to a decreased response. Historically, pre-clinical trials typically have made the assumption of a saturating curve shape [6]; however, research has shown that for both tuberculosis and influenza vaccines a peaking shape may better describe the dose-response curve [7,8]. Adenoviral vaccine dose-response curve shape has not yet been established.

We aim to provide insight into adenoviral dose-response such that vaccine developers may better optimise adenoviral vector vaccines dosing. Our objectives were:

1) assessing the prevalence of peaking/saturating dose-response curve shapes in published adenoviral vector vaccine studies.
2) assessing whether dose-response curve shape may be predicted by response type, host species, adenoviral species and route of administration (RoA).
3) assessing which of host species, adenoviral species and RoA are the most likely predictors of dose-response curve shape.

Understanding these objectives may be key in predicting likely dose-response curve shape, reducing the number or trial subjects required to determine curve shape and therefore reducing cost for adenoviral dose-response trials.

2. Materials and Methods

2.1. Data Collation and Preparation

The data were identified in a systematic review of adenoviral dose-response [9], summarised here. In summary, PubMed was searched systematically using terms related to the concepts of adenovirus vector vaccines, immunogenicity, and dose-response through 23 November 2018. Inclusion required prime response data for replication-defective adenovirus vector vaccines with intramuscular (IM) or subcutaneous (SQ) RoA. We excluded cancer models and data where the vaccine was adjuvant coadministered or was recorded post challenge or boost. Non-primate studies were excluded if there were less than 5 animals per dose, and non-human primate data excluded if there were less than 3 individuals per dose. As both representative curves had three unknown parameters, only studies with at least three non-placebo dose-response datapoints were included.

Data collation included vaccine name, vaccine backbone, host species, RoA, response type, response units, time point, whether the response was a summary statistic or individual and the endnote reference for the paper. Studies that provided dose-response data from multiple experiments, and/or multiple time points within the same experiment, were included as separate datasets. In almost all cases the titres were reported as viral particles (VPs), which are less informative in cross study comparisons than infectious units (IUs) [10]. The ratio of VP:IU ranges from 20:1 to 150:1 but is usually found in the 40–80 range.

2.2. Objective 1: Assessing the Prevalence of Peaking/Saturating Dose-Response Curve Shapes in Published Adenoviral Vector Vaccine Studies

In this objective we chose representative peaking and saturating curves and calibrated both curves to the dose-response data. Goodness-of-fit tests were used to determine which curve shape best described each trial dataset, and the prevalence of "peaking" or "saturating" curves across the whole dataset was calculated.

2.2.1. Representative Curves

A sigmoidal and gamma probability density function (PDF) were chosen as the representative dose-response curves for saturating and peaking behaviour respectively (Figure 1). See the Appendix A for the curve equations (Equations (A1) and (A2)).
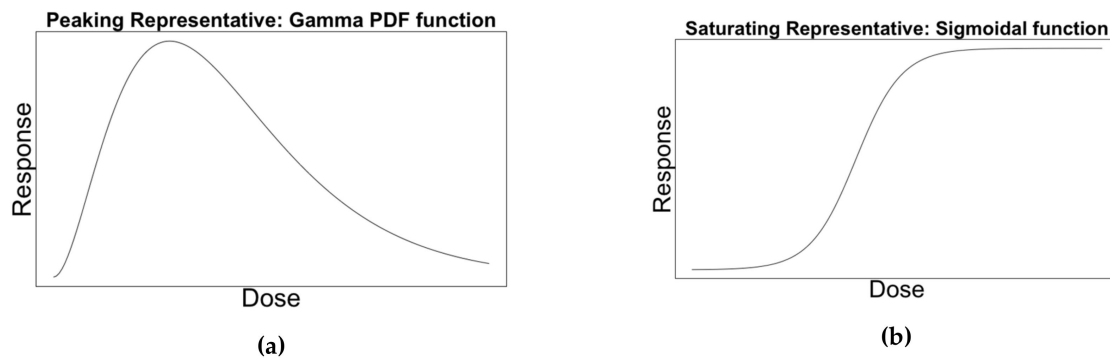
Figure 1. Examples of a representative curve for (**a**) peaking and (**b**) saturating behaviour.

2.2.2. Calibrating Curves to Data

Calibration of the curves to the data was done by finding the parameter estimates that minimise squared error iteratively. Specifically, the nls and nls2 functions in R were used [11,12], which used a brute force algorithm to find reasonable starting parameters and the nl2sol algorithm to further minimise the squared error [13].

The Akaike Information Criterion (AIC) was calculated for both the calibrated peaking and saturating curve for each dataset. The curve shape with the lower AIC was defined as best describing the dataset [14]. The absolute difference between the AIC between peaking and saturating curves for a given dataset was defined as ΔAIC. ΔAIC was used to calculate the support a dataset had for one either curve shape, as follows [14]:

- Provides no evidence, ΔAIC <2
- Positive evidence, 2 ≤ ΔAIC < 6
- Strong evidence, 6 ≤ ΔAIC < 10
- Very strong evidence 10 ≤ ΔAIC

A dataset where ΔAIC was greater than or equal to 2 was defined as a dataset "providing evidence" towards either peaking/saturating curve shape and not providing evidence otherwise. Figure 2 shows an example of calibrated peaking and saturating curves and their respective AICs for a three datasets, one for which no curve is superior to the other (ΔAIC <2), and one each for which the peaking/saturating curve was superior (ΔAIC >2)

2.2.3. Calculating Dose-Response Curve Shape Prevalence

The prevalence of datasets providing evidence for either of the curve shapes were calculated. One-sampled, two-tailed t-tests were used to estimate the 95% confidence interval for the probability of a dataset providing evidence for peaking or saturating shape. We also calculated the prevalence of curve shapes by response type, and in only human studies.

2.2.4. Exploring Potential Bias of Independence Assumption

To explore if our conclusions on prevalence of peaking vs saturation curve share were robust to removing datasets from multiple time points with the same study, we repeated the analysis with each study contributing only one dataset.
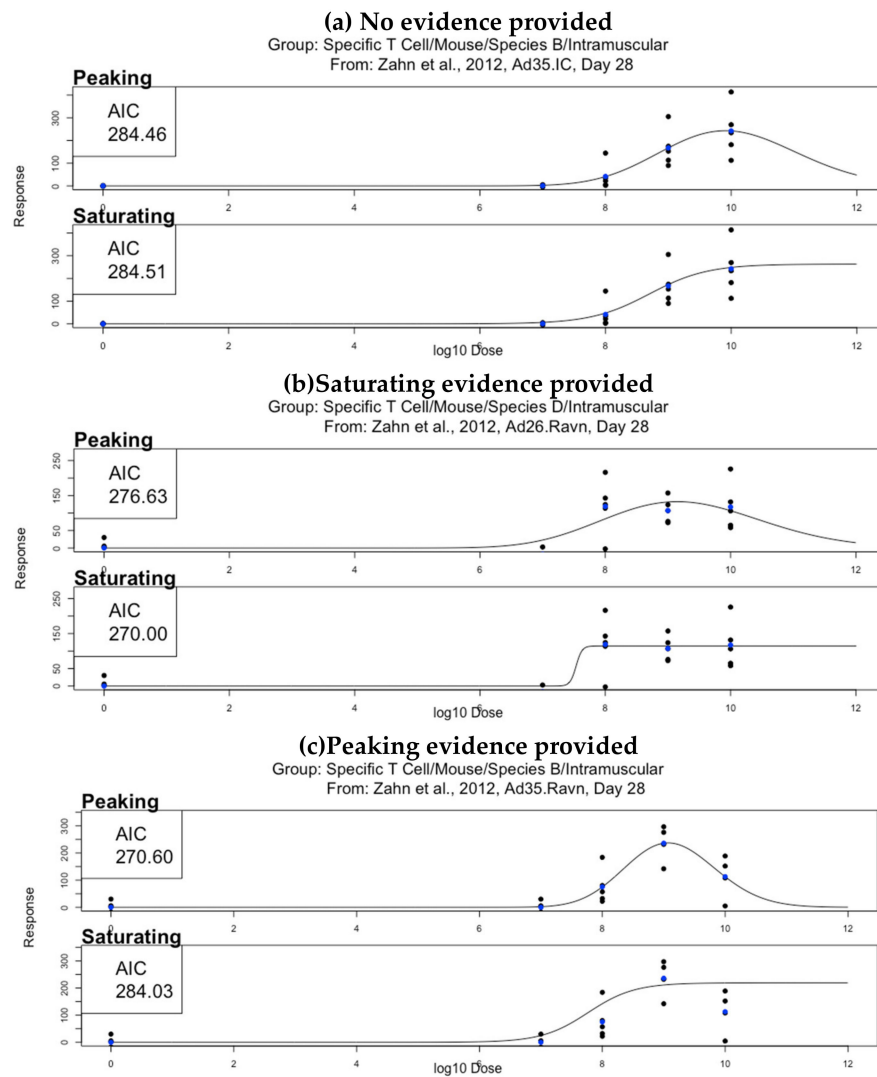
Figure 2. An example of three datasets; (**a**) does not provide evidence to support one curve shape over the other (i.e., ΔAIC = 0.05 < 2), (**b**) provides evidence to support a saturating curve shape over a peaking curve shape (i.e., ΔAIC = 6.63 > 2), and (**c**) provides evidence to support a peaking curve shape over a saturating curve shape (i.e., ΔAIC = 13.43 > 2). For each dataset, the top plot shows the calibrated peaking curve and associated Akaike Information Criterion (AIC), and the bottom plot shows the calibrated saturating curve and associated AIC. Black dots represent individual mice and blue dots represent the mean response for that dose.

2.3. Objective 2: Assessing whether Dose-Response Curve Shape may be Predicted by Response Type, Host Species, Adenoviral Species, and Route of Administration (RoA)

In this objective we grouped each dataset by response type, host species, adenoviral species and RoA. Within groups, consistency of curve shape was evaluated. This allowed us to assess whether these attributes could predict dose-response curve shape.

2.3.1. Grouping

We grouped each dataset by response type, host species, adenoviral species and RoA.

2.3.2. Evaluating Consistency

This analysis used groups with at least two datasets providing evidence (2 ≤ ΔAIC). A group was defined as consistent if all datasets within that group provided evidence for the same curve

shape, and inconsistent if not. A binomial test was used to determine the confidence interval for the probability of within group consistency.

2.4. Objective 3: Assessing which of Host Species, Adenoviral Species and RoA are the most likely Predictors of Dose-Response Curve Shape

In this objective we paired datasets that differed in one of the attributes host species, adenoviral species and RoA. Within pairs consistency of curve shape was evaluated. Pairwise consistency was evaluated stratified by the attribute that differed in the pair. This allowed us to assess whether a change in host species, adenoviral species or RoA would lead to a change in dose-response curve shape.

2.4.1. Pairing

Only datasets that provided evidence ($2 \leq \Delta AIC$) were considered in the analysis. We defined a pair as two datasets with the same response type and two of host species, adenoviral species and RoA the same (Figure 3).

		Route of Administration	
		IM	SQ
Host Species	Mouse	Group Mouse/IM	Group Mouse/SQ
	Human	Group Human/IM	Group Human/SQ
	Rat	Group Rat/IM	Group Rat/SQ
	Monkey	Group Monkey/IM	Group Monkey/SQ

Figure 3. A representation of the definition of a comparing across group, simplified to the 2 attributes of host and route of administration (RoA). For analysis of consistency of shape across groups, a dataset from group Mouse/IM (purple cell) would be paired with only datasets from group Mouse/SQ when predicting across RoA (red cells), and with only datasets from groups Human/IM, Rat/IM and Monkey/IM when predicting across host species (blue cells). This was extended to the full 4 attribute dimensions in the analysis.

2.4.2. Evaluating Consistency

A pair was defined as consistent if both datasets within that pair provided evidence for the same curve shape, and inconsistent if not. For each of the attributes of host species, adenoviral species and RoA an exact one-sided binomial test was conducted, with trials being the number of pairs for that attribute and successes being the number of consistent pairs for that attribute. This was used to determine the confidence interval for the probability that altering that attribute would not alter the dose-response curve shape. For example, an Antibody/Mouse/Species C/IM dataset could be paired with an Antibody/Mouse/Species C/SQ dataset to examine consistency when varying on RoA.

3. Results

3.1. Data

The systematic review identified 2787 references, reduced to 581 by title screening, then 300 by screening of the abstract. Screening by full text reduced the number of papers available for analysis to 35 [15–49]. These ranged across five Adenoviral species (B,C,D,E and G), six Host species

(Cattle, Human, Monkey, Mouse, Rabbit and Rat), and two routes of administration (IM and SQ). The data ranged across 12 different responses including antibody, T-cell and neutralization titre. The full list of response types considered are in Table 1, and the matrix of papers and their response type/hosts/adenoviral species/RoA can be found in Table A1. Details of the vectors in these papers, including species and origin, can be found in Supplementary S1a and S1b.

3.2. Objective 1: Assessing the Prevalence of Peaking/Saturating Dose-Response Curve Shapes in Published Adenoviral Vector Vaccine Studies

3.2.1. Overall Prevalence

In total, 191 datasets were extracted from the 35 papers (Supplementary Figure S1, Supplementary Tables S1–S12). Datasets on Antibody, T-cell, and CD8+ response data were the most common, each with 30+ datasets. 20+ datasets were available on virus neutralising titre.

Of the 191 datasets, 73.3% (140/191) did not provide evidence for either curve shape (Total that provided evidence, Table 1). Also, 22.0% (42/191) of datasets provided evidence for a peaking shape, and 4.7% (8/191) of datasets provided saturating evidence (total provided evidence = 26.7% (50/191)).

Of datasets that provided evidence for peaking or saturating curve shape, datasets were five times more probable to provide peaking evidence than saturating evidence. Using two-tailed binomial tests with 95% confidence intervals, we estimated that the true probability of a dataset providing peaking evidence across all data was 16.3% to 28.5%, versus 1.8% to 8.1% for saturating evidence.

3.2.2. Prevalence by Response Type

Similarly, the true probability for datasets providing very strong peaking evidence was in the interval 6.9% to 16.3%, compared to 0.0% to 2.9% for very strong saturating evidence [Table 1]. Antibody, T-cell, and Virus Neutralisation responses had datasets providing evidence for both peaking and saturating behaviour. All other responses provided evidence for peaking shape curve shapes only.

3.2.3. Prevalence in Human Data

37 datasets provided data on humans (Table 2). Of these, 56.8% were shown to provide no evidence for either curve shape, 43.2% provided evidence for a peaking shape and 0.0% provided evidence for a saturating shape.

3.2.4. Exploring Potential Bias of Independence Assumption

37 datasets were available after excluding multiple datasets from the same study. Our results were robust to this analysis, with the peaking to saturating evidence ratio remaining approximately 5:1.

3.3. Objective 2: Assessing whether Dose-Response Curve Shape may be Predicted by Response Type, Host Species, Adenoviral Species, and Route of Administration (RoA).

52/720 (7.2%) groups contained at least one dataset (Supplementary Tables S1–S12).

Evaluating Consistency

11/52 groups contained at least two datasets that provided evidence (ΔAIC>2) (Table 3). Of these groups 100% (11/11, 95%CI = 71.5–100%, p <0.001) were consistent, i.e., all datasets provided evidence for the same curve shape. Of the 11 groups, 18.2% (9/11) only had evidence towards a peaking shape and 81.8% (2/11) only had evidence towards a saturating shape.

Table 1. Number of datasets for each level of evidence, sorted by response type.

Response Type	Very Strong Peaking 10≤ ΔAIC	Strong Peaking 6≤ ΔAIC <10	Positive Peaking 2≤ ΔAIC <6	No Evidence ΔAIC<2	Positive Saturating 2≤ ΔAIC <6	Strong Saturating 6≤ ΔAIC <10	Very Strong Saturating 10≤ ΔAIC	Total Number of Datasets for Response
				Total Number of All Datasets by Level of Evidence and Response				
Antibody	3	0	2	46	1	1	1	54
T-Cell	1	2	0	27	3	1	0	34
CD4+	1	1	0	4	0	0	0	6
CD8+	11	3	7	42	0	0	0	63
CD4 IFNγ+	1	1	0	0	0	0	0	2
CD8 IFNγ+	1	0	0	2	0	0	0	3
CD4+ TNFα+	0	0	1	1	0	0	0	2
CD8+ TNFα+	0	0	1	0	0	0	0	1
CD4+ IL-2+	1	0	0	1	0	0	0	2
CD8+ IL-2+	0	0	1	0	0	0	0	1
CD4+ IL-17+	0	0	1	0	0	0	0	1
Virus Neutralization Titre	2	0	1	17	1	1	0	22
Total that provided evidence (%)	42 (22.0%)			140 (73.3%)	8 (4.7%)			191

The far left column indicates response for that row type, the far right column gives the response type distribution of the whole dataset. The other columns indicate type of evidence and the number of datasets found for that response type and type of evidence. The opacity of the red colouring was equal to the proportion of datasets for that row's response type for which that type of evidence was found, so a darker red indicates that for this response type this type of evidence was more prevalent than for a cell in that row with a lighter red. The opacity of the blue colouring for each response type in the far right column was equal to the proportion of total datasets which had that response type, so a darker blue indicates that this response type was more prevalent in the whole dataset than for a cell with a lighter blue. The final row summarises the number of datasets that provided peaking evidence, saturating evidence, or no evidence across all datasets.

Table 2. Number of datasets for each level of evidence, for human only data.

	Very Strong Peaking 10≤ΔAIC	Strong Peaking 6≤ΔAIC<10	Positive Peaking 2≤ΔAIC<6	No Evidence ΔAIC<2	Positive Saturating 2≤ΔAIC<6	Strong Saturating 6≤ΔAIC<10	Very Strong Saturating 10≤ΔAIC
			Total Number of Human Datasets by Associated Level of Evidence				
All responses (%)	6 (16.2%)	4 (10.8%)	6 (16.2%)	21 (56.8%)	0 (0%)	0 (0%)	0 (0%)
Total that provided evidence (%)	16 (43.2%)			21 (56.8%)	0 (0%)		

The opacity of the red colouring was proportional to the percentage of datasets that with that associated level of evidence, so a darker red indicates that this type of evidence was more prevalent than for a cell in that row with a lighter red.

Table 3. Summary of consistency of curve shape within groups.

Group (Response Type /Host/Adenoviral Species/Route of Administration)	Number of Datasets Providing Evidence Towards a Peaking Shape	Number of Datasets Providing Evidence Towards a Saturating Shape	Consistency
Antibodies/Human/B/IM	2	0	Consistent
Antibodies/Mouse/C/IM	2	0	Consistent
Antibodies/Monkey/C/IM	0	3	Consistent
T Cell/Mouse/C/IM	0	3	Consistent
CD4+/Human/B/IM	2	0	Consistent
CD8+/Human/B/IM	2	0	Consistent
CD8+/Mouse/C/IM	5	0	Consistent
CD8+/Mouse/C/SQ	6	0	Consistent
CD8+/Mouse/G/SQ	4	0	Consistent
CD4+ IFNγ+/Human/B/IM	2	0	Consistent
Virus Neutralization Titre/Mouse/C/IM	2	0	Consistent

B = species B adenoviral vector. C = species C adenoviral vector. G = species G adenoviral vector. IM = intramuscular RoA. SQ = subcutaneous RoA.

3.4. Objective 3: Assessing which of Host Species, Adenoviral Species, and RoA are the most likely Predictors of Dose-Response Curve Shape

3.4.1. Evaluating Pairwise Consistency for Host

Of the 50 datasets with evidence, we found 14 pairings such that only the host species was different (Figure 4). 5/14 pairs (27.3%) were consistent. An exact one-sided binomial test with 5 successes in 14 trials gave the 95% confidence interval for the probability of a pairing being consistent as 15.3% to 100.0% with *p*-value = 0.91. This was not considered significant evidence to support predicting curve shape across host species.
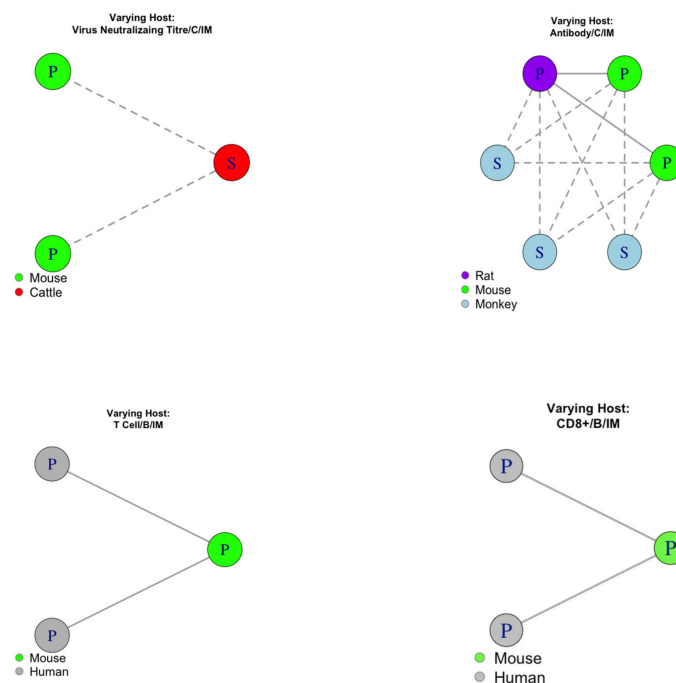
Figure 4. Pairings that differed in host species, examined for analysis of consistency between groups. Each point represents a dataset that provided evidence, where P/S denotes the dataset had evidence towards a peaking/saturating shape, respectively. Datasets were colour coded according to their host. A solid line shows consistent pairings and a dashed line shows inconsistent pairings. B = species B adenoviral vector. C = species C adenoviral vector. IM = Intramuscular route of administration.

3.4.2. Evaluating pairwise consistency for adenoviral species

Of the 50 datasets with evidence, we found 64 pairings such that only the adenoviral species was different (Figure 5), and 60/64 pairs (93.8%) were consistent. An exact one-sided binomial test with 60 successes in 64 trials gave the 95% confidence interval for the probability of a pairing being consistent as 86.3% to 100% with *p*-value < 0.0001. This was considered significant evidence to support predicting curve shape across adenoviral species.
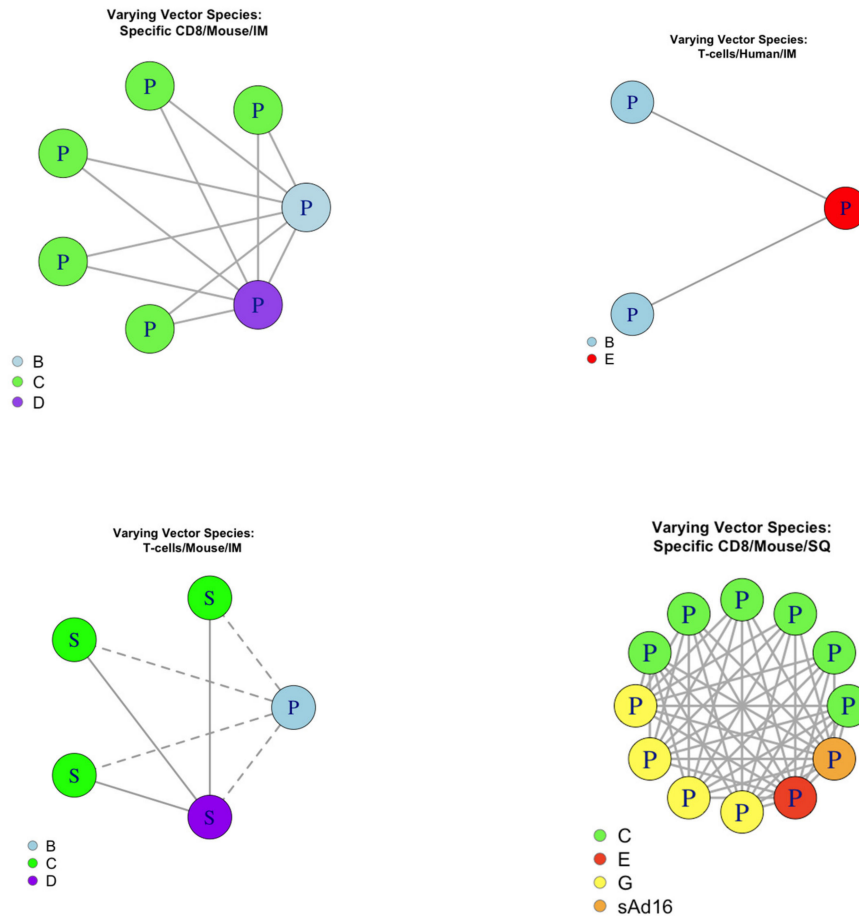
Figure 5. Pairings that differed in host species examined for analysis of consistency between groups. Each point represents a dataset that provided evidence, where P/S denotes the dataset had evidence towards a peaking/saturating shape, respectively. Datasets were colour coded according to their vector species. A solid line shows consistent pairings and a dashed line shows inconsistent pairings. B = species B adenoviral vector. C = species C adenoviral vector. D = species D adenoviral vector. E = species E adenoviral vector. G = species G adenoviral vector. sAd16 = adenoviral vector was sAd16. IM = Intramuscular route of administration. SQ = Subcutaneous route of administration.

3.4.3. Evaluating pairwise consistency for RoA

Of the 50 datasets with evidence, we found 31 pairings such that only the RoA was different (Figure 6). 30/30 of these pairings (100%) were consistent. A one-sided binomial test with 31 successes in 31 trials gave the 95% confidence interval for the probability of a pairing being consistent as 90.8% to 100%, with *p*-value < 0.001. This was considered significant evidence to support predicting curve shape across RoA.
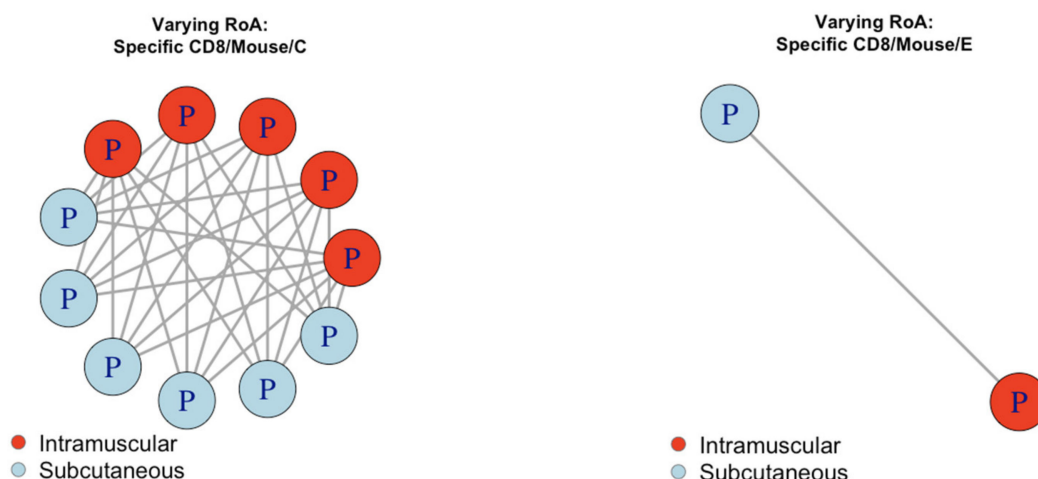
Figure 6. Pairings that differed in RoA species examined for analysis of consistency between groups. Each point represents a dataset that provided evidence, where P/S denotes the dataset had evidence towards a peaking/saturating shape, respectively. Datasets were colour coded according to their Route of Administration. A solid line shows consistent pairings and a dashed line shows inconsistent pairings. C = species C adenoviral vector. E = species E adenoviral vector.

4. Discussion

In this work, peaking and saturating curves were fit to adenoviral data to determine dose-response curve shape overall, and potential adenovirus vaccine attributes were assessed for their potential to determine curve shape. The results suggested that evidence towards a peaked adenoviral dose-response curve occurred five times more frequently than for a saturating curve. Curve shape was consistent within groups, suggesting curve shape was determined by the combined attributes of response type, host species, adenoviral species and RoA. There was strong evidence to support curve shape prediction across adenoviral species and RoA, but not host species.

This study is the first of its kind in the field of adenoviral dose optimization, exploring broad scale patterns in adenoviral dose-response across a large number of different candidate vaccines. The broadness of the data allowed for exploration of the effect of a wide range of vaccine attributes on adenoviral vaccine dose-response, which has not been previously attempted. To our knowledge, a systematic review to extract and evaluate vaccine dose-response data has never been done and this is the first adenoviral example of describing dose-response curve shape.

This work represents the use of an approach towards using quantitative methods to better optimise vaccine dosing. We have presented a rigorous statistical investigation into the prevalence of peaking and saturating curve shapes and into the potential of using vaccine study attributes to predict curve shape. Applying mathematical and statistical methods is novel in adenoviral dose optimisation and was shown to be a reasonable approach towards predicting adenoviral dose response curve shape, a potential advance from more empirical approaches to dose optimisation. Additionally, the calibration of curves and shape analysis did not require specialised software or complex mathematical models, meaning that a similar methodology could be implemented easily in future clinical trials, informed by this work.

There were weaknesses and limitations to our work. Firstly, while we used data covering a broad spectrum of published adenoviral studies, data was relatively sparse, and studies predominantly used mouse hosts and species C adenoviral vectors. Secondly, for 73% of the data, no evidence for either curve shape could be determined. A potential cause for this was that, for some of the datasets, the number of doses analysed was too few or doses too similar in magnitude. It was rare for doses to cross more than three $\log_{10}$, which is likely needed to see large differences. Even for datasets where evidence supporting one of the curve shapes was found, further empirical data collection may be

useful to confirm the model classifications. Thirdly, it may be possible that neither the simple peaking or saturating shape are optimal to describe adenovirus dose-response curves and there may exist more complex curves that better describe dose-response data for specific groups. However, these simple curves were deemed biologically plausible and chosen to analyse the consistency within and between groups in order to make predictions where possible. Further assessment of curve fitting could be considered if there were enough data sets with greater than 4 points over multiple logs of doses. Fourthly, there was a lot of heterogeneity in the data sets. For example, the extracted data structure was heterogeneous with some datasets on the individual level and others summarised across individuals. The heterogeneity limited the modelling approaches that could be applied to the data.

The methods of this work align well with the new field of Immunostimulation/Immunodynamic (IS/ID) modelling [4]. This field suggests modelling as a useful approach towards understanding how a vaccine's efficacy depends on dose. Similar to our work, other IS/ID modelling studies have fit statistical curves and mathematical models to dose-response data for novel tuberculosis vaccine, H56+IC31 [8] and parainfluenza/influenza vaccines HPIV-3 and IAV [7] to determine curve shape. For H56+IC31 and HPIV-3, peaking curve shape was found to be a better description of the dose-response curve shape compared to a saturating shape [7,8], consistent with our results for adenoviral vaccines. However, the influenza vaccine, IAV, was found to be saturating [7].

We found that for adenoviral vaccines, peaking dose-response curve shapes are more prevalent than saturating shapes. Dose escalation studies may therefore not be an appropriate method of adenoviral dose optimisation, as they typically assume a saturating dose-response curve shape and select the maximal safe dose. As such, we recommend that adenoviral dose-response studies should include analysis of immunogenic outcomes chosen by their likelihood of association with protective responses, and not just toxicology, to optimise dose.

In objectives 2 and 3 we showed it may be possible to predict adenoviral dose-response curve shape using data from similar adenoviral dose-response studies. We suggested the likely attributes that would need to be the same to justify that prediction, i.e., response type, host species, adenoviral species, and RoA. We also showed that host species and response type may be sufficient for prediction of dose-response curve shape. Hence, if the dose-response curve shape is known for one vaccine for a given host species and response type, we may be able to inform and predict dose response curve shape for another vaccine regardless of differences in adenoviral species and RoA.

We could not explore the effects of other potential attributes that could influence dose-response curve shape. Stratifying by pathogen would have reduced group size, preventing analysis of curve shape within group. There exist other attributes that some of the data did not measure, which could similarly not be used to establish impact on dose response curve shape. Whilst this is a limitation of our work, we still found groups to be 100% consistent. This may imply that attributes that were not explored in this work are not required for predicting curve shape. This would support prediction of dose-response curve shape using dose-response data from vaccine targeted against a different pathogen. We did not stratify by whether adenovirus vectors were simian or human derived. This could be future areas of adenoviral dose-response curve research.

Though we have shown that prediction of adenoviral vaccine dose-response may be possible, validating such predictions would be an important area of future research. Which attributes cause the shape to be peaking and which cause the shape to be saturating should also be determined, as we assessed consistency of shape not the causal pathway in which attributes may alter dose response curve shape. Understanding the attribute causation of curve shape could result in the development of a tool that can suggest the probability that a novel vaccine will have a peaking or saturating curve shape. Aggregation of future adenoviral dose-response studies with those analysed here would aid in developing such a tool. Even without such a tool, this work suggests that to predict curve shape all that is required is to have a previously established adenoviral dose response curve shape for that response type and species.

Whilst this work showed evidence that vaccine dose-response curve shapes may be either peaking or saturating, a method of optimally designing vaccine trials to best determine dose-response curve shape is yet unknown. This is another area where future research could lead to better dose optimisation.

Statistical models, like those in this paper, allow shape to be determined. With sufficient data, a mechanistic model, like those used in [7,8], may allow for a more accurate prediction of optimal dose and deeper understanding of longitudinal dose-response, and such approaches merit further investigation in an adenoviral context.

5. Conclusions

We found that where data were informative, dose-response was approximately five times more likely to be peaking than saturating. We also found that there was evidence that host species and response type may be sufficient for prediction of dose-response curve shape. We found that where data were informative, dose-response was approximately five times more likely to be peaking than saturating. We also found that there was evidence that host species and response type may be sufficient for prediction of dose-response curve shape.

Appendix A

Equations A1 and A2

Functions of x for the Equation (A1) sigmoidal and Equation (A2) gamma PDF functions, with parameters a,b,c, and Γ being the gamma function.

$$f(x) = \left(\frac{a}{\left(1 + e^{(b(c-x))}\right)} \right), \tag{A1}$$

$$f(x) = \frac{a}{\Gamma(b)c} \left(\frac{x}{c}\right)^{b-1} e^{-\frac{x}{c}}, \tag{A2}$$

Table A1. Table of papers analysed and the respective response types, vector species, host species and RoA analysed in those papers. Also included is a reference list for adenoviral papers from which data were extracted. Papers were stored and denoted by their identification number in the bibliography software used and are referred to as such in the supplementary document of datasets.

Papers	Antibodies	T cell	CD4+	CD8+	%CD4+ that is IFN+	%CD8+ that is IFN+	%CD4+ that is TNF+	%CD8+ that is TNF+	%CD4+ that is IL2+	%CD8+ that is Il-2+	%CD4+ that is Il-17+	Virus Neutralization Titre	B	C	D	E	G	None	Rabbit	Mouse	Human	Cattle	Monkey	Rat	SQ	IM
140 [15]	1													1						1						1
249 [16]	1													1						1						1
305 [17]		1													1						1					1
309 [18]			1	1	1	1	1		1		1		1							1						1
417 [19]	1	1													1					1						1
441 [20]	1	1			1	1	1	1	1	1			1							1						1
461 [21]	1			1								1		1						1						1
467 [22]	1											1		1						1						1
555 [23]				1									1	1	1	1	1	1		1			1			
574 [24]	1											1		1						1						1
578 [25]	1	1												1	1					1						1
594 [26]	1	1										1			1						1					1
633 [27]	1	1											1							1						1
669 [28]												1		1							1					1
686 [29]		1										1			1						1					1
744 [30]	1													1								1				1
924 [31]		1	1	1										1						1						1
936 [32]	1													1						1				1		
1039 [33]				1								1								1						1
1201 [34]			1	1									1							1						1
1269 [35]	1	1		1								1	1							1						1
1343 [36]												1			1								1			1
1474 [37]		1	1	1										1									1			1
1492 [38]	1			1									1	1						1						1
1539 [39]	1															1				1						1
1801 [40]						1										1				1						1
1877 [41]	1	1												1									1			1
2030 [42]												1	1		1					1					1	
2505 [43]												1	1							1					1	
2531 [44]	1													1						1			1			1
2841 [45]												1		1								1				1
2916 [46]		1												1						1						1
2919 [47]	1			1											1					1						1
2980 [48]				1										1						1				1		
3018 [49]				1										1						1						1
Sum	18	12	4	12	2	3	2	1	2	1	1	11	9	18	4	9	1	1	1	21	9	1	2	2	5	

References

1. Andre, F.; Booy, R.; Bock, H.; Clemens, J.; Datta, S.; John, T.; Lee, B.; Lolekha, S.; Peltola, H.; Ruff, T.; et al. Vaccination greatly reduces disease, disability, death and inequity worldwide. *Bull. World Health Organ.* **2008**, *86*, 140–146. [CrossRef]

2. Fougeroux, C.; Holst, P.J. Future Prospects for the Development of Cost-Effective Adenovirus Vaccines. *Int. J. Mol. Sci.* **2017**, *18*. [CrossRef] [PubMed]

3. Zhang, C.; Zhou, D. Adenoviral vector-based strategies against infectious disease and cancer. *Hum. Vaccines Immunother.* **2016**, *12*, 2064–2074. [CrossRef] [PubMed]

4. Rhodes, S.J.; Knight, G.M.; Kirschner, D.E.; White, R.G.; Evans, T.G. Dose finding for new vaccines: The role for immunostimulation/immunodynamic modelling. *J. Theor. Biol.* **2018**. ArXiv181104024 Q-Bio. [CrossRef] [PubMed]

5. Beals, C.R.; Railkar, R.A.; Schaeffer, A.K.; Levin, Y.; Kochba, E.; Meyer, B.K.; Evans, R.K.; Sheldon, E.A.; Lasseter, K.; Lang, N.; et al. Immune response and reactogenicity of intradermal administration versus subcutaneous administration of varicella-zoster virus vaccine: An exploratory, randomised, partly blinded trial. *Lancet Infect. Dis.* **2016**, *16*, 915–922. [CrossRef]

6. Rhodes, S.J.; Guedj, J.; Fletcher, H.A.; Lindenstrøm, T.; Scriba, T.J.; Evans, T.G.; Knight, G.M.; White, R.G. Using vaccine Immunostimulation/Immunodynamic modelling methods to inform vaccine dose decision-making. *NPJ Vaccines* **2018**, *3*, 36. [CrossRef]

7. Handel, A.; Li, Y.; McKay, B.; Pawelek, K.A.; Zarnitsyna, V.; Antia, R. Exploring the impact of inoculum dose on host immunity and morbidity to inform model-based vaccine design. *PLOS Comput. Biol.* **2018**, *14*, e1006505. [CrossRef]

8. Rhodes, S.J.; Zelmer, A.; Knight, G.M.; Prabowo, S.A.; Stockdale, L.; Evans, T.G.; Lindenstrøm, T.; White, R.G.; Fletcher, H. The TB vaccine H56+IC31 dose-response curve is peaked not saturating: Data generation for new mathematical modelling methods to inform vaccine dose decisions. *Vaccine* **2016**, *34*, 6285–6291. [CrossRef]

9. Afrough, S.; Rhodes, S.; Evans, T.; White, R.; Benest, J. Immunologic Dose-Response to Adenovirus-Vectored Vaccines in Animals and Humans: A Systematic Review of Dose-Response Studies of Replication Incompetent Adenoviral Vaccine Vectors when Given via an Intramuscular or Subcutaneous Route. *Vaccines* **2020**, *8*, 131. [CrossRef]

10. Dicks, M.D.J.; Spencer, A.J.; Edwards, N.J.; Wadell, G.; Bojang, K.; Gilbert, S.C.; Hill, A.V.S.; Cottingham, M.G. A Novel Chimpanzee Adenovirus Vector with Low Human Seroprevalence: Improved Systems for Vector Derivation and Comparative Immunogenicity. *PLoS ONE* **2012**, *7*. [CrossRef]

11. Grothendieck, G. nls2: Non-linear regression with brute force; R package version 0.2. 2013. Available online: https://CRAN.R-project.org/package=nls2 (accessed on 5 January 2019).

12. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2018.

13. Ciaburro, G. *Regression Analysis with R: Design and Develop Statistical Nodes to Identify Unique Relationships Within Data at Scale*; Packt Publishing Ltd.: Birmingham, UK, 2018; ISBN 978-1-78862-270-7.

14. Raftery, A.E. Bayesian Model Selection in Social Research. *Sociol. Methodol.* **1995**, *25*, 111–163. [CrossRef]

15. Tapia, M.D.; Sow, S.O.; Lyke, K.E.; Haidara, F.C.; Diallo, F.; Doumbia, M.; Traore, A.; Coulibaly, F.; Kodio, M.; Onwuchekwa, U.; et al. Use of ChAd3-EBO-Z Ebola virus vaccine in Malian and US adults, and boosting of Malian adults with MVA-BN-Filo: A phase 1, single-blind, randomised trial, a phase 1b, open-label and double-blind, dose-escalation trial, and a nested, randomised, double-blind, placebo-controlled trial. *Lancet Infect. Dis.* **2016**, *16*, 31–42. [PubMed]

16. Ewer, K.; Rampling, T.; Venkatraman, N.; Bowyer, G.; Wright, D.; Lambe, T.; Imoukhuede, E.B.; Payne, R.; Fehling, S.K.; Strecker, T.; et al. A Monovalent Chimpanzee Adenovirus Ebola Vaccine Boosted with MVA. *N. Engl. J. Med.* **2016**, *374*, 1635–1646. [CrossRef] [PubMed]

17. Ondondo, B.; Abdul-Jawad, S.; Bridgeman, A.; Hanke, T. Characterization of T-Cell Responses to Conserved Regions of the HIV-1 Proteome in BALB/c Mice. *Clin. Vaccine Immunol. CVI* **2014**, *21*, 1565–1572. [CrossRef]

18. Kagina, B.M.N.; Tameris, M.D.; Geldenhuys, H.; Hatherill, M.; Abel, B.; Hussey, G.D.; Scriba, T.J.; Mahomed, H.; Sadoff, J.C.; Hanekom, W.A.; et al. The novel tuberculosis vaccine, AERAS-402, is safe in healthy infants previously vaccinated with BCG, and induces dose-dependent CD4 and CD8T cell responses. *Vaccine* **2014**, *32*, 5908–5917. [CrossRef]

19. Antrobus, R.D.; Coughlan, L.; Berthoud, T.K.; Dicks, M.D.; Hill, A.V.; Lambe, T.; Gilbert, S.C. Clinical assessment of a novel recombinant simian adenovirus ChAdOx1 as a vectored vaccine expressing conserved Influenza A antigens. *Mol. Ther. J. Am. Soc. Gene Ther.* **2014**, *22*, 668–674. [CrossRef]

20. Ouédraogo, A.; Tiono, A.B.; Kargougou, D.; Yaro, J.B.; Ouédraogo, E.; Kaboré, Y.; Kangoye, D.; Bougouma, E.C.; Gansane, A.; Henri, N.; et al. A phase 1b randomized, controlled, double-blinded dosage-escalation trial to evaluate the safety, reactogenicity and immunogenicity of an adenovirus type 35 based circumsporozoite malaria vaccine in Burkinabe healthy adults 18 to 45 years of age. *PLoS ONE* **2013**, *8*, e78679. [CrossRef]

21. Vemula, S.V.; Amen, O.; Katz, J.M.; Donis, R.; Sambhara, S.; Mittal, S.K. Beta-defensin 2 enhances immunogenicity and protection of an adenovirus-based H5N1 influenza vaccine at an early time. *Virus Res.* **2013**, *178*, 398–403. [CrossRef]

22. Creech, C.B.; Dekker, C.L.; Ho, D.; Phillips, S.; Mackey, S.; Murray-Krezan, C.; Grazia Pau, M.; Hendriks, J.; Brown, V.; Dally, L.G.; et al. Randomized, placebo-controlled trial to assess the safety and immunogenicity of an adenovirus type 35-based circumsporozoite malaria vaccine in healthy adults. *Hum. Vaccines Immunother.* **2013**, *9*, 2548–2557. [CrossRef]

23. Quinn, K.M.; Da Costa, A.; Yamamoto, A.; Berry, D.; Lindsay, R.W.B.; Darrah, P.A.; Wang, L.; Cheng, C.; Kong, W.-P.; Gall, J.G.D.; et al. Comparative analysis of the magnitude, quality, phenotype, and protective capacity of simian immunodeficiency virus gag-specific CD8+ T cells following human-, simian-, and chimpanzee-derived recombinant adenoviral vector immunization. *J. Immunol. Baltim. Md 1950* **2013**, *190*, 2720–2735. [CrossRef]

24. Zhou, G.; Wang, H.; Wang, F.; Yu, L. Recombinant adenovirus expressing type Asia1 foot-and-mouth disease virus capsid proteins induces protective immunity against homologous virus challenge in mice. *Res. Vet. Sci.* **2013**, *94*, 796–802. [CrossRef] [PubMed]

25. Zahn, R.; Gillisen, G.; Roos, A.; Koning, M.; van der Helm, E.; Spek, D.; Weijtens, M.; Grazia Pau, M.; Radošević, K.; Weverling, G.J.; et al. Ad35 and ad26 vaccine vectors induce potent and cross-reactive antibody and T-cell responses to multiple filovirus species. *PLoS ONE* **2012**, *7*, e44115. [CrossRef] [PubMed]

26. Baden, L.R.; Walsh, S.R.; Seaman, M.S.; Tucker, R.P.; Krause, K.H.; Patel, A.; Johnson, J.A.; Kleinjan, J.; Yanosick, K.E.; Perry, J.; et al. First-in-human evaluation of the safety and immunogenicity of a recombinant adenovirus serotype 26 HIV-1 Env vaccine (IPCAVD 001). *J. Infect. Dis.* **2013**, *207*, 240–247. [CrossRef] [PubMed]

27. Keefer, M.C.; Gilmour, J.; Hayes, P.; Gill, D.; Kopycinski, J.; Cheeseman, H.; Cashin-Cox, M.; Naarding, M.; Clark, L.; Fernandez, N.; et al. A phase I double blind, placebo-controlled, randomized study of a multigenic HIV-1 adenovirus subtype 35 vector vaccine in healthy uninfected adults. *PLoS ONE* **2012**, *7*, e41936. [CrossRef]

28. Pandey, A.; Singh, N.; Vemula, S.V.; Couëtil, L.; Katz, J.M.; Donis, R.; Sambhara, S.; Mittal, S.K. Impact of preexisting adenovirus vector immunity on immunogenicity and protection conferred with an adenovirus-based H5N1 influenza vaccine. *PLoS ONE* **2012**, *7*, e33428. [CrossRef]

29. O'Hara, G.A.; Duncan, C.J.A.; Ewer, K.J.; Collins, K.A.; Elias, S.C.; Halstead, F.D.; Goodman, A.L.; Edwards, N.J.; Reyes-Sandoval, A.; Bird, P.; et al. Clinical assessment of a recombinant simian adenovirus ChAd63: A potent new vaccine vector. *J. Infect. Dis.* **2012**, *205*, 772–781. [CrossRef]

30. Sun, Y.; Li, H.-Y.; Tian, D.-Y.; Han, Q.-Y.; Zhang, X.; Li, N.; Qiu, H.-J. A novel alphavirus replicon-vectored vaccine delivered by adenovirus induces sterile immunity against classical swine fever. *Vaccine* **2011**, *29*, 8364–8372. [CrossRef]

31. Rhee, E.G.; Kelley, R.P.; Agarwal, I.; Lynch, D.M.; Porte, A.L.; Simmons, N.L.; Clark, S.L.; Barouch, D.H. TLR4 Ligands Augment Antigen-Specific CD8+ T Lymphocyte Responses Elicited by a Viral Vaccine Vector. *J. Virol.* **2010**, *84*, 10413–10419. [CrossRef]

32. Steitz, J.; Barlow, P.G.; Hossain, J.; Kim, E.; Okada, K.; Kenniston, T.; Rea, S.; Donis, R.O.; Gambotto, A. A Candidate H1N1 Pandemic Influenza Vaccine Elicits Protective Immunity in Mice. *PLoS ONE* **2010**, *5*. [CrossRef]

33. Richardson, J.S.; Yao, M.K.; Tran, K.N.; Croyle, M.A.; Strong, J.E.; Feldmann, H.; Kobinger, G.P. Enhanced protection against Ebola virus mediated by an improved adenovirus-based vaccine. *PLoS ONE* **2009**, *4*, e5308. [CrossRef]

34. Rodríguez, A.; Goudsmit, J.; Companjen, A.; Mintardjo, R.; Gillissen, G.; Tax, D.; Sijtsma, J.; Weverling, G.J.; Holterman, L.; Lanar, D.E.; et al. Impact of Recombinant Adenovirus Serotype 35 Priming versus Boosting of a Plasmodium falciparum Protein: Characterization of T- and B-Cell Responses to Liver-Stage Antigen 1. *Infect. Immun.* **2008**, *76*, 1709–1718. [CrossRef] [PubMed]

35. Ophorst, O.J.A.E.; Radosević, K.; Klap, J.M.; Sijtsma, J.; Gillissen, G.; Mintardjo, R.; van Ooij, M.J.M.; Holterman, L.; Companjen, A.; Goudsmit, J.; et al. Increased immunogenicity of recombinant Ad35-based malaria vaccine through formulation with aluminium phosphate adjuvant. *Vaccine* **2007**, *25*, 6501–6510. [CrossRef] [PubMed]

36. Widjojoatmodjo, M.N.; Bogaert, L.; Meek, B.; Zahn, R.; Vellinga, J.; Custers, J.; Serroyen, J.; Radošević, K.; Schuitemaker, H. Recombinant low-seroprevalent adenoviral vectors Ad26 and Ad35 expressing the respiratory syncytial virus (RSV) fusion protein induce protective immunity against RSV infection in cotton rats. *Vaccine* **2015**, *33*, 5406–5414. [CrossRef]

37. Capone, S.; Meola, A.; Ercole, B.B.; Vitelli, A.; Pezzanera, M.; Ruggeri, L.; Davies, M.E.; Tafi, R.; Santini, C.; Luzzago, A.; et al. A novel adenovirus type 6 (Ad6)-based hepatitis C virus vector that overcomes preexisting anti-ad5 immunity and induces potent and broad cellular immune responses in rhesus macaques. *J. Virol.* **2006**, *80*, 1688–1699. [CrossRef] [PubMed]

38. Ophorst, O.J.A.E.; Radošević, K.; Havenga, M.J.E.; Pau, M.G.; Holterman, L.; Berkhout, B.; Goudsmit, J.; Tsuji, M. Immunogenicity and Protection of a Recombinant Human Adenovirus Serotype 35-Based Malaria Vaccine against Plasmodium yoelii in Mice. *Infect. Immun.* **2006**, *74*, 313–320. [CrossRef] [PubMed]

39. Hashimoto, M.; Boyer, J.L.; Hackett, N.R.; Wilson, J.M.; Crystal, R.G. Induction of Protective Immunity to Anthrax Lethal Toxin with a Nonhuman Primate Adenovirus-Based Vaccine in the Presence of Preexisting Anti-Human Adenovirus Immunity. *Infect. Immun.* **2005**, *73*, 6885–6891. [CrossRef] [PubMed]

40. Pinto, A.R.; Fitzgerald, J.C.; Giles-Davis, W.; Gao, G.P.; Wilson, J.M.; Ertl, H.C.J. Induction of CD8+ T cells to an HIV-1 antigen through a prime boost regimen with heterologous E1-deleted adenoviral vaccine carriers. *J. Immunol. Baltim. Md 1950* **2003**, *171*, 6774–6779. [CrossRef]

41. Casimiro, D.R.; Chen, L.; Fu, T.-M.; Evans, R.K.; Caulfield, M.J.; Davies, M.-E.; Tang, A.; Chen, M.; Huang, L.; Harris, V.; et al. Comparative immunogenicity in rhesus monkeys of DNA plasmid, recombinant vaccinia virus, and replication-defective adenovirus vectors expressing a human immunodeficiency virus type 1 gag gene. *J. Virol.* **2003**, *77*, 6305–6313. [CrossRef]

42. Xiang, Z.; Gao, G.; Reyes-Sandoval, A.; Cohen, C.J.; Li, Y.; Bergelson, J.M.; Wilson, J.M.; Ertl, H.C.J. Novel, Chimpanzee Serotype 68-Based Adenoviral Vaccine Carrier for Induction of Antibodies to a Transgene Product. *J. Virol.* **2002**, *76*, 2667–2675. [CrossRef]

43. Xiang, Z.Q.; Yang, Y.; Wilson, J.M.; Ertl, H.C. A replication-defective human adenovirus recombinant serves as a highly efficacious vaccine carrier. *Virology* **1996**, *219*, 220–227. [CrossRef]

44. Eloit, M.; Adam, M. Isogenic adenoviruses type 5 expressing or not expressing the E1A gene: Efficiency as virus vectors in the vaccination of permissive and non-permissive species. *J. Gen. Virol.* **1995**, *76*, 1583–1589. [CrossRef] [PubMed]

45. Neilan, J.G.; Schutta, C.; Barrera, J.; Pisano, M.; Zsak, L.; Hartwig, E.; Rasmussen, M.V.; Kamicker, B.J.; Ettyreddy, D.; Brough, D.E.; et al. Efficacy of an adenovirus-vectored foot-and-mouth disease virus serotype A subunit vaccine in cattle using a direct contact transmission model. *BMC Vet. Res.* **2018**, *14*. [CrossRef] [PubMed]

46. Tang, A.; Freed, D.C.; Li, F.; Meschino, S.; Prokop, M.; Bett, A.; Casimiro, D.; Wang, D.; Fu, T.-M. Functionally inactivated dominant viral antigens of human cytomegalovirus delivered in replication incompetent adenovirus type 6 vectors as vaccine candidates. *Hum. Vaccines Immunother.* **2017**, *13*, 2763–2771. [CrossRef] [PubMed]

47. Fonseca, J.A.; McCaffery, J.N.; Kashentseva, E.; Singh, B.; Dmitriev, I.; Curiel, D.T.; Moreno, A. A prime-boost immunization regimen based on a Simian Adenovirus 36 vectored multi-stage malaria vaccine induces protective immunity in mice. *Vaccine* **2017**, *35*, 3239–3248. [CrossRef]

48. Nazerai, L.; Bassi, M.R.; Uddback, I.E.M.; Holst, P.J.; Christensen, J.P.; Thomsen, A.R. Early life vaccination: Generation of adult-quality memory CD8+ T cells in infant mice using non-replicating adenoviral vectors. *Sci. Rep.* **2016**, *6*, 38666. [CrossRef]

49. Penaloza-MacMaster, P.; Alayo, Q.A.; Ra, J.; Provine, N.M.; Larocca, R.; Lee, B.; Barouch, D.H. Inhibitory receptor expression on memory CD8 T cells following Ad vector immunization. *Vaccine* **2016**, *34*, 4955–4963. [CrossRef]

Supplementary material for paper 2

These supplementary are adapted from those that were included with the published version of this paper, which can be found online [190].

Figures showing calibrated peaking and saturating curves for all available data. Figures are arranged hierarchically by response type, vector species, host species and route of administration, and paper number, in that order. For each dataset, the time since inoculation is given, along with the vector if there were more than one used in that paper. For each dataset two plots are shown, with the left plot including the calibrated saturating curve and the right plot the calibrated peaking curve. The AIC is included for each. Blue dots represent the mean response for each dosing group. Black dots represent individual responses or upper and lower confidence intervals, depending on the data availability for that dataset. The x axis for each plot is the log10 of the given dose, and the y axis is this recorded response for that dataset, as given in the paper that the dataset was taken from.

I only show the subfigures for the first paper of the first group of data (Antibody/Vector Species B/ Host Species Mouse/RoA IM) here for demonstration, the other figures are included in the appendix of this thesis [A.D.1.Supplementary Figures]

Response Type: Antibody
Vector Species: B
Host Species: Mouse
Route of Administration: IM
Paper 578:
Day 28 - Ad35 Zaire
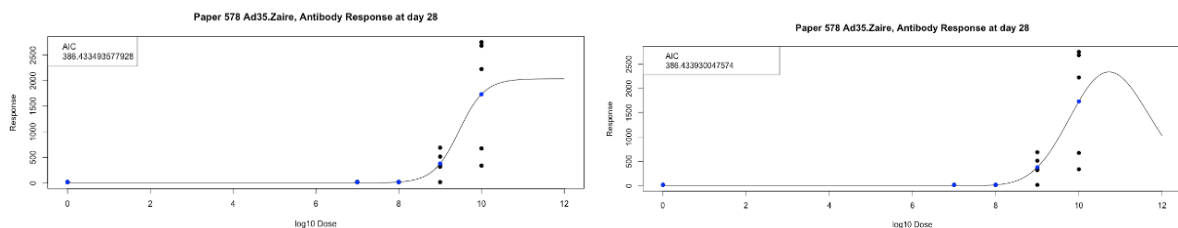
Day 28 - Ad35 Angola
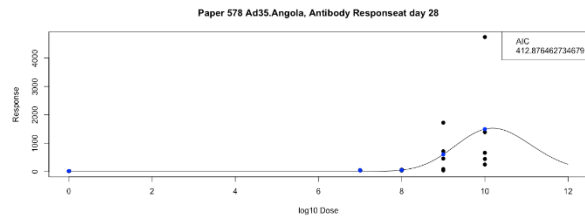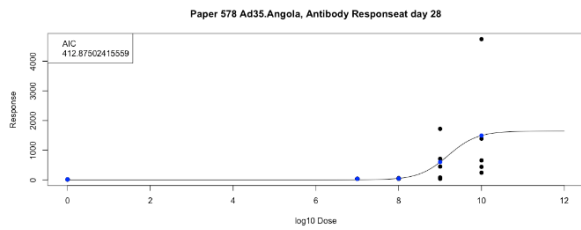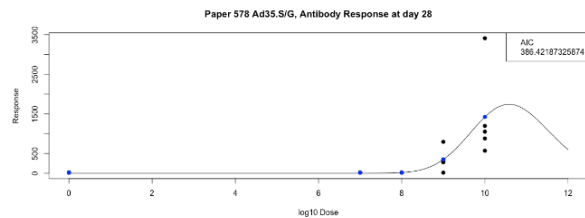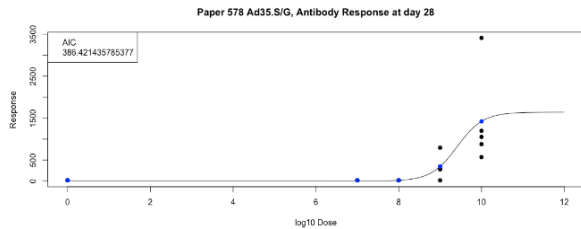
Day 28 - Ad35 S/G
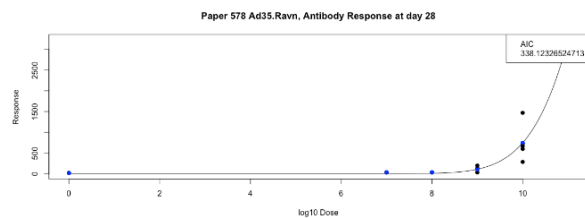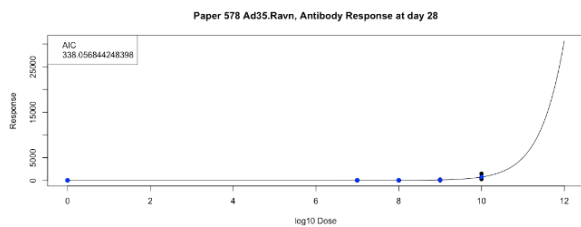
Day 28 - Ad35 Ravn

Day 28 - Ad35 I.C
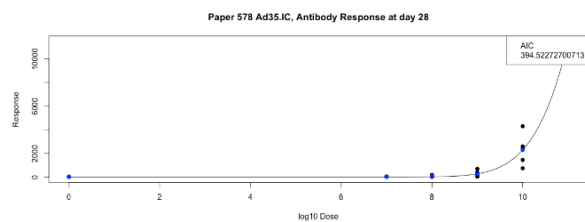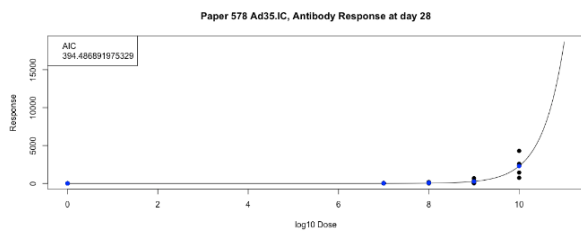
Figure S.Figures.A. A small subsection of the plotted calibrated models. Left shows the saturating curve and right shows the peaking curve.

Supplementary Tables

These tables summarise the datasets in Supplementary Figures S1 including the level of evidence towards either a peaking or saturating curve shape, determined by AIC difference. An 'x' is used to identify the level of evidence. There exists a table for each response type. Bold lines are used to show when the group the dataset belongs to changes. The left most column combined with the response type of the table gives the group. Each dataset also includes the paper number the dataset was gathered from and the days post inoculation that the responses were measured.

I only show the first table here for demonstration, the other tables are included in the appendix of this thesis [A.D.1.Supplementary Tables]

119

Antibodies									
Group (Adenoviral Species/Host Species/Administration Route)	Paper	Days post inoculation	Absolute Peaking Evidence (>10)	Strong Peaking Evidence (10-6)	Slight Peaking Evidence (2-6)	No Evidence (<2)	Slight Saturating Evidence (2-6)	Strong Saturating Evidence (10-6)	Strong Saturating Evidence (>10)
B/Mouse/IM	578	28				x			
		28				x			
		28				x			
		28				x			
		28				x			
	1269	56				x			
	1492	14				x			
B/Human/IM	441	28				x			
	467	60			x				
	633	28	x						
C/Rabbit/IM	744	28				x			
C/Mouse/IM	461	7				x			
		14				x			
	574	7				x			
		21				x			
		35			x				
		49				x			
		63				x			
		77				x			
		91				x			
		119				x			
		147				x			
		161				x			
	1492	14				x			
	2531	10	x						
		10				x			
C/Mouse/SQ	936	35				x			
		35				x			
C/Human/IM	140	28				x			
	249	14				x			
		28				x			
		180				x			
C/Monkey/IM	1877	28						x	
		56							x
		84				x			
		112					x		
		140				x			
C/Rat/IM	2531	10				x			
		10	x						
D/Mouse/IM	578	28				x			

		28				x			
		28				x			
		28				x			
		28				x			
D/Human/IM	594	14				x			
E/Mouse/IM	1539	14				x			
		28				x			
		42				x			
		14				x			
		28				x			
		42				x			
	2919	20				x			
E/Human/IM	417	14				x			
		21				x			
Count			3	0	2	46	1	1	1

Table S1. Evidence for antibody response

Vectors by Paper

Paper Number	Present Paper Name	Vector(s)
140 [15]	Use of ChAd3-EBO-Z Ebola virus vaccine in Malian and US adults, and boosting of Malian adults with MVA-BN-Filo: a phase 1, single-blind, randomised trial, a phase 1b, open-label and double-blind, dose-escalation trial, and a nested, randomised, double-blind, placebo-controlled trial	ChAd3
249 [16]	A Monovalent Chimpanzee Adenovirus Ebola Vaccine Boosted with MVA	ChAd3
305[17]	Characterization of T-Cell Responses to Conserved Regions of the HIV-1 Proteome in BALB/c Mice	ChAdV63
309 [18]	The novel tuberculosis vaccine, AERAS-402, is safe in healthy infants previously vaccinated with BCG, and induces dose-dependent CD4 and CD8T cell responses	Ad35
417 [19]	Clinical assessment of a novel recombinant simian adenovirus ChAdOx1 as a vectored vaccine expressing conserved Influenza A antigens	ChAdOx1
441 [20]	A phase 1b randomized, controlled, double-blinded dosage-escalation trial to evaluate the safety, reactogenicity and immunogenicity of an adenovirus type 35 based circumsporozoite malaria vaccine in Burkinabe healthy adults 18 to 45 years of age	Ad35
461 [21]	Beta-defensin 2 enhances immunogenicity and protection of an adenovirus-based H5N1 influenza vaccine at an early time	Ad5
467 [22]	Randomized, placebo-controlled trial to assess the safety and immunogenicity of an adenovirus type 35-based circumsporozoite malaria vaccine in healthy adults	Ad35
555[23]	Comparative analysis of the magnitude, quality, phenotype, and protective capacity of simian immunodeficiency virus gag-specific CD8+ T cells following human-, simian-, and chimpanzee-derived recombinant adenoviral vector immunization	Ad35, sAd11, sAd16, ChAd3, Ad5, Ad28, ChAd63
574 [24]	Recombinant adenovirus expressing type Asia1 foot-and-mouth disease virus capsid proteins induces protective immunity against homologous virus challenge in mice	Ad5
578 [25]	Ad35 and ad26 vaccine vectors induce potent and cross-reactive antibody and T-cell responses to multiple filovirus species	Ad35, Ad26
594[26]	First-in-human evaluation of the safety and immunogenicity of a recombinant adenovirus serotype 26 HIV-1 Env vaccine (IPCAVD 001)	Ad26
633 [27]	A phase I double blind, placebo-controlled, randomized study of a multigenic HIV-1	Ad35

121

	adenovirus subtype 35 vector vaccine in healthy uninfected adults	
669 [28]	Impact of preexisting adenovirus vector immunity on immunogenicity and protection conferred with an adenovirus-based H5N1 influenza vaccine	Ad5
686 [29]	Clinical assessment of a recombinant simian adenovirus ChAd63: a potent new vaccine vector	ChAd63
744 [30]	A novel alphavirus replicon-vectored vaccine delivered by adenovirus induces sterile immunity against classical swine fever	Ad5
924 [31]	TLR4 Ligands Augment Antigen-Specific CD8+ T Lymphocyte Responses Elicited by a Viral Vaccine Vector	Ad26
936 [32]	A Candidate H1N1 Pandemic Influenza Vaccine Elicits Protective Immunity in Mice	Ad5
1039[33]	Enhanced protection against Ebola virus mediated by an improved adenovirus-based vaccine	Ad5
1201 [34]	Impact of Recombinant Adenovirus Serotype 35 Priming versus Boosting of a Plasmodium falciparum Protein: Characterization of T- and B-Cell Responses to Liver-Stage Antigen 1	Ad35
1269 [35]	Increased immunogenicity of recombinant Ad35-based malaria vaccine through formulation with aluminium phosphate adjuvant	Ad35
1343 [36]	Recombinant low-seroprevalent adenoviral vectors Ad26 and Ad35 expressing the respiratory syncytial virus (RSV) fusion protein induce protective immunity against RSV infection in cotton rats	Ad26
1474 [37]	A novel adenovirus type 6 (Ad6)-based hepatitis C virus vector that overcomes preexisting anti-ad5 immunity and induces potent and broad cellular immune responses in rhesus macaques	Ad5, Ad6
1492 [38]	Immunogenicity and Protection of a Recombinant Human Adenovirus Serotype 35-Based Malaria Vaccine against Plasmodium yoelii in Mice	Ad5, Ad35
1539 [39]	Induction of Protective Immunity to Anthrax Lethal Toxin with a Nonhuman Primate Adenovirus-Based Vaccine in the Presence of Preexisting Anti-Human Adenovirus Immunity	AdC7
1801[40]	Induction of CD8+ T cells to an HIV-1 antigen through a prime boost regimen with heterologous E1-deleted adenoviral vaccine carriers	AdC6
1877 [41]	Comparative immunogenicity in rhesus monkeys of DNA plasmid, recombinant vaccinia virus, and replication-defective adenovirus vectors expressing a human immunodeficiency virus type 1 gag gene	Ad5
2030 [42]	Novel, Chimpanzee Serotype 68-Based Adenoviral Vaccine Carrier for Induction of Antibodies to a Transgene Product	Ad5, ChAd63
2505 [43]	A replication-defective human adenovirus recombinant serves as a highly efficacious vaccine carrier	Ad5
2531 [44]	Isogenic adenoviruses type 5 expressing or not expressing the E1A gene: efficiency as virus vectors in the vaccination of permissive and non-permissive species	Ad5
2841 [45]	Efficacy of an adenovirus-vectored foot-and-mouth disease virus serotype A subunit vaccine in cattle using a direct contact transmission model	Ad5
2916[46]	Functionally inactivated dominant viral antigens of human cytomegalovirus delivered in replication incompetent adenovirus type 6 vectors as vaccine candidates	Ad6
2919 [47]	A prime-boost immunization regimen based on a Simian Adenovirus 36 vectored multi-stage malaria vaccine induces protective immunity in mice	sAd36
2980 [48]	Early life vaccination: Generation of adult-quality memory CD8+ T cells in infant mice using non-replicating adenoviral vectors	Ad5
3018 [49]	Inhibitory receptor expression on memory CD8 T cells following Ad vector immunization	Ad5

Table S13a. Vectors by paper

Species and origin of vectors

Name	Species	Origin
Ad26	D	Human
Ad28	D	Human
Ad35	B	Human
Ad5	C	Human
Ad6	C	Human
AdC6	E	Simian
AdC7	E	Simian
Chad3	C	Simian
ChAd63	E	Simian
ChAdOx1	E	Simian
sAd11	G	Simian
sAd16	N.A.	Simian
sAd36	E	Simian

Table S13b. Species and origin of vectors

Chapter 4: A case study in multi-factor optimisation: a modelling study to maximise vaccine safety, efficacy, and affordability: Paper 3

Chapter 4 Introduction

The objective of this paper was to extend IS/ID beyond optimising vaccine dose to maximise efficacy, by considering the selection of 'optimal' vaccine dose as a multi-objective optimisation problem where trade-offs between the value of different dose-dependent objectives (i.e., efficacy, toxicity, cost) must be considered. I hypothesised that model-predicted 'optimal' doses may depend on the utility functions used to quantify these trade-offs. This chapter addresses objective 3 of this thesis.

Using data from a phase I dose-ranging study of a recombinant adenovirus type-5 COVID-19 single-dose vaccine (Ad5-nCoV), I calibrated models of dose-seroconversion and dose-grade 3 toxicity. Using the dose-seroconversion as a surrogate dose-efficacy model, I predicted optimal dose for this vaccine using three utility functions that could potentially be used as criterion for selection of optimal dose, namely i) achieving herd immunity, ii) balancing efficacy and toxicity, and iii) balancing efficacy, toxicity, and cost.

I choose to use published and publicly available data to calibrate these models. When I conducted this work, the vaccine was in development, and as such no dose had been chosen for clinical use. I also had no direct consultation with the developers or researchers of the dose-ranging study. I considered this beneficial as it meant that my research was unbiased by knowledge of which dose was found 'optimal' by the developers, and reduced potential for conflicts of interest.

I was interviewed to discuss this paper as part of its inclusion in a featured news article for the journal Nature discussing the future of computational mathematical modelling in vaccine development [191].

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed <u>for each</u> research paper included within a thesis.

SECTION A – Student Details

Student ID Number	lsh1804914	Title	Mr
First Name(s)	John Helier		
Surname/Family Name	Benest		
Thesis Title	Optimising Vaccine Dose in Inoculation against SARS-CoV-2, a Multi-Factor Optimisation Modelling Study to Maximise Vaccine Safety and Efficacy		
Primary Supervisor	Richard G. White		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?	MDPI Vaccines		
When was the work published?	22 January 2021		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	N.A		
Have you retained the copyright for the work?*	Yes	Was the work subject to academic peer review?	Yes

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	

Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	The conceptualisation of this work was my own. I also was the sole individual conducting dose-response modelling and prediction of optimal dose. I wrote the first draft of this work and performed re-drafting in line with co-author comments. All authors reviewed the paper. The interpretation of the results was my own work

SECTION E

Student Signature	████████████████
Date	11 August 2022

Supervisor Signature	████████████████
Date	12 August 2022

126

Paper 3 Title: Optimising Vaccine Dose in Inoculation against SARS-CoV-2, a Multi-Factor Optimisation Modelling Study to Maximise Vaccine Safety and Efficacy

Authors: John Benest, Sophie Rhodes, Matthew Quaife, Thomas G. Evans, Richard G. White

Permission from copyright owner to include this work:
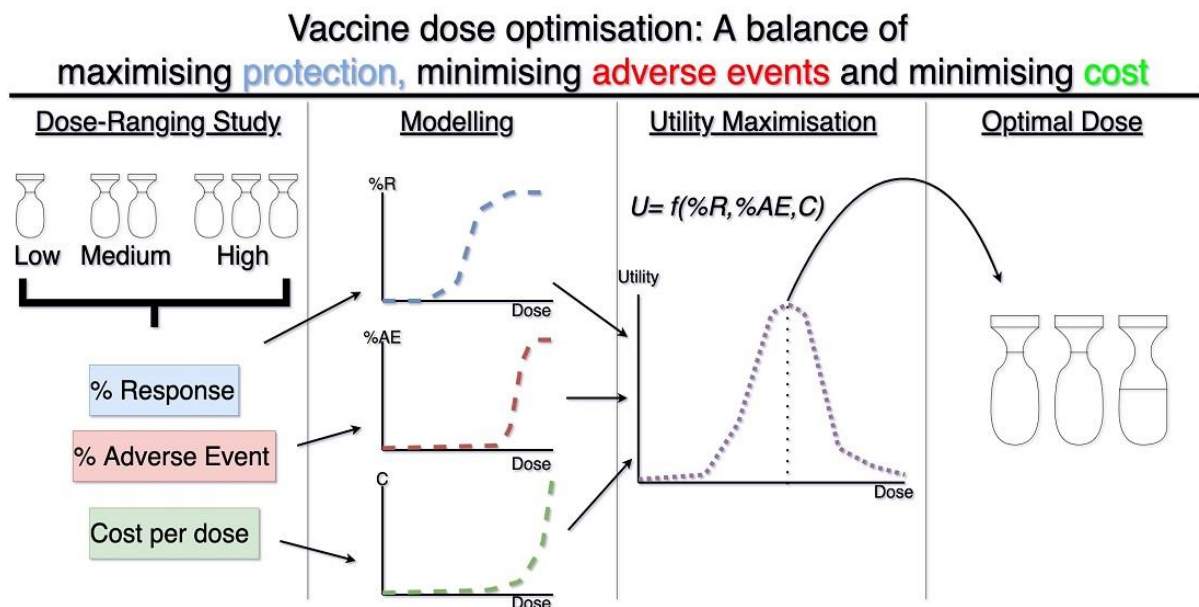
Figure GA. Graphical Abstract for Paper 3

Optimising Vaccine Dose in Inoculation against SARS-CoV-2, a Multi-Factor Optimisation Modelling Study to Maximise Vaccine Safety and Efficacy

John Benest [1],*, Sophie Rhodes [1], Matthew Quaife [1], Thomas G. Evans [2] and Richard G. White [1]

1 Department of Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK; sophie.rhodes@lshtm.ac.uk (S.R.); matthew.quaife@lshtm.ac.uk (M.Q.); Richard.White@lshtm.ac.uk (R.G.W.)
2 Vaccitech Ltd., The Schrodinger Building, Heatley Road, The Oxford Science Park, Oxford OX4 4GE, UK; tom.evans@vaccitech.co.uk
* Correspondence: john.benest@lshtm.ac.uk

Abstract: Developing a vaccine against the global pandemic SARS-CoV-2 is a critical area of active research. Modelling can be used to identify optimal vaccine dosing; maximising vaccine efficacy and safety and minimising cost. We calibrated statistical models to published dose-dependent seroconversion and adverse event data of a recombinant adenovirus type-5 (Ad5) SARS-CoV-2 vaccine given at doses 5.0×10^{10}, 1.0×10^{11} and 1.5×10^{11} viral particles. We estimated the optimal dose for three objectives, finding: (A) the minimum dose that may induce herd immunity, (B) the dose that maximises immunogenicity and safety and (C) the dose that maximises immunogenicity and safety whilst minimising cost. Results suggest optimal dose [95% confidence interval] in viral particles per person was (A) 1.3×10^{11} [0.8–7.9×10^{11}], (B) 1.5×10^{11} [0.3–5.0×10^{11}] and (C) 1.1×10^{11} [0.2–1.5×10^{11}]. Optimal dose exceeded 5.0×10^{10} viral particles only if the cost of delivery exceeded £0.65 or cost per 10^{11} viral particles was less than £6.23. Optimal dose may differ depending on the objectives of developers and policy-makers, but further research is required to improve the accuracy of optimal-dose estimates.

Keywords: dosing; dose-response; adenovirus-vectored vaccines; dose dynamics; COVID-19

1. Introduction

Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) has been an unprecedented burden on global health throughout 2020 [1]. Due to the health risks associated with infection, countries have had to implement policies of isolation and quarantine, causing global disruption to economic and health systems [2]. Vaccination is a vital tool in improving global health [3] and an effective vaccine against SARS-Cov-2 could drastically reduce the spread of this highly infectious pathogen. There is an urgent need to accelerate the development of a SARS-CoV-2 vaccine to protect the population [4]. However, maintaining safety and immunogenicity standards within vaccine development is paramount. Decisions relating to vaccine development need to be made quickly and accurately. One important decision is determining vaccine dose, defined as the quantity or magnitude of vaccine given.

Model-based drug development is commonly used to accelerate drug decision making, and the field of Immunostimulation/Immunodynamic (IS/ID) modelling has been developed to adapt these methods for vaccine development [5]. IS/ID modelling has shown promise in discovering optimal dose for TB and influenza inoculations [6,7] and for exploring dose-response trends for adenoviral-vectored vaccines [8], however, the previous works have focused entirely on optimising dose with respect to immunological response. In these studies, the modelling of dose-response and hence finding the dose that maximises response can be considered single-objective optimisation problems.

Whilst optimising response with respect to vaccine dose is essential to ensuring effective vaccination, the change in financial cost and safety of vaccination with respect to vaccine dose are also important. A vaccine should ideally maximise protective immunogenicity, minimise the risk of vaccine-related toxicity and minimise the cost of using that vaccine. Optimising dose in relation to immunogenicity, safety and cost is a multi-objective or multi-criteria-decision-analysis problem [9]. Using models to analyse the dose-response, dose-safety and dose-cost relationships can provide insight into the multi-dimensional dose-utility curve and hence optimal dose. There may also exist cases where the cost of vaccination can be ignored, for example, if the vaccine is not limited in supply. This could also be true for cases where there is a very high willingness to pay for vaccination by policymakers, which permits costs to be nearly ignored to ensure a maximum reduction in disease burden.

Approximately 180 SARS-CoV-2 vaccines are in development [10], and there is much interest in which vaccines may offer the greatest efficacy in reducing the burden of SARS-CoV-2 on global health [11–13]. Not only should the dose of any vaccine implemented be optimised with regards to safety, immunogenicity and cost, but appropriate dosing is equally important to ensure that unbiased estimate of vaccine immunogenicity and safety is used. In addition to establishing optimal dose at an individual level, the potential of a candidate vaccine to induce herd immunity in an entirely vaccinated population can also be considered.

Whilst accelerating these decisions and improving dosing is important now [14], it is also hoped that modelling and multi-objective optimisation can aid in rapid vaccine response to future epidemics.

In this work, we aimed to optimise the dose of a recombinant adenovirus type-5 (Ad5) vectored SARS-CoV-2 vaccine using IS/ID modelling and multifactorial optimisation. Our objectives were:

(1) Using published data, calibrate mathematical models to the relationship between dose and seroconversion, safety and cost of a single inoculation.
(2) Identify the minimum dose that is predicted to theoretically induce herd immunity.
(3) Identify the dose that maximises immunogenicity and safety.
(4) Identify the dose that maximises immunogenicity and safety whilst minimising cost.

2. Materials and Methods

2.1. Data

Data were extracted from a published phase 1 study of a SARS-Cov-2 adenoviral-based vaccine [15]. Data on neutralising antibody-based seroconversion and adverse events were extracted for 108 healthy human participants inoculated with the candidate vaccine on day 0. Individuals were divided into three dose groups (5×10^{10}, 1×10^{11} and 1.5×10^{11} Viral Particles (VP)) and their responses were measured at day 28.

Seroconversion was chosen as a surrogate of protective immunity and was defined as a neutralising antibody response post-vaccination of at least a four-fold increase relative to baseline, above which an individual was protected [16].

Safety was defined by the proportion of individuals that experienced adverse events within 0–28 days post-vaccination. Vaccine adverse events can be graded depending on severity (Table 1), with grade 3 or above considered severe [17,18]. As a measure of safety, we considered both the proportion that experienced "any grade" adverse events and the proportion that experienced grade 3 or above adverse events. Whilst the designation "Grade 3+" is used to be consistent with the terms used by the original authors, no adverse events above grade 3 were reported.

Table 1. Description of adverse event grading in vaccine clinical trials.

Adverse Reaction Grade	General Descriptions
1	Mild. Does not interfere with normal activity
2	Moderate. Interference with normal activity. Little or no treatment required.
3	Severe. Prevents normal activity. Requires treatment.
4	Serious or Potentially Life-Threatening. Generally requires hospitalisation and stopping of any clinical trial where this grade is observed.

2.2. Objective 1. Using Published Data, Calibrate Mathematical Models to the Relationship between Dose and Seroconversion, Safety and Cost of a Single Inoculation

2.2.1. Dose-Seroconversion Relationship

We calibrated a sigmoid function to the dose-seroconversion data using the nls function in R [19,20]. Sigmoid functions are commonly used to describe biological processes, including in both drug and vaccine response modelling. The formula is

$$sigmoid(Dose) = BaseResponse + \frac{MaxResponse - BaseResponse}{1 + e^{-Scale \times (Dose - Dose_{50})}} \quad (1)$$

where *Dose* was log10 (Viral Particles), and *BaseResponse* (the minimum output of the function), *MaxResponse* (the maximum output of the function), *Dose*50 (the dose which defines the functions midpoint) and *Scale* (which determines the steepness of the curve) were the model parameters.

We also calibrated a representative non-saturating ("peaking") curve, using the methods discussed in [8]. However, as these methods did not support the non-saturating curve providing a better description of the data, we assume that the dose-response follows a sigmoidal function (Supplementary S1).

In the absence of seroconversion data from a placebo group, a base seroconversion rate of 0% was assumed (*BaseResponse* = 0). We predicted the dose-seroconversion curve for doses of up to 10^{15} VP, to ensure previous adenoviral dosing ranges are explored [21] and state the dose that would induce 50% and 90% seroconversion.

2.2.2. Dose-Safety Relationship

We calibrated a sigmoid function (Equation (1)) to the dose—"any grade adverse event" data using the nls function in R, and another to the dose—"grade 3+ adverse event" data. Again, in the absence of a placebo group, a base adverse event rate of 0% was assumed for both curves (*BaseResponse* = 0). Further, we assumed that for sufficiently large doses, 100% of individuals would experience adverse events. We predict the probability of an adverse event for doses of up to 10^{15} VP.

Lastly, we found the doses for which the proportion of individuals that would experience grade 3+ adverse events above the thresholds of 30% and 17%. 30% of grade 3+ adverse events has been defined as a threshold for unacceptable toxicity in dose-escalation studies [22–24]. However, of the commonly CDC recommended vaccines, the largest grade 3+ adverse reaction rate is 17% for the Shingrix herpes zoster vaccine [25].

2.2.3. Dose-Cost Relationship

We assumed that the cost of a single individual receiving a vaccination can be described by

$$Cost_{Total}(Dose) = Cost_{Delivery} + Cost_{Dose-dependant}(Dose) \quad (2)$$

where $Cost_{Total}(Dose)$ was the total cost in British Pounds Sterling (GBP, £) to vaccinate one person with a given dose. We assumed that this was the sum of "*Delivery*" costs, which are independent of dose, and the dose-dependent cost. Expected costs for doses of up to 10^{15} VP were calculated using this formula and parameters in Table 2.

Table 2. Parameter values for the cost function. Where needed, a conversion rate of 0.78 U.S. Dollars per GBP was used [26], and a 10-year inflation rate was estimated as 1.35 2020 GBP per 2010 GBP Pound [27].

Name of Parameter	Value	Unit	Description	References
$Cost_{Personnel}$ (£ per vaccination) = 4.398707				
AnnualWage	30,615	£ per years	GBP per NHS Band 5 Income per annum (2020/21)	[28]
AnnualHours	1740	hours per years	Work hours per year for average UK nurse	[29]
Time per vaccination	0.25	hours per vaccination	Recommended hours per vaccination appointment	[30]
$Cost_{Storage}$ (£ per vaccination) = 0.014				
$Cost_{Storage per month}$	0.014	£ per month	GBP per vaccination per month's storage. Converted and adjusted for inflation from $0.014 2010 USD.	[31]
$Cost_{materials}$ (£ per vaccination) = 0.83				
$Cost_{Gloves}$	0.08	£ per vaccination	GBP of gloves for one vaccination. Converted and adjusted for inflation from $0.08 USD.	[31]
$Cost_{Alcohol}$	0.03	£ per vaccination	GBP of sterilising alcohol for one vaccination. Converted and adjusted for inflation from $0.03 2010 USD.	[31]
$Cost_{PFS}$	0.40	£ per vaccination	GBP of the pre-filled syringe for one vaccination. Converted and adjusted for inflation from $0.39 2010 USD.	[31]
$Cost_{Needles}$	0.32	£ per vaccination	GBP of needle for one vaccination. Converted and adjusted for inflation from $0.31 2010 USD.	[31]
Cost per viral particle (£ *per VP*) = 7.6×10^{-12}				
$Cost_{Adenoviral Batch}$	342,000	£ per Batch	GBP per single-use reference process batch (converted from 450,000 US Dollars)	[32]
Adenoviral Concentration	9×10^{13}	*VP per L*	Viral Particles per litre in single-use reference process batch	[32]
Batch volume	500	*L per Batch*	Volume of Adenovirus produced in single-use reference process batch	[32]

Specifically, the delivery costs included the cost of disposable materials used in vaccination (gloves, sterilizing alcohol, prefilled syringes, needles), storage (assumed to be 1 month of storage costs), and personnel costs (15 min of nursing time). The delivery cost formula was therefore calculated as

$$Cost_{Delivery}(Dose) = Cost_{Materials} + Cost_{Storage} + Cost_{Personnel} \qquad (3)$$

where

$$Cost_{Materials}(Dose) = Cost_{Gloves} + Cost_{Alcohol} + Cost_{PFS} + Cost_{Needles} \qquad (4)$$

$$Cost_{Storage}(Dose) = Cost_{Storage\,per\,month} \times 1(months) \qquad (5)$$

$$Cost_{Personnel}(Dose) = (AnnualWage \div AnnualHours) \times Time\,per\,vaccination \qquad (6)$$

The dose-dependent cost was the cost of the manufactured adenoviral vaccine measured in viral particles, assuming bulk production, which increased linearly with the vaccine dose. The dose-dependent cost formula was therefore calculated as

$$Cost_{Dose-dependent} = Viral\,particles\,per\,vaccination(Dose) \times Cost\,per\,viral\,particle \qquad (7)$$

With

$$Cost\,per\,viral\,particle = Cost_{Adenoviral\,Batch} \div (Adenoviral\,Concentration \times Batch\,Volume) \qquad (8)$$

These costs were a simplification of real-world costs but represent an approximate cost of vaccination. The delivery cost of vaccination ($Cost_{Delivery}$) was calculated as £5.24, and the cost per 10^{11} VP ($Cost_{Dose\text{-}dependent}$) of adenovirus was £0.76 (Table 2).

2.3. Objective 2. Identify the Minimum Dose that Is Predicted to Theoretically Induce Herd Immunity

An optimal vaccine dose should maximise response (maximise the proportion that will seroconvert), safety (minimise the proportion that experience adverse events), and affordability (minimise cost per vaccination). Increasing dose may increase seroconversion, but would also increase cost and adverse event prevalence. Vaccinating a population is often done to induce herd immunity, and so a factor in selecting the optimal dose is whether that dose could induce herd immunity in an entirely vaccinated population. Therefore, one approach to dose-optimisation is to choose a dose that can induce herd immunity within the population, and specifically to choose the minimum such dose to minimise cost and adverse event prevalence.

We used the suggested 65.5% of the population required to be protected to cause herd immunity in the United Kingdom (UK) [33] to establish whether there exist doses for this vaccine that could induce this 65.5% seroconversion and hence induce herd immunity in an entirely vaccinated population. We then identified the minimum such dose that could do so. A 95% confidence interval for optimal dose was determined using a parametric bootstrapping approach (Supplementary S2.3.1).

2.4. Objective 3. Identify the Dose that Maximises Immunogenicity and Safety

Another approach to dose optimisation is to choose the dose which maximises the proportion of individuals that "safely" seroconvert. To do this, a multifactorial utility function was derived, defined here as a mathematical formula that estimates the "worth" or utility of doses relative to each other.

Using the assumption that the probability of seroconversion, P_s, and the probability of grade 3+ adverse events, P_t, were mutually independent, the probability of a safe seroconversion was equal to

$$P_S \times (1 - P_t) \qquad (9)$$

(Figure 1). Therefore, the utility function, $U_{Costless}$, to be maximised was

$$U_{Costless}(Dose) = P_s(Dose) \times (1 - P_t(Dose)) \tag{10}$$
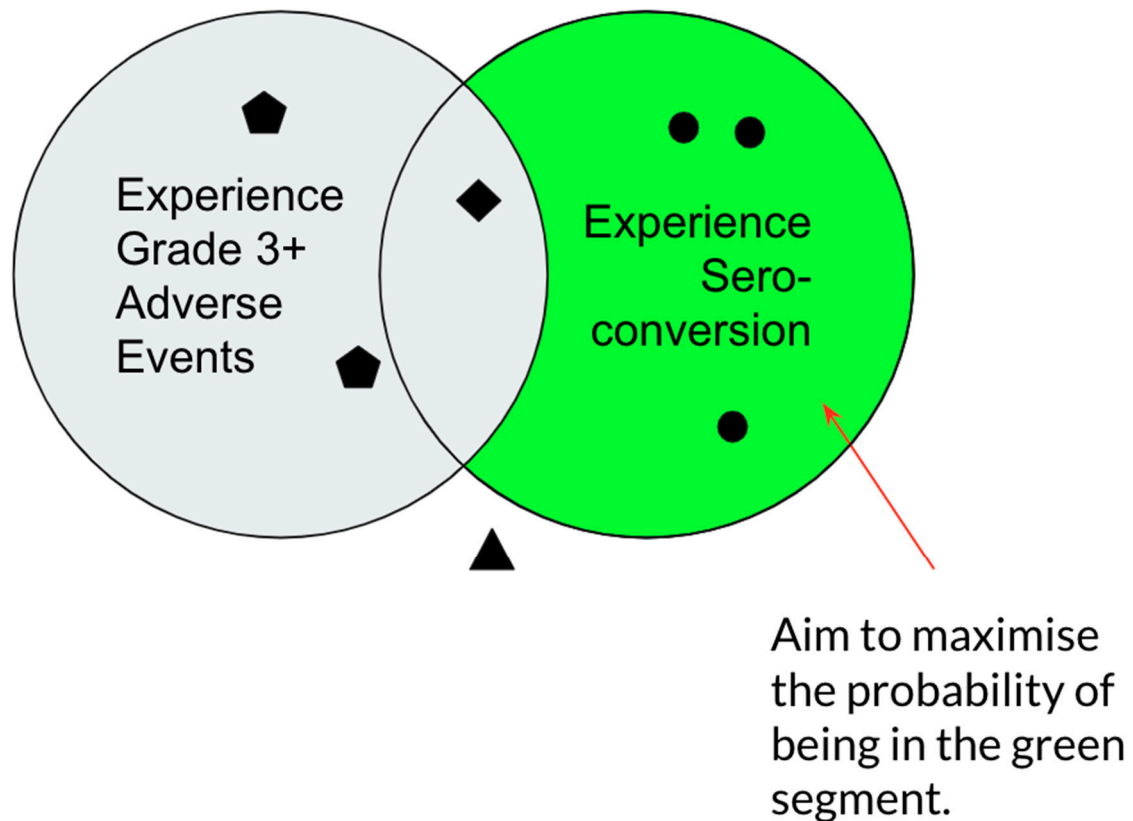
Figure 1. Venn diagram representation of possible outcomes of inoculation, where the left set includes individuals that experience grade 3+ adverse events and the right set includes individuals that experience seroconversion. We aimed to maximise the number of individuals that experience seroconversion and do not experience grade 3+ adverse events, represented in the green segment of the diagram. Black diamonds represent individuals that experience both outcomes, black pentagons represent individuals that experience grade 3+ adverse events with no seroconversion, and black triangles represent individuals that experience neither outcome.

Optimal dose was defined as the dose that maximised this function. A 95% confidence interval for optimal dose was determined using a parametric bootstrapping approach (Supplementary S2.3.2).

2.5. Objective 4. Identify the Dose that Maximises Immunogenicity and Safety Whilst Minimising Cost

We can also include the increased cost associated with an increased dose into the previous utility function (Equation (10)). In the case where cost is included as a potential limiting factor, a potential costed utility function was

$$U_{Costed}(Dose) = \frac{P_s(Dose) \times (1 - Pt(Dose))}{Cost_{Total}(Dose)} \tag{11}$$

Using this utility function, we predicted the optimised dose for maximising seroconversion and minimising adverse events and cost. Note that $U_{Costless}(Dose)$ (Equation (10)) is precisely the numerator of the $U_{Costed}(Dose)$ (Equation (11)). A 95% confidence interval for optimal dose was again determined using a parametric bootstrapping approach (Supplementary S2.3.2).

Threshold Analysis

Due to the difficulty in accurately estimating cost parameters, we conducted a threshold analysis on the parameter values of the cost model (Equations (3)–(8)). This was conducted to determine how much error would be needed in our costing model parameters to qualitatively alter the optimal predicted dose. We chose 5×10^{11}, 1×10^{11} and 5×10^{10} VP as the thresholds of interest.

To conduct a threshold analysis of parameters $Cost_{Delivery}$, we fixed all other parameters at the calibrated/literature derived value and allowed $Cost_{Delivery}$ to vary. The region over which we varied $Cost_{Delivery}$ was $+/-$ 3 orders of magnitudes of the value (£5.24) we used in the main model. In other words, we considered the effect of $Cost_{Delivery}$ being 1000 times larger or smaller (from £0.0052 per vaccination to £5240 per vaccination) on the prediction of optimal dose. This range was considered certainly to almost contain a reasonable estimate of the dose-independent costs of a single vaccination. This procedure was then repeated for *Cost per 10^{11}* viral particles, ranging from £0.00076 to £760 per 10^{11} VP.

We found the parameter ranges for which the dose that optimised $U_{Costed}(Dose)$ were above and below the stated thresholds (5×10^{11}, 1×10^{11}, 5×10^{10} VP).

3. Results

3.1. Objective 1. Using Published Data, Calibrate Mathematical Models to the Relationship between Dose and Seroconversion, Safety, and Cost of a Single Inoculation

3.1.1. Does-Seroconversion Relationship

The empirical data showed that doses of 5.0×10^{10}, 1.0×10^{11} and 1.5×10^{11} induced 50%, 50%, 75% seroconversion on day 28, respectively. The calibrated saturating dose-seroconversion curve is displayed in Figure 3.3a. 50% and 95% seroconversion were predicted at a dose of 5.9×10^{10} and 2.4×10^{12} VP, respectively. Population demographics including age, gender and pre-existing adenovirus neutralising antibody titre were described [15] (Supplementary S3).

3.1.2. Dose-Safety Relationship

The study showed that doses of 5×10^{10}, 1×10^{11} and 1.5×10^{11} VP induced 86%, 83% and 75% any grade adverse events and 6%, 6% and 17% grade 3+ adverse events, respectively. The calibrated saturating dose-adverse event curves are displayed in Figure 3.3b,c. The two thresholds of safety we previously chose were 17% and 30% grade 3+ adverse reaction proportion. The calibrated dose-adverse curve predicted that a rate of adverse events greater than 17% occurs for doses in excess of 1.58×10^{11} VP and exceeds 30% at 2.45×10^{11} VP.

3.2. Objective 2. Identify the Minimum Dose that Is Predicted to Theoretically Induce Herd Immunity

The dose-seroconversion prediction for the minimum dose that could induce theoretical herd immunity is shown in Figure 3a. Given that an estimate for the critical herd immunity threshold in the UK has been estimated as 65.5%, a dose of 1.3×10^{11} VP would be required to reach this threshold, assuming the entire UK population was vaccinated. The 95% confidence interval for optimal dose was (8.0×10^{10}, 7.9×10^{11}) (Supplementary S2.3.1). Using the dose-safety model, this dose was predicted to cause 13.5% of vaccinated individuals to have a grade 3+ adverse event.

(**a**)

(**b**)

(**c**)
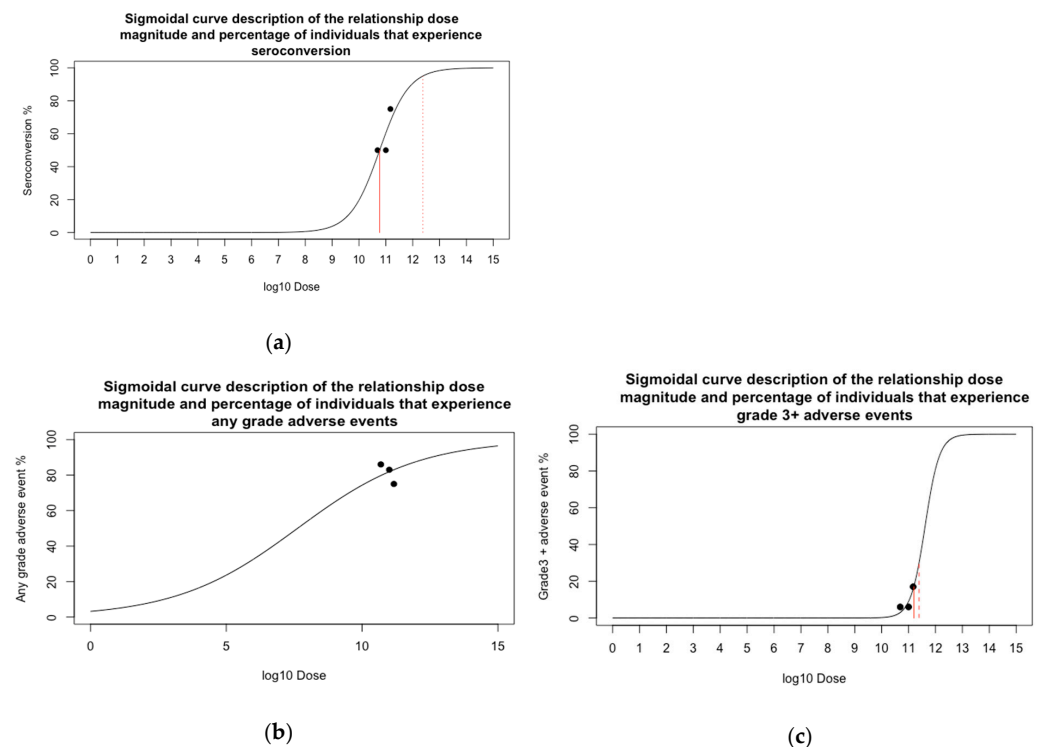
Figure 2. The three curves displaying the relationship between dose and (**a**) percentage of vaccinated individuals predicted to seroconvert, (**b**) percentage of vaccinated individuals predicted to experience any grade adverse events and (**c**) percentage of vaccinated individuals predicted to experience grade 3+ adverse events. The curves are sigmoid curves calibrated to data. Black dots represent the data the curves were calibrated to. In (**a**) the solid and dashed red lines show respectively the doses for which 50% and 90% of individuals are predicted to seroconvert. In (**c**) the solid and dashed red lines show respectively the doses for which 17% and 30% of individuals are predicted to experience grade 3+ adverse events. We note that the percentage of individuals experiencing any grade adverse events in (**b**) qualitatively decreased with increasing dose, whereas the model curve was increasing. This decreasing trend could be explained by the expected stochasticity in the data, hence the sigmoid model did not seem unreasonable (Supplementary S4).

3.3. Objective 3. Identify the Dose that Maximises Immunogenicity and Safety

The dose-utility prediction is shown in Figure 3b. The dose that optimised this function was 1.5×10^{11} VP (Figure 3b, red diamond). It was predicted that dosing at this magnitude would lead to a seroconversion rate of 67.6%, and cause 15.8% of vaccinated individuals to have a grade 3+ adverse event (83.0% any grade adverse events).

Sensitivity analysis (Supplementary S2) showed that the prediction of the optimal dose was most sensitive to variance in the $Dose_{50}$ parameter of the dose-seroconversion sigmoid function and the $Dose_{50}$ parameter of the dose-safety sigmoid function (Supplementary S2.2.1). These were respectively equal to the doses that were predicted to induce 50% of vaccinated individuals to seroconvert and 50% to experience grade 3+ adverse events (with an increase in these parameters qualitatively shifting the curves in Figure 3.3a,c to the right). The 95% confidence interval for optimal dose was $(2.9 \times 10^{10}, 5.0 \times 10^{11})$ (Supplementary S2.3.2).
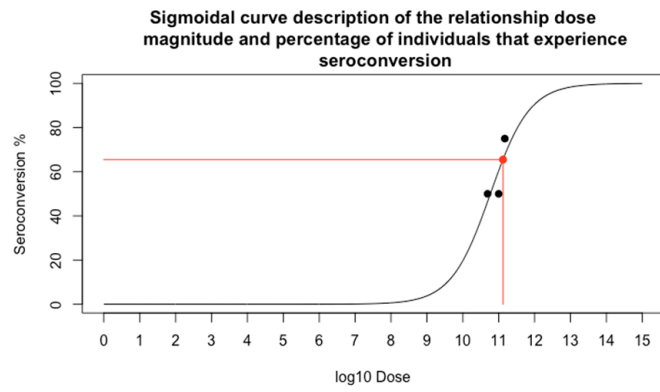
(**a**)
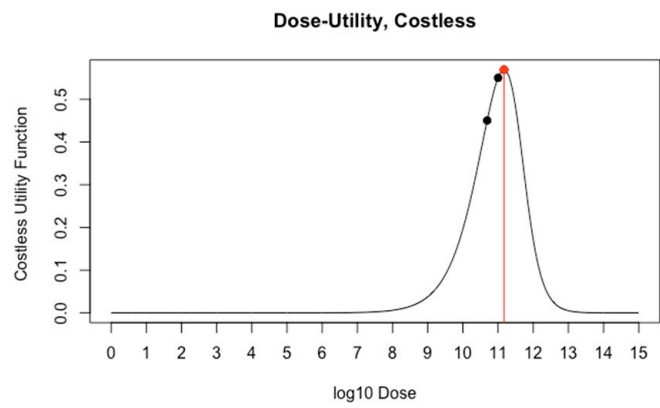
(**b**)
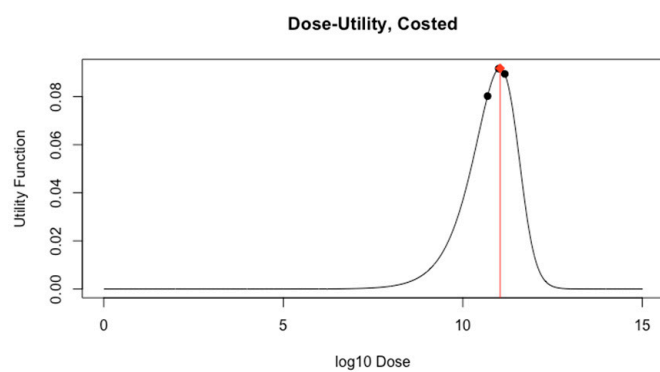
(**c**)

Figure 3. Displays of the predicted utility of doses between 10^0 and 10^{15} VP. (**a**) shows dose-seroconversion, with the horizontal red line indicating the 65.5% seroconversion threshold required for herd immunity. (**b**) shows the relationship between dose and the costless utility function and (**c**) shows the relationship between dose and the costed utility function. The black dots represent Table 1. 3×10^{11}, (**b**) 1.5×10^{11} VP and (**c**) 1.1×10^{11} VP.

3.4. Objective 4. Identify the Dose that Maximises Immunogenicity and Safety Whilst Minimising Cost

The dose-utility relation including cost is shown in Figure 3c. The dose that optimised this function was 1.1×10^{11} VP. It was predicted that dosing at this magnitude would lead to a seroconversion rate of 62.20%, cost £6.07 per dose, and cause 10.32% of vaccinated individuals to have a grade 3+ adverse event (82.2% any grade adverse events). The 1×10^{11} VP dose had the highest utility of the doses tested in the study, and both of the 5×10^{10} and 1.5×10^{11} VP doses appeared to be near-optimal. This analysis, therefore, suggested that if the cost was included in the utility function then a marginally reduced dose was found optimal relative to the costless utility function. The predicted cost is within the expected range [\$5–\$37] for a single SARS-CoV-2 vaccine dose [34].

Again, we found that the prediction of the optimal dose was most sensitive to variance in the $Dose_{50}$ parameter of the dose-seroconversion sigmoid function. This parameter is equal to the dose that we predict would induce 50% of vaccinated individuals to seroconvert (with an increase in these parameters qualitatively shifting the curve in Figure 3.3a to the right). Optimal dose was not sensitive to <10% error in the estimation of $Cost_{Delivery}$ or $Cost per viral particle$ (Supplementary S2.2.2). The 95% confidence interval for optimal dose was $(2.1 \times 10^{10}, 1.5 \times 10^{11})$ (Supplementary S2.3.3).

Threshold Analysis

For $Cost_{Delivery}$, we found that the predicted optimal dose was independent of the parameter value for large values (Figure 4). We found that the optimal dose was in excess of 1×10^{11} and 5×10^{10} VP for $Cost_{Delivery}$ values in excess of £3.79 and £0.65, respectively (hence optimal dose was only less than 5×10^{10} VP for $Cost_{Delivery}$ less than £0.65). These values were respectively 0.7 and 0.1 times the value that was used in the main analysis. We find that the optimal dose was not in excess of 5×10^{11} VP for any $Cost_{Delivery}$ values.

(**a**)　　　　　　　　　　　　　　　　　　　　　　　　(**b**)
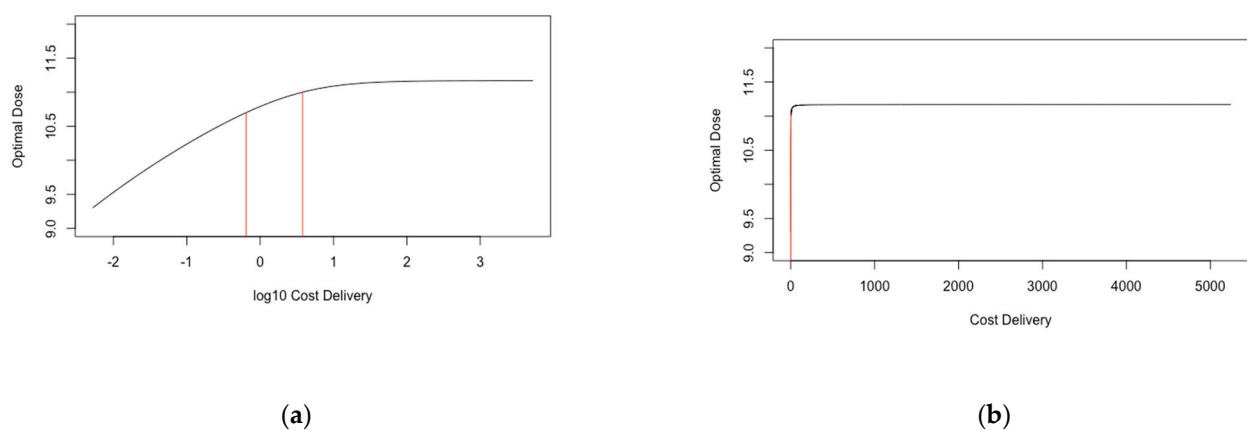
Figure 4. Optimal predicted dose for $+/-$ 3 orders of magnitude around $Cost_{Delivery}$. (**a**) has $Cost_{Delivery}$ at a log10 scale and (**b**) scaled normally. The black line represents the optimal dose, and the red lines indicate the threshold values of $Cost_{Delivery}$ for which optimal dose is 1×10^{11} and 5×10^{10} VP.

For *Cost per 10^{11} viral particles*, we found that the optimal dose was independent of the parameter value for large values (Figure 5). We found that the optimal dose was in excess of 1×10^{11} and 5×10^{10} VP for *Cost per 10^{11} viral particles* values in less than £1.06 and £6.23, respectively (hence optimal dose was only less than 5×10^{10} VP for *Cost per 10^{11} viral particles* greater than £6.24). These values were respectively 1.3 and 8.2 times the value that was used in the main analysis. We find that the optimal dose was not in excess of 5×10^{11} VP for any *Cost per 10^{11} viral particles* values.

(a)

(b)
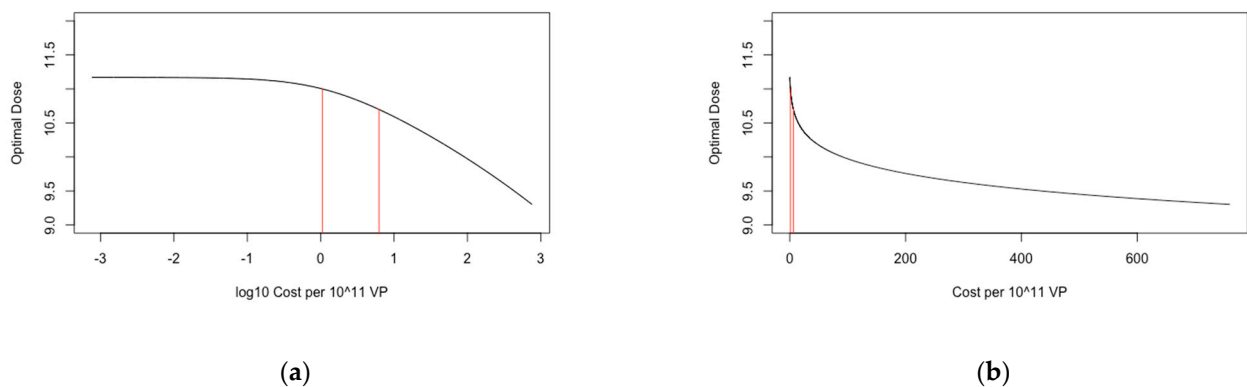
Figure 5. Optimal predicted dose (log10 scale) for $+/-3$ orders of magnitude (log10 scale) around *Cost per 10^{11} viral particles*. Table 1. *viral particles* at a log10 scale and the right is scaled normally. The black line represents the optimal dose, and the red lines indicate the threshold values of *Cost per 10^{11} viral particles* for which the optimal dose was 1×10^{11} and 5×10^{10} VP.

We additionally explored varying both parameters simultaneously (Supplementary S5). If *Cost per 10^{11} viral particles* was less than 0.2 times $Cost_{Delivery}$, then the predicted optimal dose was between 1.0×10^{11} and 1.5×10^{11} VP.

4. Discussion

Vaccination is an important part of global healthcare and disease prevention. Vaccination must be protective, safe, and affordable at a population level, and all of these factors may be impacted by dose. We used modelling and multifactorial optimisation approaches to predict the optimal dose of an adenoviral vectored vaccine against SARS-CoV-2 based on protection, safety and cost. A dose of 1.1×10^{11} VP of this vaccine was found to be optimal with respect to seroconversion, safety and cost. However, an increased dose of 1.3×10^{11} VP or 1.5×10^{11} VP could be justified depending on the objectives of developers and policymakers. These methods highlight how quantitative analysis can be used to ensure that vaccines are dosed optimally, and could aid in accelerating vaccine development.

The IS/ID methods used in this work have previously been used to analyse and optimise vaccine dose-response [7,34]. Compared to those studies, this work is novel in its inclusion of dose-safety and dose-cost models and multi-objective optimisation methods. This work used data published as part of a vaccine development protocol, which highlights how these methods do not require additional complexity in trial design. Using only the published data, we were able to hypothesize the best dosing for a candidate SARS-CoV-2 vaccine. Such methods could routinely be used to evaluate dose for other clinical and preclinical vaccine trials. Given the pandemic setting in which the SARS-CoV-2 candidates are being developed where trials are being accelerated and progressing faster than would normally be expected, modelling may be even more of an important adjunct in ensuring optimal vaccine dosing.

We had to make some assumptions in this work. Firstly, the assumed cost function was based on a simplification of a vaccine campaign cost estimate suggested by the World Health Organisation [35,36], discounting costs incurred by vaccine wastage and incremental costs of maintaining hospitals and clinics. Additionally, the exact cost per viral particle of the vaccine was unknown and would vary with production scale, however, the threshold analyses showed that the optimal-dose prediction may still be robust despite this uncertainty. As this was a financial rather than economic analysis, we did not account for "Disability Adjusted Life Years" [37] (DALYs) gained by reducing SARS-CoV-2 impact or societal costs incurred by missing work to be vaccinated. Including DALYs or similar measures into the cost function may give a greater understanding of whether this vaccine would be a cost-effective approach to controlling SARS-CoV-2. However, the focus of the work was not in producing a fully functional economic evaluation of implementing this vac-

cine. Additionally, to analyse potential DALYs averted would have required the inclusion of economic and epidemiological models that were beyond the scope of this paper.

An additional assumption of the utility model was that avoiding grade 3+ adverse events was as important to the utility function as inducing seroconversion. To address this, a weighting function may be applied if the expected discomfort of a SARS-CoV-2 infection is greatly in excess of a grade 3+ adverse event and a discomfort ratio between these two outcomes can then be determined (see Supplementary S6). Alternatively, thresholds for acceptable levels of seroconversion and adverse events could be determined, and any doses predicted to meet these thresholds considered optimal.

We found that dosing at 2.45×10^{11} VP would likely induce grade 3+ adverse events in greater than 30% of individuals vaccinated, which is a typical threshold for safety in clinical trials. Previous work has found that human-hosted adenoviral vector vaccine trials typically do not dose in excess of 2×10^{11} [21]. This suggests that adenoviral vaccine trials are being dosed at magnitudes that ensure that grade 3+ adverse reactions remain below the 30% threshold. However, for this vaccine, the available data was not sufficient to determine whether the dose-seroconversion curve shape was better described by a peaking or saturating curve shape. This implies that we cannot be confident that the percentage of individuals that seroconvert would continue to increase as dose increases beyond those empirically tested. This is likely the result of using too few doses or not dosing at a sufficiently large dose to observe peaking or saturating dose-response behaviour. We have previously shown that curve shape could not be determined for 75% of adenoviral dose-response data [8]. However, in this case, it is possible that dosing at a large enough magnitude to determine curve shape could cause an unacceptable number of grade 3+ adverse events.

There were limiting factors to our analyses. We had to assume that seroconversion implies that an individual was protected against SARS-CoV-2 infection. This seemed appropriate in the absence of a validated model for predicting SARS-CoV-2 protective immunity or a challenge study, but could be updated as a greater understanding of SARS-CoV-2 correlates of protection is developed [38–40]. Additionally, we had to assume that the base seroconversion and adverse events percentage was 0%. That is to say that individuals that received no vaccine dose would not seroconvert or experience any adverse events. This was reasonable given the lack of a placebo group in the data but may limit the predictive validity of toxicity and seroconversion at lower doses.

Further limitations are due to the non-inclusion of potential population effects and covariates in the model. The proportion of individuals that the vaccine needs to protect may change depending on the number of individuals that have been previously infected or on the extent that a prior infection provides lasting immunity. Additionally, prior adenoviral exposure or age of vaccinated individuals could impact the probabilities of seroconversion and adverse events. Individuals younger than 45 were shown to be less likely to experience fever and more likely to experience seroconversion [15] which, given that there existed some heterogeneity in age distribution between dosing groups in the data, may impact the model's future predictive validity. Finally, given that no grade 4 (serious/life-threatening) events were observed, no analyses could be done to assess the dose-varying probability of these events.

This work implies that the doses that have been trialled for this vaccine were near the theoretical optimal dose. Whilst we predicted 1.1×10^{11} VP of the vaccine would be the dose that optimises safety, cost, and protective immunity, if vial size restricts precision on which doses can be administered then, of the previously empirically tested doses, both the 1.0×10^{11} and 1.5×10^{11} doses could be reasonable. We also predicted that inducing complete herd immunity in an otherwise entirely susceptible population may be feasible with this vaccine given 100% uptake, but we predict that approximately 13.5% of vaccinated individuals would experience grade 3+ adverse events and that this would require a dose of at least 1.3×10^{11} VP. As the dose optimising the costless utility function was in excess of this threshold, but not the dose optimising the utility function with cost, this work implies

that to fully protect the UK population with this vaccine would require accepting some level of cost inefficiency.

We anticipate the following future work. Firstly, to ensure that the models used to make these suggestions are accurate and valid, further clinical trials would need to be conducted, preferably at a wider spread of doses for which empirical data do not yet exist. In particular, a placebo group could allow for adjustment of the assumption that individuals that receive no vaccine dose do not seroconvert or experience any adverse events. This may help to increase accuracy in the prediction of optimal dose. Secondly, the simple assumptions made in developing the optimisation utility functions mean the function could be applied broadly but should be adapted to the specific knowledge and needs of vaccine developers and policymakers. For example, including a specific adverse reaction threshold that has been defined in the study protocol, or by predicting protective immunity as highlighted in [41].

Thirdly, these methods applied to other candidate SARS-CoV-2 vaccines may provide a method to compare the relative utility of these candidates. Fourthly, with respect to the dose-safety model, including a weighting of the relative risks/discomforts associated with SARS-CoV-2 infection/adverse events would be informative. Fifthly, the data was gathered from individuals residing in the Wuhan region only. Similar data should be gathered from other populations to assess potential differences in response and safety. Given sufficient data, incorporating the covariates of age and pre-existing rAd5 neutralising antibody titre into the model could aid in predictive validity across various populations and in assessing dose effect.

Sixthly, this work only considers a single dose of the vaccine and response at one time-point. Further modelling could be attempted to address dose-optimisation the different time points or to consider a prime-boost paradigm. Finally, the dose-seroconversion and dose-safety models developed in this work could also be incorporated into the epidemiologic transmission and economic models to more accurately determine the health and economic impact of a given dose.

5. Conclusions

The SARS-CoV-2 pandemic has caused global health and economic issues and has led to increased pressure to rapidly develop a potentially life-saving vaccine. Dose is a key attribute in determining vaccine immunogenicity, safety and cost, and therefore dose-optimisation is an important aspect of vaccine development. Modelling and multifactorial optimisation methods allow for fast, quantitatively-based dosing decisions. Given the increased pressures for rapid vaccine development in response to pandemics, these tools should be considered a useful approach to accelerating vaccine development.

Data Availability Statement: Not applicable.

Conflicts of Interest: This work is partially funded by Vaccitech, a company that is developing novel adenoviral vector vaccines using the vectors ChAdOx1 and ChAdOx2.

References

1. Sharma, A.; Tiwari, S.; Deb, M.K.; Marty, J.L. Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2): A Global Pandemic and Treatment Strategies. *Int. J. Antimicrob. Agents* **2020**, *56*, 106054. [CrossRef] [PubMed]
2. Nicola, M.; Alsafi, Z.; Sohrabi, C.; Kerwan, A.; Al-Jabir, A.; Iosifidis, C.; Agha, M.; Agha, R. The Socio-Economic Implications of the Coronavirus Pandemic (COVID-19): A Review. *Int. J. Surg.* **2020**, *78*, 185–193. [CrossRef] [PubMed]
3. Greenwood, B. The Contribution of Vaccination to Global Health: Past, Present and Future. *Philos. Trans. R. Soc. B Biol. Sci.* **2014**, *369*, 20130433. [CrossRef] [PubMed]
4. Kaur, S.P.; Gupta, V. COVID-19 Vaccine: A Comprehensive Status Report. *Virus Res.* **2020**, *288*, 198114. [CrossRef]
5. Rhodes, S.J.; Knight, G.M.; Kirschner, D.E.; White, R.G.; Evans, T.G. Dose Finding for New Vaccines: The Role for Immunostimu-lation/Immunodynamic Modelling. *J. Theor. Biol.* **2019**, *465*, 51–55. [CrossRef]
6. Handel, A.; Li, Y.; McKay, B.; Pawelek, K.A.; Zarnitsyna, V.; Antia, R. Exploring the Impact of Inoculum Dose on Host Immunity and Morbidity to Inform Model-Based Vaccine Design. *PLoS Comput. Biol.* **2018**, *14*, e1006505. [CrossRef]
7. Rhodes, S.J.; Zelmer, A.; Knight, G.M.; Prabowo, S.A.; Stockdale, L.; Evans, T.G.; Lindenstrøm, T.; White, R.G.; Fletcher, H. The TB Vaccine H56+IC31 Dose-Response Curve Is Peaked Not Saturating: Data Generation for New Mathematical Modelling Methods to Inform Vaccine Dose Decisions. *Vaccine* **2016**, *34*, 6285–6291. [CrossRef]
8. Benest, J.; Rhodes, S.; Afrough, S.; Evans, T.; White, R. Response Type and Host Species May Be Sufficient to Predict Dose-Response Curve Shape for Adenoviral Vector Vaccines. *Vaccines* **2020**, *8*, 155. [CrossRef]
9. Frazão, T.D.C.; Camilo, D.G.G.; Cabral, E.L.S.; Souza, R.P. Multicriteria Decision Analysis (MCDA) in Health Care: A Systematic Review of the Main Characteristics and Methodological Steps. *BMC Med. Inform. Decis. Mak.* **2018**, *18*, 90. [CrossRef]
10. World Health Organisation. DRAFT Landscape of COVID-19 Candidate Vaccines. Available online: https://www.who.int/publications/m/item/draft-landscape-of-covid-19-candidate-vaccines (accessed on 9 September 2020).
11. Izda, V.; Jeffries, M.A.; Sawalha, A.H. COVID-19: A Review of Therapeutic Strategies and Vaccine Candidates. *Clin. Immunol.* **2020**, *222*, 108634. [CrossRef]
12. Oxford University Breakthrough on Global COVID-19 Vaccine | University of Oxford. Available online: https://www.ox.ac.uk/news/2020-11-23-oxford-university-breakthrough-global-covid-19-vaccine (accessed on 30 November 2020).
13. Mahase, E. COVID-19: Vaccine Candidate May Be More than 90% Effective, Interim Results Indicate. *BMJ* **2020**, *371*, m4347. [CrossRef] [PubMed]
14. Mullard, A. COVID-19 Vaccine Development Pipeline Gears Up. *Lancet* **2020**, *395*, 1751–1752. [CrossRef]
15. Zhu, F.-C.; Li, Y.-H.; Guan, X.-H.; Hou, L.-H.; Wang, W.-J.; Li, J.-X.; Wu, S.-P.; Wang, B.-S.; Wang, Z.; Wang, L.; et al. Safety, Tolerability, and Immunogenicity of a Recombinant Adenovirus Type-5 Vectored COVID-19 Vaccine: A Dose-Escalation, Open-Label, Non-Randomised, First-in-Human Trial. *Lancet* **2020**, *395*, 1845–1854. [CrossRef]
16. Manners, C.; Larios Bautista, E.; Sidoti, H.; Lopez, O.J. Protective Adaptive Immunity Against Severe Acute Respiratory Syndrome Coronaviruses 2 (SARS-CoV-2) and Implications for Vaccines. *Cureus* **2020**, *12*, e8399. [CrossRef]
17. US Food and Drug Administration. *Guidance for Industry: Toxicity Grading Scale for Healthy Adult and Adolescent Volunteers Enrolled in Preventive Vaccine Clinical Trials*; US Department of Health and Human Services, Food and Drug Administration: Silver Spring, MD, USA, 2007.
18. Sibille, M.; Patat, A.; Caplain, H.; Donazzolo, Y. A Safety Grading Scale to Support Dose Escalation and Define Stopping Rules for Healthy Subject First-Entry-into-Man Studies. *Br. J. Clin. Pharmacol.* **2010**, *70*, 736–748. [CrossRef]
19. Grothendieck, G. *Nls2: Non-Linear Regression with Brute Force*, R package version 0.2; 2013. Available online: https://CRAN.R-project.org/package=nls2 (accessed on 5 January 2019).
20. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2018.
21. Afrough, S.; Rhodes, S.; Evans, T.; White, R.; Benest, J. Immunologic Dose-Response to Adenovirus-Vectored Vaccines in Animals and Humans: A Systematic Review of Dose-Response Studies of Replication Incompetent Adenoviral Vaccine Vectors When Given via an Intramuscular or Subcutaneous Route. *Vaccines* **2020**, *8*, 131. [CrossRef]
22. Wang, C.; Rosner, G.L.; Roden, R.B.S. A Bayesian Design for Phase I Cancer Therapeutic Vaccine Trials. *Stat. Med.* **2019**, *38*, 1170–1189. [CrossRef]
23. Wages, N.A.; Slingluff, C.L. Flexible Phase I–II Design for Partially Ordered Regimens with Application to Therapeutic Cancer Vaccines. *Stat. Biosci.* **2020**, *12*, 104–123. [CrossRef]
24. Le Tourneau, C.; Lee, J.J.; Siu, L.L. Dose Escalation Methods in Phase I Cancer Clinical Trials. *JNCI J. Natl. Cancer Inst.* **2009**, *101*, 708–720. [CrossRef]
25. Safety Information by Vaccine | CDC. Available online: https://www.cdc.gov/vaccinesafety/vaccines/index.html (accessed on 15 October 2020).
26. Average for the Year to 31 March 2020—GOV.UK. Available online: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/885947/Spot.csv/preview (accessed on 16 October 2020).

27. Inflation Calculator. Available online: http://www.bankofengland.co.uk/monetary-policy/inflation/inflation-calculator (accessed on 16 October 2020).
28. Agenda for Change—Pay Rates. Available online: https://www.healthcareers.nhs.uk/working-health/working-nhs/nhs-pay-and-benefits/agenda-change-pay-rates (accessed on 16 October 2020).
29. Annual Leave and Holiday | Advice Guides | Royal College of Nursing. Available online: /get-help/rcn-advice/annual-leave-and-holiday-pay (accessed on 16 October 2020).
30. *Vaccinating Adults: A Step-by-Step Guide*; Centers for Disease Control and Prevention: St. Paul, MI, USA, 2017.
31. Pereira, C.C.; Bishai, D. Vaccine Presentation in the USA: Economics of Prefilled Syringes versus Multidose Vials for Influenza Vaccination. *Expert Rev. Vaccines* **2010**, *9*, 1343–1349. [CrossRef]
32. GE Healthcare. Scalable Process for Adenovirus Production 2018. Available online: https://cdn.cytivalifesciences.com/dmm3bwsv3/AssetStream.aspx?mediaformatid=10061&destinationid=10016&assetid=27018 (accessed on 15 October 2020).
33. Kwok, K.O.; Lai, F.; Wei, W.I.; Wong, S.Y.S.; Tang, J.W.T. Herd Immunity—Estimating the Level Required to Halt the COVID-19 Epidemics in Affected Countries. *J. Infect.* **2020**, *80*, e32–e33. [CrossRef] [PubMed]
34. Raja, A.T.; Alshamsan, A.; Al-jedai, A. Current COVID-19 Vaccine Candidates: Implications in the Saudi Population. *Saudi Pharm. J. SPJ* **2020**, *28*, 1743–1748. [CrossRef] [PubMed]
35. Rhodes, S.J.; Guedj, J.; Fletcher, H.A.; Lindenstrøm, T.; Scriba, T.J.; Evans, T.G.; Knight, G.M.; White, R.G. Using Vaccine Immunostimulation/Immunodynamic Modelling Methods to Inform Vaccine Dose Decision-Making. *NPJ Vaccines* **2018**, *3*, 36. [CrossRef] [PubMed]
36. Newall, A.T.; Chaiyakunapruk, N.; Lambach, P.; Hutubessy, R.C.W. WHO Guide on the Economic Evaluation of Influenza Vaccination. *Influenza Other Respir. Viruses* **2018**, *12*, 211–219. [CrossRef]
37. World Health Organistaion. *Guidelines for Estimating Costs of Introducing New Vaccines into the National Immunization System*; World Health Organisation: Geneva, Switzerland, 2002.
38. Murray, C.J. Quantifying the Burden of Disease: The Technical Basis for Disability-Adjusted Life Years. *Bull. World Health Organ.* **1994**, *72*, 429–445.
39. Mercado, N.B.; Zahn, R.; Wegmann, F.; Loos, C.; Chandrashekar, A.; Yu, J.; Liu, J.; Peter, L.; McMahan, K.; Tostanoski, L.H.; et al. Single-Shot Ad26 Vaccine Protects against SARS-CoV-2 in Rhesus Macaques. *Nature* **2020**, *586*, 583–588. [CrossRef]
40. Oland, G.A.; Ovsyannikova, I.G.; Kennedy, R.B. SARS-CoV-2 Immunity: Review and Applications to Phase 3 Vaccine Candidates. *Lancet* **2020**, *396*, 1595–1606. [CrossRef]
41. Deng, W.; Bao, L.; Liu, J.; Xiao, C.; Liu, J.; Xue, J.; Lv, Q.; Qi, F.; Gao, H.; Yu, P.; et al. Primary Exposure to SARS-CoV-2 Protects against Reinfection in Rhesus Macaques. *Science* **2020**, *369*, 818–823. [CrossRef]

Supplementary material for paper 3

These supplementary are adapted from those that were included with the published version of this paper, which can be found online [192].

Any references in these supplementary documents use the reference numbers of paper 3.

S1. Peaking vs Saturating

We calibrated both the sigmoid function and a representative peaking curve as we have previously described in [8], and calculated AIC (Akaike Information Criterion) of both. As the difference in AIC between the models was less than 2.0, there was not considered to be evidence to choose the peaking curve over the saturating curve. Thus, the saturating (sigmoidal) curve was assumed to describe the dose-seroconversion dynamics.
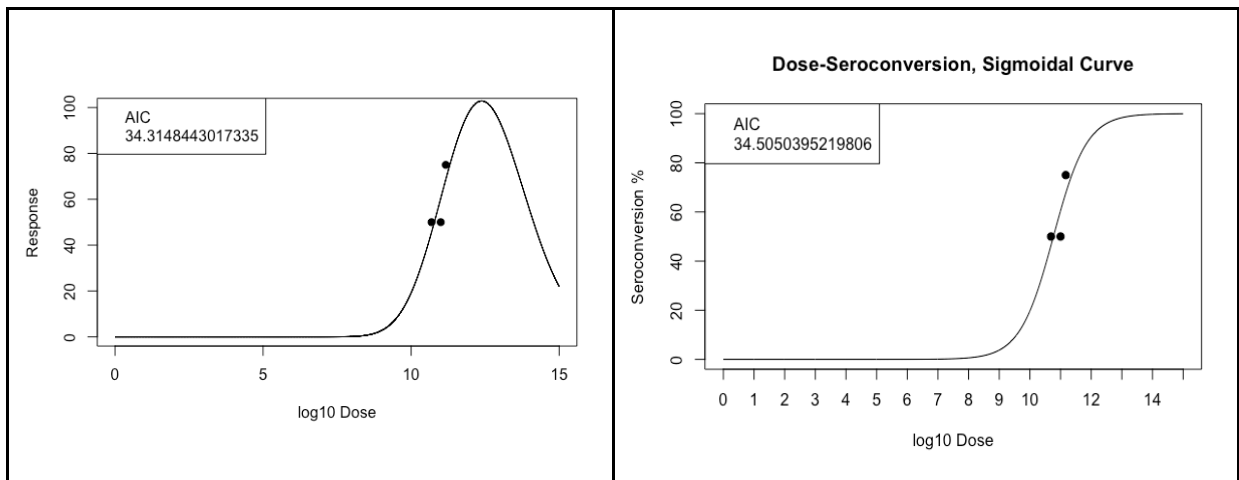
Figure S1. Calibrated peaking (left) and saturation/sigmoidal(right) curves. The y axis is predicted seroconversion, and black points are the available data. AICs are given and were within 2 of each other.

S2. Sensitivity

We attempted to account for uncertainty in the data and models.

143

Available seroconversion data gave the number of individuals per dosing group and the number that was seroconverted. Following a Bayesian perspective of the data and a parametric bootstrapping approach, we consider each dosing group as being sampled from a binomial distribution with n = 36 and the unknown true probability of seroconversion p. The likelihood distribution of p was calculated for each of the dosing groups, and we can consider that the true probability of seroconversion follows these likelihood distributions for each group. This was repeated for the adverse event data (Figure S2).
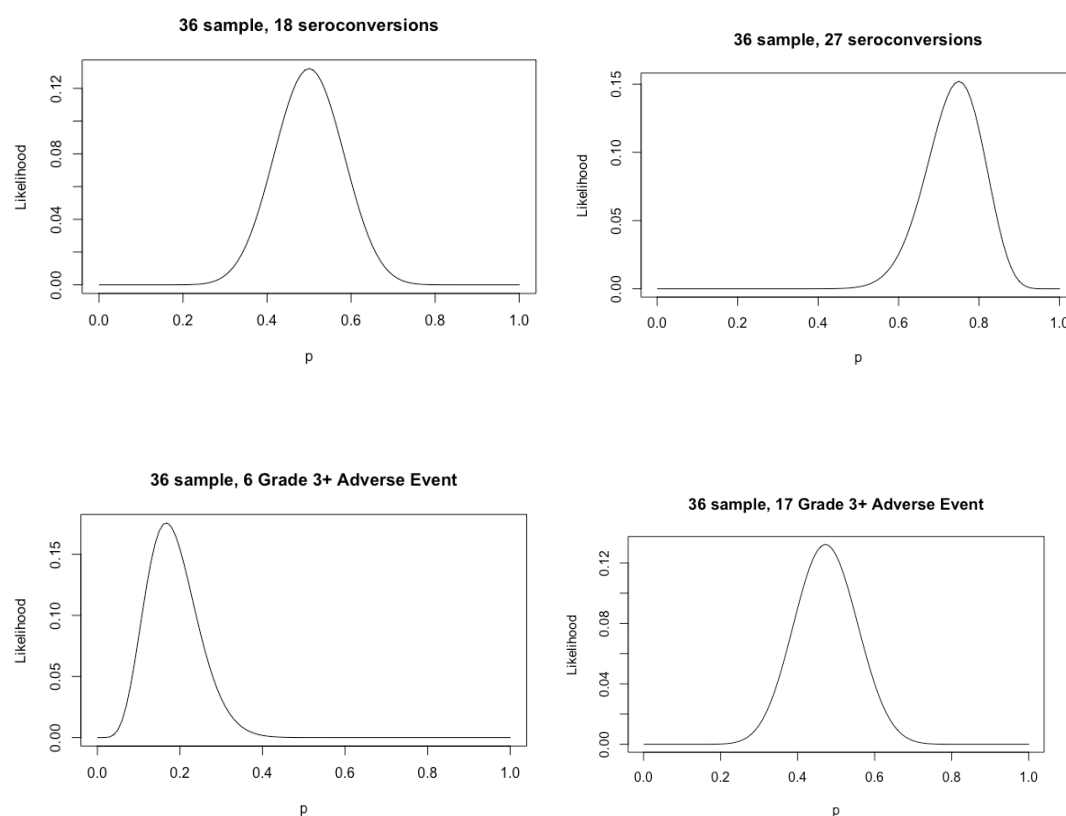
Figure S2. Likelihood distribution of true seroconversion (top) or true grade 3+ adverse reaction (bottom) probabilities for binomial processes with 36 trials and S successes. Top left shows the distribution given S = 18 (dose = 5 x 10^{10}, 1 x 10^{11}), top right shows the distribution given S = 27 (dose = 1.5 x 10^{11}), bottom left shows the distribution given S = 6 (dose = 5 x 10^{10}, 1 x 10^{11}), bottom left shows the distribution given S = 17 (dose = 1.5 x 10^{11}).

We sampled from each of these likelihood distributions 5000 times to create 5000 bootstrap dose-response data sets. For each of these data sets, we calibrated a sigmoid curve and recorded *MaxResponse, Scale,* and *Dose$_{50}$* for each. This gave a

144

distribution of the values of *MaxResponse, Scale,* and *Dose*$_{50}$ for the seroconversion that were reasonable giving the observed data. This was repeated for the adverse event data to give a distribution of the values of *Scale* and *Dose*$_{50}$ for the safety curve (Figure S3).
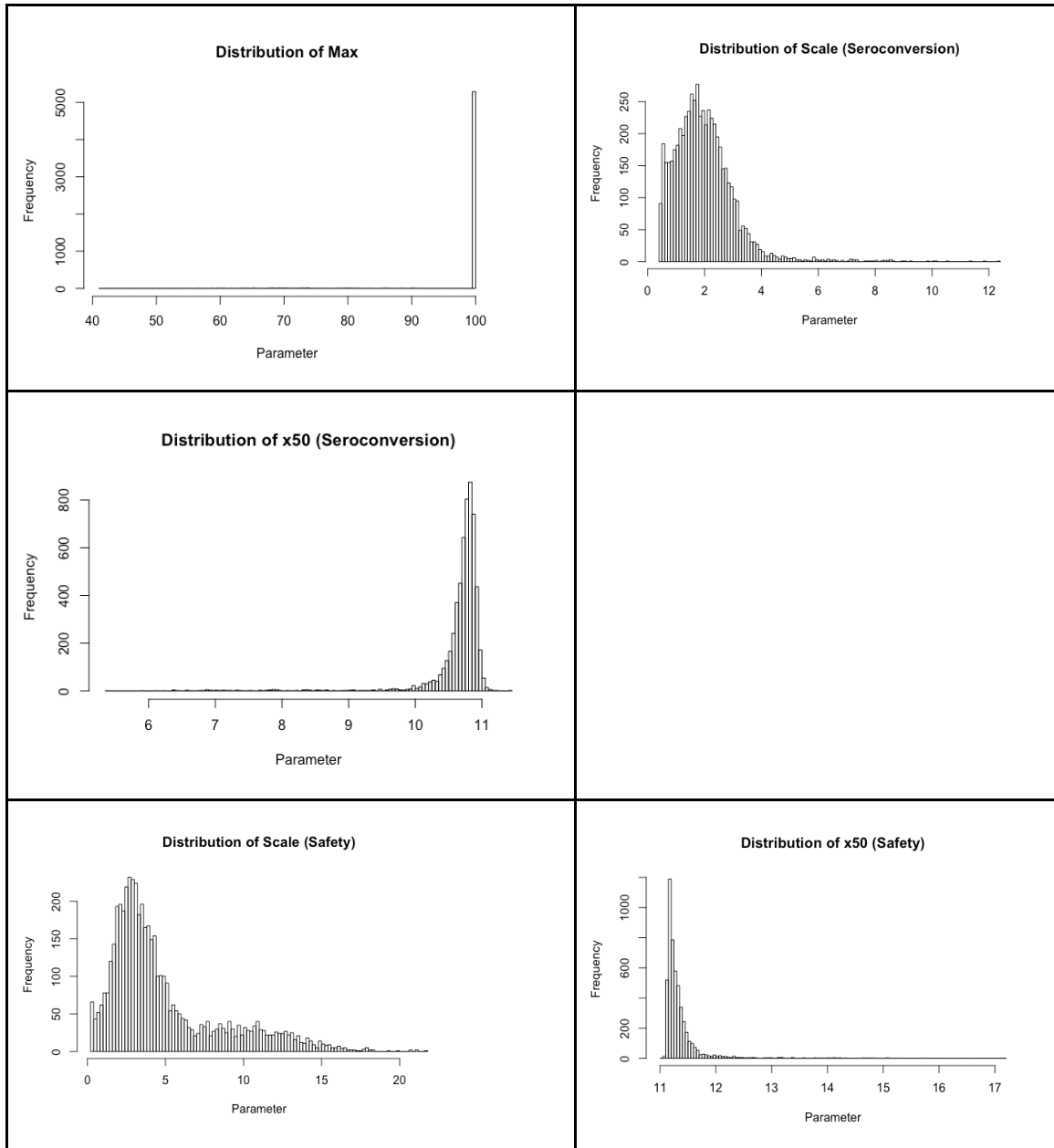
Figure S3. Distribution of model parameters following bootstrapping process with 5000 samples.

The two *Scale* parameters appeared to be the least well identified, and *MaxResponse* appeared well identified. We calculated a non-parametric 95% confidence interval for parameters by finding the 2.5th and 97.5th percentile of the

145

parameter distributions. This is given in the varying range column of Table S1. As the parameters of the cost model were literature derived, we instead allowed them to vary in a 10% region around the value derived from literature. We write *Cost per 10^{11} viral particles* rather than *Cost per viral particle* due to the small size of the *Cost per viral particle* parameter. Note that this does not alter the analysis.

Parameters	Model	Description	Value	Varying range
MaxResponse	Seroconversion	Predicted seroconversion given infinite dose.	100	[70.17,100]
Scale (Seroconversion)	Seroconversion	Gradient of seroconversion sigmoid function	1.83145	[0.53, 4.31]
Dose$_{50}$ (Seroconversion)	Seroconversion	Dose required to reach 50% of *MaxResponse* seroconversion	10.767948	[9.49,10.97]
Scale (Safety)	Safety	Gradient of safety sigmoid function	3.76339	[0.67, 14.17]
Dose$_{50}$ (Safety)	Safety	Dose required to reach 50% of individuals experiencing grade 3+ adverse events	11.6145	[11.12, 12.35]
Cost$_{Delivery}$	Cost	Dose independent vaccination costs	5.24	[4.72, 5.77]
Cost per 10^{11} viral particles	Cost	Dose dependent vaccination costs (per 10^{11}VP)	0.76	[0.68, 0.84]

Table S1. Parameters that were explored through sensitivity analysis. The 'value' column gives the numerical value which were used in objectives 2 and 3 and the 'varying range' column gives a reasonable bound on the parameter values as in supplementary S2.1.

S2.2. Parameter Sensitivity

These parameters define the utility function. To determine the sensitivity of the optimal dose prediction to a parameter, θ, we fix all other parameters at the calibrated/literature derived value and allow θ to vary in the region around it that we just defined [Table S1]. The optimal dose for each of these varying θ values were

146

calculated and plotted. We did this analysis for both the costless and cost utility functions.

S2.2.1. Costless

The utility function was most sensitive to variance in the $Dose_{50}$ *(Seroconversion)* and *Scale (Safety)* parameters, but some uncertainty in optimal dose may also be caused by variance in the estimated *Scale (Seroconversion)* parameter.
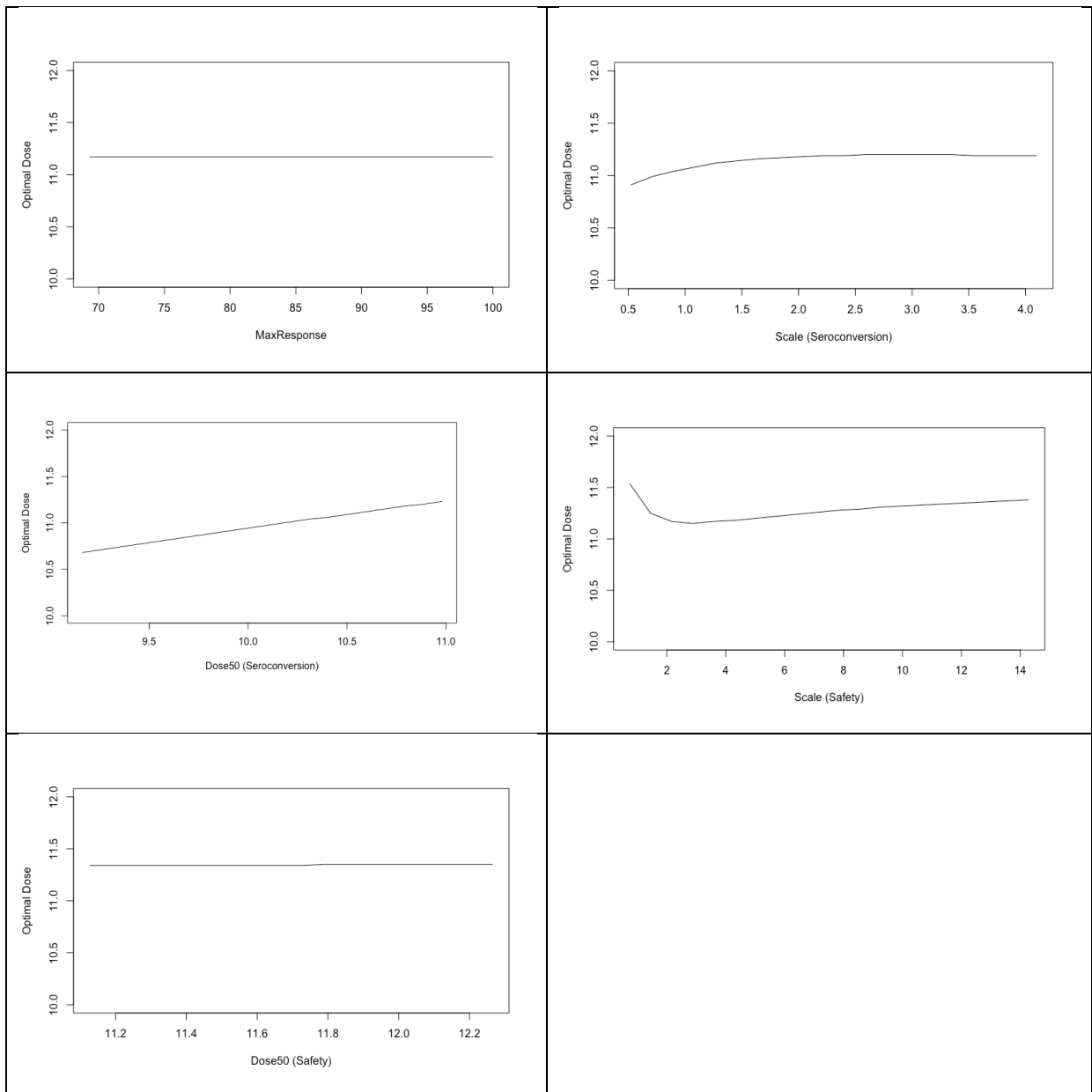
Figure S4a. Sensitivity of costless optimal dose by the model parameters of MaxResponse, Scale (Seroconversion), Dose$_{50}$ (Seroconversion), Scale (Safety), Dose$_{50}$ (Safety).

S2.2.2. Costed

The utility function was again most sensitive to variance in the *Dose$_{50}$(Seroconversion)* parameter but was less sensitive to uncertainty in the *Scale (Safety)* parameter. Again, optimal dose may also be caused by variance in the estimated *Scale (Seroconversion)* parameter. Optimal dose did not appear to be sensitive to variance in either cost parameter.
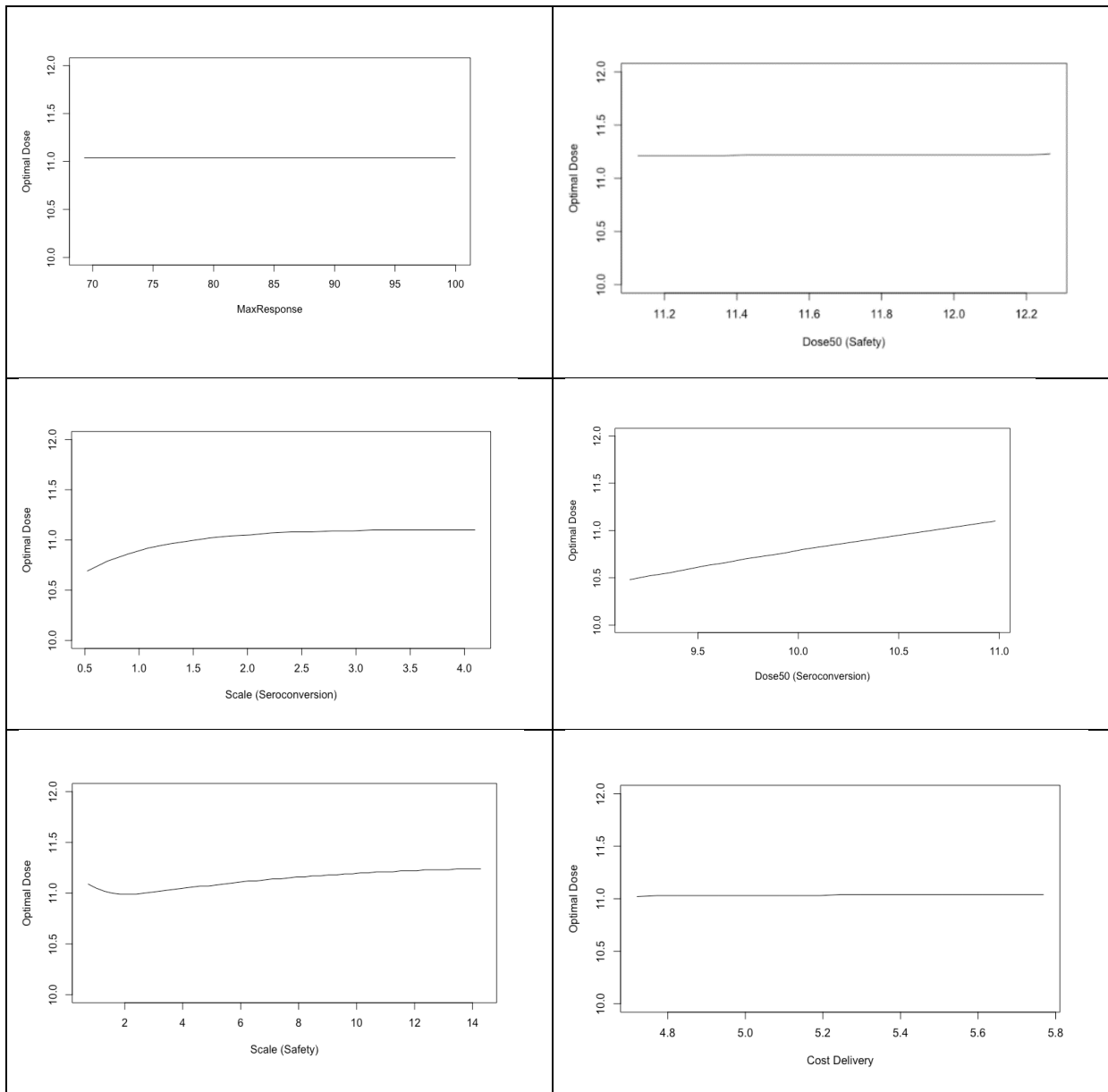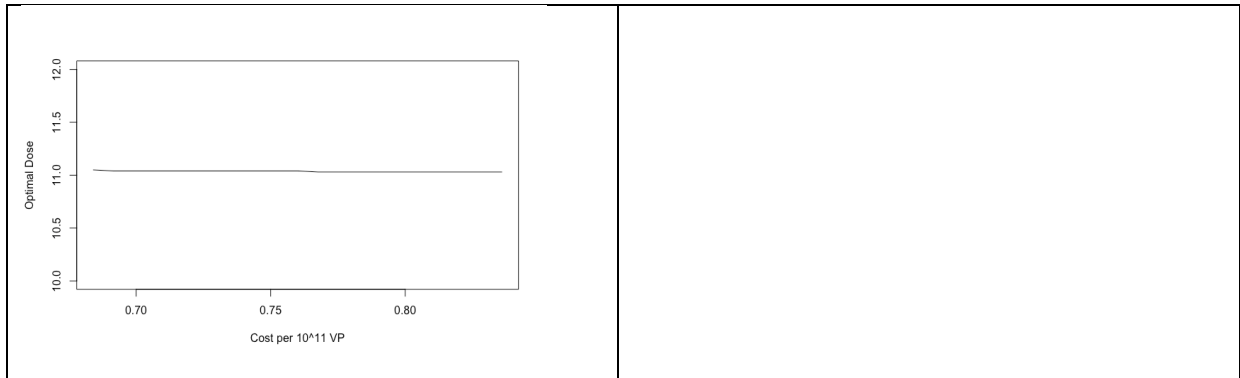
Figure S4b. Sensitivity of costed optimal dose by the model parameters of MaxResponse, Scale (Seroconversion), Dose$_{50}$ (Seroconversion), Scale (Safety), Dose$_{50}$ (Safety). CostDelivery and cost per 10^{11} viral particles

S2.3. Optimal dose Confidence Interval

We resampled with replacement from the bootstrap dose-response data calibrated parameter sets. We did this for a combined 10,000 seroconversion/safety parameter sets. For these we calculated the optimal dose as defined by all utility functions. This was used to calculate an approximate 95% confidence interval for the optimal dose of both utility functions.

S2.3.1. Herd Immunity

Optimal dosing follows an approximate skewed normal distribution (figure S4), with the qualitative peak being approximately $10^{11.11}$ (=1.3 x 10^{11}). The calculated 2-tailed 95% distribution of the data had lower bound $10^{10.90}$ (= 0.8 x 10^{11}) and upper bound $10^{11.90}$ (= 7.9 x 10^{11}).

149

Histogram of Optimal Doses, Optimise for herd immunity
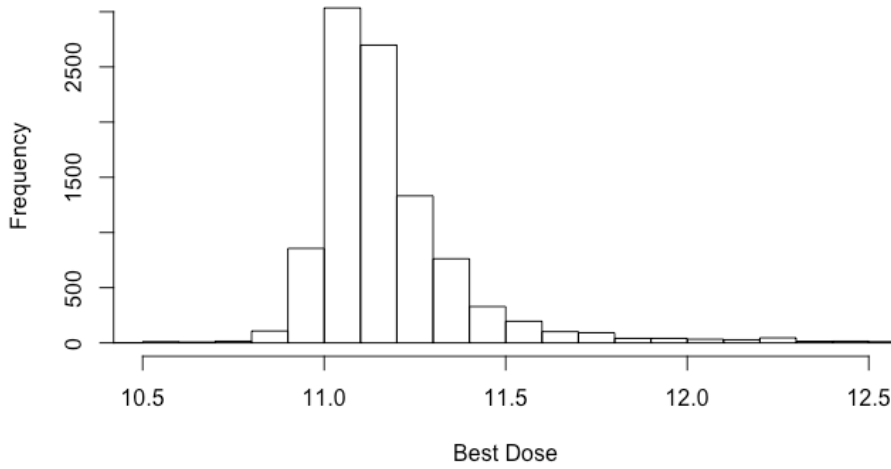
Figure S5. Distribution of optimal dose based on herd immunity from parametric bootstrapping of the data.

S2.3.2. Costless

Optimal dosing follows an approximate normal distribution (figure S4), with the qualitative peak being approximately 10^{11} (=1.0 x 10^{11}). The calculated 2-tailed 95% distribution of the data had lower bound $10^{10.46}$ (= 0.29 x10^{11}) and upper bound $10^{11.67}$ (= 5.0 x 10^{11}).

Histogram of Optimal Doses, No Cost
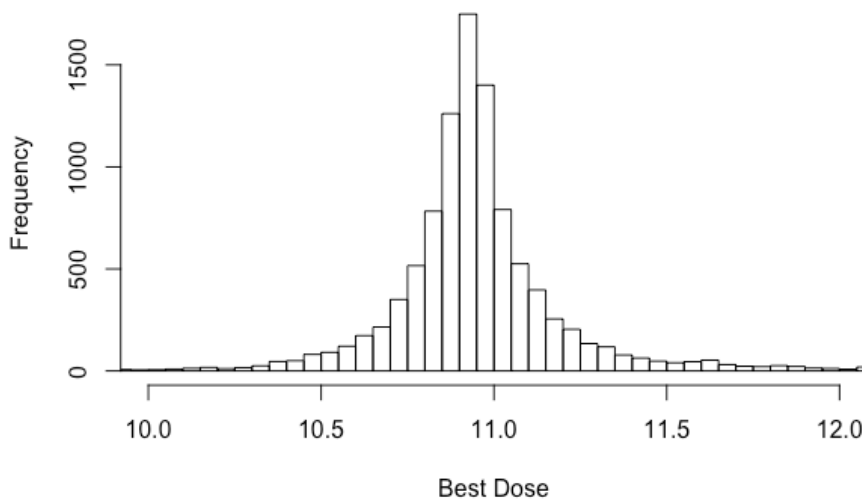
Figure S6. Distribution of optimal dose without including cost from parametric bootstrapping of the data.

S2.3.3. Costed

Optimal dosing follows an approximate left-skewed normal distribution (figure S5), with the qualitative peak at approximately $10^{10.9}$ (= .75 x 10^{11}). The calculated 2-tailed 95% distribution of the data had lower bound $10^{10.32}$ (= 0.21 x 10^{11}) and upper bound $10^{11.18}$ (= 1.54 x 10^{11}).
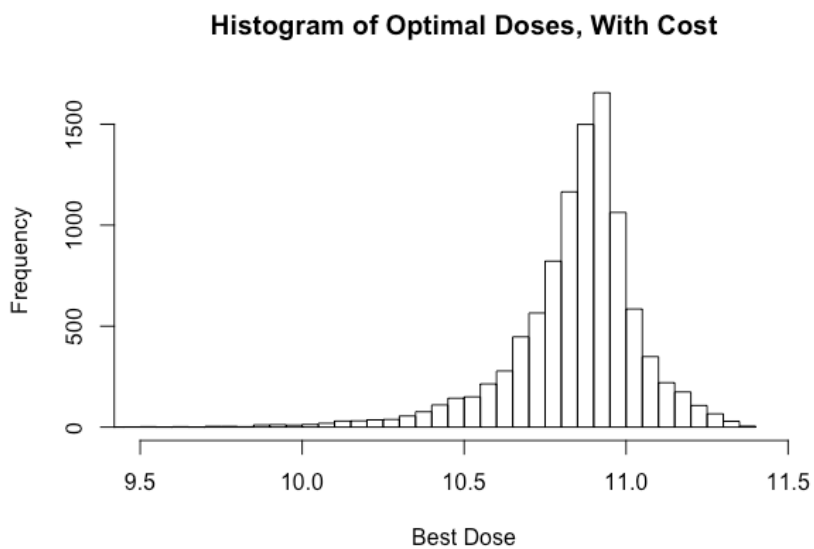
Figure S7. Distribution of optimal dose (log10 scale) including cost from parametric bootstrapping of the data.

S3. Population demographics

Population demographics of age, gender, and pre-existing adenovirus neutralising antibody titre as described in the body of work the data were extracted from.

151

	Dose (VP)		
	5.0×10^{10}	1.0×10^{11}	1.5×10^{11}
Age			
18–29	9 (25%)	12 (33%)	10 (28%)
30–39	13 (36%)	14 (39%)	15 (42%)
40–49	8 (22%)	3 (8%)	7 (19%)
50–60	6 (17%)	7 (19%)	4 (11%)
Mean age (years)	37.2 (sd =10.7)	36.3 (sd = 11.5)	35.5 (sd = 10.1)
Sex			
Male	18 (50%)	19 (53%)	18 (50%)
Female	18 (50%)	17 (47%)	18 (50%)
Pre-existing adenovirus type-5 neutralising antibody			
≤200, titre	16 (44%)	17 (47%)	20 (56%)
>200, titre	20 (56%)	19 (53%)	16 (44%)
Mean geometric mean titre	168.9 (13.9)	149.5 (10.5)	115.0 (13.4)

Table S2. Distribution of sample covariates for each dosing group. Data are given as number (percentage).

S4. Variability in the data.

We note that in plot 2b the data shows that for the three dosing groups (5.0×10^{10}, 1.0×10^{11} and 1.5×10^{11}), 86%, 83%, and 75% of individuals experienced any grade adverse events respectively. This represents respectively that for each of the three dosing groups of size N=36 (31,30, and 27) individuals of individuals experienced

any grade adverse events. There is a qualitative downwards trend, which our strictly increasing sigmoid model would be unable to model.

We considered using this data whilst taking the interpretation that individuals are independent samples of an underlying Bernoulli process, we can calculate the 95% confidence interval on the true probability of experiencing any grade adverse events, using a similar approach to that described in [S2]. These are:

- For dose 5.0 x10^{10}; 86% (71%,95%)
- For dose 1.0 x 10^{11}; 83% (67%,94%)
- For dose 1.5 x 10^{11}; 75% (58%,88%)

As these confidence intervals do overlap, we did not believe that there was sufficient justification to consider the possibility that an increased dose could reduce the number of adverse events experienced, even given the downward trend observed. We believe it more likely that all three data points have similar probabilities of any grade adverse events.

To illustrate this point please consider the below plot, which shows the data described overlaid with the 95% confidence intervals for Bernoulli trials assuming that our underlying model is correct. As all of the points are within these bounds, again this model seems reasonable with the available data.
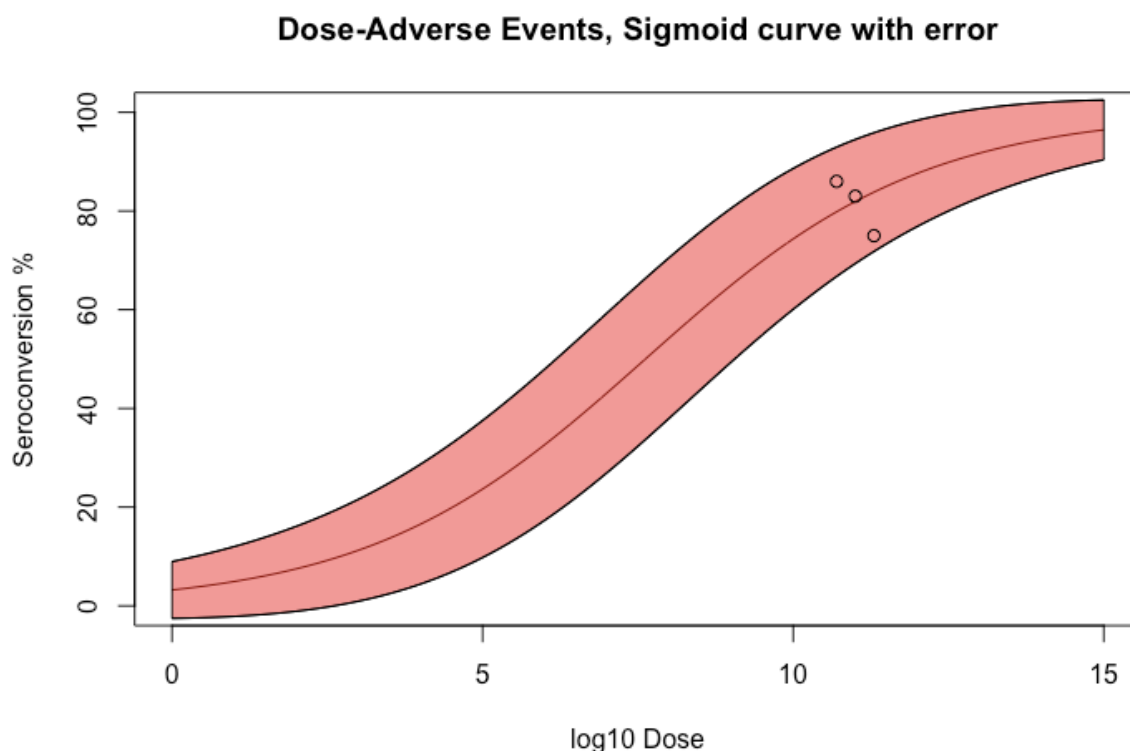
Figure S8. A plot of the expected any-grade adverse event data compared to observed data.
The black line plots the calibrated curve. The red area plots the 95% confidence interval for the
percentage of individuals that would experience any-grade adverse events in a 36-per-group trial
assuming that this is the true model (for example, if the true probability of an adverse event for a
given dose is 0.5, then approximately 95% of trials of that dose with size 36 would have between
13(=36%) and 23(=63%) individuals experiencing adverse events)

However, further investigation into the relationship between dose and proportion of
individuals experiencing adverse events would be useful if there was sufficient data.

S5. Threshold analysis, Bivariable

We considered varying both of the $Cost_{Delivery}$ and *Cost per* 10^{11} *viral particles*
parameters in the +/- 3 orders of magnitude range at the same time, and found that
for high values of $Cost_{Delivery}$ the optimal dose was independent of *Cost per* 10^{11}*viral
particles* (Figure S6). If these plots are censored to include only points where the
predicted optimal dose is less than 10^{11}, 5×10^{10} and 10^{10} VP, we find the behaviour
observed in Figures S7, S8, and S9 respectively. A clear linear separation is
observed for all three plots. By finding the line between (0,0) and the point with the
maximum $Cost_{Delivery}$, we can approximate these decision boundaries as respectively

154

10^{11}: Cost per 10^{11} viral particles= 0.2 x Cost$_{Delivery}$

5 x 10^{10}: Cost per 10^{11} viral particles= 1.3 x Cost$_{Delivery}$

10^{10}: Cost per 10^{11} viral particles= 17.9 x Cost$_{Delivery}$

Hence, we can suggest that, assuming no uncertainty in the safety and seroconversion related model parameters:

- If the cost per 10^{11} VP is greater than 0.2 times the cost per vaccination that is independent of dose, optimal dose is less than 10^{11} VP.
- If the cost per 10^{11} VP is greater than 1.3 times the cost per vaccination that is independent of dose, the optimal dose is less than 5 x 10^{10} VP.
- If the cost per 10^{11} VP is greater than 17.9 times the cost per vaccination that is independent of dose, the optimal dose is less than 10^{10} VP.
- In all other cases optimal dose is greater than 10^{11}, with the largest recommended dose across all costing parameters was 1.5 x 10^{11} VP, which is the dose recommended by the results in objective 2.
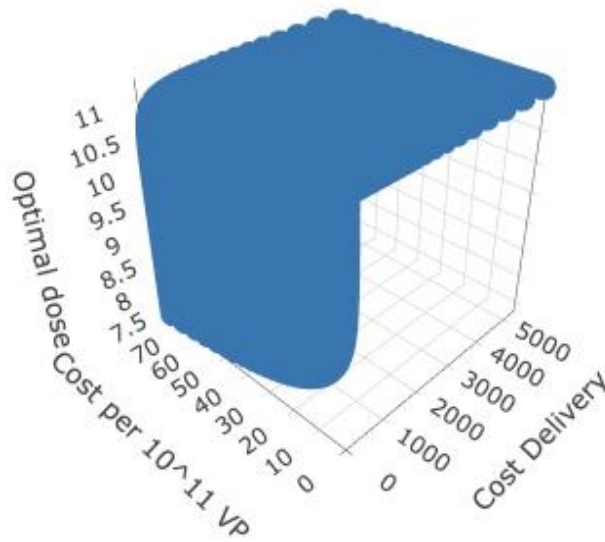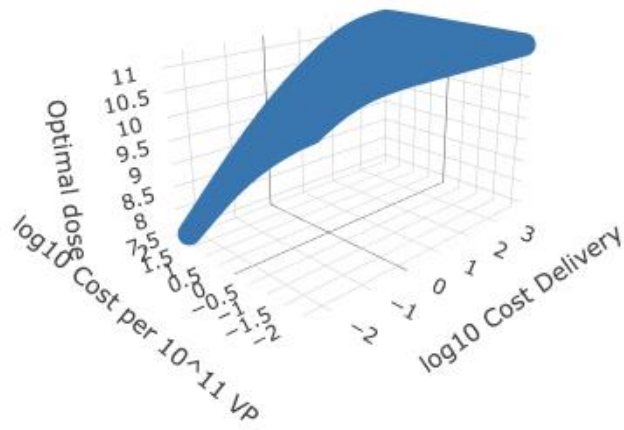
Figure S9. Optimal predicted dose for +/- 3 orders of magnitude (log10 scale) around *Cost per 10^{11} viral particles* and *Cost_{Delivery}*. The top has *Cost per 10^{11} viral particles* at a log10 scale, and the bottom scaled normally.

Optimal Dose less than 10^11
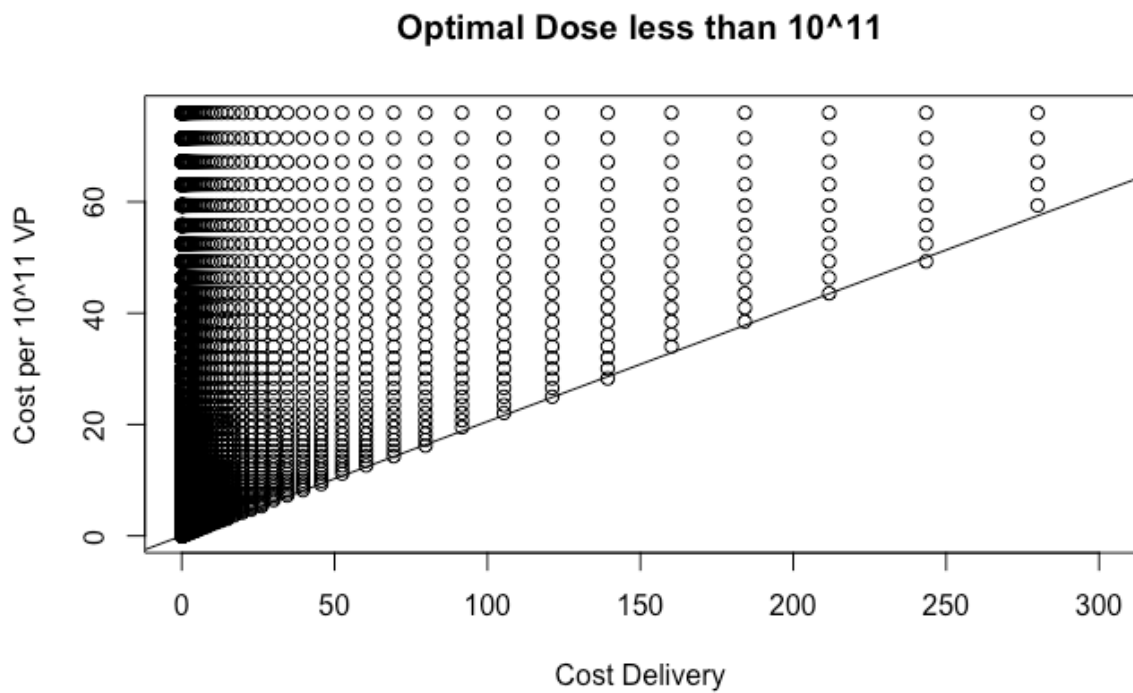
Figure S10. Pairs of Cost per 10^{11} viral particles and CostDelivery for which the optimal predicted dose was less than 10^{11} VP. Black line represents the estimated decision boundary.

Optimal Dose less than 5 x 10^10
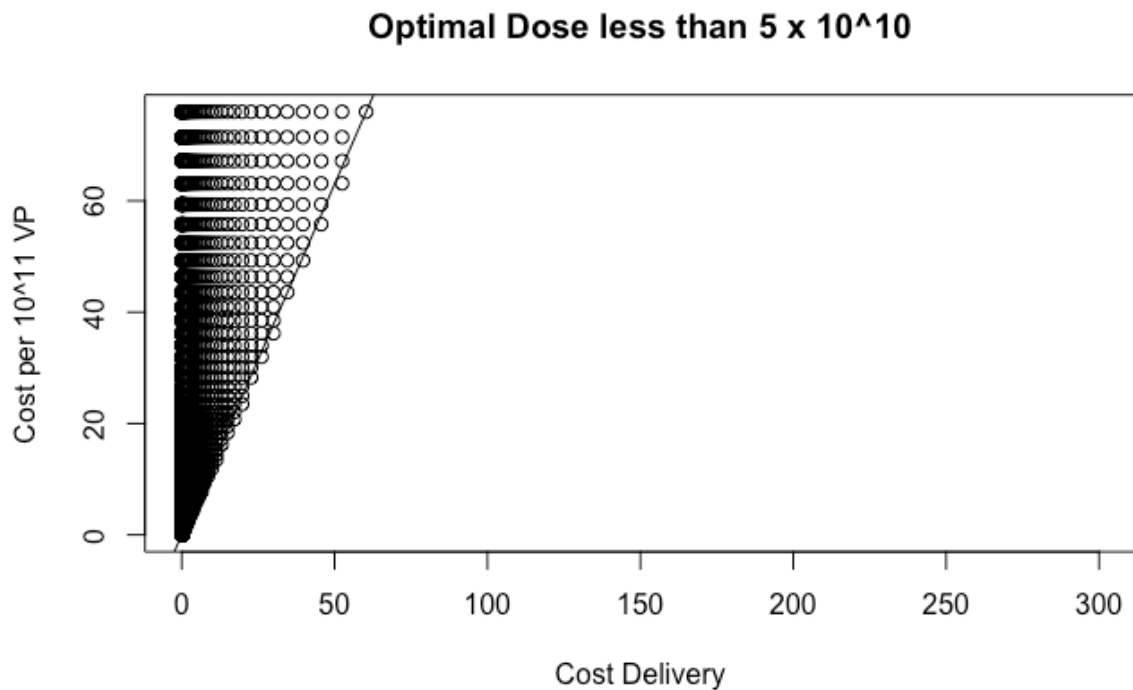
Figure S11. Pairs of Cost per 10^{11} viral particles and Cost$_{Delivery}$ for which the optimal predicted

157

dose was less than 5 x 10^{10} VP. Black line represents the estimated decision boundary.

Optimal Dose less than 10^10
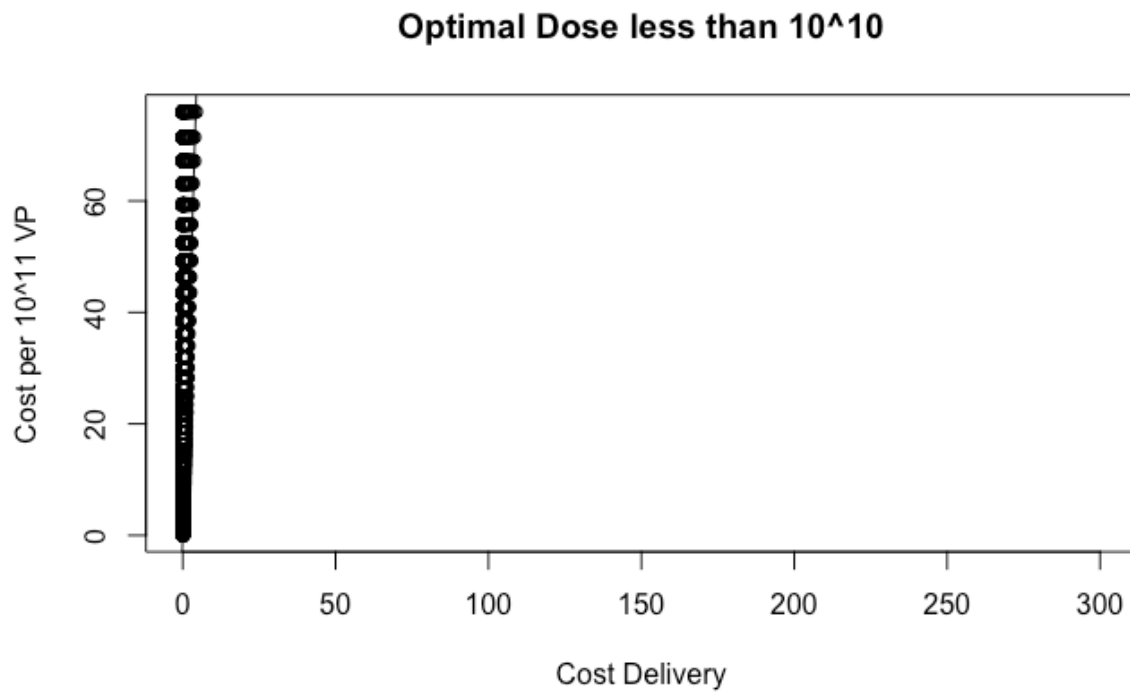
Figure S12. Pairs of *Cost per 10^{11} viral particles* and *Cost$_{Delivery}$* for which the optimal predicted dose was less than 10^{10} VP. Black line represents the estimated decision boundary.

S6. Weighted Utility Functions

In this section I considered a further two utility functions similar to that used in chapter 5. See appendix [A.D.2.].

Chapter 5: Theoretical analysis of mathematical modelling for vaccine dose optimisation: efficacy curve shape, trial size, and trial dose selection method: Paper 4

Chapter 5 Introduction

In chapter 4 I investigated vaccine dose-optimisation as a multi-objective optimisation problem and predicted potential optimal doses for three different utility functions using mathematical modelling of publicly available clinical trial data. Whether these predicted optimal doses for that vaccine were indeed optimal for their respective utility functions could not be validated without conducting extensive clinical trials that were outside the scope of this PhD or having perfect knowledge of the true vaccine dose-response curve. Additionally, in both chapter 4 and chapter 3 I had found that adenoviral vector vaccine clinical trial data are often not sufficient for modellers to be able to discriminate between peaking/saturating dose-response curve shape.

Following these observations, in chapter 5, I investigated the true optimality of doses that are predicted optimal using mathematical models, and into whether model misspecification could lead to suboptimal vaccine dosing. In order to do this, I used simulation study methodology, which meant that I had access to perfect knowledge of the true vaccine dose-response curve. In doing so, I simulated vaccine clinical trials, and therefore had to decide the trial designs that these simulated trials would follow. In particular, I hypothesised that a weighted model averaging technique could limit the potential for model misspecification to cause suboptimal dose selection. Additionally, I hypothesised that a trial design with a 'uniform' method of trial dose selection with many dosing groups would be a reasonable trial design to support modelling. I also used this as an opportunity to investigate the use of modelling to inform adaptive trial designs, which is considered beneficial in drug dose-finding studies. As part of this investigation, I also considered that trial size may influence

the effectiveness of mathematical modelling methods of vaccine dose selection. The aspects of modelling for optimal vaccine dose selection that were investigated in this work were therefore:

1. The assumed model of vaccine dose-efficacy (peaking, saturating, or a weighted average)
2. The trial design, consisting of:
    a. Trial size (10-100 trial participants)
    b. Method of trial dose selection (uniform with retrospective modelling, or one of three example adaptive design/continual modelling methods)

Whilst the above topics form the backbone of this paper's place in this thesis and within IS/ID modelling, there were three other topics that were included in this work that I believed warranted investigation. Firstly, whilst the CRM style of dose-response modelling is considered as an effective method of conducting dose-finding trials in drugs, I hypothesised that only investigating the dose that is predicted optimal during trial dose selection could lead to suboptimal final dose selection due to the exploration/exploitation trade-off. I believed this due to previous findings in the separate field multi-armed bandit problems and model-based reinforcement learning. Thall [193] also suggested that 'greedy' trial dose selection may be suboptimal but did not provide a demonstration of this phenomena.

Secondly, I hypothesised that the error in predictions of utility at a model predicted optimal dose could be assumed to be symmetrically distributed around the true utility value of utility at that dose if the assumed efficacy and toxicity models were appropriate. I primarily conducted the analysis of inaccuracy to investigate whether the variance of this error term was altered by trial size/continual modelling. The results of this paper contradict this hypothesis, which may have ramifications for future IS/ID modelling.

Thirdly, I modelled toxicity under an ordinal grading system. In paper 3, I suggest that this was possible with sufficient data, and therefore believed that it would be appropriate to highlight such a model.

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT

T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed <u>for each</u> research paper included within a thesis.

SECTION A – Student Details

Student ID Number	lsh1804914	**Title**	Mr
First Name(s)	John Helier		
Surname/Family Name	Benest		
Thesis Title	Mathematical Modelling for Optimal Vaccine Dose Finding: Maximising Efficacy and Minimising Toxicity		
Primary Supervisor	Richard G. White		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?	MDPI Vaccines		
When was the work published?	11 May 2022		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	N.A		
Have you retained the copyright for the work?*	Yes	Was the work subject to academic peer review?	Yes

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	
Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	The conceptualisation of this work was my own. I also was the sole individual developing codebase that was used for the dose-response modelling and simulation of clinical trials. All analysis of the data that was generated by these simulations were also conducted by me. I wrote the first draft of this work and performed re-drafting in line with co-author comments. All authors reviewed the paper. The interpretation of the results was my own work

SECTION E

Student Signature	███████████████
Date	11 August 2022

Supervisor Signature	███████████████
Date	12 August 2022

162

Paper 4 Title: Mathematical Modelling for Optimal Vaccine Dose Finding: Maximising Efficacy and Minimising Toxicity

Authors: John Benest, Sophie Rhodes, Thomas G. Evans, Richard G. White

Permission from copyright owner to include this work:

Article

Mathematical Modelling for Optimal Vaccine Dose Finding: Maximising Efficacy and Minimising Toxicity

John Benest [1,*] **, Sophie Rhodes** [1] **, Thomas G. Evans** [2] **and Richard G. White** [1]

1   Department of Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine,
    Keppel Street, London WC1E 7HT, UK; sophie.rhodes@lshtm.ac.uk (S.R.); richard.white@lshtm.ac.uk (R.G.W.)
2   Vaccitech Ltd., The Schrodinger Building, Heatley Road, The Oxford Science Park, Oxford OX4 4GE, UK;
    tom.evans@vaccitech.co.uk
*   Correspondence: john.benest@lshtm.ac.uk

Abstract: Vaccination is a key tool to reduce global disease burden. Vaccine dose can affect vaccine efficacy and toxicity. Given the expense of developing vaccines, optimising vaccine dose is essential. Mathematical modelling has been suggested as an approach for optimising vaccine dose by quantitatively establishing the relationships between dose and efficacy/toxicity. In this work, we performed simulation studies to assess the performance of modelling approaches in determining optimal dose. We found that the ability of modelling approaches to determine optimal dose improved with trial size, particularly for studies with at least 30 trial participants, and that, generally, using a peaking or a weighted model-averaging-based dose–efficacy relationship was most effective in finding optimal dose. Most methods of trial dose selection were similarly effective for the purpose of determining optimal dose; however, including modelling to adapt doses during a trial may lead to more trial participants receiving a more optimal dose. Clinical trial dosing around the predicted optimal dose, rather than only at the predicted optimal dose, may improve final dose selection. This work suggests modelling can be used effectively for vaccine dose finding, prompting potential practical applications of these methods in accelerating effective vaccine development and saving lives.

Keywords: dosing; dose response; modelling; clinical trials; adaptive design; continual modelling

1. Introduction

Vaccination is a key tool in global disease burden reduction and disease prevention. However, developing a vaccine for clinical use is an expensive and time-consuming process. As the magnitude of vaccine dose amount (hereafter 'dose') can affect the efficacy, toxicity, and cost of administering the vaccine, finding optimal vaccine dose is important. It is important to ensure that the chosen dose best balances maximal efficacy and minimal toxicity [1]. Preclinical and early phase 1/2 dose-finding trials aim to achieve this, typically through direct comparison of the efficacy and toxicity profiles of a small number of doses [2]. However, if none of these small number of doses is the optimal dose, then the vaccine will proceed to further study or clinical use with a suboptimal dose. This could reduce the potential for disease burden reduction, either due to reduced vaccine efficacy or decreased vaccine uptake arising from increased risk of vaccine-related adverse events. Hence, choosing from only a small number of doses may cost lives and be wasteful, given the expense of vaccine development. However, generating data on a larger number of doses requires larger and more expensive trials.

Mathematical-modelling-based approaches for vaccine dose optimisation have been explored previously and represent a solution for identifying optimal dose amongst a large number of possible doses without greatly increasing the size of trials [3–6]. Under these approaches, rather than comparing the efficacy and toxicity data directly between dosing groups, the data are used to inform models that attempt to describe the dose–efficacy and dose–toxicity relationships. These models are then combined and used to inform

vaccine dose decision making, similar to the idea of 'model-based drug development', which is prevalent in selecting optimal drug dose [7]. These 'models' are equations or systems of equations that are used to describe the relationship between vaccine dose and vaccine response. These models can either be mechanistic, leveraging knowledge of immunodynamics to describe an approximation of the dose-dependent immune system dynamics [3,4], or statistical, using a simpler set of assumptions about the general nature of the relationship between dose and efficacy/toxicity [8–11]. Although both approaches have been used to explore vaccine dose optimisation, neither approach has been fully validated, accepted, and used. We will be discussing statistical models of dose efficacy and dose toxicity throughout the remainder of this work.

Given that modelling is not consistently used in finding optimal vaccine dose, there are a number of questions that arise concerning its implementation. Firstly, which types of mathematical models would be most useful for determining optimal dose? Secondly, how many individuals in the trial population are required for modelling to generate reliable evidence? Thirdly, how should the trial population be dosed to improve the model's ability to determine optimal dose? This final question include is whether modelling should be used only retrospectively (as has been done in the past [3,4,11]) or whether it should be used continually at interim timepoints to guide dose selection throughout the trial (in the style of adaptive design or continual reassessment modelling [12–14]) in combination with retrospective modelling. Continual modelling/dose recommendation approaches have previously been suggested to be a more ethical approach to conducting dose-ranging studies in drugs [14,15].

Although modelling has previously been applied in vaccine dose optimisation using real-world data, such data are often noisy, and true underlying dose–efficacy and dose–toxicity relationships are unknown. This means that whether the doses that have been selected by these dose-optimisation approaches are truly optimal is unknown. By simulating clinical trial data, where the underlying dose–efficacy and dose–toxicity relationships are known, a 'simulation study' [15–17] allows for analysis of these dose-optimsation approaches not hampered by noisy data [17,18].

In this work, we aimed to use simulation of dose-finding clinical trials to assess the capability of statistical mathematical models to determine optimal dose. To answer the questions posed above, we investigated modelling-based dose-optimisation approaches, which were defined by:

i.   Assumed statistical efficacy model.
ii.   Trial size.
iii. Method of trial dose selection.

In order to perform this analysis, we used a number of qualitatively different 'scenarios', each representing a different 'true' vaccine dose–efficacy and dose–toxicity relationship. We considered three metrics of the quality of a dose-optimsation approach: not only the quality of the final selected dose but also the accuracy of predictions and benefit to trial participants.

Specifically, our objectives were to investigate, through simulation studies over many qualitatively different scenarios:

1.   When the method of trial dose selection is fixed, how dose-optimisation approaches are affected by the assumed statistical efficacy model and trial size.
2.   When trial size is fixed, how dose-optimisation approaches are affected by the assumed statistical efficacy model and method of trial dose selection.

2. Materials and Methods

Here, we summarise the simulation study used in this approach, then detail the component parts.

2.1. Overview of Simulation Study Methodology

We would like to evaluate how (i) assumed statistical efficacy model, (ii) trial size, and (iii) method of trial dose selection affect how well optimal dose is determined and use a simulation study to do this (Figure 1). Optimal dose was defined as a function that aims to maximise efficacy and minimise toxicity (3.2). We propose 'dose-optimsation approaches', which vary in (i–iii) (3.4). These varying dose-optimsation approaches are tested by simulating clinical trials. These clinical trials are simulated using a number of 'scenarios' representing theoretical vaccines that could be optimised (3.3). Each clinical trial was a pairing of a dose-optimsation approach and a scenario and therefore represents how well that dose-optimisation approach could optimise the vaccine represented in that scenario.
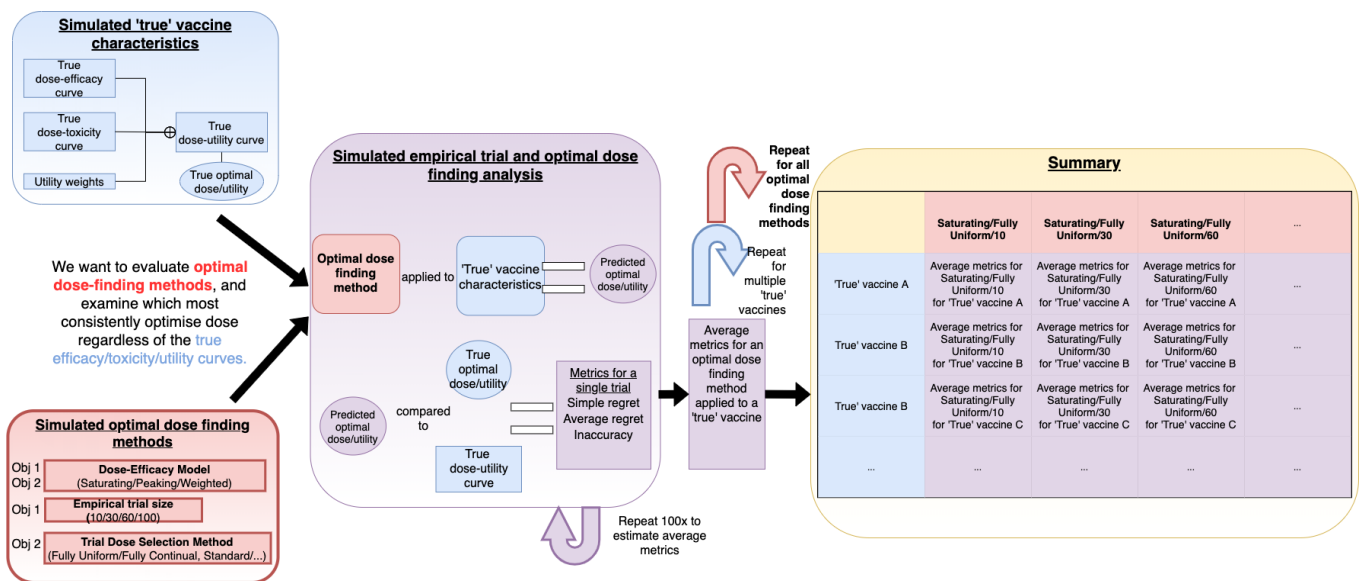
Figure 1. Visual depiction of the process of conducting simulation studies used in this work to assess mathematical-modelling-based dose-optimisation approaches. The aim was to evaluate dose-optimisation approaches (red), in particular the effect of changing the assumed dose–efficacy model, trial size, and trial dose-selection method. These were tested by simulating clinical trials (purple) based on 'scenarios' (blue). Repeated simulation of clinical trials was conducted for different dose-optimisation approach/scenario pairs, and metrics related to how effectively optimal dose was located were calculated. These were tabulated and compared to assess whether the assumed dose–efficacy model, trial size, and trial dose-selection method influence the consistency of dose optimsation.

The fact that the 'true' dose–efficacy and dose–toxicity curves are known in these scenarios allows these dose-optimisation approaches to be assessed. By repeatedly simulating different dose-optimisation approaches/scenarios we can evaluate the effect of varying (i–iii). Using different scenarios reduces the probability that we would recommend a dose-optimisation approach that does not optimise dose well in general, despite optimising dose well in simulations.

2.2. Efficacy, Toxicity, and Utility

We introduce the concept of dose-utility as a function of dose efficacy and dose toxicity and define the mathematical models that we used to describe these relationships.

2.2.1. Dose Efficacy

Vaccine efficacy or protection can be defined by many clinical endpoints, for example, reduced risk of infection, reduced risk of symptoms, reduced risk of severe symptoms, or reduced risk of hospitalisation [19]. Without the use of challenge studies or larger phase 3 studies determining relative reduction in disease, the probability of protection can be

difficult to determine [20]. Instead, immunological data are typically used in early trials as an anticipated surrogate of protection [21].

Although immunological data may be continuous in nature, a predictive model between immunological readout and probability of efficacy is often unknown [22]. Hence, it is common to define a threshold and consider individuals with immune response in excess of that threshold to have experienced an efficacious response [23]. Therefore, the actual desired endpoint (e.g., protection/survival) is likely binary in nature, and surrogates are often also binary. For simplicity and to aid in general usability, we therefore assumed that for dose-ranging studies, there would exist a binary efficacy outcome that can be measured and that the probability of this binary efficacy outcome is aimed to be maximised.

Even under these assumptions, there was a further challenge in modelling dose-efficacy. Whereas for many drugs, we can assume that an increased dose increases efficacy, for vaccines, this may not be the case. It is possible that there exists some dose for which the probability of efficacy is maximised and that increasing this dose decreases the probability of efficacious response [3,24–26]. Below, we define approaches for modelling vaccine efficacy. We chose a sigmoidal curve to represent the former "saturating" dose–response curve shape (Figure 2a) and a latent quadratic function to represent the latter "peaking" dose–response curve shape (Figure 2b). These equations are presented below and have previously been suggested in the literature [27,28]:

$$\text{Saturating}(\text{Dose}) = \frac{\text{maximum}}{1 + e^{(\text{gradient}(\text{midpoint}-\text{Dose}))}} \tag{1}$$

$$\text{Peaking}(\text{Dose}) = \frac{1}{1 + e^{(\text{base}+\text{gradient1}\times\ \text{Dose}+\text{gradient2}\times\text{Dose}^2)}} \tag{2}$$
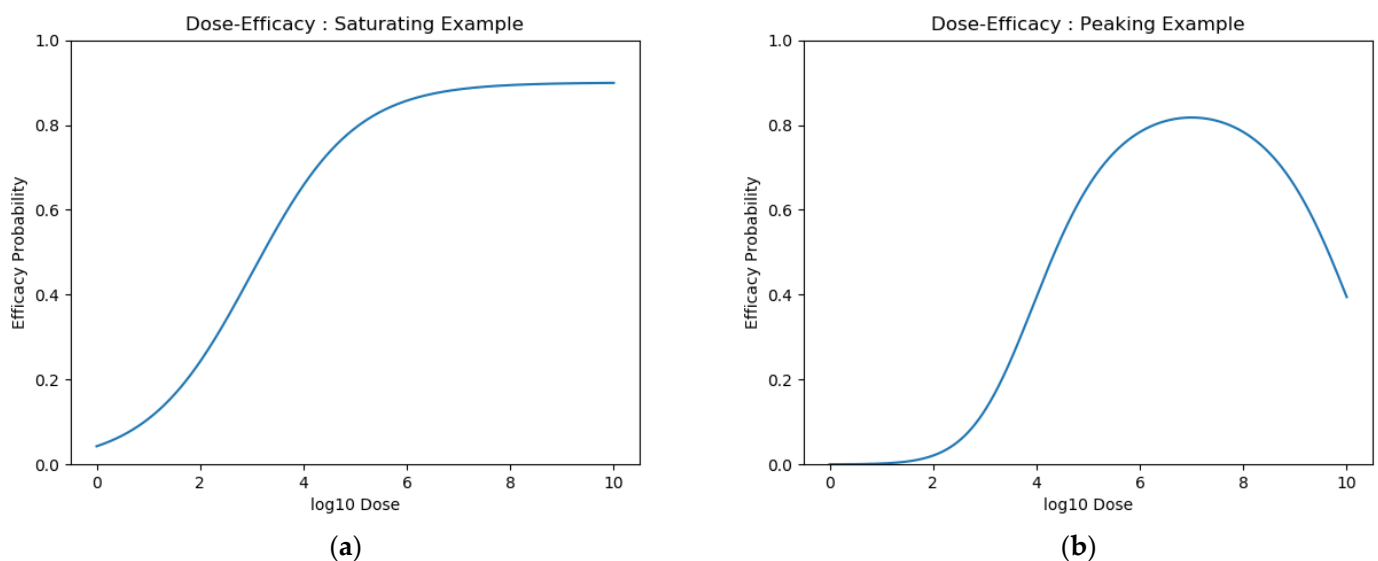
Figure 2. Example curves for (**a**) saturating and (**b**) peaking dose efficacy.

Further details of these models can be found in Supplementary S1.

In the case of uncertainty in the true dose–efficacy shape, a model averaging technique could also be considered [29]. Here, the saturating model and peaking models make predictions and are then combined based on how well each model describes the data. The mathematics behind this are discussed in Supplementary S2 and [22] and a visual depiction is presented in (Figure 3).
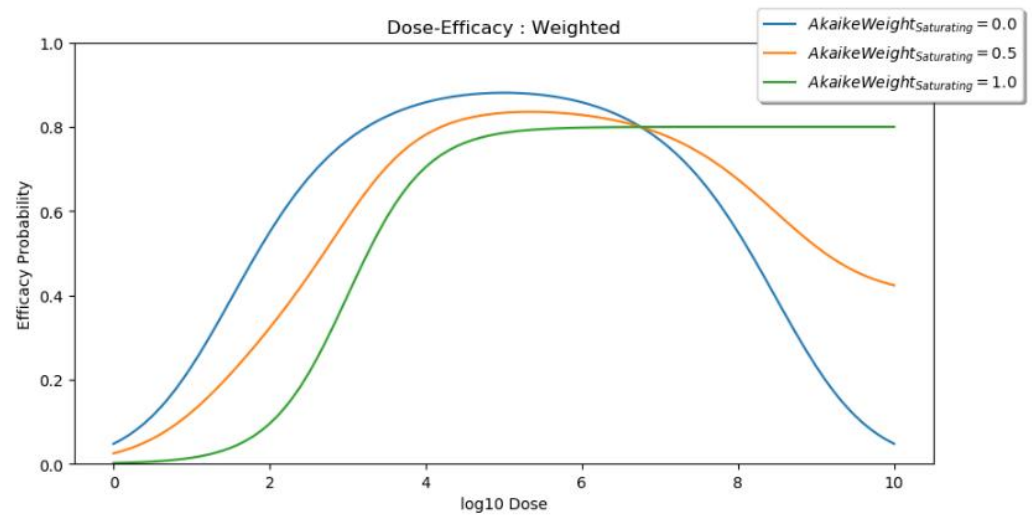
Figure 3. Visual example of model averaging. When the saturating Akaike weight is 0, the predicted efficacy curve is defined entirely by the peaking model (blue). When the saturating Akaike weight is 1, the predicted efficacy curve is defined entirely by the saturating model (green). If both models are equally as likely, given the available data, then the saturating Akaike weight and the peaking Akaike weight are both 0.5, and the predicted efficacy curve is the midpoint of the saturating and peaking curves (orange).

2.2.2. Dose Toxicity

As vaccine adverse events are typically less severe than adverse events for drugs and vaccines are preventive rather than therapeutic, we decided that only modelling higher-grade adverse events was unrealistic. It is likely that in vaccine dose-optimisation, minimising lower-grade adverse events may be preferable and relevant to vaccine uptake. Hence, we modelled vaccine dose toxicity using the ordinal dose-toxicity model [15]. Here, toxicity was described by four grades using a four-level toxicity grading system (Table 1).

Table 1. Description of the assumed grades of adverse events. These follow the gradings described in [30,31].

Adverse Reaction Grade	General Description
0	None.
1	Mild. Does not interfere with normal activity.
2	Moderate. Interference with normal activity. Little or no treatment required.
3	Severe. Prevents normal activity. Requires treatment.

We modelled the relationship between dose and the ordinal toxicities using the probit method described in [15] and discussed further in Supplementary S1. A visual description of an example ordinal model is given in Figure 4. Four parameters were needed to define this model. Three parameters defined the dose thresholds for which at least 50% of individuals experience greater than grade 0, grade 1, and grade 2 adverse events. The final parameter defined the steepness of these thresholds.
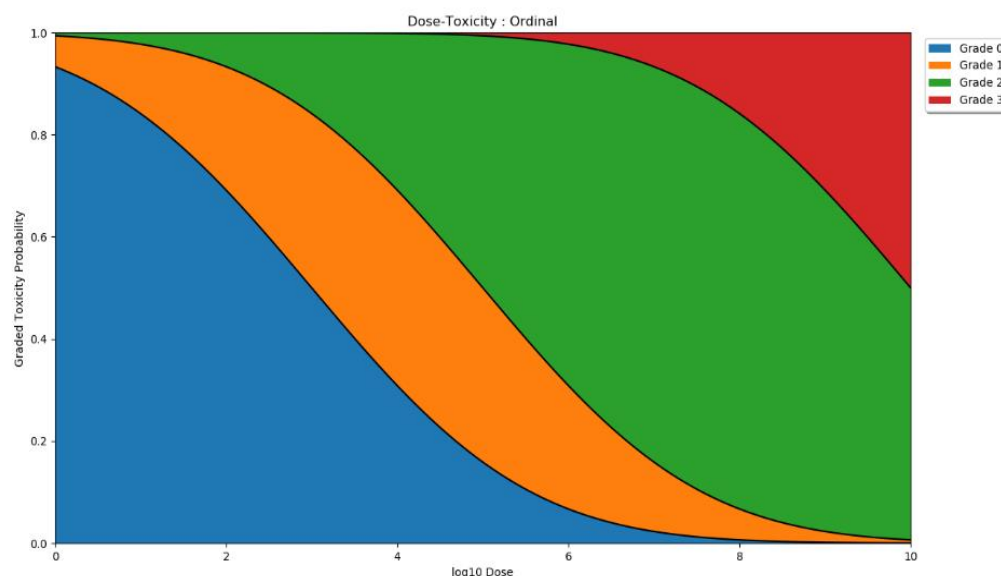
Figure 4. Visual example of ordinal dose toxicity. The plot shows the proportion of individuals that would experience different adverse event grades for each dose. In this example, at low doses, grade 0 (blue) adverse events are most likely. By dose 6, grade 1 (yellow) and grade 2 (green) adverse events are likely but grades 0 and 3 are also possible. By the maximum dose, approximately 50% of individuals would experience a grade 3 adverse event, and almost all others would experience grade 2 events.

2.2.3. Dose Utility

Optimising vaccine dose can be considered a multi-objective optimisation problem, in which we aim to maximise efficacy and minimise toxicity. To better define this problem, we made use of a utility function that attempts to balance maximising efficacy and minimising toxicity in a manner that should be clinically meaningful (Supplementary S3). Although many utility functions might be reasonable, to reduce complexity, a simple and interpretable dose-utility function was chosen [32].

For each dose, we assumed that there is some (predicted or true) probability of efficacy, P(Efficacy | Dose). Additionally, we assume that there are probabilities for each grade of toxicity, P(Toxicity = 0 | Dose), P(Toxicity = 1 | Dose), P(Toxicity = 2 | Dose), and P(Toxicity = 3 | Dose). We then defined utility weights, which were:

- $\text{Weight}_{\text{Efficacy}}$
- $\text{DisabilityWeight}_{\text{Toxicity0}}$
- $\text{DisabilityWeight}_{\text{Toxicity1}}$
- $\text{DisabilityWeight}_{\text{Toxicity2}}$
- $\text{DisabilityWeight}_{\text{Toxicity3}}$

These were measures of how beneficial an efficacious response was relative to the detrimental effect of the different adverse event grades. For example, if $\text{Weight}_{\text{Efficacy}} > \text{DisabilityWeight}_{\text{Toxicity2}}$, then the protection that may be gained from an efficacious vaccine response would outweigh the discomfort of the grade 2 event. Conversely, if $\text{Weight}_{\text{Efficacy}} < \text{DisabilityWeight}_{\text{Toxicity3}}$, then the protection that may be gained from an efficacious vaccine response would be outweighed by the discomfort of the grade 3 event. The disability weight for each grade was increasing (i.e., a grade 2 adverse event was worse than a grade 1 adverse event) (Table 2).

The dose-utility function is given by:

$$\text{Utility}(\text{Dose}) = \text{Weight}_{\text{Efficacy}} \times \text{P}(\text{Efficacy}|\text{Dose}) - \text{WeightedToxicity}(\text{Dose}) \tag{3}$$

$$\text{WeightedToxicity(Dose)} = \sum_{\text{Grade}=0}^{3} \text{P(Toxicity} = \text{Grade|Dose)} \times \text{DisabilityWeight}_{\text{Grade}} \quad (4)$$

A similar idea of vaccine risk/benefit is discussed in relation to the recent COVID-19 AstraZeneca vaccine [33]. $\text{Weight}_{\text{Efficacy}}$ would vary depending on the disease's severity, prevalence, and level of confidence in the surrogate of protection. Hence, in this work, we chose $\text{Weight}_{\text{Efficacy}}$ to be similar relative to $\text{DisabilityWeight}_{\text{Toxicity3}}$ (Table 2). This ensures that both maximising efficacy and minimising toxicity are important and prevents the optimal dose from being one that is optimal with regards to only one of these goals. Practically, $\text{Weight}_{\text{Efficacy}}$ could be chosen based on epidemiological models [34].

Table 2. Disability and efficacy weights for the utility functions.

Weight	Value	Source
$\text{Weight}_{\text{Efficacy}}$	0.133 or 0.266	Chosen to be equal to either $\text{DisabilityWeight}_{\text{Toxicity3}}$ or twice $\text{DisabilityWeight}_{\text{Toxicity3}}$
$\text{DisabilityWeight}_{\text{Toxicity0}}$	0.000	Chosen to be 0, as no discomfort/toxicity is caused
$\text{DisabilityWeight}_{\text{Toxicity1}}$	0.006	[35]
$\text{DisabilityWeight}_{\text{Toxicity2}}$	0.051	[35]
$\text{DisabilityWeight}_{\text{Toxicity3}}$	0.133	[35]

2.3. Scenarios

We considered it preferable to ensure that any dose-optimisation approaches that are used in clinical practice are 'consistent', which is to say that they optimise dose well for any vaccine they are applied to [36]. The opposite possibility would be for a dose-optimisation approach to be 'overly specific', which is to say that the approach would optimise dose very well for a small number of possible vaccines but would fail to choose a good dose for the majority of possible vaccines. To test whether these dose-optimisation approaches were 'consistent', we generated a number of qualitatively different 'scenarios' that dose-optimsation approaches could be tested on, similar to the study designs used in other dose-optimisation modelling studies [17,28].

Scenarios can be considered as simulated potential 'truths' for future vaccine dose/toxicity/response characteristics. Here, a 'scenario' was defined by a dose–efficacy curve, a dose–toxicity curve, and utility weights, i.e., the dose–utility curve resulting from these three scenarios (Figure 1, blue box). These scenarios were defined in order to be qualitatively different from each other, covering a broad range of potential dose/toxicity/response characteristics, not based on historical data.

We created and then tested our approaches on 14 such scenarios. For their true dose–efficacy curves, five scenarios used the sigmoid saturating curve, another five scenarios used the latent quadratic peaking curve, and the remaining four scenarios used curves that deviate from the parametric form of those two curves. Visualisations for three of the scenarios are shown in Figure 5, and further visualisation and parameterisation for all 14 scenarios can be found in Supplementary S4.

Scenario Name (Optimal Dose)	Efficacy Curve	Toxicity Curve	Utility Curve
Saturating 1 (8.0)			
Peaking 1 (6.8)			
Other 4 (4.8)			
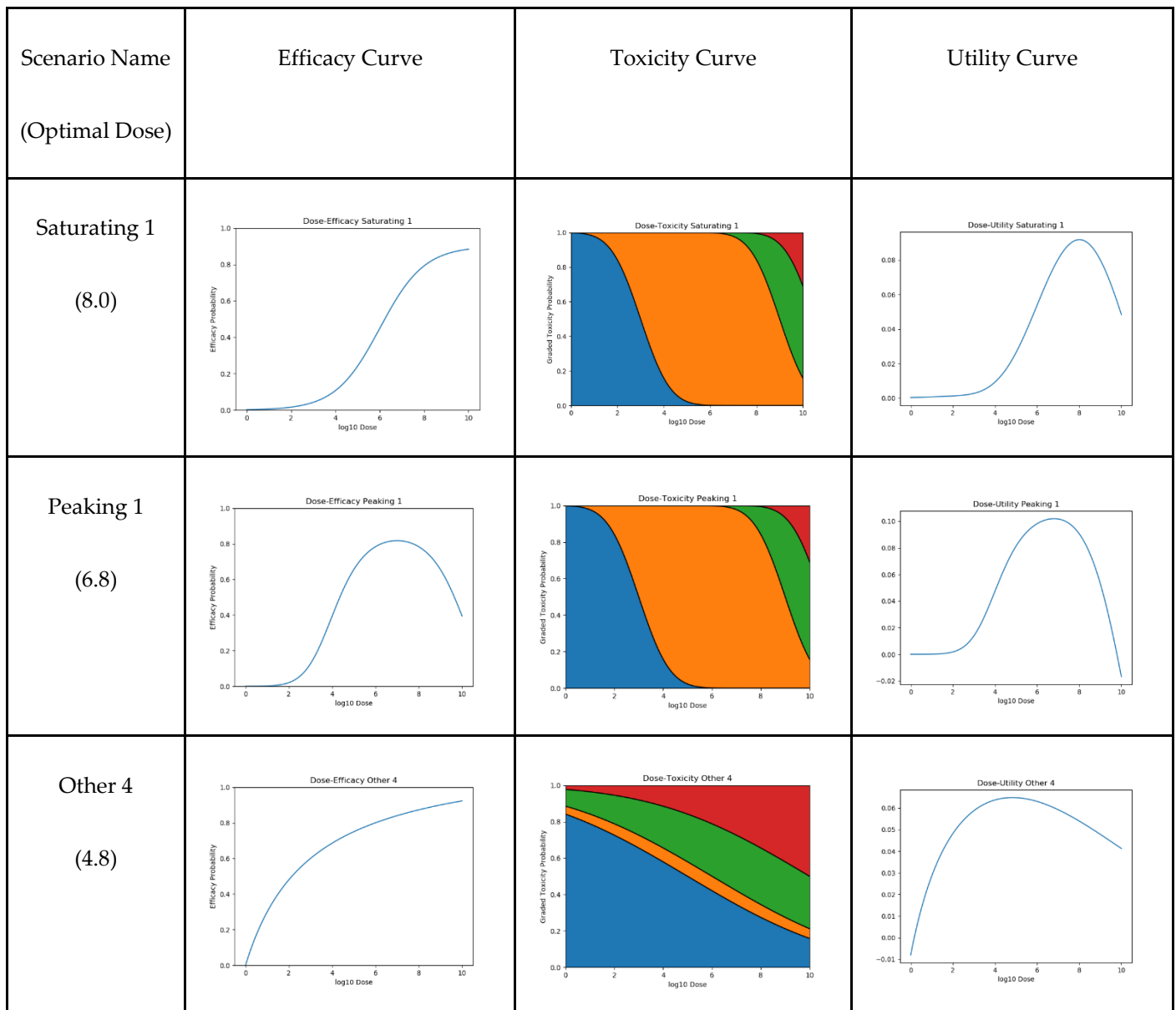
Figure 5. Three examples of the 14 tested scenarios. For each scenario, we show dose efficacy, dose toxicity, and the resultant dose–utility plots. Optimal dose is also given. For the toxicity plots, grade 0, 1, 2, and 3 adverse event probabilities are represented by blue, orange, green, and red, respectively.

2.4. Dose-Optimisation Approaches

A dose-optimisation approach can be considered as the combined approach by which a vaccine dose-finding study is conducted, data are gathered, and an 'optimal' dose is chosen based on these data. Although there are many possible considerations for doing so, we only considered a subsection of modelling-based dose-optimisation approaches. Therefore (Figure 1, red boxes), for the purposes of this work, a dose-optimisation approach was defined as a combination of:

i.   An assumed efficacy model (saturating, peaking, or weighted);
ii.   A trial size (10/30/60/100);
iii.   A method of trial dose selection (with either retrospective or continual modelling).

Objective 1 focuses on i and ii, and objective 2 focuses on i and iii.

2.5. Additional Details

Throughout this work, we considered dose on a $\log_{10}$ scale, although we did not otherwise assume units. For viral vector vaccines, these units would likely be viral particles or infectious units. Additionally, we consistently used a dose range of 0–10 on the $\log_{10}$ scale. This was purely for convenience and could be rescaled to the minimum and maximum possible dose for any given vaccine. This is referred to as the 'dosing space'.

For all models, parameter estimation was conducted by minimising negative log likelihood. This was done using the simplex method of Nelder and Mead [37] with the SciPy optimisation package in python [38]. Bounds were placed on parameters to ensure biological plausibility, see Supplementary S1.

2.6. Objective 1: When the Method of Trial Dose Selection Is Fixed, How Dose-Optimisation Approaches Are Affected by the Assumed Statistical Efficacy Model and Trial Size

We first assessed the use of dose-optimisation approaches using the three models of dose efficacy discussed above (saturating, peaking, and weighted) with regards to retrospective modelling with various trial sizes. Using the definition of a dose-optimisation approach outlined above, we assessed the following approaches:

i. Efficacy model: saturating, peaking, or weighted;
ii. Trial dose-selection method: full uniform exploration;
iii. Trial size: 10, 30, 60, or 100.

The method of dose selection for this objective was 'full uniform exploration'. This method distributes trial participants uniformly over the dosing space. For example, if there were only 6 available trial participants over the [0–10] $\log_{10}$-scale dosing space, we would have assigned test doses at 0, 2, 4, 6, 8, and 10. This method of dose selection is reasonable as a naive method, as it would ensure that all areas of the dosing space were evenly explored. As these data would then be a representative sample of all possible doses, this should have allowed for good model calibration and hence a good suggestion of optimal dose.

We assessed 4 different trial sizes explored in this objective. These were 10, 30, 60, and 100 individuals, representing reasonable sizes for vaccine phase I and II trials [39–41]. Hence, there were 12 (=4 × 3) dose-optimisation approaches, reflecting a combination of the 4 trials sizes and 3 assumed efficacy models. Each scenario/approach pairing was simulated 100 times for a total of 16800 (=12 × 14 × 100) simulated trials and 840,000 simulated individuals.

2.6.1. Metrics for Comparison between Approaches

We compared dose-optimisation approaches by calculating 'simple regret', 'percentage simple regret', 'inaccuracy', 'absolute inaccuracy', 'average regret', and 'percentage average regret' for each simulation (Figure 6).

Simple Regret

Simple regret in this setting was defined by the true utility score of the predicted optimal dose compared to the true optimal utility for the given vaccine. Ideally, this should be minimised. This is shown in Figure 6a and given by the following formula:

$$\text{Simple Regret} = \text{Utility}_{\text{TrueOptimal}} - \text{Utility}_{\text{Chosen}} \tag{5}$$

As the maximum and minimum possible utilities varied between scenarios, we also used the percentage simple regret (PSR) metric to allow for meaningful comparison across combinations of scenarios. PSR is given by the following formula:

$$\text{PSR} = 100 \times \frac{\text{Utility}_{\text{TrueOptimal}} - \text{Utility}_{\text{Chosen}}}{\text{Utility}_{\text{TrueOptimal}} - \text{Utility}_{\text{TrueLeastOptimal}}} \tag{6}$$

where a PSR of 100 implies the least optimal dose was chosen, and a PSR of 0 implies that the optimal dose was chosen.
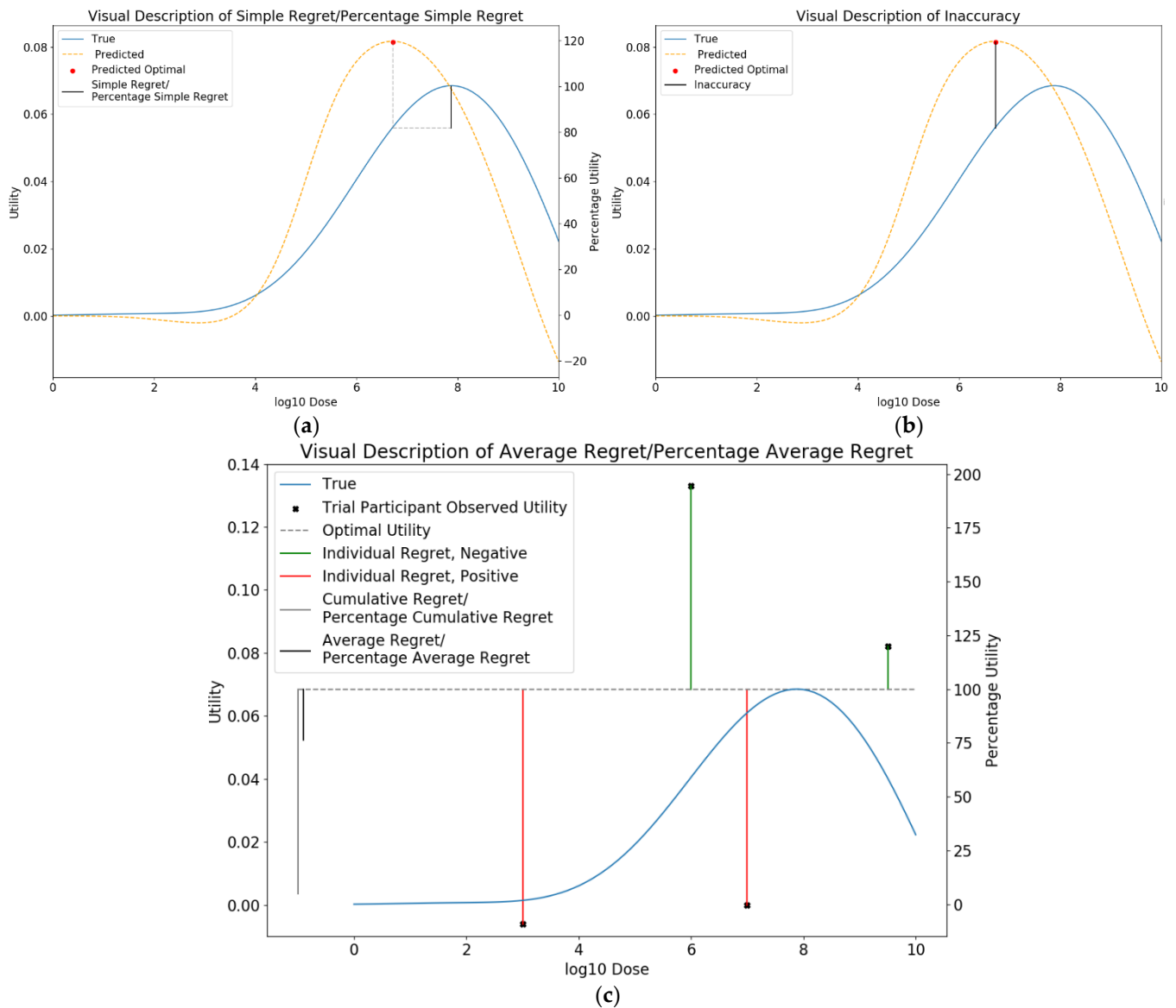
Figure 6. Visual description of (**a**) simple regret, (**b**), inaccuracy, and (**c**) average regret.

Inaccuracy

Inaccuracy in this setting was defined by the predicted utility score of the predicted optimal dose compared to the true utility at that dose. This is shown in Figure 6b and given by the following formula:

$$\text{Inaccuracy} = \text{PredictedUtility}_{\text{Chosen}} - \text{Utility}_{\text{Chosen}} \tag{7}$$

Ideally, this should be as close to zero as possible, which is equivalent to minimising the metric of absolute inaccuracy, which is given by:

$$\text{Absolute Inaccuracy} = \max(\text{Inaccuracy}, -\text{Inaccuracy}) \tag{8}$$

Average Regret

Each trial individual experiences a certain level of utility from receiving a vaccine. This utility can be subtracted from the true optimal utility to determine the 'regret' for that

individual. Average regret in this setting was defined by the utility that the average trial individual experienced relative to the true utility at the true optimal dose. Ideally, this should be minimised. This is shown in Figure 6c and given by the following formula:

$$\text{Average Regret} \ = \ \text{Cumulative Regret} \ / \ \text{n} \tag{9}$$

where n is the number of trial participants and

$$\text{Cumulative Regret} = \sum_{\text{individual}=1}^{\text{n}} \text{Regret}_{\text{Individual}} \tag{10}$$

where

$$\text{Regret}_{\text{Individual}} \ = \ \text{Utility}_{\text{Individual}} \ - \ \text{Utility}_{\text{True Optimal}} \tag{11}$$

We further defined percentage average regret to again enable comparison between scenarios.

$$\text{Percentage Average Regret} = \frac{\text{Average Regret}}{\text{Utility}_{\text{True Optimal}} \ - \text{Utility}_{\text{True Least Optimal}}} \tag{12}$$

2.7. Objective 2: When Trial Size Is Fixed, How Dose-Optimisation Approaches Are Affected by the Assumed Statistical Efficacy Model and Method of Trial Dose Selection

For this objective, we assessed different methods of trial dose selection in combination with the three efficacy models. In addition to the full uniform exploration described in objective 1, which was retrospective, we considered three continual-modelling-based methods of trial dose selection. Using the definition of a dose-optimisation approach outlined above, we investigated the following approaches:

i.   Efficacy model: saturating, peaking, or weighted;
ii. Trial size: 30;
iii. Trial dose-selection method: full uniform exploration, standard fully continual modelling, balanced exploration (softmax) fully continual modelling, or three-stage (softmax).

Hence, there were 12 (=4 $\times$ 3) dose-optimisation approaches, reflecting a combination of the 4 methods of trial dose selection and 3 assumed efficacy models. Each scenario/approach pairing was simulated 100 times for a total of 16800 (=12 $\times$ 14 $\times$ 100) simulated trials and 840,000 simulated individuals.

Although the 'full uniform exploration' trial design assessed in objective 1 seemed a reasonable design for improving model calibration, there are drawbacks to this design. Many individuals may be trialled with a suboptimal dose due to the uniform nature of the design. Modelling is also performed retrospectively; therefore, the generated data are not used to improve trial dosing. Hence, for this objective, we considered approaches that use continual-modelling-based methods of trial dose selection, which have been proposed to lead to more ethical trials [13]. These essentially repeat a cycle of:

1.   Conducting a small trial on a select set of doses;
2.   Gathering efficacy and toxicity data from this experiment;
3.   Updating the efficacy and toxicity models based on these data;
4.   Using the models to select either the next set of doses to test or to select the final dose to predict as 'optimal'.

2.7.1. Fully Continual Standard

The standard fully continual method is the simplest continual modelling dose-selection method. Each 'experiment' consists of one individual tested with the model-predicted optimal dose.

2.7.2. Fully Continual, Balanced Exploration (Softmax)

The standard fully continual dose-selection method above has previously been shown to be potentially useful in drug dose optimisation; however, analysis of optimisation

problems outside of dose finding have shown that testing only the predicted optimal may not be beneficial [42]. Being willing to 'explore' doses that are not predicted to be optimal may ultimately improve the final selected dose. As such, we considered softmax selection [43,44], where doses with higher predicted utilities were more likely to be selected; however, the selected trial doses were not always exactly at the predicted optimal. The degree of exploration was controlled by an exploration parameter, and further detail is given in Supplementary S5.

2.7.3. Three-Stage (Softmax)

Whereas the fully continual modelling process has been shown to be effective in drug dose optimisation, typically in that setting, the time between treatment and measurement of effect is short. In the vaccine setting, the time between vaccination and measurement of effect (immunological response) could be days, weeks, months, or even years. Hence, the application of a fully continual modelling process could take much longer than is feasible. We therefore considered a dose-selection method that contained elements of both the fully continual and fully retrospective modelling designs.

There are many ways this could be implemented. We considered a three-stage approach as follows:

1. Stage 1.

   a. $^1/_3$ of the trial population is dosed following the full uniform exploration approach outlined in objective 1.
   b. Efficacy and toxicity models are calibrated using these data and pseudo-data [3.7.5].

2. Stage 2.

   a. The second $^1/_3$ of the population is dosed according to the utility predictions of the combined efficacy and toxicity models, using the softmax selection method with relatively high exploration.
   b. Efficacy and toxicity models are calibrated using these data, data from step one, and downweighted pseudo-data.

3. Stage 3.

   a. The final $^1/_3$ of the population is dosed according to the utility predictions of the combined efficacy and toxicity models, using the softmax selection method with relatively low exploration.
   b. Efficacy and toxicity models are calibrated using all collected data, with pseudo-data being ignored. The predicted optimal dose is selected according to the utility predictions of the combined efficacy and toxicity models.

2.7.4. Dose-Escalation/De-Escalation Rules

We also included a simple escalation/de-escalation rule for the fully continual dose-selection methods, which is typically suggested for such continual modelling dose-selection methods. The first dose was always 5 on the $\log_{10}$ scale (that is to say the middle dose). A dose could not be in excess of $\frac{1}{2}$ a log above of the maximum previously tested dose or more than $\frac{1}{2}$ a log below the minimum previously tested dose. For example, dose 10 (10^{10}) could not be tested unless a dose of at least 9.5 ($10^{9.5}$) had been previously tested. This was suggested to reduce the risk of unexpected higher-grade toxicities.

As the first stage of the three-stage softmax approach included the smallest and largest allowed doses in the dosing space, the dose escalation/de-escalation rules would have no effect.

2.7.5. Pseudo-Data

Such continual modelling approaches can be implemented when insufficient data are available. Calibration with a small amount of data can be unstable; hence, pseudo-data were used to stabilise the calibration, as suggested in [15]. We used minimally informative

pseudo-data, which was quickly outweighed by real data and was ignored in the calibration step prior to final dose selection. Full details can be found in Supplementary S6.

2.7.6. Comparison between Approaches/Trial Designs

As in objective 1, percentage simple regret, inaccuracy, absolute inaccuracy, average regret, and percentage average regret were calculated. We used the Copeland method to identify a quantitative ranking of these approaches for their simple regret, absolute inaccuracy, and average regret outcomes [45,46], see Supplementary S7. Sum of ranks and mean of Copeland metrics across simple regret, absolute inaccuracy, and average regret were also obtained.

3. Results

3.1. Objective 1: When the Method of Trial Dose Selection Is Fixed, How Dose-Optimisation Approaches Are Affected by the Assumed Statistical Efficacy Model and Trial Size

A clear relationship between trial size and percentage simple regret (PSR) was observed (Figure 7), with a reduction in PSR as trial size increased, indicating that a more optimal dose was selected when trial size was larger. This was true regardless of whether a saturating, peaking, or weighted efficacy model was used and suggests an increased trial size improved final dose selection. However, the PSR aggregated across all scenarios was lower for the peaking and weighted approaches than for the approaches with saturating efficacy models (Figure 7), suggesting that using either a peaking or weighted model increased the average utility of the final selected dose. For almost all scenarios and trial sizes, it was better to assume a peaking curve than a saturating curve to minimise PSR, with a few exceptions, see Supplementary S8.
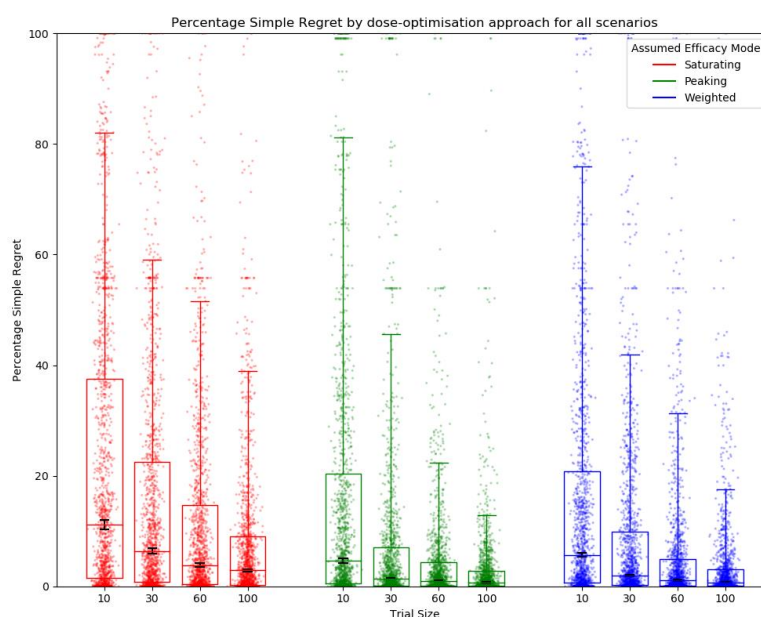
Figure 7. Percentage simple regret (PSR) for all scenarios by assumed efficacy model and trial size. Trial dose selection method was full uniform exploration. A lower PSR denotes a more optimal final dose. Individual points represent PSR for a single simulated clinical trial using one dose-optimisation approach for one of the 14 scenarios. The middle line of each boxplot is the median value; the box marks the 25th and 75th percentiles, and the whiskers mark the 5th and 95th percentiles of the data. Black lines represent the 95% confidence interval for the median of each distribution [47]. The majority of these distributions of PSR were different to a statistically significant extent at the $p = 0.05$ threshold according to the Kolmogorov–Smirnov test due to the large number of simulations conducted (100 per approach/scenario pairing). For further details on statistical significance see Supplementary S12.

Similarly, the accuracy of the predicted utility at the predicted optimal dose increased with increasing trial size (decreased inaccuracy) (Figure 8). An 'optimistic bias' was observed (positive inaccuracy) (Figure 8a), with predicted utility typically higher than the true utility for the given dose, as is often expected in optimisation problems [48,49] (Supplementary S9). There was no difference in inaccuracy between efficacy models for any trial size.

(**a**)   (**b**)
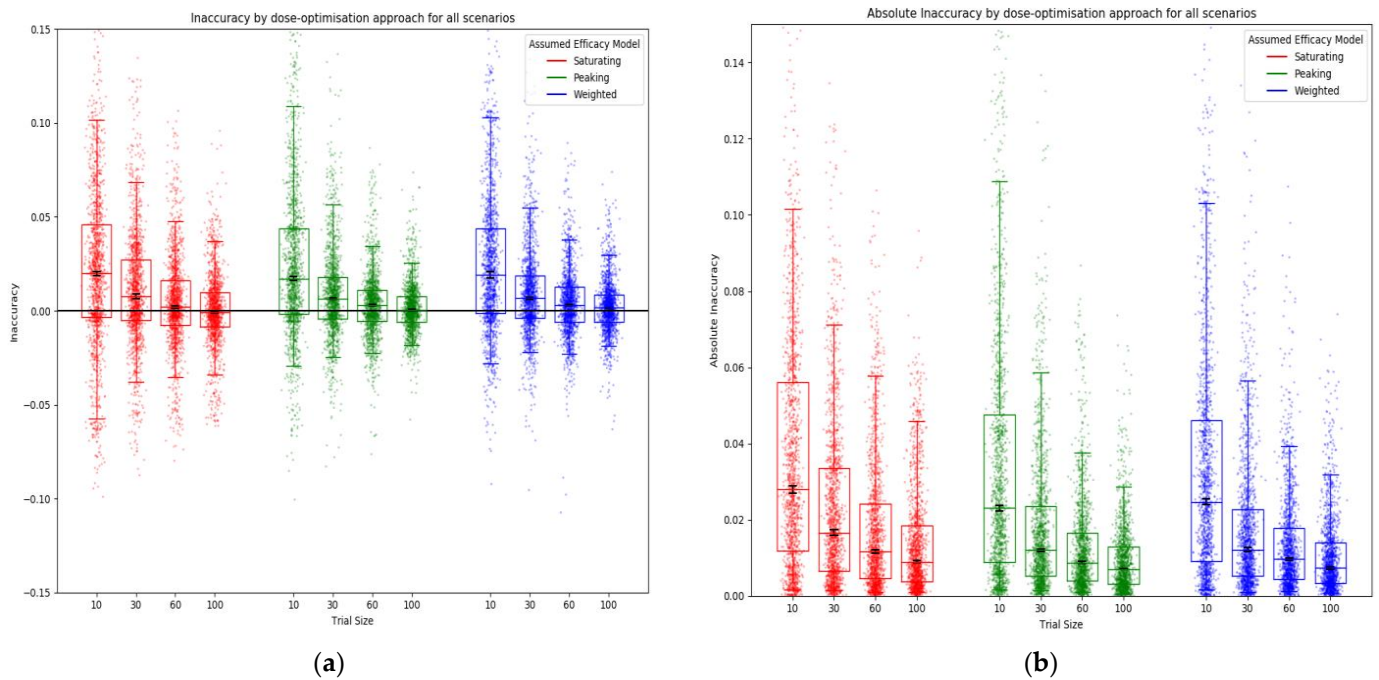
Figure 8. Inaccuracy (**a**) and absolute inaccuracy (**b**) for all scenarios by assumed efficacy model and trial size. Trial dose-selection method was full uniform exploration. The closer the inaccuracy/absolute accuracy was to 0, the more accurate the prediction of utility was at the predicted optimal dose. Individual points represent inaccuracy/absolute inaccuracy for a single simulated clinical trial using that dose-optimisation approach for one of the 14 scenarios. The middle line of each boxplot is the median value; the box marks the 25th and 75th percentiles, and the whiskers mark the 5th and 95th percentiles of the data. Black lines represent the 95% confidence interval for the median of each distribution [47]. The majority of these distributions of absolute inaccuracy were different to a statistically significant extent at the $p = 0.05$ threshold according to the Kolmogorov–Smirnov test due to the large number of simulations conducted (100 per approach/scenario pairing). For further details on statistical significance, see Supplementary S12.

There was no difference in median percentage average regret between efficacy model and trial size (Figure 9). This was to be expected, as all used the same method of trial dose selection of full uniform exploration with no continual modelling to allow later trial participants to benefit from early trial data.

All plots for PSR, inaccuracy, and percentage average regret for each scenario are shown in Supplementary S10 and S11. Analysis of the distributions of PSR, absolute inaccuracy, and PAR for statistical significance are given in the Supplementary S12.
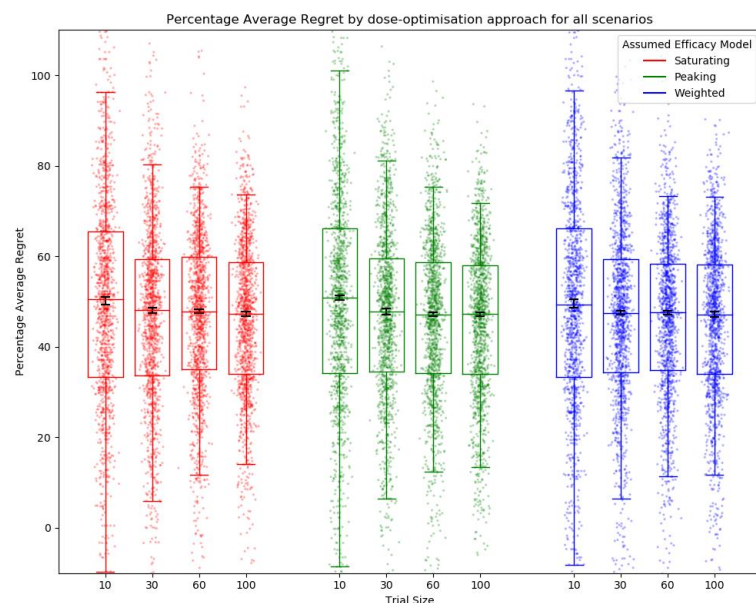
Figure 9. Percentage average regret for all scenarios by assumed efficacy model and trial size. Trial dose-selection method was full uniform exploration. Individual points represent percentage average regret for a single simulated clinical trial using that dose-optimisation approach for one of the 14 scenarios. The middle line of each boxplot is the median value; the box marks the 25th and 75th percentiles, and the whiskers mark the 5th and 95th percentiles of the data. Black lines represent the 95% confidence interval for the median of each distribution [47]. The majority of these distributions of PAR were not different to a statistically significant extent at the $p = 0.05$ threshold according to the Kolmogorov–Smirnov test. For further details on statistical significance, see Supplementary S12.

3.2. Objective 2: When Trial Size Is Fixed, How Dose-Optimisation Approaches Are Affected by the Assumed Statistical Efficacy Model and Method of Trial Dose Selection

3.2.1. Qualitative Analysis

With a trial of size 30, we found that using the peaking or weighted efficacy model still typically led to more optimal dose selection when compared to the saturating model (as shown by decreased PSR) (Figure 10). Neither the full uniform exploration modelling approaches nor the continual modelling approaches consistently showed a reduced PSR relative to one another. For some scenarios (saturating 5, Supplementary S13) we found that PSR was reduced by using continual modelling approaches. For others (peaking 1, peaking efficacy curve assumed, Supplementary S13), we found that the full uniform exploration approach appeared to best reduce PSR. This may suggest that the benefits of high levels of exploration or continual modelling for reducing PSR depend on the scenario. In general, the fully continual balanced exploration modelling approaches and the three-stage softmax approach appeared to lead to a slight reduction in PSR across the 14 scenarios relative to the standard fully continual modelling approach, suggesting that exploration may be important in consistent dose optimisation.

With a trial size of 30, there was minimal difference in inaccuracy and absolute inaccuracy across the approaches (Figure 11). This may suggest that the accuracy of utility predictions at the model predicted optimal vaccine dose was not dramatically improved by using a continual modelling method of dose selection. There was still an optimistic bias, although this was slightly reduced in the three-stage approaches relative to the standard fully continual, balanced exploration fully continual, and full uniform exploration approaches. Again, this was minimal relative to the differences that were observed when changing trial size in objective 1.
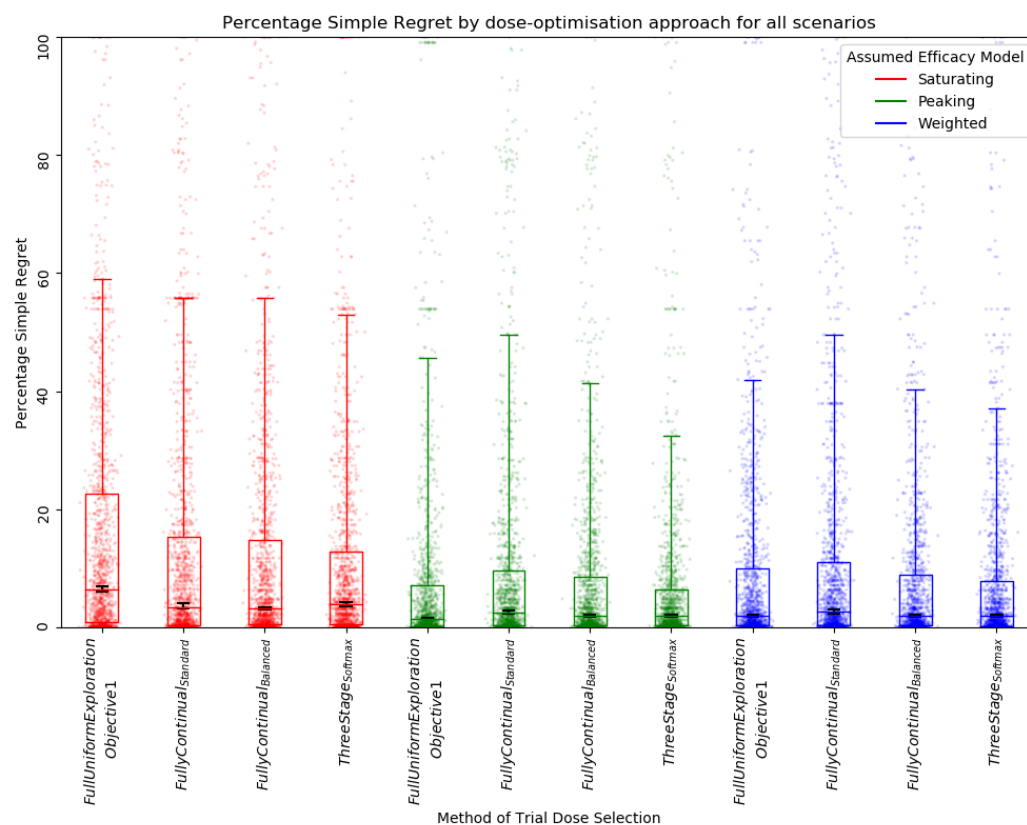
Figure 10. Percentage simple regret (PSR) for all scenarios by assumed efficacy model and trial dose-selection method. Trial size was 30. Individual points represent PSR for a single simulated clinical trial using that dose-optimisation approach for one of the 14 scenarios. The middle line of each boxplot is the median value; the box marks the 25th and 75th percentiles, and the whiskers mark the 5th and 95th percentiles of the data. A lower PSR denotes a more optimal final dose. Black lines represent the 95% confidence interval for the median of each distribution [47]. The distributions of PSR for the approaches that assumed a saturating model were different to the distributions of the approaches that assumed a peaking or weighted efficacy mode to a statistically significant extent at the *p* = 0.05 threshold according to the Kolmogorov–Smirnov test. For further details on statistical significance, see Supplementary S12.

With a trial size of 30, the results suggest that fully continual modelling (both standard and balanced) and three-stage approaches identify optimal dose with a greater net benefit to trial participants than the retrospective full uniform exploration approaches (as shown by decreased average regret) (Figure 12). The balanced exploration variant of the fully continual modelling dose-selection method appeared to have a marginally increased percentage average regret compared to approaches with standard fully continual modelling dose selection, but average regret was still significantly reduced relative to approaches using the three-stage softmax or full uniform exploration methods of trial dose selection. The three-stage softmax approaches showed a reduced average regret relative to full uniform exploration but a greater average regret relative to the fully continual approaches. These findings were the same regardless of the assumed efficacy model.

Similar plots for each individual scenario are shown in Supplementary S10 and S13. Analysis of the distributions of PSR, absolute inaccuracy, and PAR for statistical significance are given in the Supplementarys S12.

(**a**)   (**b**)
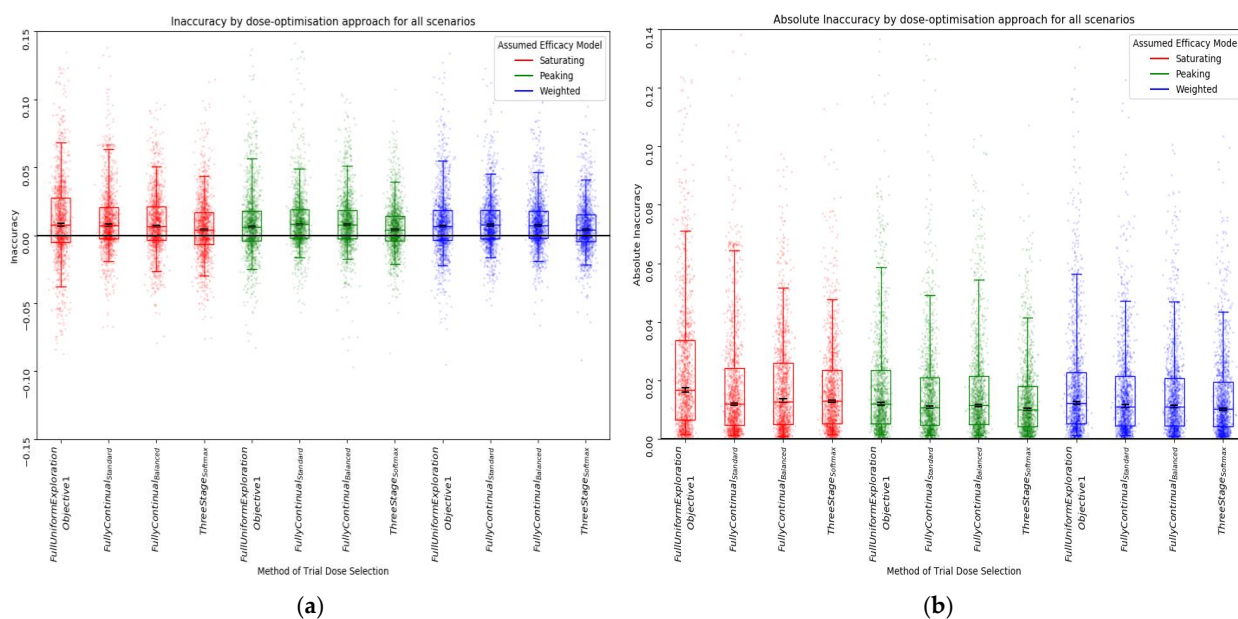
Figure 11. Inaccuracy (**a**) and absolute inaccuracy (**b**) for all scenarios by assumed efficacy model and trial dose-selection method. Trial size was 30. Individual points represent inaccuracy/absolute inaccuracy for a single simulated clinical trial using that dose-optimisation approach for one of the 14 scenarios. The middle line of each boxplot is the median value; the box marks the 25th and 75th percentiles, and the whiskers mark the 5th and 95th percentiles of the data. The closer inaccuracy/absolute accuracy is to 0, the more accurate the prediction of utility is at the predicted optimal dose. Black lines represent the 95% confidence interval for the median of each distribution [47]. The majority of these distributions of absolute inaccuracy were not different to a statistically significant extent at the $p = 0.05$ threshold according to the Kolmogorov–Smirnov test. For further details on statistical significance, see Supplementary S12.
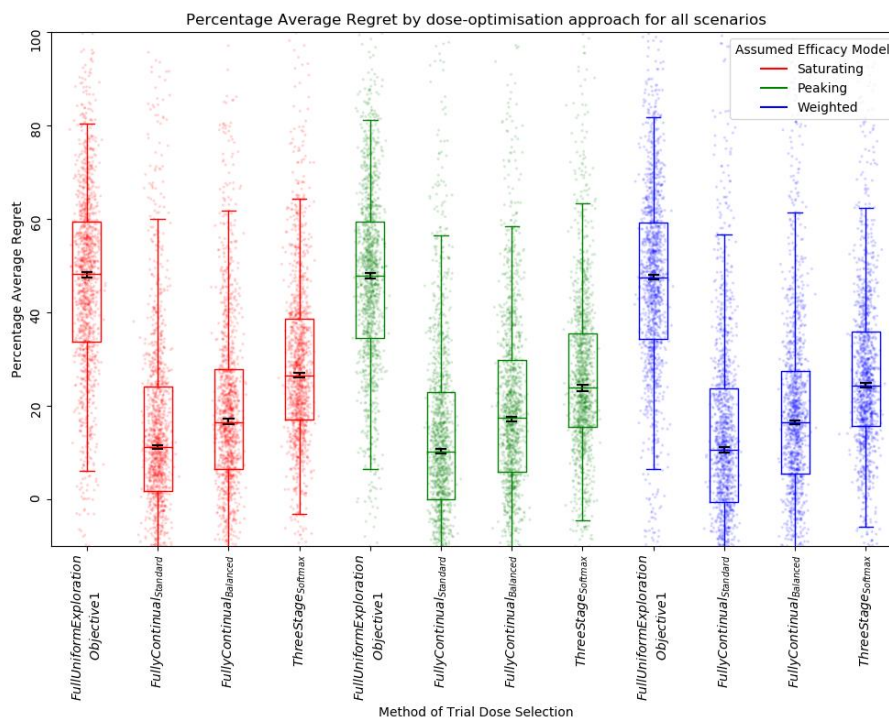
Figure 12. Percentage average regret for all scenarios by assumed efficacy model and trial dose-selection method. Trial size was 30. Individual points represent percentage average regret for a single simulated clinical trial using that dose-optimisation approach for one of the 14 scenarios. The middle

line of each boxplot is the median value; the box marks the 25th and 75th percentiles, and the whiskers mark the 5th and 95th percentiles of the data. A lower percentage average regret denotes better outcomes for trial participants. Black lines represent the 95% confidence interval for the median of each distribution [47]. The majority of these distributions of PAR were not different to a statistically significant extent at the $p = 0.05$ threshold according to the Kolmogorov–Smirnov test. For further details on statistical significance, see Supplementary S12.

3.2.2. Quantitative Ranking

For minimising PSR, the approach assuming a weighted efficacy curve and using fully continual modelling with balanced exploration was most consistent across the scenarios that we tested (Table 3). The fully continual modelling with balanced exploration approaches outranked the standard fully continual modelling approaches for each efficacy model. The three-stage softmax approaches also performed well, along with the approach with full uniform exploration with an assumed peaking efficacy curve. This may suggest that when assuming a peaking curve shape, exploration improves final dose selection.

For minimising average regret, the standard fully continual modelling approach assuming a peaking efficacy curve was most consistent across the scenarios that we tested. The shape of the model's efficacy curve was less important than the method of trial dose selection for minimising average regret, with the order from worst to best being full uniform exploration, three stage softmax, balanced fully continual modelling, and standard fully continual modelling. This may suggest that for small trial sizes (30), the standard fully continual modelling approach is most ethical, as the average regret was lowest. Therefore, the reduction in simple regret observed when including exploration may come at the cost of increased average regret for such small trial sizes.

For minimising absolute inaccuracy, the three-stage softmax approach assuming a peaking efficacy curve was most consistent across these scenarios, suggesting that dosing trial participants both near the predicted optimal and further away from the predicted optimal may reduce inaccuracy. The full uniform exploration approaches ranked lowest.

The dose-optimisation approach with an assumed weighted efficacy curve and fully continual modelling with balanced exploration had the best sum of ranks, which suggests that this approach should be chosen if simple regret, inaccuracy, and average regret are all equally valued. Copeland tables for each scenario are given in Supplementary S14.

Table 3. Copeland scores and rankings for all approaches with a trial size of 30 across all scenarios. Ordering is by aggregate rank. Aggregate rank was calculated as the sum of ranks for simple regret, absolute inaccuracy, and average regret. Aggregate score was the mean of scores for simple regret, inaccuracy, and average regret.

Approach	Aggregate of Simple Regret, Absolute Inaccuracy, and Average Regret		Simple Regret		Absolute Inaccuracy		Average Regret	
	Rank	Score	Rank	Score	Rank	Score	Rank	Score
Weighted, Fully Continual, Balanced	8	0.570	1	0.564	3	0.522	4	0.625
Peaking, Fully Continual, Standard	12	0.572	7	0.498	4	0.517	1	0.701
Peaking, Softmax Three Stage	12	0.536	4	0.552 [1]	1	0.556	7	0.500
Peaking, Fully Continual, Balanced	14	0.557	3	0.552 [1]	6	0.510	5	0.610
Weighted, Fully Continual, Standard	15	0.565	8	0.485	5	0.514	2	0.698
Weighted, Softmax Three Stage	15	0.528	5	0.541	2	0.549	8	0.493
Saturating, Fully Continual, Standard	20	0.543	10	0.447	7	0.492	3	0.691
Peaking, Full uniform exploration	24	0.414	2	0.563	10	0.480	12	0.201
Saturating, Fully Continual, Balanced	24	0.519	9	0.463	9	0.486	6	0.609
Saturating, Softmax Three Stage	28	0.465	11	0.442	8	0.489	9	0.465
Weighted, Full uniform exploration	28	0.400	6	0.516	11	0.480	11	0.203
Saturating, Full uniform exploration	34	0.330	12	0.378	12	0.406	10	0.205

[1] Scores are rounded to three decimal places, but ranks were calculated before rounding.

4. Discussion

In this work, we used simulation studies to evaluate mathematical-modelling-based approaches to optimising vaccine dose, maximising efficacy while minimising toxicity. We found that doses selected using these methods were improved with increased trial size, particularly for studies with at least 30 trial participants. Using a peaking model or a weighted model averaging approach for modelling dose efficacy was generally most effective for determining optimal dose. Identification of optimal dose was minimally affected by the method of trial dose selection. However, using modelling at interim timepoints to select trial doses led to trial participants receiving more optimal doses. Dosing only at the predicted optimal dose during a clinical trial may lead to less optimal dose selection relative to dosing around the predicted optimal. This work suggests modelling can be used effectively for vaccine dose finding, accelerating effective vaccine development and saving lives.

There were a number of strengths in our work. We included ordinal toxicity, which is highly relevant in vaccines due to their general safety profiles and potential for prophylactic use. Whereas we have previously seen vaccine dose optimisation applied using real-world data [3,24], simulation studies allow for an increased understanding of the potentials and pitfalls of dose-finding methodologies because the "truth" is known. By explicitly defining scenarios, we were able to accurately test metrics such as PSR and inaccuracy for these dose-optimisation approaches. Additionally, the scenarios we chose explored a wide range of curve shapes so that a multitude of potential real-life dosing scenarios could be reflected.

We chose to consider optimisation over a large number of potential doses, whereas previous dose-optimisation simulation studies have typically focussed on choosing between a small number of dosing levels [50]. Using a small number of doses may not be appropriate if none of the selected doses achieves optimal vaccine utility. Additionally, we chose to use 'simple regret'-based metrics rather than 'percentage best arm identification', which considers an approach to have been successful in optimising dose in a simulated clinical trial if and only if the true optimal dose was predicted to be optimal [51]. Using 'simple regret' as a metric is more appropriate for vaccine dose-optimisation, as multiple closely spaced arms may have similar utility. Selecting a dose with approximately equal clinical value to that of the true optimal dose would be considered to be an improvement to selecting an inferior dose under simple regret, but both would be considered as failures in terms of optimising dose under the percentage best arm identification metric.

Additionally, we discussed the concept of 'exploration', which is rarely considered in dose-optimisation work but is instrumental to the wider class of 'multi-armed bandit problems', which dose-optimisation can be considered to be part of [52]. The analysis of full uniform exploration and weighted modelling approaches was also novel in this setting. Finally, we also evaluated the concept of accuracy and inaccuracy in dose-optimisation modelling approaches, which is typically not well researched. This seems relevant, given the overestimation bias observed across all of these approaches to dose optimisation, which could lead to overestimations in the potential of vaccine utility and incorrectly guide policy.

There were weaknesses with both this work and the dose-optimisation approaches evaluated in general. It is likely that the 14 scenarios we chose may not represent all possible real-life dose–efficacy, dose–toxicity, and dose–utility curves. However, the 14 scenarios we used were qualitatively different and sufficient to cover plausible prior belief for any specific vaccine. When designing dose-finding trials for a future vaccine, it may be reasonable to consider only the findings for scenarios that are most similar to the clinician's prior beliefs about a given vaccine's likely dose–efficacy/dose–toxicity curves. Additionally, we performed only 100 clinical trial simulations for each approach/scenario pairing. This is in excess of the minimum of 10 that has been suggested [53], and we believe a larger number of simulated trials would not have impacted the results. Although not a weakness of the work, the observed overestimation bias appears to be a weakness of these optimisation approaches, as it decreases with increased trial size. We note that an overestimation bias is expected in both model-based [48,49] and traditional comparative dose-selection [54]

methodologies (Supplementary S9). Methods to remedy this have been suggested but were not addressed in this work [49,55].

Our work is consistent with previous modelling findings. These studies also showed that continual modelling approaches increased the average quality of clinical trial drug dosing (e.g., decreased average regret) [12]. It has previously been shown that for similar optimisation problems, exploration is key to maximising utility, but the level of exploration varies based on scenario and sample size [56]. This is consistent with our finding that for a sample size of 30, whether approaches without exploration could outperform explorative approaches depended on the scenario. Specifically, in previous work on optimisation of drug dose (using a small number of potential dosing levels, Bayesian methodologies, and only an assumed saturating efficacy curve), it was found that including exploration was beneficial for drug dose optimisation [57]. We also found, with regards to efficacy curves, that the peaking latent quadratic curve often outperformed the saturating sigmoid curve in cases where the true scenario curve was sigmoid saturating. Although we might expect that using the model that best describes true dose–response dynamics would be preferable for optimisation purposes, previous studies suggested that in some cases, models that do not well approximate the true dynamics can be preferable for optimisation purposes [58].

There were limitations to this work and the approaches discussed. We assumed that a binary measure for vaccine efficacy is known; however, for many diseases, a surrogate (binary or otherwise) of protection is not known. We also excluded many complicating factors that have been discussed in previous continual modelling literature. For example correlation in the probabilities of efficacy and toxicity [59], multiple toxicity subtypes (e.g., pain and nausea) [60], stopping rules [13], and placebo doses [61] were not considered. Given that modelling-based vaccine dose optimisation is a large topic that is still in the proof-of-concept stage, these were omitted to simplify the work. Additionally, we did not address cost and time requirements for trials. The time taken by continual modelling approaches in a vaccine clinical trial setting may not be justified by the resulting improvement in outcome for trial participants, which reflected by a decrease in average regret. We also did not compare these approaches to model-free approaches, such as 3 + 3 design [62,63], as such approaches are inherently designed to choose between a small number of doses and therefore have similar problems associated with selecting from a small number of doses that we discussed above.

We also did not include a fifth or sixth toxicity grade, which would typically represent a serious adverse event resulting in hospitalisation or death, respectively. As these events are likely to be rare in most vaccine trials [1] and would typically require the trial to be stopped, we excluded these gradings. Finally, we did not consider potential confounders (for example, age or sex) that may occur in practical dose-ranging trials. This would have further increased the complexity of this work.

There is much future work to be done on this topic. Additional scenarios should be tested to investigate any further shortcomings of these approaches. Only one simple utility function was considered in this work. This could be made more complex by including dose–cost relationships or modelling dose–response curves for multiple different efficacy or toxicity responses [11,60,64]. Creating a meaningful utility function is non-trivial [65,66] and is key to effective optimisation. Previous work has shown that both Bayesian methodologies and the frequentist methodologies discussed in this work perform similarly for some continual modelling approaches [67], but this should be further tested and validated. Optimising the degree of exploration that should occur could potentially decrease simple regret and average regret, but the optimal amount of exploration almost certainly depends on the scenario and trial size [56]. Additionally, using mechanistic models for dose efficacy could be beneficial if there is a good understanding of the immunodynamics relating to the vaccine [3,4,68]. However, this would likely introduce more complexity to the modelling process and to the utility function.

Although simulation demonstrates that these dose-optimisation approaches could be used when designing trials to optimise vaccine dose, these approaches clearly need

to be tested in a real-world setting to evaluate their practical implementation. Although the continual modelling approaches reduced average regret (improved trial participant outcomes) relative to the retrospective approaches, there is clearly a trade-off regarding whether this is worth the increased time requirements (which dramatically increase when using fully continual approaches) or additional complexity (particularly in approaches that use softmax selection). Hence, when using modelling for vaccine dose optimisation, there appears to be a balance between improved trial participant experience and the cost of increased time to clinical use. Whether this is ethically justified is a matter for further discussion. There may also be discussion of whether the potential for greater information efficiency from modelling may reduce trial size relative to standard dose-finding trial design and justify the increased time requirements. Furthermore, there should be consideration of how to approach confounding variables. Clinical trial design randomisation typically aims to ensure that populations in different dosing groups are homogenous [69]. The approaches discussed here assume that individuals are independent of each other; therefore, randomised sampling from a homogenous population should still be used to minimise risk of confounding variables (for example, avoiding correlation between dose and age of trial participants). Additionally, choosing optimal dose for prime-boost paradigm vaccines may require more complicated mathematical modelling methods, as efficacy and toxicity outcomes may be dependent on both prime and boost dose [70].

In drug development, mathematical modelling methodologies have led to improved drug efficacy and toxicity profiles, as well as a reduction in the cost of clinical trials. Despite the limitations and open questions discussed above, the application of mathematical modelling methodologies in vaccine pharmacological and biotechnology industries could allow for more quantitative and informed decision making.

5. Conclusions

Choosing the optimal vaccine dose is a complicated endeavour. Through this work, we evaluated model-based dose-optimisation approaches, along with trial design, to utilise these methodologies. Model-based dose-optimisation approaches may be effective for making vaccine dose decisions, which may increase efficacy and decrease toxicity, both during clinical trials and upon vaccine implementation. We hope that this work leads to future research and practical application of modelling methods in selecting vaccine doses. This may accelerate effective vaccine development and save lives.

References

1. Kaur, R.J.; Dutta, S.; Bhardwaj, P.; Charan, J.; Dhingra, S.; Mitra, P.; Singh, K.; Yadav, D.; Sharma, P.; Misra, S. Adverse Events Reported From COVID-19 Vaccine Trials: A Systematic Review. *Indian J. Clin. Biochem.* **2021**, *36*, 427–439. [CrossRef] [PubMed]
2. Afrough, S.; Rhodes, S.; Evans, T.; White, R.; Benest, J. Immunologic Dose-Response to Adenovirus-Vectored Vaccines in Animals and Humans: A Systematic Review of Dose-Response Studies of Replication Incompetent Adenoviral Vaccine Vectors When Given via an Intramuscular or Subcutaneous Route. *Vaccines* **2020**, *8*, 131. [CrossRef] [PubMed]
3. Handel, A.; Li, Y.; McKay, B.; Pawelek, K.A.; Zarnitsyna, V.; Antia, R. Exploring the Impact of Inoculum Dose on Host Immunity and Morbidity to Inform Model-Based Vaccine Design. *PLOS Comput. Biol.* **2018**, *14*, e1006505. [CrossRef] [PubMed]
4. Rhodes, S.J.; Knight, G.M.; Kirschner, D.E.; White, R.G.; Evans, T.G. Dose Finding for New Vaccines: The Role for Immunostimu-lation/Immunodynamic Modelling. *J. Theor. Biol.* **2019**, *465*, 51–55. [CrossRef] [PubMed]
5. Boissel, J.-P.; Pérol, D.; Décousus, H.; Klingmann, I.; Hommel, M. Using Numerical Modeling and Simulation to Assess the Ethical Burden in Clinical Trials and How It Relates to the Proportion of Responders in a Trial Sample. *PLoS ONE* **2021**, *16*, e0258093. [CrossRef]
6. Rigaux, C.; Sébastien, B. Evaluation of Non-Linear-Mixed-Effect Modeling to Reduce the Sample Sizes of Pediatric Trials in Type 2 Diabetes Mellitus. *J. Pharmacokinet. Pharmacodyn.* **2020**, *47*, 59–67. [CrossRef]
7. Kim, T.H.; Shin, S.; Shin, B.S. Model-Based Drug Development: Application of Modeling and Simulation in Drug Development. *J. Pharm. Investig.* **2018**, *48*, 431–441. [CrossRef]
8. Gillespie, W.R. Noncompartmental Versus Compartmental Modelling in Clinical Pharmacokinetics. *Clin. Pharmacokinet.* **1991**, *20*, 253–262. [CrossRef]
9. Gabrielsson, J.; Weiner, D. Non-Compartmental Analysis. In *Computational Toxicology: Volume I*; Reisfeld, B., Mayeno, A.N., Eds.; Methods in Molecular Biology; Humana Press: Totowa, NJ, USA, 2012; pp. 377–389. ISBN 978-1-62703-050-2.
10. Bonate, P.L. *Pharmacokinetic-Pharmacodynamic Modeling and Simulation*; Springer Science & Business Media: Berlin, Germany, 2011; ISBN 978-1-4419-9485-1.
11. Benest, J.; Rhodes, S.; Quaife, M.; Evans, T.G.; White, R.G. Optimising Vaccine Dose in Inoculation against SARS-CoV-2, a Multi-Factor Optimisation Modelling Study to Maximise Vaccine Safety and Efficacy. *Vaccines* **2021**, *9*, 78. [CrossRef]
12. O'Quigley, J.; Pepe, M.; Fisher, L. Continual Reassessment Method: A Practical Design for Phase 1 Clinical Trials in Cancer. *Biometrics* **1990**, *46*, 33–48. [CrossRef]
13. Pallmann, P.; Bedding, A.W.; Choodari-Oskooei, B.; Dimairo, M.; Flight, L.; Hampson, L.V.; Holmes, J.; Mander, A.P.; Odondi, L.; Sydes, M.R.; et al. Adaptive Designs in Clinical Trials: Why Use Them, and How to Run and Report Them. *BMC Med.* **2018**, *16*, 29. [CrossRef] [PubMed]
14. Wheeler, G.M.; Mander, A.P.; Bedding, A.; Brock, K.; Cornelius, V.; Grieve, A.P.; Jaki, T.; Love, S.B.; Odondi, L.; Weir, C.J.; et al. How to Design a Dose-Finding Study Using the Continual Reassessment Method. *BMC Med. Res. Methodol.* **2019**, *19*, 18. [CrossRef] [PubMed]
15. Van Meter, E.M.; Garrett-Mayer, E.; Bandyopadhyay, D. Dose-Finding Clinical Trial Design for Ordinal Toxicity Grades Using the Continuation Ratio Model: An Extension of the Continual Reassessment Method. *Clin. Trials Lond. Engl.* **2012**, *9*, 303–313. [CrossRef] [PubMed]
16. James, G.D.; Symeonides, S.; Marshall, J.; Young, J.; Clack, G. Assessment of Various Continual Reassessment Method Models for Dose-Escalation Phase 1 Oncology Clinical Trials: Using Real Clinical Data and Simulation Studies. *BMC Cancer* **2021**, *21*, 1–10. [CrossRef] [PubMed]
17. Morris, T.P.; White, I.R.; Crowther, M.J. Using Simulation Studies to Evaluate Statistical Methods. *Stat. Med.* **2019**, *38*, 2074–2102. [CrossRef]
18. Takahashi, A.; Suzuki, T. Bayesian Optimization Design for Dose-Finding Based on Toxicity and Efficacy Outcomes in Phase I/II Clinical Trials. *Pharm. Stat.* **2021**, *20*, 422–439. [CrossRef]
19. Andrews, N.; Tessier, E.; Stowe, J.; Gower, C.; Kirsebom, F.; Simmons, R.; Gallagher, E.; Thelwall, S.; Groves, N.; Dabrera, G.; et al. Duration of Protection against Mild and Severe Disease by Covid-19 Vaccines. *N. Engl. J. Med.* **2022**, *386*, 340–350. [CrossRef]
20. Nauta, J. Statistics in Clinical Vaccine Trials: Estimating The Protection Curve. In *Statistics in Clinical Vaccine Trials*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 108–109. ISBN 978-3-642-44191-2.
21. Ward, B.J.; Pillet, S.; Charland, N.; Trepanier, S.; Couillard, J.; Landry, N. The Establishment of Surrogates and Correlates of Protection: Useful Tools for the Licensure of Effective Influenza Vaccines? *Hum. Vaccines Immunother.* **2018**, *14*, 647–656. [CrossRef]
22. Nauta, J. Statistics in Clinical Vaccine Trials: Standard Statistical Methods for the Analysis of Immunogenicity Data. In *Statistics in Clinical Vaccine Trials*; Springer: Berlin/Heidelberg, Germany, 2011; p. 41. ISBN 978-3-642-44191-2.
23. Voysey, M.; Sadarangani, M.; Pollard, A.J.; Fanshawe, T.R. Computing Threshold Antibody Levels of Protection in Vaccine Clinical Trials: An Assessment of Methodological Bias. *PLoS ONE* **2018**, *13*, e0202517. [CrossRef]
24. Rhodes, S.J.; Zelmer, A.; Knight, G.M.; Prabowo, S.A.; Stockdale, L.; Evans, T.G.; Lindenstrøm, T.; White, R.G.; Fletcher, H. The TB Vaccine H56+IC31 Dose-Response Curve Is Peaked Not Saturating: Data Generation for New Mathematical Modelling Methods to Inform Vaccine Dose Decisions. *Vaccine* **2016**, *34*, 6285–6291. [CrossRef]
25. Wages, N.A.; Slingluff, C.L. Flexible Phase I–II Design for Partially Ordered Regimens with Application to Therapeutic Cancer Vaccines. *Stat. Biosci.* **2020**, *12*, 104–123. [CrossRef] [PubMed]

26. Benest, J.; Rhodes, S.; Afrough, S.; Evans, T.; White, R. Response Type and Host Species May Be Sufficient to Predict Dose-Response Curve Shape for Adenoviral Vector Vaccines. *Vaccines* **2020**, *8*, 155. [CrossRef] [PubMed]
27. O'Quigley, J.; Iasonos, A.; Bornkamp, B. Dose-Response Functions. In *Handbook of Methods for Designing and Monitoring Dose Finding Trials*; CRC Press: Boca Raton, FL, USA, 2019; p. 199. ISBN 978-0-367-33068-2.
28. Thall, P.F.; Cook, J.D. Dose-Finding Based on Efficacy–Toxicity Trade-Offs. *Biometrics* **2004**, *60*, 684–693. [CrossRef] [PubMed]
29. Symonds, M.R.E.; Moussalli, A. A Brief Guide to Model Selection, Multimodel Inference and Model Averaging in Behavioural Ecology Using Akaike's Information Criterion. *Behav. Ecol. Sociobiol.* **2011**, *65*, 13–21. [CrossRef]
30. Sibille, M.; Patat, A.; Caplain, H.; Donazzolo, Y. A Safety Grading Scale to Support Dose Escalation and Define Stopping Rules for Healthy Subject First-Entry-into-Man Studies. *Br. J. Clin. Pharmacol.* **2010**, *70*, 736–748. [CrossRef]
31. Food and Drug Administration. Guidance for Industry: Toxicity Grading Scale for Healthy Adult and Adolescent Volunteers Enrolled in Preventive Vaccine Clinical Trials. Available online: https://fda.gov/media/73679/download (accessed on 7 March 2022).
32. Talbi, E.-G. Aggregation Method. In *Metaheuristics: From Design to Implementation*; John Wiley & Sons: Hoboken, NJ, USA, 2009; pp. 324–326. ISBN 978-0-470-27858-1.
33. European Mediciens Agency. Visual Risk Contextualisation for Vaxzeria Art.5.3 Referral. Available online: https://www.ema.europa.eu/en/documents/chmp-annex/annex-vaxzevria-art53-visual-risk-contextualisation_en.pdf (accessed on 7 March 2022).
34. Moore, S.; Hill, E.M.; Tildesley, M.J.; Dyson, L.; Keeling, M.J. Vaccination and Non-Pharmaceutical Interventions for COVID-19: A Mathematical Modelling Study. *Lancet Infect. Dis.* **2021**, *21*, 793–802. [CrossRef]
35. Salomon, J.A.; Haagsma, J.A.; Davis, A.; de Noordhout, C.M.; Polinder, S.; Havelaar, A.H.; Cassini, A.; Devleesschauwer, B.; Kretzschmar, M.; Speybroeck, N.; et al. Disability Weights for the Global Burden of Disease 2013 Study. *Lancet Glob. Health* **2015**, *3*, e712–e723. [CrossRef]
36. Lattimore, T.; Szepesvári, C. Instance-Dependent Lower Bounds. In *Bandit Algorithms*; Cambridge University Press: Cambridge, NY, USA, 2020; p. 171. ISBN 978-1-108-57140-1.
37. Nelder, J.A.; Mead, R. A Simplex Method for Function Minimization. *Comput. J.* **1965**, *7*, 308–313. [CrossRef]
38. Virtanen, P.; Gommers, R.; Oliphant, T.E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **2020**, *17*, 261–272. [CrossRef]
39. David, S.; Kim, P.Y. Drug Trials. In *StatPearls*; StatPearls Publishing: Treasure Island, FL, USA, 2022.
40. The 5 Stages of COVID-19 Vaccine Development: What You Need to Know about How a Clinical Trial Works | Johnson & Johnson. Available online: https://www.jnj.com/innovation/the-5-stages-of-covid-19-vaccine-development-what-you-need-to-know-about-how-a-clinical-trial-works (accessed on 7 March 2022).
41. Food and Drug Administration. Step 3: Clinical Research. Available online: https://www.fda.gov/patients/drug-development-process/step-3-clinical-research (accessed on 7 March 2022).
42. Villar, S.S.; Bowden, J.; Wason, J. Multi-Armed Bandit Models for the Optimal Design of Clinical Trials: Benefits and Challenges. *Stat. Sci. Rev. J. Inst. Math. Stat.* **2015**, *30*, 199–215. [CrossRef]
43. Reverdy, P.; Leonard, N.E. Parameter Estimation in Softmax Decision-Making Models with Linear Objective Functions. *IEEE Trans. Autom. Sci. Eng.* **2016**, *13*, 54–67. [CrossRef]
44. Vamplew, P.; Dazeley, R.; Foale, C. Softmax Exploration Strategies for Multiobjective Reinforcement Learning. *Neurocomputing* **2017**, *263*, 74–86. [CrossRef]
45. Saari, D.G.; Merlin, V.R. The Copeland Method: I.: Relationships and the Dictionary. *Econ. Theory* **1996**, *8*, 51–76. [CrossRef]
46. Talbi, E.-G. Ordinal Data Analysis. In *Metaheuristics: From Design to Implementation*; John Wiley & Sons: Hoboken, NJ, USA, 2009; p. 65. ISBN 978-0-470-27858-1.
47. Conover, W.J. *Practical Nonparametric Statistics*; Wiley Series in Probability and Statistics; Chapter 3; Applied Probability and Statistics Section; Wiley: New York, NY, USA, 1999; ISBN 978-0-471-16068-7.
48. Hobbs, B.F.; Hepenstal, A. Is Optimization Optimistically Biased? *Water Resour. Res.* **1989**, *25*, 152–160. [CrossRef]
49. Ito, S.; Yabe, A.; Fujimaki, R. Unbiased Objective Estimation in Predictive Optimization. In Proceedings of the 35th International Conference on Machine Learning, Stockholm, Sweden, 3 July 2018.
50. Diniz, M.A.; Tighiouart, M.; Rogatko, A. Comparison between Continuous and Discrete Doses for Model Based Designs in Cancer Dose Finding. *PLoS ONE* **2019**, *14*, e0210139. [CrossRef]
51. Jamieson, K.; Nowak, R. Best-Arm Identification Algorithms for Multi-Armed Bandits in the Fixed Confidence Setting. In Proceedings of the 2014 48th Annual Conference on Information Sciences and Systems (CISS), Princeton, NJ, USA, 19–21 March 2014; IEEE: Princeton, NJ, USA, 2014; pp. 1–6.
52. Kaibel, C.; Biemann, T. Rethinking the Gold Standard With Multi-Armed Bandits: Machine Learning Allocation Algorithms for Experiments. *Organ. Res. Methods* **2021**, *24*, 78–103. [CrossRef]
53. Talbi, E.-G. Statistical Analysis. In *Metaheuristics: From Design to Implementation*; John Wiley & Sons: Hoboken, NJ, USA, 2009; pp. 63–64. ISBN 978-0-470-27858-1.
54. Morton, V.; Torgerson, D.J. Effect of Regression to the Mean on Decision Making in Health Care. *BMJ* **2003**, *326*, 1083–1084. [CrossRef]
55. van Hasselt, H.; Guez, A.; Silver, D. Deep Reinforcement Learning with Double Q-Learning. In Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, Phoenix, AZ, USA, 12–17 February 2016; AAAI Press: Palo Alto, CA, USA, 2016; pp. 2094–2100.

56. Tokic, M.; Palm, G. Value-Difference Based Exploration: Adaptive Control between Epsilon-Greedy and Softmax. In *KI 2011: Advances in Artificial Intelligence*; Lecture Notes in Computer Science; Bach, J., Edelkamp, S., Eds.; Springer: Berlin/Heidelberg, Germany, 2011; Volume 7006, pp. 335–346. ISBN 978-3-642-24454-4.

57. Aziz, M.; Kaufmann, E.; Riviere, M.-K. On Multi-Armed Bandit Designs for Dose-Finding Trials. *J. Mach. Learn. Res.* **2021**, *22*, 1–38.

58. Audet, C.; Hare, W. Optimization Using Surrogates and Models. In *Derivative-Free and Blackbox Optimization*; Springer Series in Operations Research and Financial Engineering; Springer International Publishing: Cham, Switzerland, 2017; p. 235. ISBN 978-3-319-68913-5.

59. Brock, K.; Billingham, L.; Copland, M.; Siddique, S.; Sirovica, M.; Yap, C. Implementing the EffTox Dose-Finding Design in the Matchpoint Trial. *BMC Med. Res. Methodol.* **2017**, *17*, 112. [CrossRef]

60. Lee, S.M.; Cheng, B.; Cheung, Y.K. Continual Reassessment Method with Multiple Toxicity Constraints. *Biostat. Oxf. Engl.* **2011**, *12*, 386–398. [CrossRef]

61. Cai, C.; Rahbar, M.H.; Hossain, M.M.; Yuan, Y.; Gonzales, N.R. A Placebo-Controlled Bayesian Dose Finding Design Based on Continuous Reassessment Method with Application to Stroke Research. *Contemp. Clin. Trials Commun.* **2017**, *7*, 11–17. [CrossRef] [PubMed]

62. Le Tourneau, C.; Lee, J.J.; Siu, L.L. Dose Escalation Methods in Phase I Cancer Clinical Trials. *JNCI J. Natl. Cancer Inst.* **2009**, *101*, 708–720. [CrossRef] [PubMed]

63. Tighiouart, M.; Cook-Wiens, G.; Rogatko, A. A Bayesian Adaptive Design for Cancer Phase I Trials Using a Flexible Range of Doses. *J. Biopharm. Stat.* **2018**, *28*, 562–574. [CrossRef]

64. Du, Z.; Wang, L.; Pandey, A.; Lim, W.W.; Chinazzi, M.; Piontti, A.P.Y.; Lau, E.H.Y.; Wu, P.; Malani, A.; Cobey, S.; et al. Modeling Comparative Cost-Effectiveness of SARS-CoV-2 Vaccine Dose Fractionation in India. *Nat. Med.* **2022**, *28*, 1–5. [CrossRef] [PubMed]

65. Emmerich, M.T.M.; Deutz, A.H. A Tutorial on Multiobjective Optimization: Fundamentals and Evolutionary Methods. *Nat. Comput.* **2018**, *17*, 585–609. [CrossRef]

66. Shavarani, S.M.; López-Ibáñez, M.; Knowles, J. Realistic Utility Functions Prove Difficult for State-of-the-Art Interactive Multiobjective Optimization Algorithms. In Proceedings of the Genetic and Evolutionary Computation Conference, Lille, France, 26 June 2021; ACM: Lille, France; pp. 457–465.

67. O'Quigley, J.; Shen, L.Z. Continual Reassessment Method: A Likelihood Approach. *Biometrics* **1996**, *52*, 673–684. [CrossRef]

68. Zarnitsyna, V.I.; Handel, A.; McMaster, S.R.; Hayward, S.L.; Kohlmeier, J.E.; Antia, R. Mathematical Model Reveals the Role of Memory CD8 T Cell Populations in Recall Responses to Influenza. *Front. Immunol.* **2016**, *7*, 165. [CrossRef]

69. Randomisation. *Fundamentals of Clinical Trials*; Friedman, L.M., Ed.; Springer: Cham, Switzerland, 2015; pp. 123–145. ISBN 978-3-319-18538-5.

70. McDonald, I.; Murray, S.M.; Reynolds, C.J.; Altmann, D.M.; Boyton, R.J. Comparative Systematic Review and Meta-Analysis of Reactogenicity, Immunogenicity and Efficacy of Vaccines against SARS-CoV-2. *NPJ Vaccines* **2021**, *6*, 1–14. [CrossRef]

71. Glass, E.J. Genetic Variation and Responses to Vaccines. *Anim. Health Res. Rev.* **2004**, *5*, 197–208. [CrossRef]

72. Hodges, J.L. The Significance Probability of the Smirnov Two-Sample Test. *Ark. För Mat.* **1958**, *3*, 469–486. [CrossRef]

73. Mann, H.B.; Whitney, D.R. On a Test of Whether One of Two Random Variables Is Stochastically Larger than the Other. *Ann. Math. Stat.* **1947**, *18*, 50–60. [CrossRef]

74. Weisstein, E.W. Bonferroni Correction. Available online: https://mathworld.wolfram.com/ (accessed on 11 April 2022).

Supplementary material for paper 4

These supplementary are adapted from those that were included with the published version of this paper, which can be found online [194].

Saturating - Sigmoid

The following model is a classical dose-response model for modelling dose-efficacy in drugs.

$$Saturating(Dose) = \frac{1}{1 + e^{gradient \times (midpoint - dose)}}$$

This may be recognisable as the model for logistic regression. This has an efficacy of 0 as dose tends to negative infinity, and an efficacy of 1 as dose tends to positive infinity. In this work we made the adjustment of multiplying this by a parameter 'maximum' that is between 0 and 1. This was because for a given vaccine candidate assuming that efficacy can become 100% for sufficiently large doses seemed unreasonable. It seemed possible that an otherwise immunogenic vaccine may not be immunogenic in some individuals regardless of dose, for example due to variations in major histocompatibility complexes [1].

Peaking - Latent Quadratic

The latent quadratic model has been commonly used for the purposes of modelling non-monotonically increasing dose response. It is similar to the saturating model, other than it includes quadratic terms to allow qualitative dose-response to vary between small and large doses. By differentiation it can be shown that the maxima point within the parameter bounds used is at:

$$dose = -\frac{gradient1}{2 \times gradient2}$$

Please see Supplementary weighted model averaging.

Probit

An excellent description of the probit model for ordinal toxicities can be found in [2]. In short, we consider the probability of observing toxicity less than grade j given dose $P(Y = j|dose)$ by:

$$P(Y = 0|dose) = P(Y < 1|dose) - P(Y < 0|dose) = P(Y < 1|dose) - 0$$

$$P(Y = 1|dose) = P(Y < 2|dose) - P(Y < 1|dose)$$

$$P(Y = 2|dose) = P(Y < 3|dose) - P(Y < 2|dose)$$

$$P(Y = 3|dose) = P(Y < 4|dose) - P(Y < 3|dose) = 1 - P(Y < 3|dose)$$

$P(Y < 0|dose)$ and $P(Y < 4|dose)$ are respectively 0 and 1 (0 and 3 are assumed to be the lowest and highest adverse event gradings possible). Otherwise $P(Y < j|dose)$ is given by $\varphi(threshold_j - gradient \times dose)$, where $\varphi()$ is the cumulative density function with $\sigma = 1$. Hence at $dose = \frac{threshold_j}{gradient}$, $P(Y < j|dose) = 0.5$

Parameters and bounding

Within the calibration of these models, certain bounds were used to aim the calibration and for biological plausibility.

Parameter	Bounds	Notes
		Saturating - Sigmoid
gradient	0 to 6	It seems biologically unlikely that a vaccine dose-efficacy curve is discontinuous or that a relatively small change in dose could be responsible for a massive change in vaccine efficacy. Therefore, the gradient was bounded to prevent an overly steep curve. The maximum gradient of 6 would mean that a single log increase in dose (only 1/10th of the dosing space) could be responsible for 90% of the change in vaccine efficacy. None of the scenarios had a gradient steeper than this, so if it is believed that 1/10th of the dosing space could be responsible for >90% of the change in vaccine efficacy a higher value should be used. A lower bound of 0 prevents efficacy from being decreasing w.r.t. increasing dose, which is one of the assumptions of saturating dose-response.
midpoint	0 to infinity	This ensures that the model cannot predict that efficacy is saturating below the lowest dose in the dosing space. This assumption was true for all scenarios but could be relaxed if it is believed that efficacy may have already begun saturating at the lowest dose in the dosing space.
maximum	0 to 1	Having a greater than 1 or less than 0 probability of efficacy is biologically and probabilistically impossible.
		Peaking - Latent Quadratic
base	-infinity to infinity	Unbounded to allow other parameters to well define the model.
gradient1	0 to 6	See notes for the gradient parameter of the saturation sigmoid model. Both used the same bounds to ensure fairness in comparison.
gradient2	-infinity to 0	Bounding to be less than zero is to ensure that an increasing dose will eventually lead to a decreasing efficacy, an assumption of the peaking model.
		Toxicity Probit
gradient	0 to 6	See notes for the gradient parameter of the saturation sigmoid model.
threshold1	-infinity to infinity	Unbounded to allow high or low levels of this toxicity grade at either the lowest or highest doses in the dosing space.
threshold2	-infinity to infinity	As above.
threshold3	-infinity to infinity	As above.

Table S1.1. Parameter bounds for dose-response models.

Supplementary 2. Weighted model averaging

Model averaging can be a useful method for inference when the current model form is uncertain. We use the method of model weighting via the Akaike Information Criterion (AIC) model averaging [3]. We outline this here.

Suppose that we have two competing models M_1 and M_2. After gathering data and calibrating models we have respective likelihoods L_1 and L_2 for the two models. Then AIC_1 and AIC_2 can be calculated as

$$AIC_i = -2\ln(L_i) + 2k_i$$

Where k_1 is the number of parameters for model i. Let AIC_{best} be the minimum of AIC_1 and AIC_2. and $\nabla_i = AIC_i - AIC_{best}$ for each model.

The Akaike weights W_1 and W_2 can be calculated as

$$W_i = \frac{e^{(-\nabla_i)}}{e^{(-\nabla_1)} + e^{(-\nabla_2)}}$$

Finally, the predictions of the weighted model at some input value x are hence given by

$$M_w(dose) = W_1 M_1(dose) + W_2 M_2(dose)$$

Where $M_1(dose)$ and $M_2(dose)$ are respective predictions for M_1 and M_2 at dose.

Supplementary 3. Pareto Optimality

This section has been adapted and lengthened, see appendix A.B.

Supplementary 4. Scenarios

We aimed for the scenarios to be qualitatively different in both shape and optimal dose.

Scenario Saturating 1

Qualitatively this scenario had a saturating efficacy curve and a high optimal dose, with a relatively steep utility curve.

	Parameter	Value
Efficacy	gradient	1.000
	midpoint	6.000
	maximum	0.900
Toxicity	gradient	1.000
	threshold1	3.000
	threshold2	9.000
	threshold3	10.500
Utility Weights	Weight$_{Efficacy}$	0.133

Table S.4.1.Saturating 1. Parameters for the scenario Saturating 1
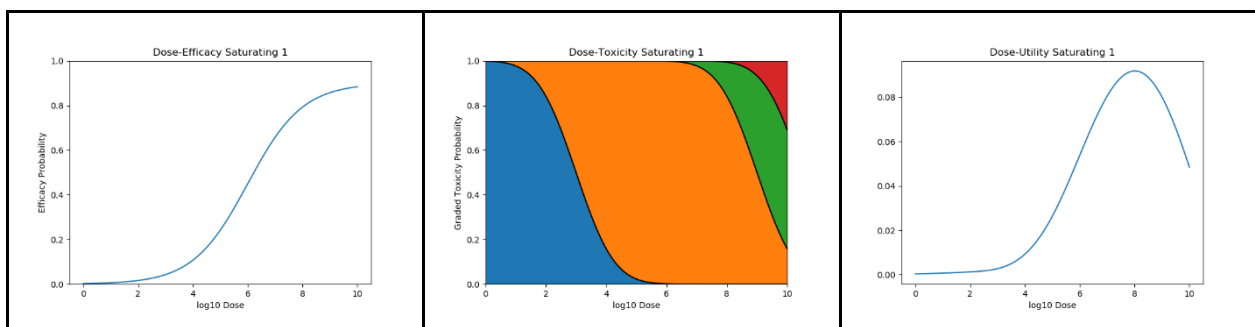
Figure.S.4.1. Saturating 1. Dose-efficacy, dose-toxicity, and dose-utility plots for the scenario Saturating 1

192

Qualitatively this scenario had a peaking efficacy curve and a low optimal dose, with a relatively steep utility curve.

	Parameter	Value
Efficacy	base	-6.000
	gradient1	5.000
	gradient2	-0.750
Toxicity	gradient	0.500
	threshold1	1.000
	threshold2	3.000
	threshold3	5.000
Utility Weights	Weight$_{Efficacy}$	0.133

Table.S.4.8.Peaking 3. Parameters for the scenario Peaking 3
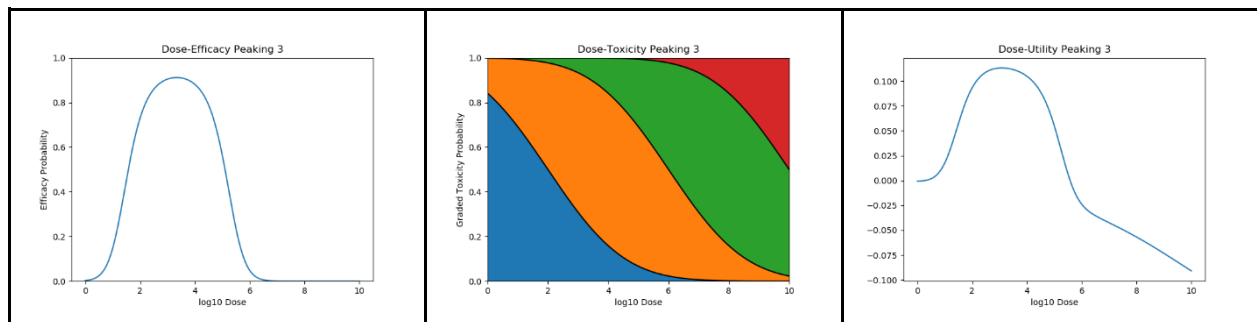
Figure.S.4.8.Peaking 3. Dose-efficacy, dose-toxicity, and dose-utility plots for the scenario Peaking 3

Other 2 represents a vaccine for which dose-efficacy is fundamentally saturating but follows a different and more complicated biphasic parametric form to the sigmoid saturating model assumed elsewhere in this paper. The efficacy model is given as

$$Biphasic(dose) = \frac{maximum \times fraction}{1 + e^{gradient1 \times (midpoint1 - dose)}} + \frac{maximum \times (1 - fraction)}{1 + e^{gradient2 \times (midpoint2 - dose)}}$$

	Parameter	Value
Efficacy	gradient1	0.500
	gradient2	3.000
	midpoint1	4.000
	midpoint2	6.000
	maximum	0.900
	fraction	0.500
Toxicity	gradient	0.500
	threshold1	1.000
	threshold2	3.000
	threshold3	5.000
Utility Weights	Weight$_{Efficacy}$	0.133

Table.S.4.12.Other 2. Parameters for the scenario Other 2
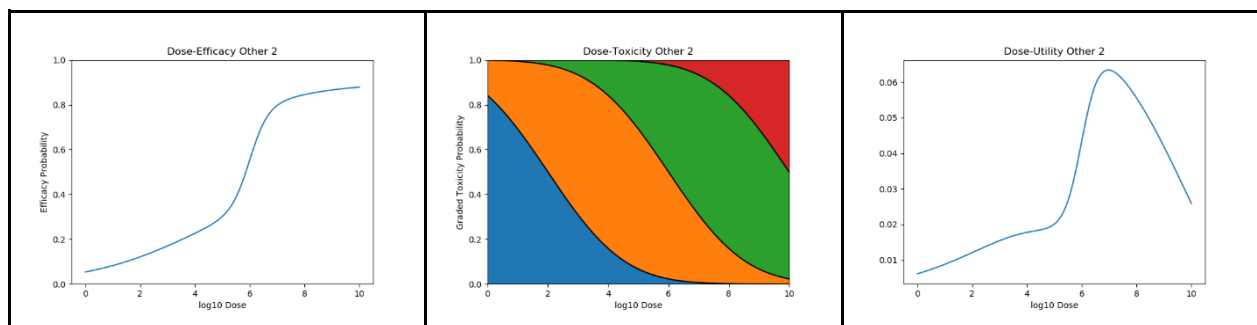
Figure.S.4.12.Other 2. Dose-efficacy, dose-toxicity, and dose-utility plots for the scenario Other 2

Method

SoftMax selection is a method of action selection used commonly in both multi-armed bandit problems and reinforcement learning. We provide a description of action/dose selection under this method. Let A_1, A_2, …., A_n be the n possible actions available to be taken, each with respective predicted utility U_1, U_2,..., U_n. Then an action A_i is selected to test (dose selected to trial) with probability

$$Probability\ of\ selecting\ action\ A_i = \frac{e^{inversetemperature \times U_i}}{\sum_{j=1}^{n} e^{inversetemperature \times U_j}}$$

where inverse_temperature is a hyperparameter which controls the degree of exploration. An increased inverse temperature leads to actions with lower predicted utility having lower probability of selection. For inverse_temperature = 0, which is the lowest possible inverse_temperature, all actions are selected with equal probability (1/n). As inverse_temperature tends to infinity, this selection method tends to only select the action(s) with the maximum predicted utility. A random number generator is used to select an action with these probabilities.

inverse_temperature values

For the balanced fully continuous trial dose selection method we used inverse_temperature = 69. This was chosen such that a utility difference of 0.01 would have a doubled probability of selection. This is shown

$$2e^{inversetemperature \times U_i} = e^{inversetemperature \times (U_i + 0.01)}$$

$$= e^{inversetemperature \times (U_i)} e^{inversetemperature \times (0.01)}$$

$$2 = e^{inversetemperature \times (0.01)}$$

$$\ln(2) = 0.69 = 0.01 \times inversetemperature$$

$$69 = inversetemperature$$

For the three stage SoftMax dose selection method we used inverse_temperature = (58.9,294) for selecting the (2nd, 3rd) stage of doses. This was chosen such that for

utility difference of (0.05, 0.01) would lead to the dose that was predicted better being selected 95% of the time. This is shown in the 0.05 case by

$$95e^{inversetemperature \times U_i} = 5e^{inversetemperature \times (U_i + 0.05)}$$

$$= 5e^{inversetemperature \times (U_i)} e^{inversetemperature \times (0.05)}$$

$$19 = e^{inversetemperature \times (0.05)}$$

$$\ln(19) = 2.94 = 0.05 \times inversetemperature$$

$$58.9 = inversetemperature$$

And similarly for 0.01.

These may not have been optimal values, but optimal values are likely to vary depending on the scenario.

Supplementary 6. Pseudodata

Pseudodata or anchor points are used to stabilise and inform models for which little real data are available. In short, we pretend that there exist data which does not actually exist but have these data points being less important than real data, and further downweight or ignore these data as more data are gathered. We aim to use minimally informative pseudodata. For all approaches pseudodata are fully ignored for final dose selection.

Efficacy Models

For efficacy modelling, pseudo data were of the form in table S.Pseudodata.1. Thus, there were 300 pseudo individuals divided evenly over 3 doses.

For the standard and balanced fully continual approaches the weight of a pseudodata-point was 0.01 of regular datapoint. Thus, the effective sample size of the pseudodata was 3 (=300x0.01), which is quickly minimal relative to the amount of real data.

For the 3 stage SoftMax approaches the weights of a pseudodata-point were 0.01 and 0.001 for the second and third trial dose selections respectively. Thus, the

196

effective sample size of the pseudodata were 3 and 0.3. Thus, for the third dose selection the pseudodata represented only 1.47% (=0.3/ (20+0.3)) of the data.

Dose	Non-efficacy response	Efficacy Response
1	90	10
5	50	50
9	10	90

Table S.6.1. Efficacy pseudodata

Toxicity Model

For toxicity modelling, pseudo data were of the form in table S/Pseudodata.2. Thus, there were 200 pseudo individuals divided evenly over 2 doses.

Again, for the standard and balanced fully continual approaches the weight of a pseudo data point was 0.01 of regular datapoint. Thus, the effective sample size of the pseudodata was 2, which is quickly minimal relative to the amount of real data.

For the 3 stage SoftMax approaches the weights of a pseudo datapoint were 0.01 and 0.001 at the second and third trial dose selections respectively. Thus, the effective sample size of the pseudodata were 2 and 0.2. Thus, for the third dose selection the pseudodata represented only 0.99% (=0.2/20.2) of the data.

Dose	Grade 0 Response	Grade 1 Response	Grade 2 Response	Grade 3 Response
1	45	35	10	10
9	2	3	5	90

Table S.6.2. Toxicity pseudodata

Supplementary 7. Copeland

Copeland's method is a method of ranking that effectively asks the question "how often would we have preferred to have used this option over a different option". The process is to compare the metrics of each simulation, and see which approach did 'best'.

In a comparison between the ith simulation for approach A and and the jth simulation for approach B:

If A_i was better that B_i: A scores +1, B scores +0

If A_i was worse that B_i: A scores +0, B scores +1

If A_i and B_i were the same: Both score +0.5

In any case the count of A comparisons (n_A) and count of B comparisons (n_B) also increase by 1.

Comparisons are conducted for all i and j between all approaches, then divided by the count of comparisons for that approach

It is well discussed, but here we show a concrete example for a tiny dataset. Recall that a lower PSR is preferable.

	Approach A	Approach B	Approach C
Respective PSR from the first and second simulation of each approach.	5	9	7
	6	1	5

Table S.7.1. Toy scenario 1 Data

The first simulation of approach A got a PSR of 5.

The first simulation of approach B got a PSR of 9. As this is greater (and therefore less preferable) than the PSR for that simulation of approach A, score A increases by 1 and score B says the same.

The second simulation of approach B got a PSR of 1. As this is lower (and therefore preferable) than the PSR for that simulation of approach A, score B increases by 1 and score A says the same.

198

The first simulation of approach C got a PSR of 7. As this is greater (and therefore less preferable) than the PSR for that simulation of approach A, score A increases by 1 and score C says the same.

The second simulation of approach C got a PSR of 5. As this is equal to the PSR for that simulation of approach A, score A and score C increase by 0.5.

Thus after making comparisons for simulation 1 of approach A we have score A = 2.5 and n_A = 4.

The second simulation of approach A got a PSR of 6. Repeating the steps leads to score A increasing to 4.5(=2.5+2) and n_A=8(=4+4). Thus, the total Copeland score is 0.5625(=4.5/8).

This can be repeated for all 3 approaches, and approaches ranked by their Copeland Score to give the following table.

	Approach A	Approach B	Approach C
Copeland Score	0.5625	0.5	0.4375
Copeland Rank	1	2	3

Table S.7.2. Toy scenario 1 Copeland Table

Suppose that there was a second scenario that these approaches were tested on, with data as below.

	Approach A	Approach B	Approach C
Respective simple regret from the first and second simulation of each approach.	8	3	6
	10	5	4

Table S.7.3. Toy scenario 2 Data

This could give the table below, representing the Copelands scores and ranking for that scenario.

	Approach A	Approach B	Approach C
Copeland Score	0	0.875	0.625
Copeland Rank	3	1	2

Table S.7.4. Toy scenario 2 Copeland Table

We could combine these by summing the scores and comparisons for both scenarios. So, for example, in scenario 1 we had $score_A = 4.5$ and $n_A = 8$. In scenario 2 we had $score_A = 0$ and $n_A = 8$. Therefore, we have a total $score_A = 4.5$ and $n_A = 16$ for a total Copeland score of 0.2815. Repeating this gives aggregate Copeland metrics in the below table.

	Approach A	Approach B	Approach C
Copeland Score	0.28125	0.6875	0.53125
Copeland Rank	3	1	2

Table S.7.5. Aggregate Copeland table for both toy scenarios.

Note that this can be read as approach B 'winning' 68.75% of comparisons across both scenarios. From this we could say that Approach B seems to be most effective for minimising PSR across both of these scenarios.

Supplementary 8. Exceptions

These exceptions were (n=60,100, scenario = Saturating 3), and (n=60,100, scenario = Other 2) [supplementary Objective 1 plots Saturating 3, Other 2]. For all other scenarios, the peaking latent-quadratic performed similarly or better than the sigmoid saturating curve. We suggest that, for the Saturating 3 scenario, the peaking curve was unable to approximate well the efficacy curve saturating at a low dose and remaining high across the rest of the dosing-space. We suggest that, for the Other 2 scenario, the steady increase in efficacy with a large jump in efficacy near the middle

200

could have been the feature of the scenario that inhibited the peaking latent-quadratic model.

In this work we found that predicted utility of the dose that was predicted optimal was often higher than the true utility at that dose. This was shown by the inaccuracy metric typically being greater than zero and is referred to as an 'optimistic bias'. This was true for all of the dose-optimisation approaches, please consider the problems of 'regression to the mean' and the 'Stein Paradox' for examples of why this should be expected even in direct comparison approaches to dose optimisation. The amount of optimism was decreased by increasing trial size. This 'optimistic bias' is clearly undesirable. However, we believe that this is neither an issue with the models or calibration, nor unique to modelling-based optimisation. In this supplementary section we show that

- Similar issues are observed in binomial direct comparison dose-optimisation approaches.
- Similar issues are observed in continuous direct comparison dose-optimisation approaches.
- Similar issues are observed in a heavily simplified modelling setting.
- Similar issues are observed in a simple physical modelling setting.

We also note that similar problems have previously been highlighted both in over modelling-based optimisation problems, and in the continuous self-optimisation problems referred to as reinforcement learning. Methods for solving these problems involve either double-q learning [4] or using only half the data for optimisation and half for prediction, neither of which seem entirely reasonable given the small amount of data and parametric model forms involved in vaccine dose-optimisation.

Direct comparison: Binomial

Consider testing k vaccine doses and attempting to choose that which maximises a binary measure of vaccine efficacy. For each of the k doses, n individuals receive

that dose and a binary efficacy outcome is recorded depending on the true probability of efficacy for that dose, $\bar{p} = (p_1, p_2, ..., p_k)$.

These are used to estimate the probability of efficacy for each dose, $ep_1, ep_2, ..., ep_k$. Predicted optimal dose is then d_i such that $ep_i >= ep_j$ for all j in 1 to k, breaking ties at random. This is the basic direct comparison approach.

After conducting this procedure, one of 3 things can happen.

$p_i < ep_i$ (overestimation)

$p_i = ep_i$ (accurate estimation)

$p_i > ep_i$ (under estimation)

We show that typically the first ($p_i < ep_i$) can be most common, which is to say that optimistic bias is observed when simulated. As an explicit example, consider k = 2, $\bar{p}$ = (0.5, 0.5), n = 10. A reasonable example efficacy observation for these doses respectively given these probabilities could be (6/10, 4/10). The dose with 6/10 efficacies would be predicted to be most likely to be optimal, and the best estimate of efficacy probability would be 0.6, an overestimation. Hence, whilst these observations were unbiased, the estimate for the optimal dose was optimistically biased.

See below tables for different simulation observations. Note that this bias is made worse by increasing the number of possible doses, and by the query doses having more similar true probabilities

k	p̄	n	Overestimation	Accurate Estimation	Underestimation
2	(0.5, 0.5)	10	61999	24774	14027
2	(0.5, 0.5)	100	70910	8017	21073
2	(0.6, 0.7)	10	59076	25649	15275
2	(0.6, 0.7)	100	53007	8579	38414
2	(0.1, 0.9)	10	34975	38748	26277
2	(0.1, 0.9)	100	45254	13238	41508
3	(0.5, 0.5, 0.5)	10	75665	18956	5379
3	(0.5, 0.5, 0.5)	100	75665	18956	5379
3	(0.5,0.6, 0.7)	20	65322	17433	17245
3	(0.1, 0.1, 0.5)	10	40845	24764	34391
3	(0.0, 0.0, 0.5)	10	37552	24683	37765

Table S9.1. Overestimation/Underestimation results from 100,000 simulated clinical trials.

This can also be considered a similar phenomenon to that of the multiple comparison problem.

Direct comparison: Continuous

Consider testing k vaccine doses and attempting to choose that which maximises a continuous, normally distributed measure of vaccine efficacy. For each of the k doses individuals receive that dose and the continuous efficacy outcome is recorded from the true probability density function of the efficacy response. This is defined by normal distributions with respective means and standards deviations, m = $(m_1, m_2, ..., m_k)$ and s = $(s_1, s_2, ..., s_k)$.

These are used to estimate the mean value of the true efficacy variable for each dose, $em_1, em_2, ..., em_k$. Predicted optimal dose is then di such that $em_i >= ep_m$ for all j in 1 to k, breaking ties at random. This is the basic direct comparison approach.

After conducting this procedure, one of 3 things can happen.

203

- $m_i < em_i$ (overestimation)
- $m_i = em_i$ (accurate estimation)- effectively impossible unless censored to a number of significant places.
- $m_i > em_i$ (under estimation)

Again, we see that the first is most likely. This is because we have for each dose di

$$em_i = m_i + error_i$$

$$error_i \sim N(0, errorscale)$$

Hence for the selection process (maximising em_i) we are not only maximising m_i, but also $error_i$, hence we are more likely to select doses which are overestimated.

See below tables for different observations. Note again that this bias is made worse by these doses having more similar true probabilities. We only show k=2 examples here, but again increasing k would exacerbate the issue.

k	m	s	n	Overestimation	Underestimation
2	(10,10)	(2,2)	10	75113	24887
2	(10,10)	(2,2)	100	75126	24874
2	(10,9)	(2,2)	10	61422	38578
2	(10,9)	(2,2)	100	50177	49823
2	(10,10)	(4,4)	10	75113	24887
2	(10,10)	(4,4)	100	75126	24874
2	(10,9)	(4,4)	10	70636	29364
2	(10,9)	(4,4)	100	53819	46181
2	(50,20)	(2,2)	10	49919	50081
2	(50,20)	(2,2)	100	50152	49848

Table S9.2 Overestimation/Underestimation results from 100,000 simulated clinical trials.

Simplified modelling

Please see appendix A.D.3.
.
Physical Modelling

Please see appendix A.D.3.

Supplementary 10. Plotted Clinicals Trial Results

Each of the below plots show the true utility curves and predicted optimal/dose response for 100 simulations of each approach for each scenario. Approaches used a saturating (left), peaking(middle), or weighted (right) efficacy curve. From top to bottom, trial size is 10, 30, 60, 100, 30, 30, 30. Method of trial dose selection from top to bottom in uniform, uniform, uniform, uniform, SoftMax 3 stage, Standard fully continual, and balanced fully continual. Red points show the model predicted optimal dose and utility prediction for that dose for a single clinical trial. Ideally these points should be near the peak of the true dose utility curve

I only show the figures for three scenarios here for demonstration, the other figures are included in the appendix of this thesis [A.D.3.]

Scenario Saturating 1
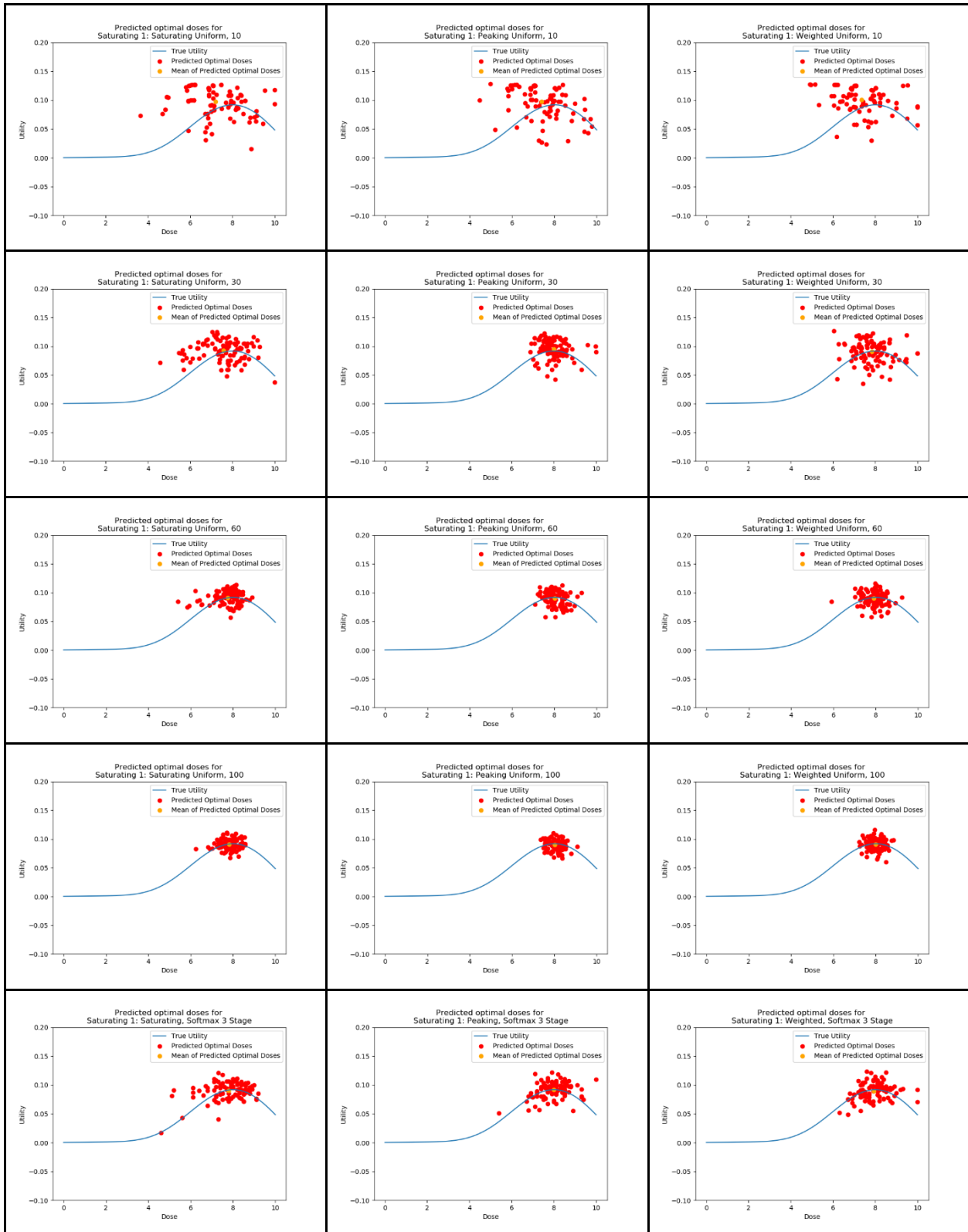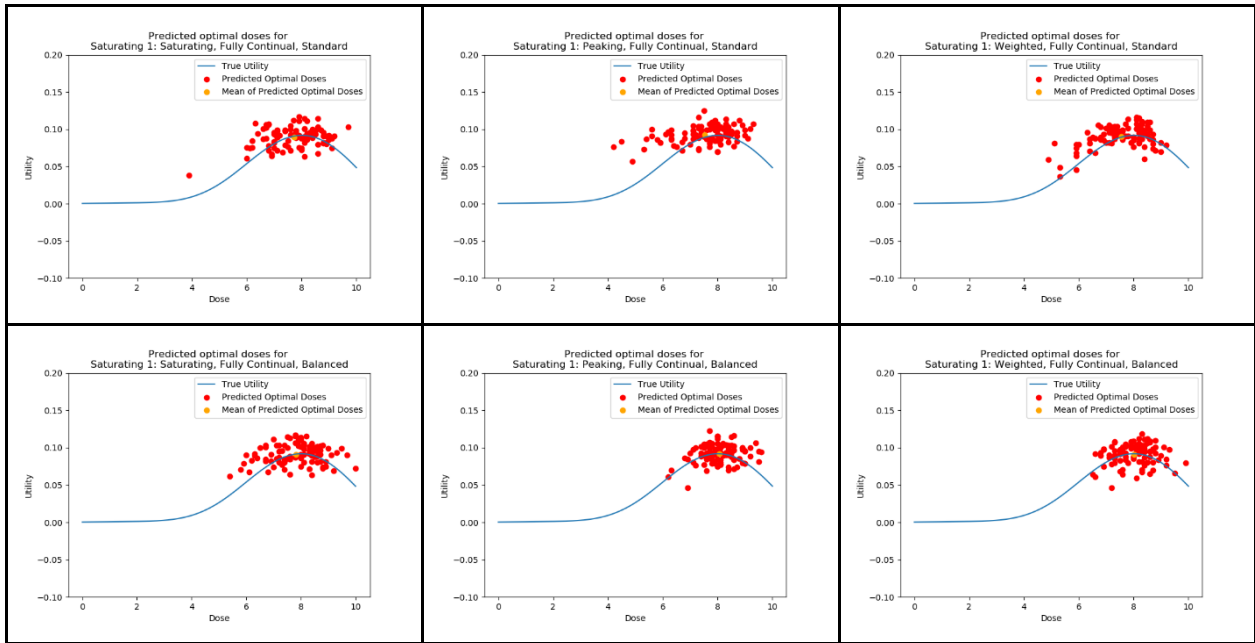
Figure S.10.1. Clinical trials by dose optimisation approach for scenario Saturating 1.

Scenario Peaking 3
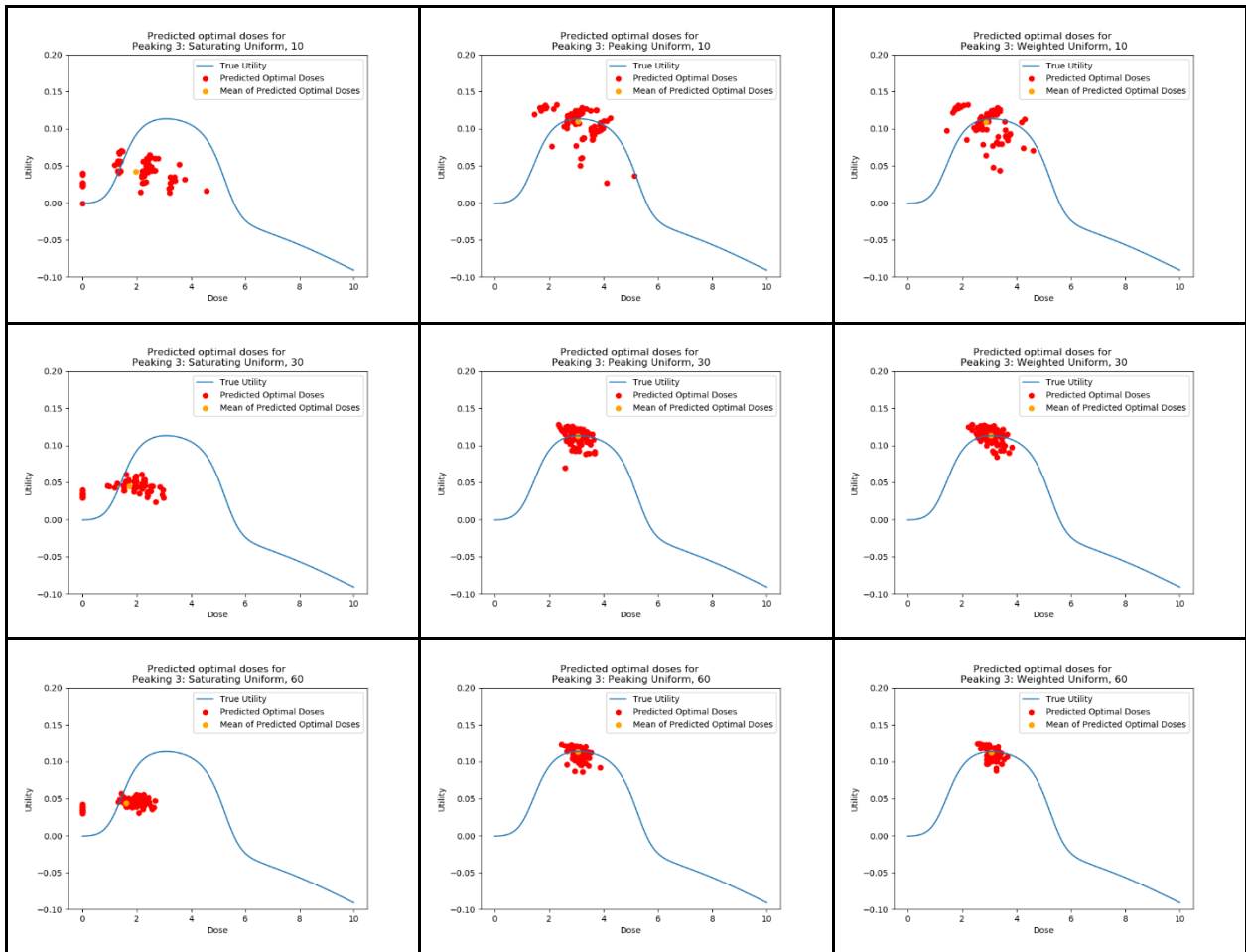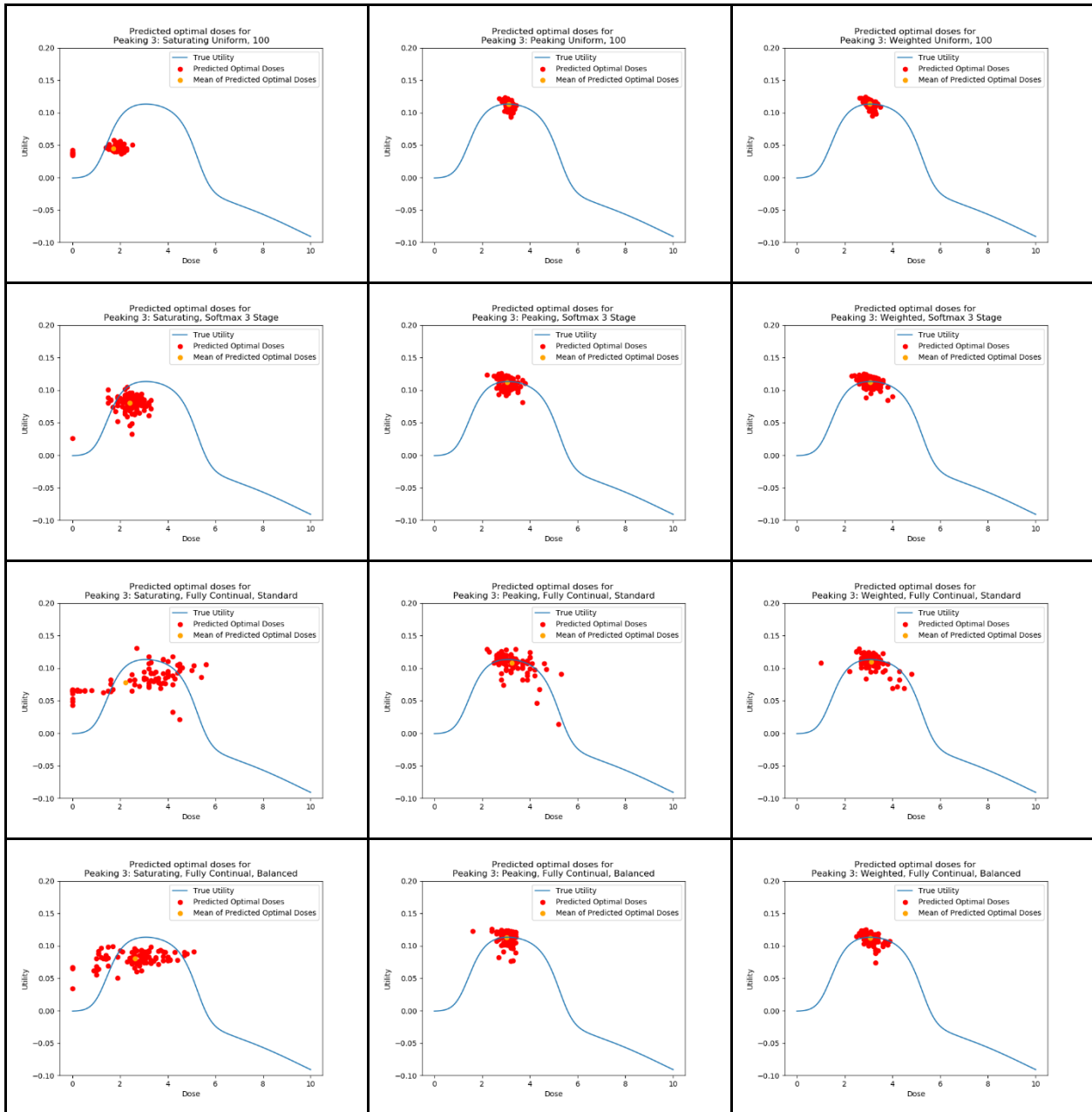
Figure S.10.8. Clinical trials by dose optimisation approach for scenario Peaking 3.

Scenario Other 2

208
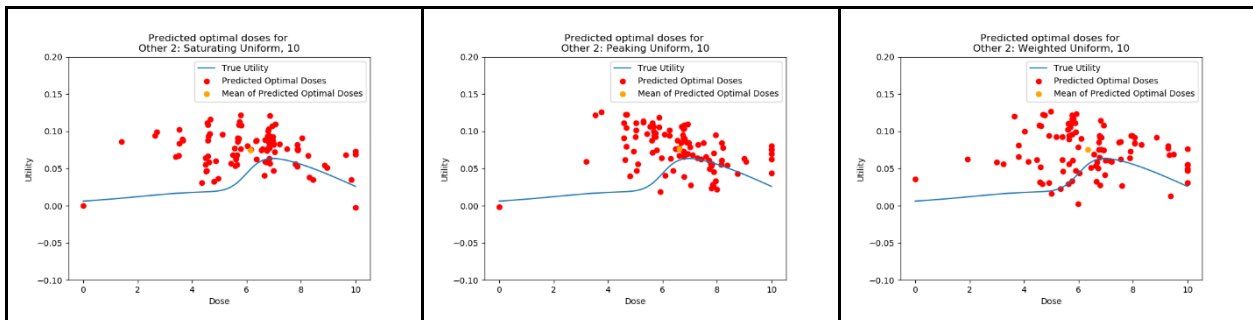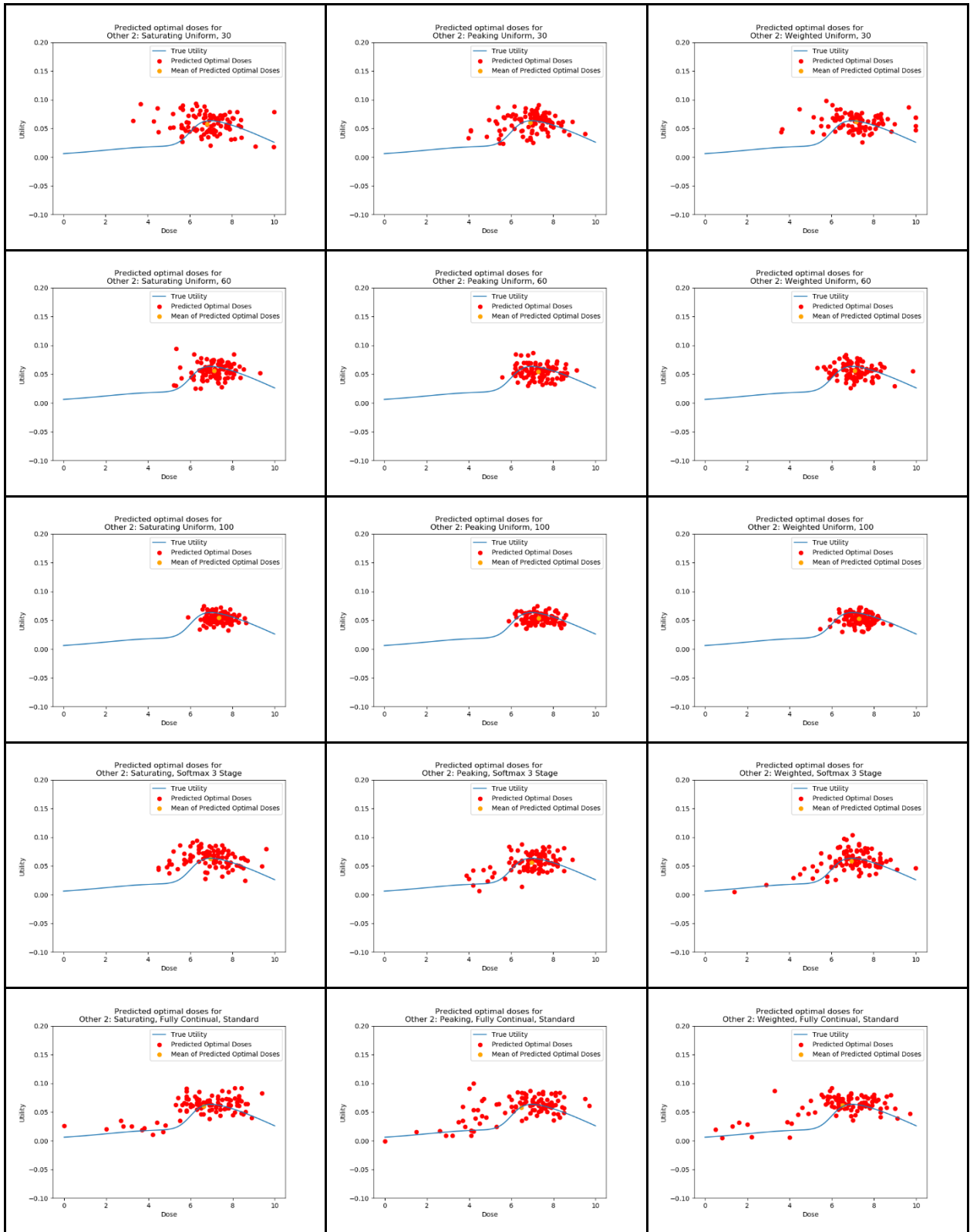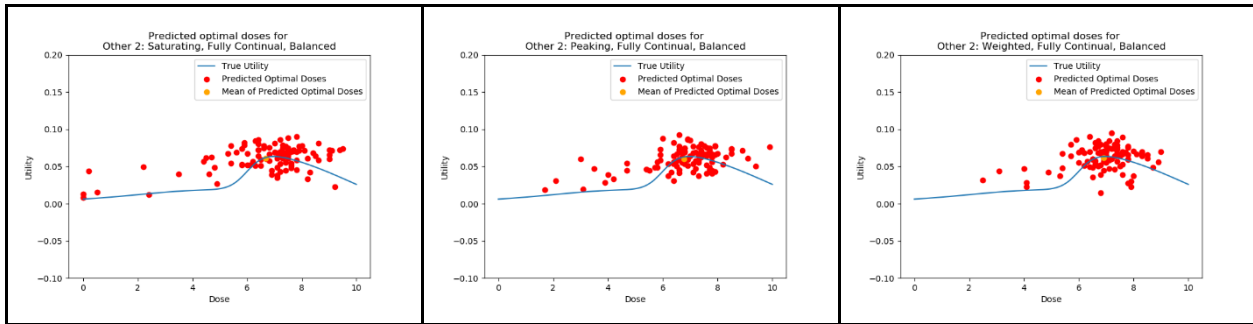
Figure S10.12 Clinical trials by dose optimisation approach for scenario Other 2.

Supplementary 11. Objective 1 Plots

The below plots show metrics from simulations for dose-optimisation approaches in objective 1 for each scenario. For each figure, the shown metrics are simple regret (top left), percentage simple regret (top right), inaccuracy (middle left), absolute inaccuracy (middle right), average regret (bottom left), and percentage average regret (bottom right).

I only show the figures for Scenario S1 here for demonstration, the other figures are included in the appendix of this thesis [A.D.3.]
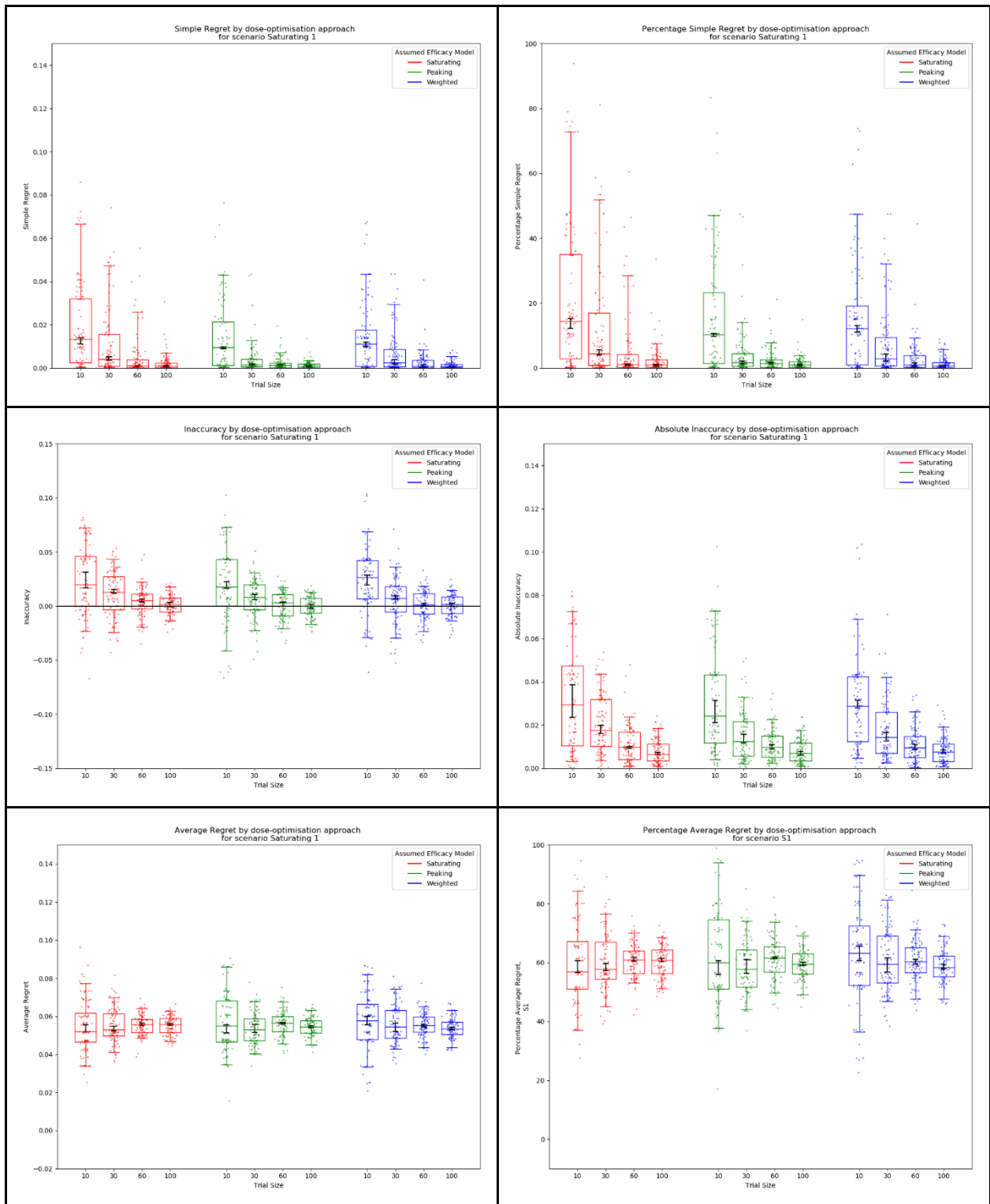
Figure S.11.1 Metrics by dose-optimisation approach for objective 1 for scenario Saturating 1.

Here we provide a statistical analysis of the data presented in the main body and supplementary sections of this work. We provide analysis in two ways.

1. The Kolmogorov–Smirnov Test is used to determine whether there is evidence to support two samples being drawn from the same distribution or from different distributions [5].
2. The One-sided Mann-Whitney U Test is used to evaluate whether values in one sample tend to be larger/smaller than another [6].

For each test we present a heatmap of p-values, with each cell in the heat map containing the p-value for that test for the comparison between the sample metrics for the dose-optimisation approach in the respective row/column. Cells in the table with a light pink hue represent the test statistic for that comparison would be significant under the threshold $p<0.05$, cells with a red hue represent the test statistic for that comparison would be significant under the threshold $p<0.05$ with Bonferroni multiple comparison correction [7]. This threshold was $p<0.00076 = 0.05/66$ (66 = 12C2 numbers of unique pairings of 12 approaches in each objective).

We believe that the qualitative analysis and Copeland metrics presented in the body of the work and Supplementary sections 10-13 are more relevant for showing practical differences between the dose-optimisation approaches than the statistical analysis presented in this section. These analyses are only presented for interest.

S12.1. Objective 1 Total Analysis

Here we show the p-values for objective 1 for the metrics of PSR, Absolute Inaccuracy, and PAR. These are the metrics for the combined data of all scenarios. This is the data in figures 7, 8b and 9 respectively.

For interpretation, the Kolmogorov–Smirnov heatmaps are symmetric, with significance representing evidence that the true distribution for the approach-scenario test metrics of PSR, Absolute Inaccuracy, and PAR differ between the two dose-optimisation approaches across all scenarios. The One-sided Mann-Whitney U

test heatmaps are not symmetric, with significance for the cell in row A and column B representing 'statistically significant' evidence that approach A was preferable to approach B with regards to that metric (e.g., lower PSR, lower Absolute Inaccuracy, Lower PAR).
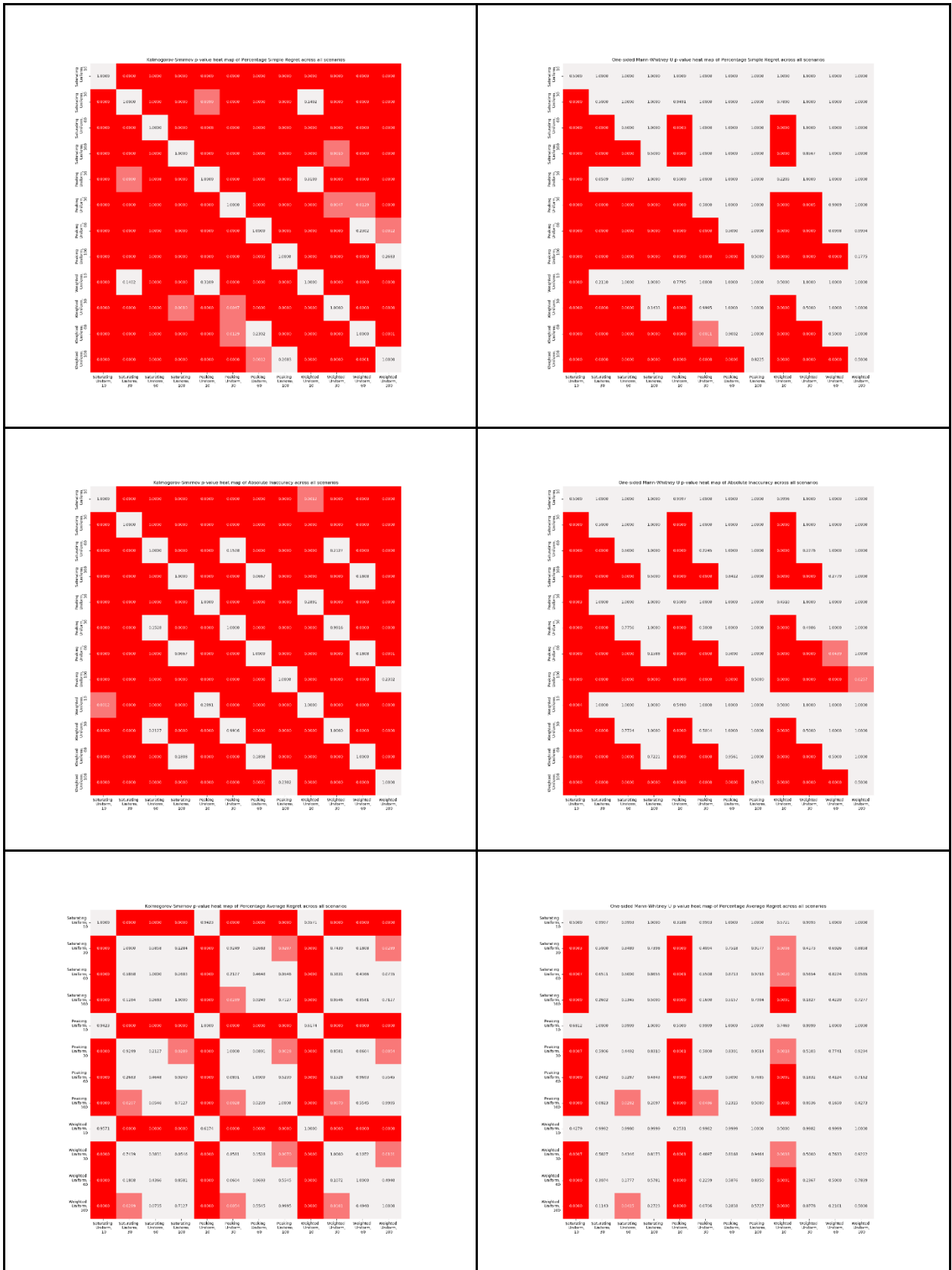
Figure S.12.1. Kolmogorov–Smirnov (left) and Mann-Whitney U (right) heatmaps of p-values for objective 1 across all scenarios. These are for the metrics of PSR (top), Absolute Inaccuracy (middle) and PAR (bottom). Cells with a light pink hue represent the test statistic for that comparison would be significant under the threshold p<0.05, cells with a red hue represent the test statistic for

214

that comparison would be significant under the threshold p<0.05 with Bonferroni multiple comparison correction.

S12.2. Objective 1 Scenario Specific

Please see appendix A.D.3.

S12.3. Objective 2 Total Analysis

Here we show the p-values for objective 2 for the metrics of PSR, Absolute Inaccuracy, and PAR. These are the metrics for the combined data of all scenarios. This is the data in figures 10, 11b and 12 respectively.

For interpretation, the Kolmogorov–Smirnov heatmaps are symmetric, with significance representing evidence that the true distribution for the approach-scenario test metrics of PSR, Absolute Inaccuracy, and PAR differ between the two dose-optimisation approaches across all scenarios. The One-sided Mann-Whitney U test heatmaps are not symmetric, with significance for the cell in row A and column B representing 'statistically significant' evidence that approach A was preferable to approach B with regards to that metric (e.g., lower PSR, lower Absolute Inaccuracy, Lower PAR).
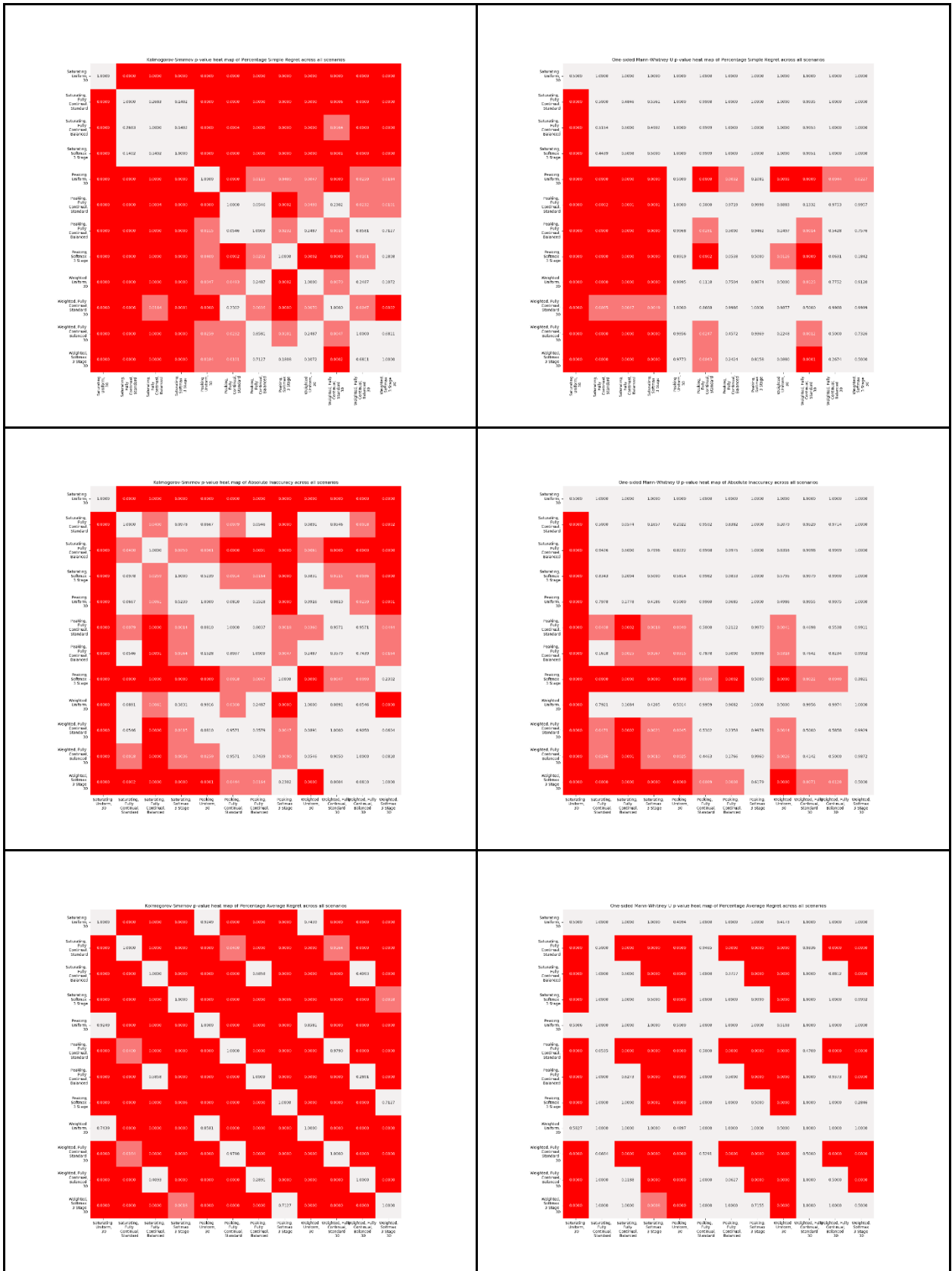
Figure S.12.3. Kolmogorov–Smirnov (left) and Mann-Whitney U (right) heatmaps of p-values for objective 2 across all scenarios. These are for the metrics of PSR (top), Absolute Inaccuracy (middle) and PAR (bottom). Cells with a light pink hue represent the test statistic for that comparison would be significant under the threshold p<0.05, cells with a red hue represent the test statistic for

216

that comparison would be significant under the threshold $p < 0.05$ with Bonferroni multiple comparison correction.

S12.4. Objective 2 Scenario Specific

Please see appendix A.D.3.

Supplementary 13. Objective 2 Plots

The below plots show metrics from simulations for dose-optimisation approaches in objective 2 for each scenario. For each figure, the shown metrics are simple regret (top left), percentage simple regret (top right), inaccuracy (middle left), absolute inaccuracy (middle right), average regret (bottom left), and percentage average regret (bottom right).

I only show the figures for Scenario S1 here for demonstration, the other figures are included in the appendix of this thesis [A.D.3.]
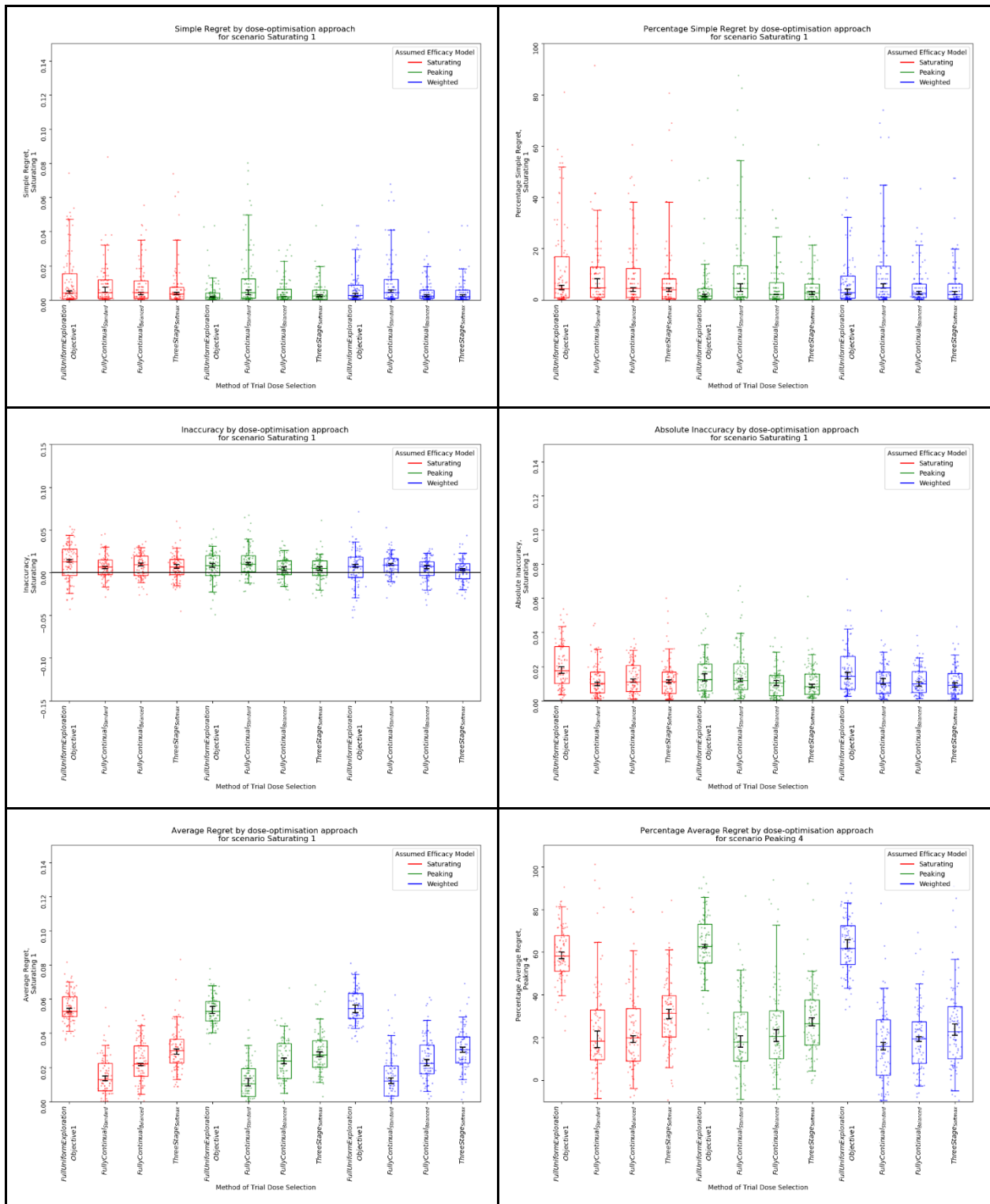
Figure S13.1. Metrics by dose-optimisation approach for objective 2 for scenario S1.

The below tables show Copeland metrics as in 3.2.2 of the paper, by scenario.

I only show the table for Scenario S1 here for demonstration, the other figures are included in the appendix of this thesis [A.D.3]

Scenario Saturating 1

	AR		IA		SR	
Approach	Rank	Score	Rank	Score	Rank	Score
Saturating Uniform, 30	11	0.113	12	0.33	12	0.426
Peaking Uniform, 30	10	0.128	10	0.457	1	0.605
Weighted Uniform, 30	12	0.107	11	0.419	6	0.498
Peaking, SoftMax 3 Stage	7	0.523	2	0.57	4	0.559
Saturating, SoftMax 3 Stage	8	0.491	5	0.539	7	0.486
Weighted, SoftMax 3 Stage	9	0.482	3	0.558	3	0.562
Saturating CRM, Fully Continual, Standard	3	0.768	4	0.549	10	0.431
Peaking CRM, Fully Continual, Standard	1	0.81	9	0.462	9	0.453
Weighted CRM, Fully Continual, Standard	2	0.772	6	0.528	11	0.43
Saturating CRM, Fully Continual, Balanced	4	0.617	8	0.492	8	0.464
Peaking CRM, Fully Continual, Balanced	5	0.602	1	0.582	2	0.566
Weighted CRM, Fully Continual, Balanced	6	0.587	7	0.514	5	0.518

Table S.14.1. Copeland metrics for scenario Saturating 1.

Supplementary References

1. Glass, E.J. Genetic Variation and Responses to Vaccines. *Anim. Health Res. Rev.* **2004**, *5*, 197–208, doi:10.1079/ahr200469.
2. O'Quigley, J.; Iasonos, A.; Bornkamp, B. *Handbook of Methods for Designing and Monitoring Dose Finding Trials*; 2019; ISBN 978-0-367-33068-2.
3. Symonds, M.R.E.; Moussalli, A. A Brief Guide to Model Selection, Multimodel Inference and Model Averaging in Behavioural Ecology Using Akaike's Information Criterion. *Behav. Ecol. Sociobiol.* **2011**, *65*, 13–21, doi:10.1007/s00265-010-1037-6.
4. van Hasselt, H.; Guez, A.; Silver, D. Deep Reinforcement Learning with Double Q-Learning. *ArXiv150906461 Cs* **2015**.
5. Hodges, J.L. The Significance Probability of the Smirnov Two-Sample Test. Ark. För Mat. 1958, 3, 469–486, doi:10.1007/BF02589501.

6. Mann, H.B.; Whitney, D.R. On a Test of Whether One of Two Random Variables Is Stochastically Larger than the Other. Ann. Math. Stat. 1947, 18, 50–60, doi:10.1214/aoms/1177730491.
7. Weisstein, E.W. Bonferroni Correction Available online: https://mathworld.wolfram.com/ (accessed on 29 April 2022).

Chapter 6: Evaluation of a novel non-parametric modelling approach to optimisation of vaccine dose: Paper 5

Chapter 6 Introduction

For chapter 6, I designed a novel vaccine dose-optimisation approach which I called the 'Correlated Beta' dose optimisation approach. This used a non-parametric model to describe vaccine dose-efficacy and dose-response in an adaptive trial design setting. I conducted a second simulation study to evaluate how well it locates optimal vaccine dose and maximises benefit to trial participants. This chapter addresses objective 5 of this thesis.

In chapters 3 through 5 of this thesis, I have discussed that it may not be reasonable to assume a specific curve shape for vaccine dose-efficacy. Similarly, there is not, at present, a well validated model for dose-efficacy for prime/boost administration vaccines. Whilst in chapter 5, I discussed the potential use of weighted model averaging, the field of model-based drug development has previously suggested that models which do not assume a specific parametric form for dose-efficacy should be considered when does-efficacy relationships cannot be assumed a-priori. These are sometimes called 'non-parametric' or 'curve-free' models. Whilst these have so far only been investigated for single-administration drug dose optimisation, I hypothesised that these may also be effective for modelling prime/boost vaccine dose-response.

The investigations of Takahashi on the non-parametric 'Latent Gaussian Process' model for dose optimisation showed that it was comparable to dose-optimisation approaches that used parametric models [75]. However, Goetschalckx et al. suggested that 'Continuous Correlated Beta Process' (CCBP) models may be simpler, more computationally efficient, and more interpretable than Latent Gaussian Process models [195]. CCBP models have also been shown to be effective for modelling the probability of binary outcomes in multi-dimensional spaces, suggesting that CCBP models could be used for modelling prime/boost vaccine dose-efficacy. It

was also suggested that CCBP models should be investigated in the context of solving multi-armed bandit problems, suggesting that these models might be effective for model based adaptive trial design/continual modelling.

I hypothesised that the CCBP model could be effectively used to model binary outcome vaccine dose-response relationships, both for single-administration and multiple-administration vaccines. I hypothesised that a 'Correlated Beta' dose-optimisation approach combining this CCBP model with the Thompson Sampling algorithm as a method of trial dose selection would be an effective approach for selecting optimal vaccine dose which could be used without requiring the assumption of specific dose-response curve shapes. I hypothesised that this would be true regardless of whether optimal dose was defined by the maximally efficacious dose or by a utility function of efficacy and toxicity probabilities. In this paper I investigated these hypotheses, evaluating the correlated beta dose-optimisation approach against other dose-optimisation approaches that used a parametric model for dose-efficacy or dose-toxicity. I also decided to evaluate dose-optimisation approaches which did not use mathematical modelling, as these represent more common approaches for dose-optimisation of vaccines.

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed <u>for each</u> research paper included within a thesis.

SECTION A – Student Details

Student ID Number	lsh1804914	**Title**	Mr
First Name(s)	John Helier		
Surname/Family Name	Benest		
Thesis Title	The Correlated Beta Dose Optimisation Approach: Optimal vaccine dose selection using mathematical modelling and adaptive trial design		
Primary Supervisor	Richard G. White		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?			
When was the work published?			
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	Choose an item.	Was the work subject to academic peer review?	Choose an item.

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	MDPI Vaccines
Please list the paper's authors in the intended authorship order:	John Benest, Sophie Rhodes, Thomas G. Evans, Richard G. White

Stage of publication	**Submitted**

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	The conceptualisation of this work was my own. I also was the sole individual developing codebase that was used for the dose-response modelling and simulation of clinical trials. All analysis of the data that was generated by these simulations were also conducted by me. I also performed all visualisation. I wrote the first draft of this work and performed re-drafting in line with co-author comments. All authors reviewed the paper. The interpretation of the results was my own work

SECTION E

Student Signature	
Date	11 August 2022

Supervisor Signature	
Date	12 August 2022

224

Paper 5 Title: The Correlated Beta Dose Optimisation Approach: Optimal vaccine dose selection using mathematical modelling and adaptive trial design

Authors

John Benest, Sophie Rhodes, Thomas G. Evans, and Richard G. White

Abstract:

Mathematical modelling methods and adaptive trial design are likely to be effective for optimising vaccine dose but are not yet commonly used. This may be due to uncertainty with regards to the correct choice of parametric model for dose-efficacy or dose-toxicity. Non-parametric models have previously been suggested to be potentially useful in this situation. We propose a novel approach for locating optimal vaccine dose based on the non-parametric Continuous Correlated Beta Process model and adaptive trial design. We call this the 'Correlated Beta' or 'CoBe' dose optimisation approach. We evaluated the CoBe dose optimisation approach compared to other vaccine dose optimisation approaches using a simulation study. Despite using simpler assumptions than other modelling-based methods, we found that the CoBe dose optimisation approach was able to effectively locate the maximum efficacy dose for both single and prime/boost administration vaccines. The CoBe dose optimisation approach was also effective in finding a dose that maximises vaccine efficacy and minimises vaccine-related toxicity. Further, we found that these modelling methods can benefit from the inclusion of expert knowledge, which has been difficult for previous parametric modelling methods. This work further shows that using mathematical modelling and adaptive trial design is likely to be beneficial to locating optimal vaccine dose, ensuring maximum vaccine benefit and disease burden reduction, ultimately saving lives

1. Introduction

Vaccines are an effective tool in global disease burden reduction. The amount of vaccine given to an individual (the 'dose') is a key decision in vaccine development to ensure an effective vaccine campaign. Dose can affect the efficacy, toxicity and cost associated with vaccine rollout [1–3]. However, selecting optimal dose ('dose optimisation') is non-trivial [4–6]. Vaccine dose-ranging trials are typically small (<100 individuals) [7–10], limiting the amount of data that can be used for dose decision making. In addition, vaccine dose-ranging clinical trials need to be conducted such that not only are useful data gathered, but also such that the interests and safety of the trial participants are respected [11].

In order to select optimal vaccine dose within the constraints of small trial sizes and ethical trial design, mathematical modelling and adaptive clinical trial design have been suggested. Previous work into mathematical modelling has shown promise for accelerating and improving dose decision making in vaccine development [2,12,13]. Whilst making dosing decisions based on modelling is common in drug development, these methodologies are not yet utilised to the same extent within vaccine development [12,14,15]. Further, adaptive trial design has also been suggested to be effective for the purpose of selecting optimal doses [16–18]. Here modelling or statistical analysis is conducted at interim time points to maximise the proportion of trial participants that receive near optimal doses. Adaptive design may lead to more optimal dose selection and more ethical clinical trials.

Vaccine dose-response mathematical models are systems of equations that are used to describe the relationship between vaccine dose and vaccine response. This requires making assumptions regarding which models can accurately describe vaccine dose-response. Previous work has shown that for some vaccines an increase in dose leads to increased efficacy responses, but that for other vaccines there is a maximum efficacy dose after which an increased dose leads to decreased vaccine efficacy [13,19,20].This means that there may be uncertainty in the correct models to use. Selecting optimal vaccine dose using models which are 'misspecified', meaning they are not appropriate for describing the dose-response

relationship for the purposes of selecting optimal dose, could lead to suboptimal vaccine dosing, decreasing efficacy or increasing toxicity [21–23].

We have previously discussed the use of model averaging to account for this uncertainty [21]. Alternatively, others have suggested that non-parametric models can be effective for locating optimal dose in the case of model uncertainty [24–26].Whilst the assumption of parametric models is that vaccine dose-response follows some pre-specified equation/shape, non-parametric models do not assume a predefined equation/shape.

One type of non-parametric model is the Continuous Correlated Beta Process (CCBP) model [27]. This is a form of non-parametric mathematical model that has previously been discussed for automated stroke rehabilitation and modelling of genetic ancestry [28,29]. CCBP models have the properties of being simple to implement, interpret and update based on available data, and do not require the assumption of a specific dose-response shape. The modelling assumption is instead that "similar" doses will cause "similar" responses. We hypothesised that the application of CCBP models in an adaptive trial design setting may be an effective approach for conducting clinical trials to select optimal vaccine dose. We call this Correlated Beta (CoBe) dose optimisation.

We evaluated this novel dose-optimisation approach in potential application to four potential open topics in mathematical modelling for optimal vaccine dose selection. Firstly, selection of a maximally efficacious vaccine dose given uncertainty in dose-efficacy curve shape. Secondly, how to locate the maximally efficacious doses for prime-boost paradigm vaccines. Thirdly, optimal vaccine dose selection that includes multiple objectives, such as both maximising efficacy and minimising toxicity. Fourthly, how can expert knowledge be incorporated into vaccine dose modelling.

In this work we aimed to use simulation of dose-finding clinical trials to assess the use of the 'Correlated Beta dose optimisation approach' in selecting optimal vaccine dose. To answer the questions posed above, we investigated the CoBe dose optimisation approach relative to three other dose optimisation approaches (DOAs).

- A 'Parametric' DOA that used parametric modelling and adaptive trial design
- An 'Adaptive Naive' DOA that used adaptive trial design but not modelling.
- A 'Uniform Naive' DOA that used neither adaptive trial design nor modelling.

To perform this analysis, we simulated a large number of clinical trials for a large number of qualitatively different 'scenarios', each representing different 'true' dose-efficacy or dose-efficacy and dose-toxicity relationships. We considered not only the quality of the final selected dose but also the benefit to clinical trial participants for all four DOAs for clinical trials with between 6-300 total trial participants.

Specifically, to address to above questions, our objectives were to:

1. Evaluate the Correlated Beta Dose Optimisation Approach for optimising vaccine efficacy for a single dose administration.
2. Evaluate the Correlated Beta Dose Optimisation Approach for optimising vaccine efficacy for a prime-dose/boost-dose administration.
3. Evaluate the Correlated Beta Dose Optimisation Approach for optimising vaccine utility, maximising efficacy, and minimising toxicity.

We also include a fourth objective which considered only the CoBe DOA

4. Evaluate the use of expert knowledge informed Continuous Correlated Beta Process priors for vaccine dose-optimisation.

2. Materials and Methods

In very high-level summary, we used a simulation study methodology [30–33] to evaluate the novel Correlated Beta (CoBe) dose optimisation approach with regards to several open topics in vaccine dose optimisation and provide a comparative evaluation relative to other potential dose-optimisation approaches that could be used to select optimal vaccine dose. This work is summarised in figure 1.
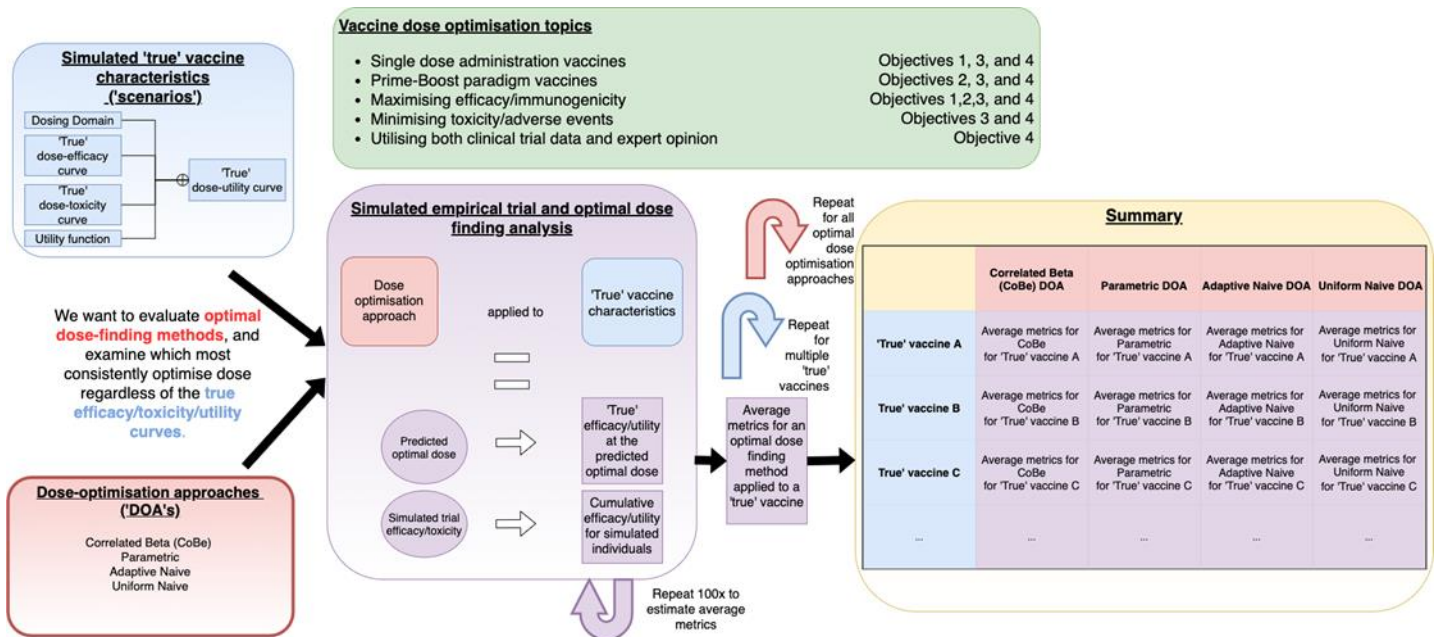
Figure 1. A visual depiction of the process of conducting simulation studies used in this work to evaluate the Correlated Beta (CoBe) and other potential approaches of vaccine dose optimisation (red). These were tested by simulating clinical trials (purple) based on 'scenarios' (blue). Repeated simulation of clinical trials was conducted for different dose-optimisation approach/scenario pairs, and metrics related to how effectively optimal dose was located were calculated. These were tabulated and compared to assess these approaches. These analyses were used with considerations of several open topics in vaccine dose-optimisation (green).

This methods section is split into four sections. In section 1, we defined the concept of 'optimal vaccine dose' and of 'dose optimisation approaches' (DOAs). In section 2, we defined and described the Correlated Beta (CoBe) DOA that was the focus of this work, along with the three other DOAs that were investigated in this work in comparison to the CoBe DOA. In section 3, we describe the simulation study methodology that was used to evaluate and compare these DOAs. Section 3 also contains description of the metrics used to evaluate the DOAs with regards to their potential effectiveness for optimisation of vaccine dose and benefit to trial participants, and details of the simulation study that would be required to replicate this work.

229

Finally, in section 4 we describe how we investigated the four objectives of this work using the concepts and terminology developed in the previous three sections.

2.1. Section 1. Definition of the concepts of 'optimal vaccine dose' and 'dose-optimisation approaches'

2.1.1. Definition of 'optimal vaccine dose'

In this work optimal dose was defined as the dose that maximises some utility function $U(p_{eff}, p_{tox})$ where p_{eff} and p_{tox} are binary efficacy and toxicity probabilities that are dependent on vaccine dose.

Throughout this work we will consider only two utility functions, one that aims to maximise efficacy ('Maximum Efficacy') versus dose and literature informed utility function that balances maximising efficacy and minimising toxicity ('Utility Contour', as used in [34,35]) versus dose. Formally these are

Maximum Efficacy:

$$U(p_{eff}, p_{tox}) = p_{eff}, \qquad (1)$$

Utility Contour:

$$U(p_{eff}, p_{tox}) = 1 - \left(\left(\frac{1 - p_{eff}}{1 - anchor_{eff}}\right)^{rho} - \left(\frac{p_{tox}}{anchor_{tox}}\right)^{rho}\right)^{\frac{1}{rho}}, \qquad (2)$$

Where $anchor_{eff}$, $anchor_{tox}$ and rho are parameters defined by clinicians to weight the relative importance of efficacy to toxicity (see S1 and [35] for more detail). Optimal dose was constricted to the dosing domain.

2.1.1.1. Dosing domain

The possible doses that can be selected for testing or predicted as optimal was called the 'dosing domain'. Dosing domains are generally continuous in nature, though are often discretized to a finite number of possible doses for the purpose of optimisation and due to potential practical limitations [36]. We will only consider discretized dosing domains in this work.

Previous work has investigated mathematical modelling for the selection of optimal dose with regards to a single-administration vaccine [2,13,21]. In this work we would also like to consider optimising dose 'prime/boost' paradigm vaccines, which are vaccines that are administered as two or more doses at separate time points [37,38]. Here, doses in the dosing domain are possible combinations of possible doses for each prime or boost administration.

2.1.2. Definition of a 'dose-optimisation approach'

A dose-optimisation approach (DOA) is the combination of methods used to design clinical trials/choose the doses that trial participants will receive, along with the methods used to select 'optimal' dose based on the resulting data. We focus here on 'continual modelling' DOAs, where modelling is conducted at interim stages of the trial and used to guide selection of the next trial doses.

For this work, a DOA consists of

- A model for vaccine dose-efficacy and/or dose-toxicity.
- A method of trial dose selection: How doses are chosen during the trial.
- A method of final dose selection: How to choose the dose that would be continued forward to further research or clinical use.
- A choice of how to discretize the dosing domain: Whether there was a small or large number of doses that could be tested, further detail in 2.3.3.1. This was previously discussed by [36,39].

2.2.1. Model for Vaccine dose-efficacy and/or toxicity: Continuous Correlated Beta Processes

The CoBe DOA uses Continuous Correlated Beta Process (CCBP) models [27] to model vaccine dose-efficacy/toxicity. These are not only simple to implement but can be extended to prime/boost dose-response problems (Objective 2) or extended to include expert prior predictions (Objective 4). In this section we discuss the intuition and implementation of Continuous Correlated Beta Processes (CCBP).

In contrast to parametric models, which assume some curve shape can describe vaccine dose-efficacy or dose-toxicity, the CCBP models defined here do not assume a specific shape, and instead make a simpler assumption; 'similar doses yield similar responses'. CCBP have been described previously in detail in the context of modelling for automatic stroke rehabilitation [27–29]. The two main elements of a CCBP model are Beta distributions and correlation kernel functions.

2.2.1.1. Beta distributions

Beta distributions describe a probability distribution of probabilities for binary outcomes [40]. Suppose that we would like to know the probability of some response (efficacy or toxicity in this work) being observed for a vaccine administered at a pre-chosen dose. We call this probability $p_{response}$, and we have no prior expectation for what the true value of $p_{response}$ is. Suppose that after a trial of 1 individual we have observed 1 responder (and hence 0 non-responders). The maximum likelihood estimate of $p_{response}$ given by these data would be $p_{response} = 1.0$. However, $p_{response} = 0.9$ would also intuitively be a reasonable guess and $p_{response} = 0.1$ would be much less probable [figure 2]. Beta distributions allow for a formalised description of the probability of a certain probability of response given the observed data.
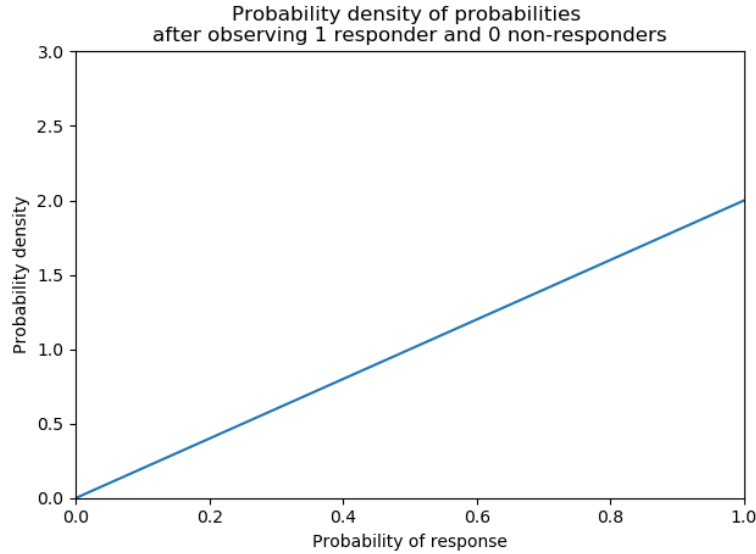
Figure 2. Probability density of probabilities of response after observing one responder and zero non-responders with no prior knowledge regarding probability of response. The higher the probability density is for a given probability of response, the more likely that it is the true probability of response given the data. The above is formally a Beta(2,1) distribution.

A beta distribution is defined by two parameters, α and β . We write

$$p_{i,r}^{n} \sim Beta(\alpha_{i,r}^{n}, \beta_{i,r}^{n}), \qquad (3)$$

to say that the probability of observing response r for some dose d_i based on the first n data points can can be described by a beta distribution with parameters $\alpha_{i,r}^{n}$ and $\beta_{i,r}^{n}$. Increasing $\alpha_{i,r}^{n}$ shifts the beta distribution towards higher $p_{i,r}^{n}$ and increasing $\beta_{i,r}^{n}$ shifts the beta distribution towards lower $p_{i,r}^{n}$. Increasing either of these parameters reduces the confidence intervals of the probability distribution. See figure 3 for a visualisation of this. In this work, response r can be efficacy or toxicity.
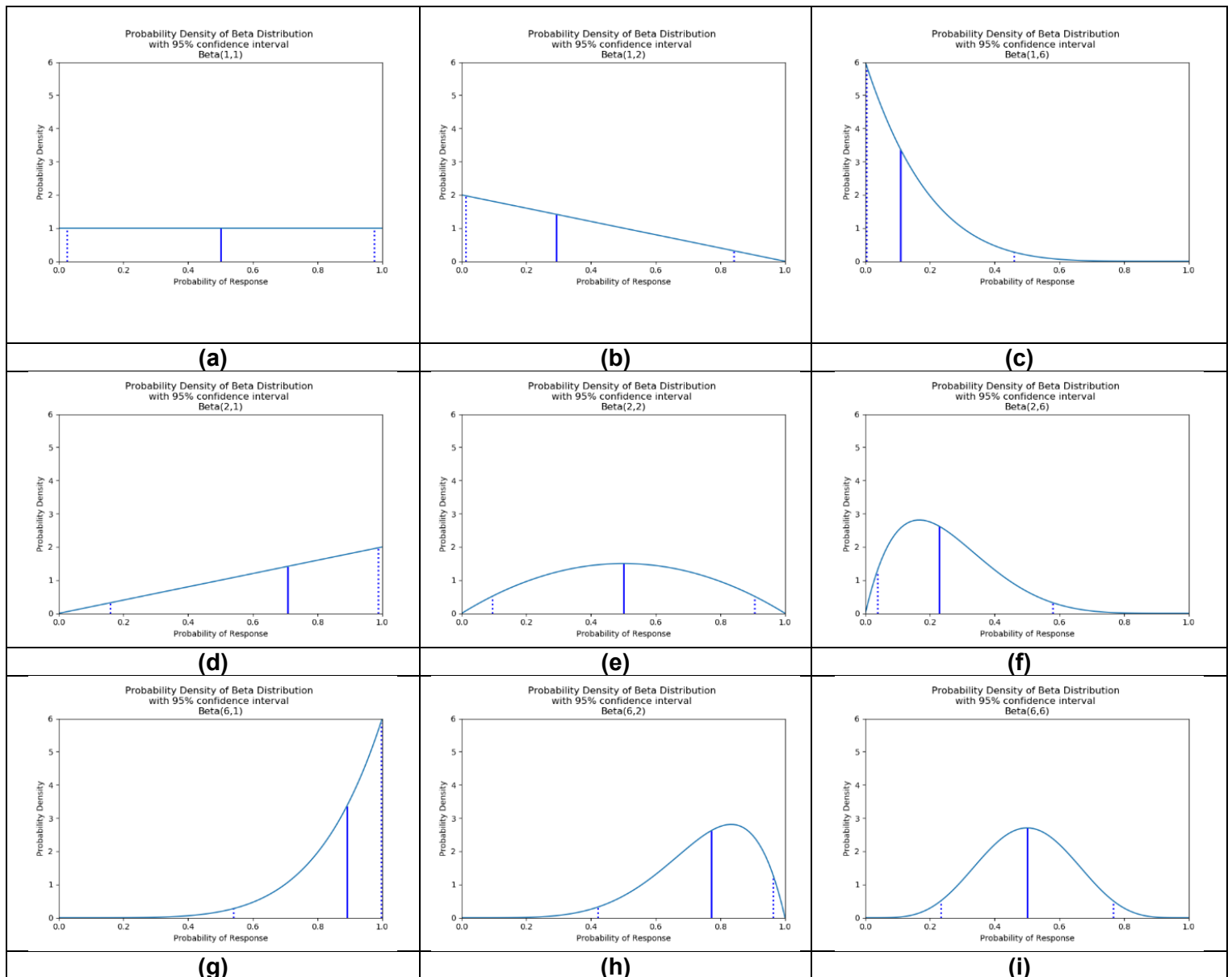
Figure 3. Probability density for beta distributions $Beta(\alpha, \beta)$ with differing values of α and β.
The vertical blue line represents the median probability of response, and the dashed vertical lines represent the 95% confidence interval. If using an uninformative prior (a), these represent the probability density after observing: (a) 0 responders and 0 non-responders, (b) 0 responders and 1 non-responders, (c) 0 responders and 5 non-responders, (d) 1 responders and 0 non-responders, (e) 1 responder and 1 non-responder, (f) 1 responder and 5 non-responders, (g) 5 responders and 0 non-responders, (h) 5 responders and 1 non-responder and (i) 5 responders and 5 non-responders.

2.2.1.1.1 Updating Beta Distributions

As we aim to run multiple trials over time, we can use the data gathered to update our beta distributions to give us a better idea of optimal dose. Algorithm 1 shows the update rule for updating the α and β parameters of beta distributions after observing data.

234

Note that this update rule means our understanding of the probability of some response for a given dose is only improved when we test at exactly that dose. Therefore, these are *uncorrelated beta distributions.*

Algorithm 1. Update rule for uncorrelated Beta distributions

This rule is for updating the beta distribution for the probability of observing response r for some dose d_i based on the $(n+1)th$ data point. Let this $(n+1)th$ data point had dose d_j.

BEGIN ALGORITHM

If $d_i = d_j$

    If response r was observed for individual $n+1$

        Set $\alpha_{i,r}^{n+1} = \alpha_{i,r}^{n} + 1$

        Set $\beta_{i,r}^{n+1} = \beta_{i,r}^{n}$

    Else (response r was not observed for individual $n+1$)

        Set $\alpha_{i,r}^{n+1} = \alpha_{i,r}^{n}$

        Set $\beta_{i,r}^{n+1} = \beta_{i,r}^{n} + 1$

Else ($d_i \neq d_j$)

        Set $\alpha_{i,r}^{n+1} = \alpha_{i,r}^{n}$

        Set $\beta_{i,r}^{n+1} = \beta_{i,r}^{n}$

END ALGORITHM

2.2.1.1.2. Priors and uninformative priors

Initial values of α and β must be chosen. Typically, if there is no prior knowledge for which response probabilities are most reasonable, it is best to use an uninformative prior. For this, the initial values of α and β for each dose d_i for response r are set to 1 [28]. That is

$$\alpha_{i,r}^{0} = 1 \tag{4}$$

$$\beta_{i,r}^{0} = 1 \tag{5}$$

This is typically a reasonable choice, as prior to data being collected this leads to equal probability for each possible value of $p_{i,r}^0$ [figure 2 a]. If there if is prior understanding about the probability of response r for dose d_i alternative values of $\alpha_{i,r}^0$ and $\beta_{i,r}^0$ can be used. A method for choosing these is discussed in 2.2.1.4. and the implications of this in conducted dose-finding trials investigated in objective 4.

2.2.1.2. Kernel Functions

Above we noted that uncorrelated beta distributions do not allow for information about response probabilities from one dose to inform understanding of response probability at any other dose. The CCBP model allows information about response probability for one dose to inform understanding of response probability for 'similar' doses. We describe what it means for doses to be 'similar' using a similarity function, $K(d_i, d_j)$, traditionally called a 'kernel' function. This is a function that takes two doses as input and returns a number between 0 and 1 that represents how similar those doses are.

In the context of vaccine dose optimisation, a kernel function $K(d_i, d_j)$ follows these rules for all doses d_i and d_j in the dosing domain:

$$0 \leq K(d_i, d_j) \leq 1 \tag{6}$$

$$K(d_i, d_j) = 1 \text{ if and only if the doses are completely similar} \tag{7}$$

$$K(d_i, d_j) = 0 \text{ if and only if the doses are completely dissimilar} \tag{8}$$

$$K(d_i, d_j) = K(d_j, d_i), \text{ so similarity is symmetrical} \tag{9}$$

$$K(d_i, d_i) = K(d_j, d_j) = 1, \text{ so a dose must be completely self-similar} \tag{10}$$

We can then use kernels to inform beta distributions for multiple 'similar' doses based on data gathered for a specific dose. Using a kernel function makes these beta processes 'correlated beta processes'.

The beta distribution update rule described in algorithm 1 is then changed to that showed in Algorithm 2.

Algorithm 2. (Continuous) Correlated Beta Process Update Rule

---

This rule is for updating the beta distribution for the probability of observing response r for some dose d_i based on the $(n + 1)th$ data point. Let this $(n + 1)th$ data point have been at dose d_j.

BEGIN ALGORITHM

Calculate $K(d_i, d_j)$

If response r was observed for individual $n + 1$

   Set $\alpha_{i,r}^{n+1} = \alpha_{i,r}^n + K(d_i, d_j)$

   Set $\beta_{i,r}^{n+1} = \beta_{i,r}^n$

Else (response r was not observed for individual $n + 1$)

   Set $\alpha_{i,r}^{n+1} = \alpha_{i,r}^n$

   Set $\beta_{i,r}^{n+1} = \beta_{i,r}^n + K(d_i, d_j)$

END ALGORITHM

An example of this update process is now given. Say doses, d_1 and d_2, have efficacy probabilities initially described by a flat prior, that is;

$$p_{1,eff}^0 \sim Beta(\alpha_{1,eff}^0, \beta_{1,eff}^0) = Beta(1,1), \tag{11}$$

$$p_{2,eff}^0 \sim Beta(\alpha_{2,eff}^0, \beta_{2,eff}^0) = Beta(1,1), \tag{12}$$

Say dose d_1 is tested and a positive efficacy response observed. If d_1 and d_2 are 50% similar (efficacy kernel $K(d_1, d_2) = K(d_2, d_1) = 0.5$), then

$$p^1_{1,eff} \sim Beta\left(\alpha^0_{1,eff} + K(d_1, d_1), \beta^0_{1,eff}\right) = Beta(1+1, 1) = Beta(2, 1), \quad (13)$$

$$p^1_{2,eff} \sim Beta\left(\alpha^0_{2,eff} + K(d_1, d_2), \beta^0_{2,eff}\right) = Beta(1 + 0.5, 1)$$
$$= Beta(1.5, 1), \quad (14)$$

In this work, we chose to use the squared exponential kernel suggested in [27,29] defined as

$$K(d_i, d_j) = e^{-\frac{(d_i - d_j)^2}{l^2}} \quad (15)$$

where l is a length hyperparameter that can be chosen to adjust the range for which doses are considered similar (examples in Figure 4). For small l, the data only influences model prediction near the tested dose [figure 4]. For larger l, the data influences model prediction at a greater distance. In this work length parameter l=0.2 for modelling single-administration vaccine dose-response.
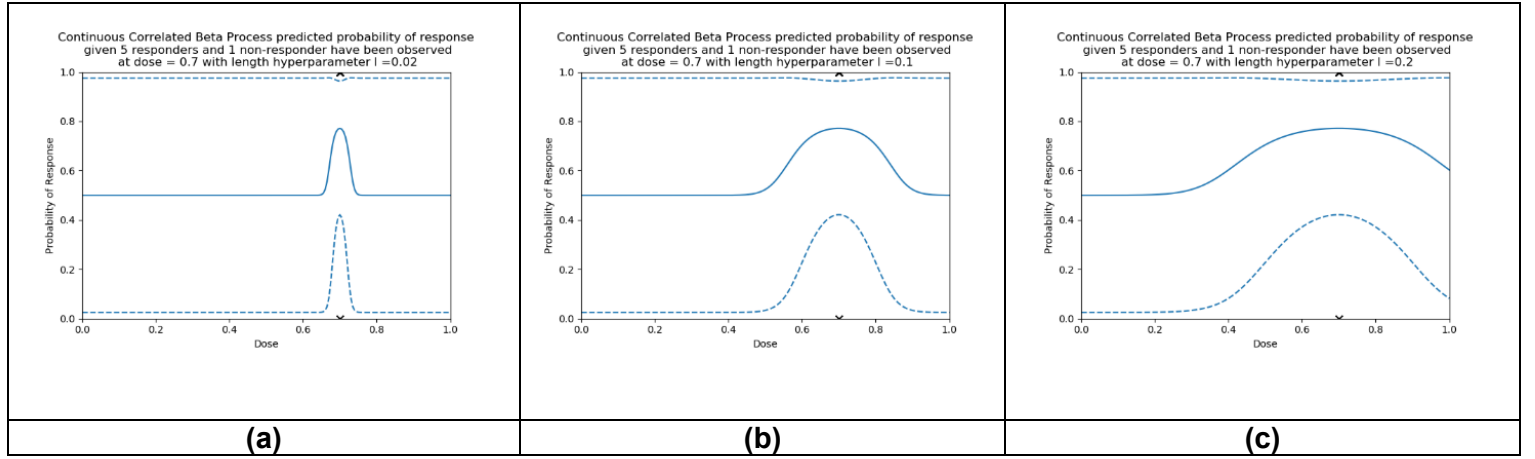
Figure 4. An example of three different CCBP models with squared exponential kernels for different length hyperparameters, using the flat prior for all doses, and 6 observed responses at dose = 0.7 (5 responders, 1 non-responder, as per figure 1h). Length parameters are (a) 0.02, (b) 0.1 and (c) 0.2. The solid line is the median predicted probability, and the dashed lines show the 95% confidence interval.

This choice of kernel is means similarity is defined continuously for any two doses in the dosing domain, making these beta processes *continuous correlated beta processes* (CCBPs).

2.2.1.3. Modelling prime/boost dose response

Extending Continuous Correlated Beta Process (CCBP) models to modelling prime/boost dose response requires only a change to the kernel function. Since a squared exponential kernel was chosen, this change is intuitive. For doses d_i and d_j, where dose d_i has prime dose d_{i1} and boost dose d_{i2} the 2-dose kernel function would be

$$K^2\big(d_i, d_j\big) = e^{-\frac{(d_{i1}-d_{j1})^2}{l_1^2}-\frac{(d_{i2}-d_{j2})^2}{l_2^2}} \tag{16}$$

Similarly, for modelling prime/boost/second-boost dose response the 3-dose kernel function would be

$$K^3\big(d_i, d_j\big) = e^{-\frac{(d_{i1}-d_{j1})^2}{l_1^2}-\frac{(d_{i2}-d_{j2})^2}{l_2^2}-\frac{(d_{i3}-d_{j3})^2}{l_3^2}} \tag{17}$$

239

This pattern can be generalised to considering to considering H doses

$$K^H(d_i, d_j) = e^{-\sum_{o=1}^{H} \frac{(d_{io} - d_{jo})^2}{l_o^2}} \tag{18}$$

In this work the length parameters $l = 0.2$ was used for modelling single-administration dose-response, $l_1 = l_2 = 0.25$ were used for modelling prime/boost dose-response, and length parameters $l_1 = l_2 = l_3 = 0.4$ were used for modelling prime/boost/second-boost dose-response [S2]. Length parameters do not need to be equal but were equal here for simplicity.

2.2.1.4. Utilising expert knowledge to inform Continuous Correlated Beta Process model priors

It is possible that including expert knowledge into the modelling process may improve optimality of the final selected dose, leading to more effective early trial doses. Methods for including expert knowledge to inform the modelling process for parametric models have been previously discussed [41,42] but are non-trivial. Expert knowledge is integrated into the CCBP model by choosing different initial values for $\alpha_{i,r}^0$ and $\beta_{i,r}^0$ which we call the expert informed prior. For each dose d_i and response r, an expert informed prior can be defined using the expert's prediction of the most likely probability of response for that dose, $p_{i,r}^{expert}$, and level of confidence in that probability, $c_{i,r}^{expert} \geq 0$. These values could be based on previous knowledge of the vaccine or a similar product. $c_{i,r}^{expert}$ can be considered as the number of individuals worth of data that is required before the data influences the model prediction more than the expert knowledge. Incorporating expert priors in an initial Beta distribution for dose d_i for response r is done by setting

$$\alpha_{i,r}^0 = p_{i,r}^{expert} \times c_{i,r}^{expert} + 1, \tag{19}$$

$$\beta_{i,r}^0 = (1 - p_{i,r}^{expert}) \times c_{i,r}^{expert} + 1 \tag{20}$$

Then the mode of the relevant Beta distribution for each d_i will be $p_{i,r}^{expert}$ [43].

2.2.2. Method of trial dose selection

The method for trial dose selection in the CoBe DOA is Thompson sampling. Thompson Sampling involves choosing clinical trial doses proportionally to the probability that they are optimal, given the available data and model. This is described in detail in [44–47], but the principle is to sample from Beta distributions for each dose, and then select the optimal dose based on the value of the utility function for each sample.

As a continuation of our earlier example, doses d_1 and d_2, had efficacy probabilities described respectively as

$$p_{1,eff}^1 \sim Beta(2, 1), \tag{21}$$

$$p_{2,eff}^1 \sim Beta(1.5, 1), \tag{22}$$

and we are using the maximum efficacy utility function. We can randomly sample efficacy probabilities $\hat{p}_{1,eff}$ and $\hat{p}_{2,eff}$ from these Beta distributions using statistical software. Then the values of the utility function for d_1 and d_2 based on these samples are $\hat{U}_1 = \hat{p}_{1,eff}$ and $\hat{U}_2 = \hat{p}_{2,eff}$. If $\hat{U}_1 > \hat{U}_2$ then d_1 is selected as the next dose to test, otherwise we select d_2. This process of sampling and selecting can be repeated to select as many trial doses as required for each sampling cohort (6 individuals per sampling cohort was used in this work).

2.2.3. Method of final dose selection

The method of final dose selection is to select the dose with maximum utility as given by the median probability prediction of response for each dose. For each dose d_i, the median probability of efficacy $\bar{p}_{i,eff}$ and $\bar{p}_{i,tox}$ are calculated. Then a dose d_i is

predicted as optimal if $U(\bar{p}_{i,eff}, \bar{p}_{i,tox}) \geq U(\bar{p}_{j,eff}, \bar{p}_{j,tox})$ for all doses dj, with ties broken randomly.

2.2.4. Discretisation

Due to the continuous nature of the CCBP model, the CoBe dose optimisation approach can be applied when choosing between a potentially large number of doses, therefore the dosing domain can be discretised to include many doses.

2.2.5. Full Correlated Beta (CoBe) dose optimisation approach and an example trial

The complete CoBe dose optimisation approach is shown in Algorithm 3, and three time-points of an example simulated clinical trial is depicted in figure 5.

Algorithm 3. Correlated Beta (CoBe) Dose Optimisation Algorithm

---

BEGIN ALGORITHM

1. Initialisation:

    a. Choose in collaboration with clinicians and experts

        i. Total trial participants available, N

        ii. Sampling cohort size, c (=6 in this work)

        iii. Determine whether a single-administration, prime/boost, or prime/boost/second-boost paradigm is being used.

        iv. Determine all potential doses, d_i, in the discretized dosing domain, see [Discretization]).

        v. Choose length parameter(s) for the efficacy similarity kernel ($l = 0.2$, $l_1 = l_2 = 0.25$, $l_1 = l_2 = l_3 = 0.4$ in this work)

        vi. Choose length parameter(s) for the toxicity similarity kernel (the same as for efficacy in this work)

        vii. Query experts to determine any potential priors.

2. Initialization of Beta distributions - *in silico*

    a. Initialise description of efficacy probability distribution for each dose d_i as $p_{i,eff}^0 \sim Beta(\alpha_{i,eff}^0, \beta_{i,eff}^0)$

    b. Initialise description of toxicity probability distribution for each dose d_i as $p_{i,tox}^0 \sim Beta(\alpha_{i,tox}^0, \beta_{i,tox}^0)$

3. Thompson sampling for dose selection - *in silico*

    a. For each dose d_i, sample $\hat{p}_{1,eff}$ and $\hat{p}_{2,eff}$ from the relevant Beta distributions.

---

b. Select for trialing dose d_i such that $\hat{U}_i > \hat{U}_j$ for all doses d_j, where $U_i(p_{i,eff}, p_{i,tox})$ is the utility function to be maximised.

4. Repeat step 3 until sampling cohort is full (c repeats total)

5. Trialing and data collection – *practical*

    a. Conduct a trial of c individuals respectively at the c doses chosen in steps 3 and 4. This was simulated in this work but would be practical lab work in real life application.

    b. Record c data points consisting of {dose given, whether efficacy was observed, whether toxicity was observed}

6. Model Updating - *in silico*

    a. Update $\alpha_{i,eff}^{n-c}, \beta_{i,eff}^{n-c}, \alpha_{i,tox}^{n-c}, \beta_{i,tox}^{n-c}$ to $\alpha_{i,eff}^{n}, \beta_{i,eff}^{n}, \alpha_{i,tox}^{n}, \beta_{i,tox}^{n}$ using:

        i. Update $\alpha_{i,eff}^{n-c}, \beta_{i,eff}^{n-c}, \alpha_{i,tox}^{n-c}, \beta_{i,tox}^{n-c}$ to $\alpha_{i,eff}^{n-c+1}, \beta_{i,eff}^{n-c+1}, \alpha_{i,tox}^{n-c+1}, \beta_{i,tox}^{n-c+1}$ using Algorithm 2 with a data point gathered in step 5.

        ii. Repeat for all other data points gathered in step 5 (order does not matter)

7. Prediction of optimal dose – *in silico*

    a. For each dose di, calculate the median response probabilities $\bar{p}_{i,eff}$ and $\bar{p}_{i,tox}$

    b. The predicted optimal dose is d_i such that $U(\bar{p}_{i,eff}, \bar{p}_{i,tox}) \geq U(\bar{p}_{j,eff}, \bar{p}_{j,tox})$ where $U_i(p_{i,eff}, p_{i,tox})$ is the utility function to be maximised.

8. Repeat steps 3-7 until all N trial participants have been utilised.

END ALGORITHM

(a)

(b)

(c)
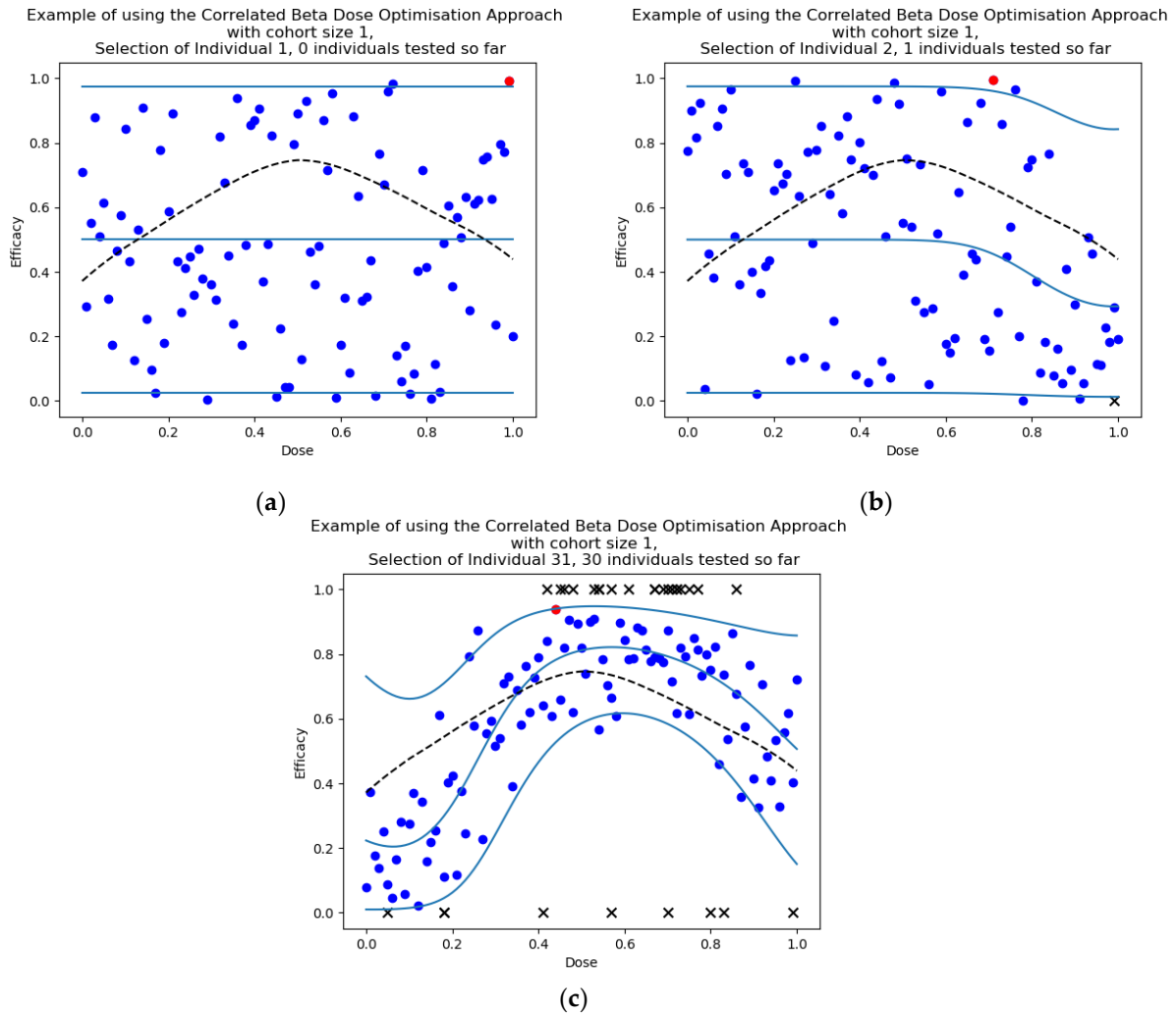
Figure 5. Three timepoints of an example dose-finding study using the CoBe DOA for selection of optimal dose (sampling cohort size c = 1) with the 'Efficacy Maximisation' utility function. At each time point, the dotted black line represents the true underlying dose-efficacy curve, and the black crosses represent observed data by that time point. The light blue lines represent the median and 95%CI for the predictions of the CCBP model, the blue dots represent efficacy prediction samples for each dose, the red dot represents the maximum of such samples which corresponds to the dose that would be selected to be tested next. At selection of individual 1 (a) efficacy samples varied uniformly between 0 and 1. After the first individual received a high dose and efficacy was not observed, (b) the median, 95%CI and samples for similar doses were lower when selecting the dose for individual 2. After the 30th individual (c), most samples had been selected near the true optimal and the model approximated the true curve (particularly near the predicted optimal dose).

2.2.6. Other dose-optimisation approaches

We consider multiple other dose-optimisation approaches other than the CoBe Dosing approach. These were the 'Parametric', 'Adaptive Naive', and 'Uniform Naive' DOAs.

244

2.2.6.1. Parametric dose-optimisation approach

The 'Parametric' DOA uses parametric models to describe dose-efficacy and dose-toxicity, as described in [21,35]. Specifically, we used the latent quadratic model [35,48] for modelling dose-efficacy for single-administration dose-optimisation problems. This is given by

$$p_{eff}(d_i) = latentquadratic(d_i) = logit(a + bd_i - cd_i^2), \tag{23}$$

With

$$logit(z) = \frac{1}{1 + e^{-z}}, \tag{24}$$

And parameters a, b, c.

We extended this model for prime/boost and prime/boost/second-boost dose-optimisation problems, respectively as

$$\begin{aligned} p_{eff}(d_i) &= latentquadratic_{2D}(d_i) \\ &= logit(a + b_1 d_{i1} - c_1 d_{i1}^2 + b_2 d_{i2} - c_2 d_{i1}^2), \end{aligned} \tag{25}$$

$$\begin{aligned} p_{eff}(d_i) &= latentquadratic_{3D}(d_i) \\ &= logit(a + b_1 d_{i1} - c_1 d_{i1}^2 + b_2 d_{i2} - c_2 d_{i1}^2 + b_3 d_{i3} \\ &\quad - c_3 d_{i3}^2), \end{aligned} \tag{26}$$

With parameters $a, b_1, c_1, b_2, c_2, b_3, c_3$ and dose d_i having prime dose d_{i1}, boost dose d_{i2}, and potential second-boost dose d_{i3}. This is similar to the approach used in [49] but extended to allow for non-monotonicity in the dose-efficacy relationships.

We used the latent linear model [35,50] for modelling dose-toxicity for single-administration dose-optimisation problems. This is given by

$$p_{tox}(d_i) = latentlinear(d_i) = logit(a + bd_i), \qquad (27)$$

with parameters a, b

We similarly extended this model for prime/boost dose-optimisation problems, respectively as

$$p_{tox}(d_i) = latentlinear_{2D}(d_i) = logit(a + b_1 d_{i1} + b_2 d_{i2}), \qquad (28)$$

With parameters a, b_1, b_2,

These models can be calibrated to the available dose-response data by determining the maximum likelihood estimate of the parameters given the available data. Pseudo-data were used to aid stability of the model calibration, as described in [[30], S3] These calibrated models were then used to predict dose-utility. The method of trial dose selection for each cohort was the SoftMax selection method described in [S4, [21,51,52]]. The method of final dose selection was to choose the dose with the maximum utility according to the predictions of the calibrated model. Due to being a modelling method, for this DOA the discretized dosing domain could include a large number of potential doses.

2.2.6.2. Adaptive Naive dose-optimisation approach

The 'Adaptive Naive' DOA has been well discussed in the past for conducting trials comparing treatments and doses [45,46]. Like the CoBe DOA, the probability of efficacy/toxicity for each potential dose is described by a beta distribution, the method of trial dose selection is Thompson sampling, and the method of final dose selection is the maximised median prediction. Unlike the CoBe DOA, however, this DOA does not make use of a similarity kernel or other modelling methods, so prediction of efficacy/toxicity for any given dose is determined by only considering previous data for that specific dose. Hence this DOA is 'adaptive', but the predictions of efficacy/toxicity for any given dose are 'naive'. Due to this lack of modelling, this DOA discretizes the dosing domain to only a small number of doses.

2.2.6.3. Uniform Naive dose-optimisation approach

The 'Uniform Naive' DOA is perhaps the most common DOA used for selecting optimal vaccine dose, though it is typically not named as such. This is the same as the Adaptive Naive DOA, except that the method of trial dose selection is to divide all clinical trial participants evenly amongst the discretized dosing domain. Commonly all sampling cohorts would be conducted at the same time given that there is no adaptive design. There is no adaptive design or modelling, so this DOA is 'uniform' in its method of trial dose selection and 'naive' in its predictions of efficacy/toxicity for each dose. Again, like the Adaptive Naive DOA, this DOA discretizes the dosing domain to only a small number of doses.

2.3. Section 3. Definition of the simulation study methodology and details of the implementation of this methodology

2.3.1. Definition of a simulation study

When conducting real life dose-finding studies we have the capacity to generate data through practical experiments. Trial individuals can be given vaccine doses, immunological/toxicity data responses recorded, and then these data can be used to make decisions regarding continued trial dose according to the DOA that is being used. In simulation studies we mimic this process, simulating clinical trial data generation according to 'true' vaccine dose-efficacy/dose-toxicity curves that we have defined and are hence known. We can define various different 'true' underlying dose-response curves to define different 'scenarios', which in turn allow the theoretical capacity of effective dose optimisation to be evaluated.

2.3.2. Definition of a scenario

In this work, a scenario consisted of:

- A dosing domain: Whether these scenarios consider a single dose or combinations of doses, and the range for which possible doses that could be tested or predicted as optimal, as described above. For simplicity, we considered that doses of vaccine (whether single administration or prime dose, or a boost dose) to have been scaled to be between 0 and 1, as

described in [[53–55], S5]. Thus, a zero dose does not necessarily correspond to no vaccine being given, but instead corresponds to the smallest dose that clinicians/developers may be willing to consider. This scaling was purely for convenience.

- A utility function: To weigh the relative benefit of efficacy, toxicity, or any other dose related outcome a utility function is needed. For this work we use either the 'maximum efficacy' or 'utility contour' utility functions defined in 2.1.1.

- Efficacy probabilities for all possible doses: For each dose in the dosing domain, there was some true probability of efficacy for each dose that was defined for the scenario.

- Toxicity probabilities for all possible doses: If our aim is to minimise toxicity as well as to maximise efficacy, as in the 'utility contour' utility function, there was some true probability of toxicity for each dose in the dosing domain that was defined for the scenario.

For details of scenario creation see S6.

2.3.3. Simulation study parameters

2.3.3.1. Discretisation

Specifically in this work,

- For all scenarios involving single-administration paradigm vaccine dose-response, for the CoBe and Parametric DOAs we discretized the dosing domain to 101 doses (0.00, 0.01, 0.02, ..., 0.99, 1.00) and for the Adaptive Naive and Uniform Naive DOAs we discretised the dosing domain to 6 doses (0.0, 0.2, 0.4, 0.6, 0.8, 1.0).

- For all scenarios involving prime/boost paradigm vaccine dose response, for the CoBe and Parametric DOAs we discretized the dosing domain to 411 doses (a 21-by-21 grid of (0.00, 0.05, ..., 0.95, 1.00)) and for the Adaptive Naive and Uniform Naive DOAs we discretised the dosing domain to 9 doses (a 3-by-3 grid of (0.0, 0.5, 1.0).

- For the scenario involving prime/boost/second-boost paradigm vaccine dose response, for the CoBe and Parametric DOAs we discretized the dosing

domain to 1331 doses (an 11-by-11-by-11 grid of (0.00, 0.10, ..., 0.90, 1.00)) and for the Adaptive Naive and Uniform Naive DOAs we discretised the dosing domain to 27 doses (a 3-by-3-by-3 grid of (0.0, 0.5, 1.0).

2.3.3.2. Trial size/sampling cohort size

As the number of individuals available for conducting a dose-ranging trial may vary in real life, we assessed the performance of the DOAs for trial sizes from 6 to 300 individuals. These are sizes reasonable given typical phase I/early phase II vaccine trial sizes [9].

Additionally, as we consider adaptive DOAs, we had to specify the size of the sampling cohort. This was the number of individuals that were tested in each round of modelling/trialling (CoBe Dose Optimisation Approach Algorithm box steps 4, 5). The CoBe, Parametric, and Adaptive Naive DOAs used a sampling cohort size of six in this work for all scenarios. The Uniform Naive DOA used a sampling cohort size equal to the number of doses in the discretised dosing domain (either 6, 9 or 27).

2.3.4. Metrics to evaluate dose-optimisation approaches

We used two metrics to evaluate the DOAs described in this work; one was related to optimal dose selection and one related to ethical trial design. Either of these may be considered to be important considerations for conducting vaccine dose-ranging trials.

- True efficacy/utility of predicted optimal dose: After each cycle of trial/modelling (each sampling cohort), each DOA can recommend a dose that is predicted optimal given the current data. As this was a simulation study, we were aware of the true efficacy/utility at that selected dose. This true efficacy/utility of the selected doses was averaged across trial simulations to assess the ability of a dose finding approach to locate optimal dose.
- Cumulative sum of efficacy/utility: Each individual in a trial may have an efficacious response and may experience vaccine-related toxicity. The cumulative number of efficacious responses (or cumulative utility if both

efficacy and toxicity are being optimised for) was averaged across simulations to assess the ability of a dose finding approach to maximise trial efficacy/utility.

The formula for 'cumulative sum of efficacy' after the first n individuals was

$$cumulative\ efficacy(n) = count_{efficacy}(n) \qquad (29)$$

and the formula for 'cumulative sum of utility' after the first n individuals was

$$cumulative\ utility(n) = n \times U(\frac{count_{efficacy}(n)}{n}, \frac{count_{toxicity}(n)}{n}) \quad (30)$$

With $U()$ being the utility contour function defined in section 1 and $count_{efficacy}(n)$ /$count_{toxicity}(n)$being the number of individuals that had experienced efficacy/toxicity in the first n individuals.

2.3.5. Implementation

Each scenario/approach pairing was simulated 100 times. The mean for both evaluation metrics for these 100 simulated clinical trials was calculated.

The simulation study was conducted in Python, using SciPy for statistical inference, for implementation of Beta distributions, and for calibration of the parametric models for the Parametric DOA [56].

2.4. Section 4. Description of the use of the concepts defined above in evaluating the Correlated Beta dose-optimisation approach in the context of our objectives

2.4.1. Objective 1. Evaluate the Correlated Beta Dose Optimisation Approach for optimising vaccine efficacy for a single dose administration.

In this objective we aimed to evaluate the vaccine dose-optimisation ability of the CoBe DOA when compared to other DOAs when choosing a single administration

dose that maximises efficacy. Using the definition of scenarios above, we consider scenarios with:

- Dosing domain: Single-administration
- Utility function: Maximise Efficacy
- Efficacy curve: Is defined for each scenario
- Toxicity curve: Not defined/not of interest

Seven scenarios were used to explore this objective (Figure 6). These scenarios reflect cases for which vaccine dose efficacy may be

1. gently saturating [Figure 6a]
2. sharply saturating [Figure 6b]
3. gently peaking [Figure 6c]
4. sharply peaking [Figure 6d]
5. decreasing [Figure 6e]
6. undulating [Figure 6f]
7. flattened peaking [Figure 6g]
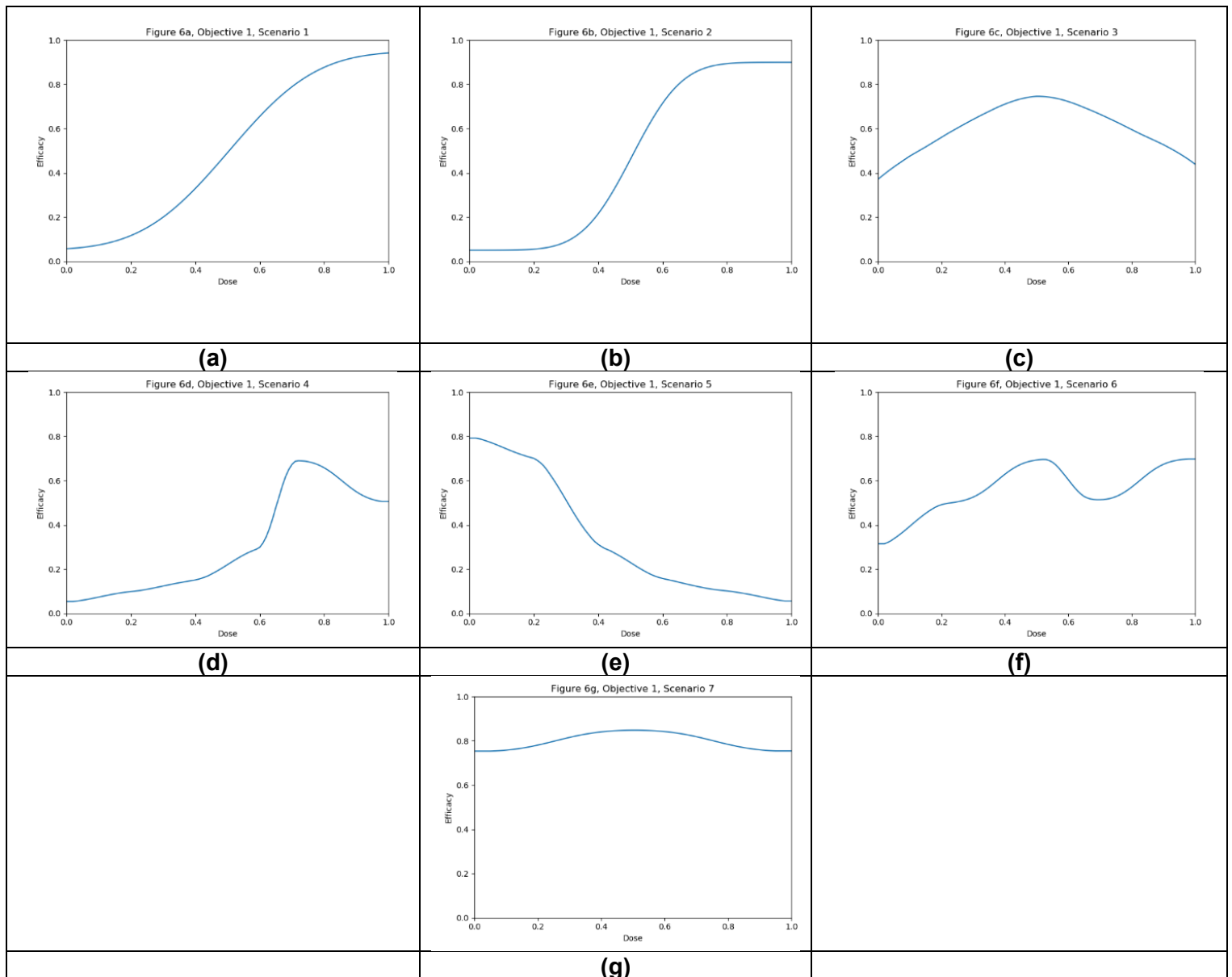
Figure 6. Dose-efficacy plots for the seven objective 1 scenarios. These were (a) scenario 1, (b) scenario 2, (c) scenario 3, (d) scenario 4, (e) scenario 5, (f) scenario 6, and (g) scenario 7 . Purple represents areas of higher efficacy. The blue line represents the probability of efficacious response for each dose.

For each scenario, we simulated 100 clinical trials conducted under each of the four DOAs (100x7x4 simulated trials total). Cohorts were of size 6, and we simulated 50 cohorts for each clinical trial (300 total individuals per simulated trial). For each clinical trial, after each cohort an optimal dose was predicted and used to calculate 'true efficacy at predicted optimal dose'.

The cumulative number of efficacious responses that had occurred up to and including that cohort in the simulated trials was also calculated as the 'cumulative

252

sum of efficacy'. The mean value of these two metrics after each cohort across the hundred simulations were then calculated for each scenario for each DOA. A 95% confidence interval for the true mean values of these metrics were also calculated. These were plotted to qualitatively show the expected 'true efficacy at the predicted optimal' and 'cumulative sum of efficacy' for each scenario for clinical trials conducted using each DOA after each cohort. This allowed comparison between the DOAs, and also a comparison to the theoretical true optimal that a DOA could achieve.

2.4.2. Objective 2: Evaluate the Correlated Beta Dose Optimisation Approach for optimising vaccine efficacy for a prime-dose/boost-dose administration.

In this objective we aim to assess the dose-optimisation ability of the CoBe dose-optimisation approach compared to other dose-optimisation approaches when choosing doses for a multiple administration vaccine that maximises efficacy. In this objective, we use scenarios where:

- Dosing domain: Prime/boost (scenarios 1-5) or prime/boost/second-boost (scenarios 6, 7)
- Utility function: Maximise Efficacy
- Efficacy curve: Is defined for each scenario
- Toxicity curve: Not defined/not of interest

Seven scenarios were used to explore this objective (Figure 5). The prime/boost scenarios reflect cases for which vaccine dose efficacy may be

1. Peaking with respect to both doses and where the combination of both vaccine doses increases their efficacy [figure 7a]
2. Saturating with regards to both doses but where the combination of both vaccine doses decreases their efficacy [figure 7b]
3. Saturating with respect to both doses and where the combination of both vaccine doses increases their efficacy [figure 7c]

4. Saturating with respect to both doses and where the combination of both vaccines increases their efficacy, but maximally dosing both vaccines causes decreased efficacy. [figure 7d]

5. Saturating with respect to both doses and where the combination of both vaccines increases their efficacy, but where one of the doses is significantly more important to maximising efficacy. [figure 7e]

Scenarios 6 and 7 are prime/boost/second-boost. These scenarios:

6. Represents a case where there is a maximally efficacious dose for each, and any increase/decrease in any of these doses decreases efficacy regardless of the other doses. Thus, the optimal dose for each of the prime/boost/second-boost was independent of what other doses were selected [figure 7f]

7. Represent a case where a maximal dose of any two of the three doses produces a highly efficacious response, but a maximal dose of all three does not produce a highly efficacious response [figure 7g].

1.
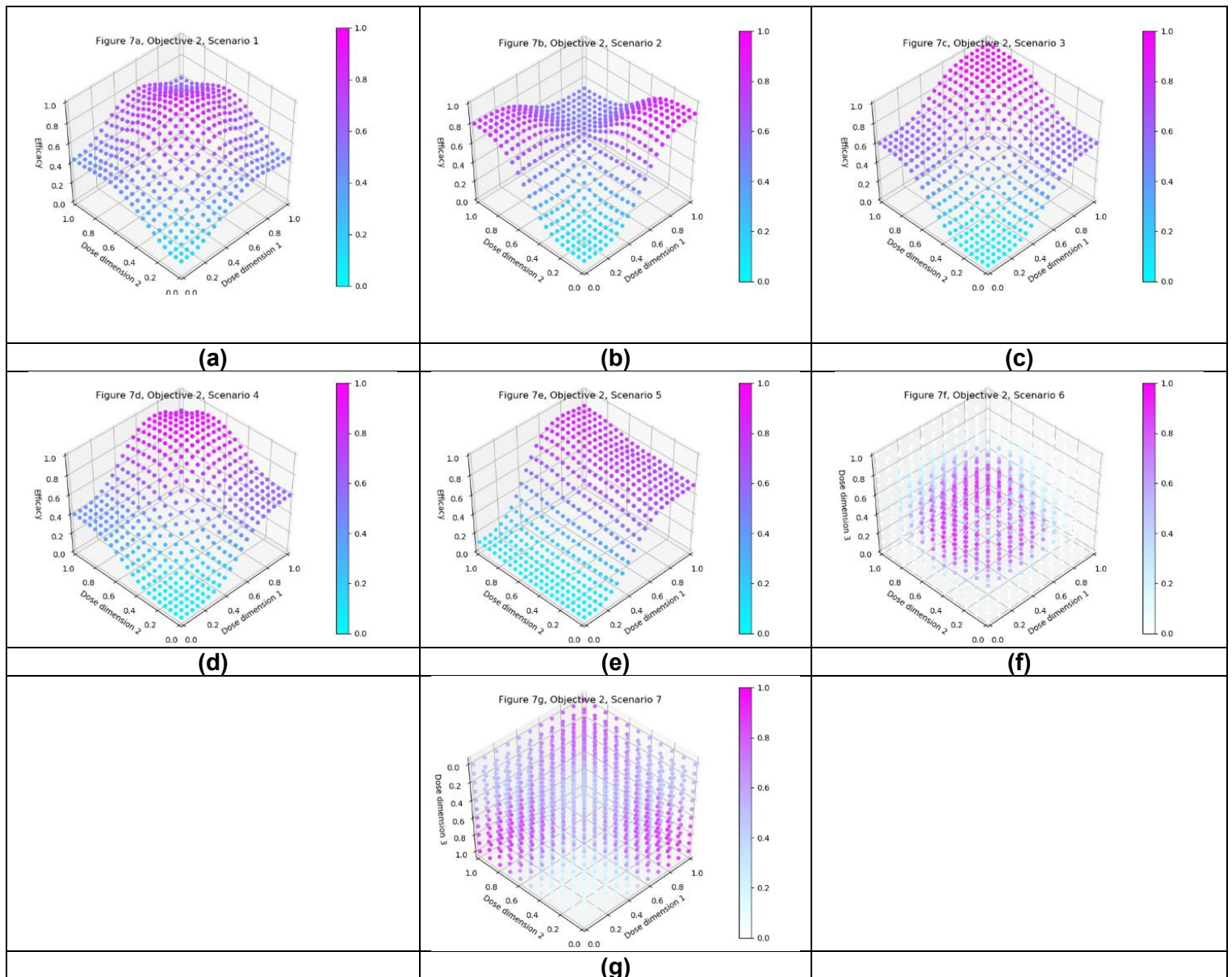
Figure 7. Dose efficacy plots for the seven objective 2 scenarios. These were (a) scenario 1, (b) scenario 2, (c) scenario 3, (d) scenario 4, (e) scenario 5, (f) scenario 6, and (g) scenario 7 . Purple represents areas of higher efficacy. Note the z-axis for (g) is inverted to better show the 3-dimensional dose-efficacy relationship.

For each scenario, we simulated 100 clinical trials conducted under each of the four DOAs (100x7x4 simulated trials total). For the CoBe, Parametric and Adaptive Naive DOAs, cohorts were of size 6, and we simulated 50 cohorts for each clinical trial (300 total individuals per simulated trial). For scenarios 1-5, the Uniform Naive DOA used cohorts of size 9, and we simulated 33 cohorts for each clinical trial (297 total individuals per simulated trial). For scenarios 6 and 7, the Uniform Naive DOA used cohorts of size 27, and we simulated 11 cohorts for each clinical trial (297 total individuals per simulated trial).

255

For each clinical trial, after each cohort an optimal dose was predicted and used to calculate 'true efficacy at predicted optimal dose'. The cumulative number of efficacious responses that had occurred up to and including that cohort in the simulated trials was also calculated as the 'cumulative sum of efficacy'. The mean value of these two metrics after each cohort across the hundred simulations were then calculated for each scenario for each DOA. A 95% confidence interval for the true mean values of these metrics were also calculated. These were plotted to qualitatively show the expected 'true efficacy at the predicted optimal' and 'cumulative sum of efficacy' for each scenario for clinical trials conducted using each DOA after each cohort. This allowed comparison between the DOAs, and also a comparison to the theoretical true optimal that a DOA could achieve.

2.4.3. Objective 3. Evaluate the Correlated Beta Dose Optimisation Approach for optimising vaccine utility, maximising efficacy, and minimising toxicity.

In this objective we aim to assess the dose-optimisation ability of the CoBe dose-optimisation approach compared to dose-optimisation approaches when choosing doses for single or multiple administration vaccines for which an optimal balance of efficacy and toxicity must be achieved. In this objective we use scenarios where:

- Dosing domain: Single-administration (scenarios 1-4) or prime/boost administration (scenarios 5-6)
- Utility function: Utility Contour
- Efficacy curve: Is defined for each scenario
- Toxicity curve: Is defined for each scenario

Six scenarios were used in this objective (Figure 8). The single-administration scenarios reflect cases for which vaccines

1. have gradually increasing efficacy and toxicity with dose [figure 8 a, b, c]
2. have sharply peaking efficacy and gradually increasing toxicity with dose [figure 8 d, e, f]

256

3. have gradually increasing efficacy and sharply increasing toxicity with dose [figure 8 g, h, i]

4. have sharply peaking efficacy and sharply increasing toxicity with dose [figure 8 j, k, l]

The prime/boost administration scenarios reflect cases for which vaccines:

5. have efficacy as per objective 2, scenario 3, and toxicity increasing for high doses of either vaccine [figure 8 m, n, o]

6. have efficacy as per objective 2, scenario 2, and toxicity increasing for high doses of either vaccine [figure 8 p, q, r]

(a)

(b)

(c)

(d)

(e)

(f)

(g)

(h)

(i)
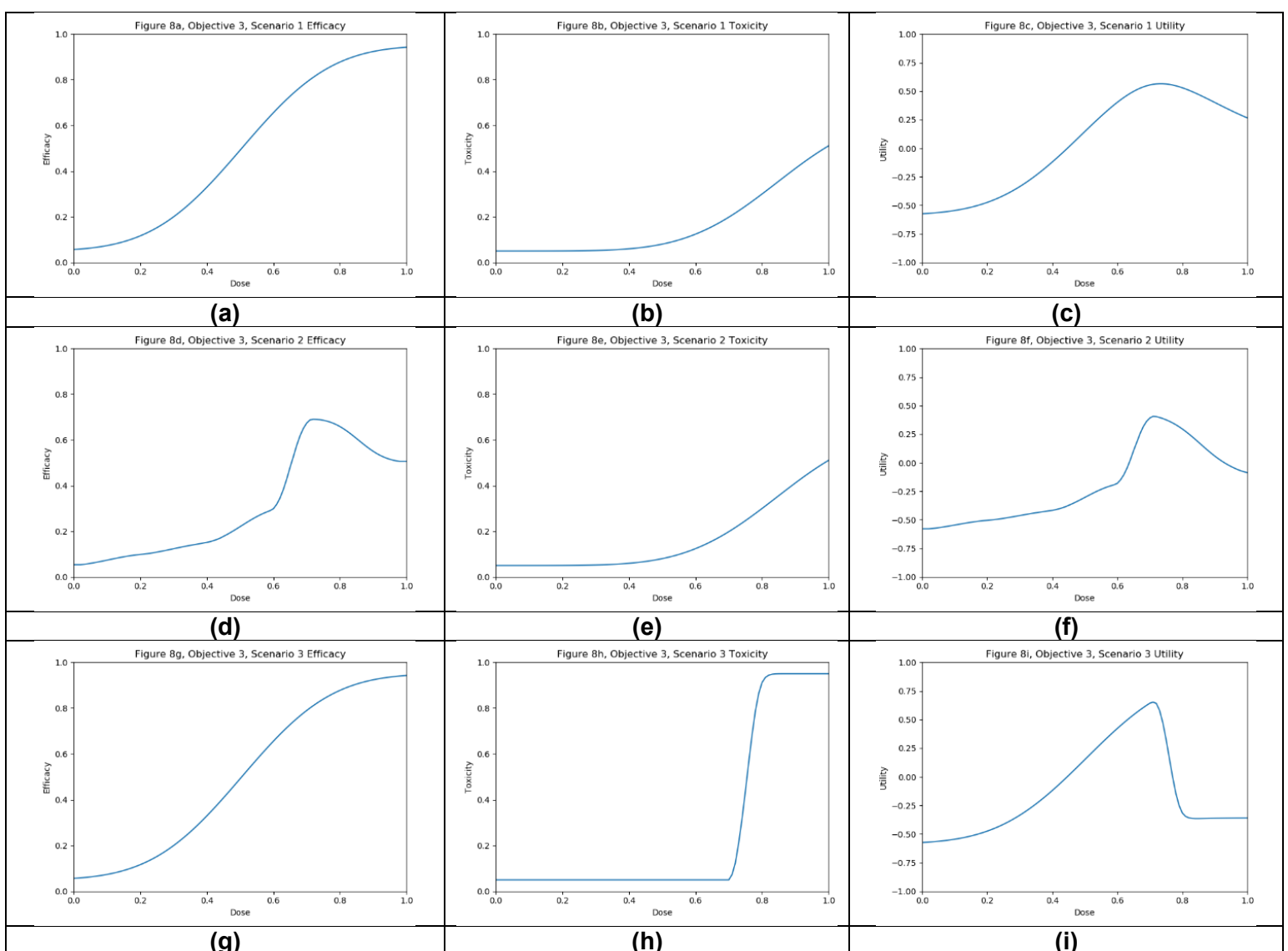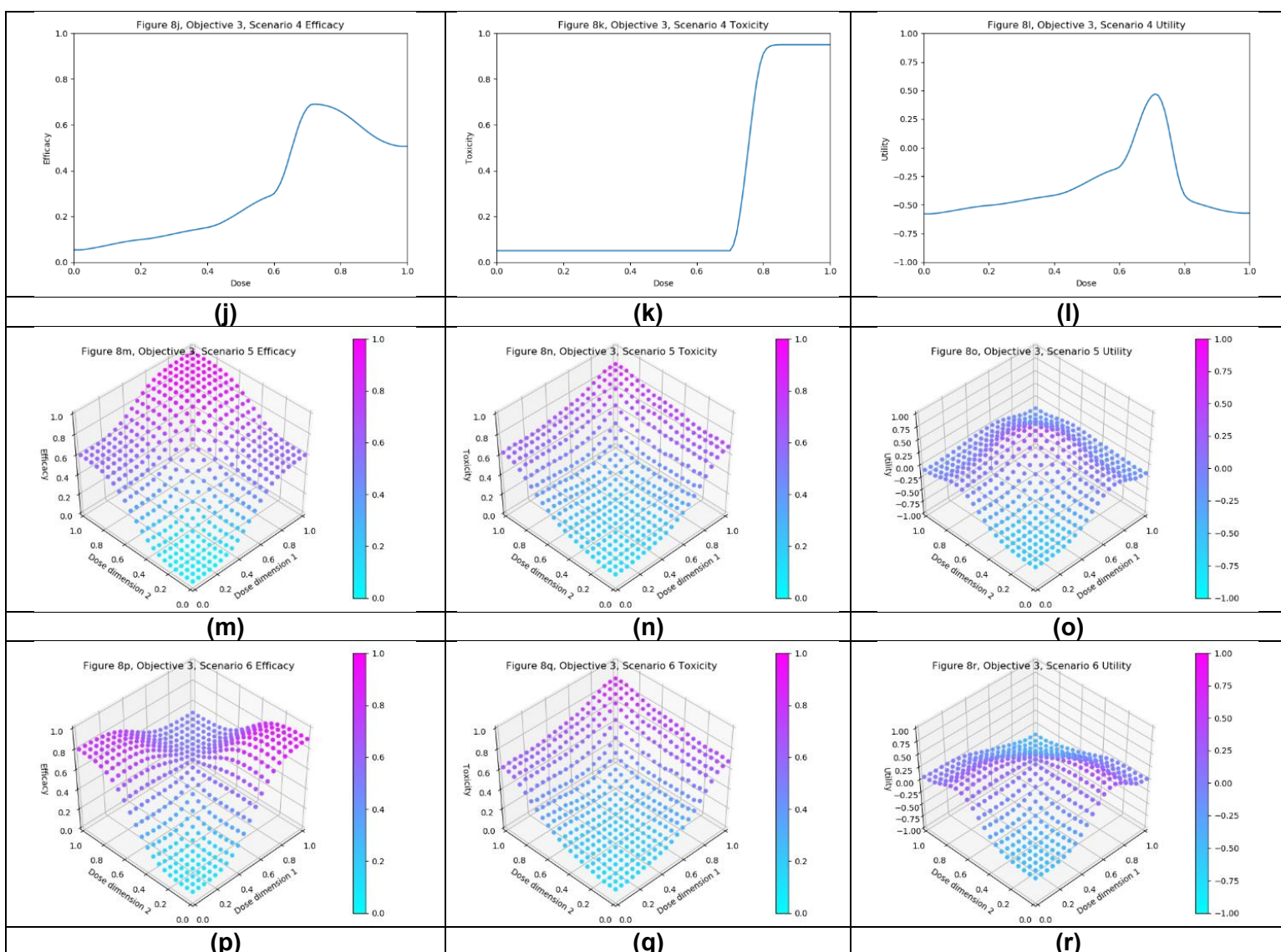
Figure 8. Dose efficacy/toxicity/utility plots for the six objective 3 scenarios. These were (a-c) scenario 1, (d-f) scenario 2, (g-i) scenario 3, (j-l) scenario 4, (m-o) scenario 5, and (p-r) scenario 6. Left/middle/right plots were respectively for efficacy/toxicity/utility. (a-l) The blue line represents the probability of efficacious/toxic response and utility for each dose. (m-o) Purple represents areas of higher efficacy/toxicity/utility.

For each scenario, we simulated 100 clinical trials conducted under each of the four DOAs (100x6x4 simulated trials total). For the CoBe, Parametric and Adaptive Naive DOAs, cohorts were of size 6, and we simulated 50 cohorts for each clinical trial (300 total individuals per simulated trial). This was also true for the Uniform Naive DOA for scenarios 1-4. For scenarios 5 and 6, the Uniform Naive DOA used cohorts of size 9, and we simulated 33 cohorts for each clinical trial (297 total individuals per simulated trial).

258

For each clinical trial, after each cohort an optimal dose was predicted and used to calculate 'true utility at predicted optimal dose'. The cumulative utility up to and including that cohort in the simulated trials was also calculated as the 'cumulative sum of utility'. The mean value of these two metrics after each cohort across the hundred simulations were then calculated for each scenario for each DOA. A 95% confidence interval for the true mean values of these metrics were also calculated. These were plotted to qualitatively show the expected 'true utility at the predicted optimal' and 'cumulative sum of utility' for each scenario for clinical trials conducted using each DOA after each cohort. This allowed comparison between the DOAs, and also a comparison to the theoretical true optimal that a DOA could achieve.

2.4.4. Objective 4. Evaluate the use of expert knowledge informed Continuous Correlated Beta Process priors for vaccine dose-optimisation

In this objective we aim to assess the dose-optimisation ability of the CoBe dose-optimisation approach when including either accurate or inaccurate expert information priors, and to what extent such priors improve or are detrimental to CoBe DOA performance.

We compared the CoBe DOA with 5 different 'priors'

- Very strong, correct
- Strong, correct
- No prior
- Strong, incorrect
- Very strong, incorrect

These priors were implemented as defined in 2.2.1.4. The 'No prior' approach had $p_{i,r}^0 \sim Beta\left(\alpha_{i,eff}^0, \beta_{1i,eff}^0\right) = Beta(1,1)$ for all doses and is the CoBe DOA used in the previous objectives. The 'Very strong, correct' and 'Strong, correct' priors assume the expert knowledge is entirely correct top the 'true' vaccine dose response, with suggested probability $p_{i,r}$ equal to the true probability of efficacy/toxicity for all doses d_i (so if the true probability of efficacy for some dose was 0.2, the 'expert' would predict an efficacy of 0.2). The 'Very strong, incorrect' and 'Strong, incorrect' priors

represent the expert being largely incorrect, with suggested probability $p_{i,r}$ equal to one minus the true probability of efficacy/toxicity for all doses (so if the true probability of efficacy for some dose was 0.2, the 'expert' would predict an efficacy of 0.8). The 'Strong, correct' and 'Strong, incorrect' priors used $c_i = 3$, which was deemed appropriate based on previous results in parametric dose-optimisation [57,58] and the 'Very strong, correct' and 'Very strong, incorrect' priors used $c_i = 20$ for all doses which represented having extreme confidence in the expert prior.

In this objective we use scenarios from the above 3 objectives (Figure 9). In this objective we use scenarios where:

- Dosing domain: Single-administration (scenarios 1, 2), prime/boost administration (scenarios 3,4, 7), or prime/boost/second-boost administration (scenarios 5,6)
- Utility function: Maximise Efficacy (Scenarios 1-6), Utility Contour (Scenario 7)
- Efficacy curve: Is defined for each scenario
- Toxicity curve: Is defined for only scenario 7

Specifically, these are:

1. Objective 1, Scenario 1 [Figure 9a]
2. Objective 1, Scenario 4 [Figure 9b]
3. Objective 2, Scenario 1 [Figure 9c]
4. Objective 2, Scenario 2 [Figure 9d]
5. Objective 2, Scenario 6 [Figure 9e]
6. Objective 2, Scenario 7 [Figure 9f]
7. Objective 3, Scenario 6 [Figure 9 g, h, i]
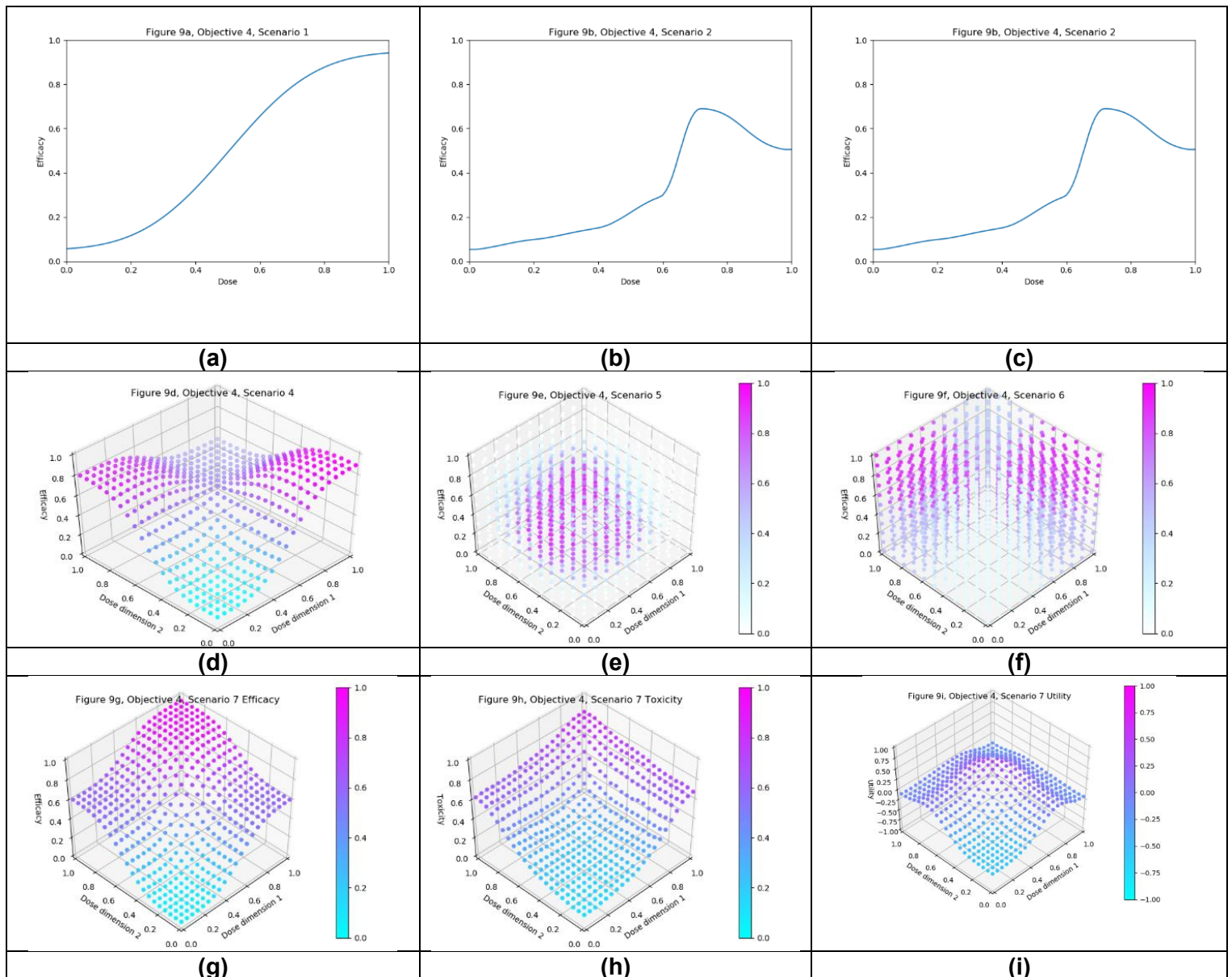
Figure 9. Dose efficacy and efficacy/toxicity/utility plots for the seven objective 4 scenarios. Dose-efficacy plots were (a) scenario 1, (b) scenario 2, (c) scenario 3, (d) scenario 4, (e) scenario 5, (f) scenario 6. For (g-i) scenario 7, these were respectively for efficacy/toxicity/utility. (a, b) The blue line represents the probability of efficacious/toxic response and utility for each dose. (c-i) Purple represents areas of higher efficacy/toxicity/utility.

For each scenario, we simulated 100 clinical trials conducted under each of the four DOAs (100x7x4 simulated trials total). Cohorts were of size 6, and we simulated 50 cohorts for each clinical trial (300 total individuals per simulated trial).

For each clinical trial, after each cohort an optimal dose was predicted and used to calculate 'true efficacy at predicted optimal' for scenarios 1-6 and 'true utility at predicted optimal dose' for scenario 7. The cumulative efficacy utility up to and

261

including that cohort in the simulated trials was also calculated as the 'cumulative sum of efficacy' for scenarios 1-6, similarly the 'cumulative sum of utility' was calculated for scenario 7. The mean value of these metrics after each cohort across the hundred simulations were then calculated for each scenario for each DOA. A 95% confidence interval for the true mean values of these metrics were also calculated. These were plotted to qualitatively show the expected 'true efficacy/utility at the predicted optimal' and 'cumulative sum of efficacy/utility' for each scenario for clinical trials conducted using each DOA after each cohort. This allowed comparison between the DOAs, and a comparison to the theoretical true optimal that a DOA could achieve.

3. Results

3.1. Objective 1. Evaluate the Correlated Beta Dose Optimisation Approach for optimising vaccine efficacy for a single dose administration.

3.1.1. True efficacy at predicted optimal dose

The DOA (dose-optimisation approach) that best maximised 'true efficacy at predicted optimal dose' (from here called 'true efficacy') for a given scenario and trial size was considered to be the 'best' DOA for the aim of selecting an optimal dose for that scenario. The left-hand side of figure 10 shows the change in mean true efficacy with increasing numbers of trial participants for each of the four DOAs for each scenario, averaged across 100 simulated clinical trials. For each of these plots the upper and lower brown lines respectively show the maximal and minimal efficacy possible for that scenario. A mean true efficacy for a DOA being closer to the upper brown line relative to a second DOA indicates the first DOA being on average better at selecting a highly efficacious dose. Equivalently, a mean true efficacy being closer to the lower brown line would represent a DOA being of average worse at selecting a highly efficacious dose.

For all scenarios, the CoBe (Correlated Beta) DOA had similar or greater mean true efficacy relative to the other DOAs for all trial sizes [figure 10, LHS]. The mean true efficacy curve typically plateaued with between 30 and 90 trial participants. Scenario 7 was an exception, which may be due to the flat nature of the scenario's dose-efficacy curve, which may have necessitated more data in order to discern a statistical difference in dose-efficacy for different doses (note the scale of the y-axis).

The CoBe DOA had a lower/worse mean true efficacy than the Parametric DOA for scenarios 1 and 2 for low numbers of trial participants. This may have been because the parametric model for the Parametric DOA was able to easily approximate the scenario 1 and 2 dose response curves with minimal data and correctly identify that the largest dose was maximally efficacious. However, the CoBe DOA had a greater mean true efficacy than the Parametric DOA for scenarios 4 and 6 for most trial sizes.

The CoBe DOA had similar mean true efficacy to the Adaptive Naive DOA for most scenarios. However, for scenarios 3 and 6 the Adaptive Naive DOA plateaued at a lower true efficacy. For scenarios where one of the potential doses in the discretised dosing domain of the Adaptive Naive DOA was near optimal (1,2,4,5, and 7), the Adaptive DOA had similar mean true efficacy to the CoBe DOA. However, when none of these potential doses were near optimal, the Adaptive Naive DOA had a lower/worse mean true efficacy than the CoBe DOA. For example, for scenarios 3 and 6, none of the six doses in the discretised dosing domain were truly optimal for those scenarios.

The uniform naive approach had a lower/worse mean true efficacy compared to the other approaches, particularly when the number of trial participants was small. The only exception was that the mean true efficacy for the Uniform Naive DOA was greater than that of the Parametric DOA for scenario 6. The mean true efficacy of the Uniform Naive DOA was typically equalled or surpassed by that of the Adaptive Naive DOA for all scenarios and numbers of trial participants.

3.1.2. Cumulative sum of efficacy

Cumulative sum of efficacy measures a DOA's capacity to maximise the benefit to trial participants. After a certain number of trial participants, a DOA with a higher cumulative efficacy would be considered 'more ethical' than a DOA with a lower cumulative efficacy, as it would reflect those trial participants having received on average more efficacious dosing. The right-hand side of figure 10 shows the change in mean cumulative sum of efficacy with increasing numbers of trial participants for each of the four DOAs for each scenario, averaged across 100 simulated clinical trials. For each of these plots the upper and lower brown lines respectively show the theoretical maximal and minimal mean cumulative efficacy possible for that scenario. A mean true efficacy for a DOA being closer to the upper brown line relative to a second DOA reflects that the trial participants for simulated clinical trials using the first DOA on average received more efficacious doses. If the relationship between number of trial participants and mean cumulative sum of efficacy for a DOA becomes parallel to the upper brown line after some number of trial participants, then that represents that the DOA gave a near optimal dose to all trial participants after that

point. No DOA could exceed this upper brown line, as this upper brown line reflects a DOA for which every trial participant receives the dose that is truly most efficacious.

The relationship between the number of trial participants and cumulative sum of efficacy for DOAs that used adaptive trial design (CoBe, Parametric, Adaptive Naive DOAs) was non-linear [figure 10, RHS]. For small numbers of trial participants the gradient of this curve was less steep than the line for the theoretical maximal (upper brown line, figure 10 RHS). When the number of trial participants was small, there was little data available to guide the adaptive design of these DOAs and so trial doses were more likely to be suboptimal. As the number of trial participants increased, more data was available to inform selection of trial doses that were likely to be efficacious, increasing the steepness of the curve.

The CoBe DOA typically had a similar or greater cumulative sum of efficacy relative to the other DOAs for all scenarios [figure 10, RHS]. In many scenarios the mean cumulative efficacy was only slightly below the theoretical maximum for that scenario for the number of trial participants. Given that this upper bound is only achievable by dosing all trial participants at the true optimal dose for that scenario (which is likely not known a-priori in a dose-finding trial), for these scenarios the CoBe DOA was found to be highly ethical. The Uniform Naive DOA had a lower/worse cumulative sum of efficacy than the other DOAs for all scenarios, especially at large trial size. This was likely because this DOA did not use any adaptive design, which is what allowed other DOAs to choose more promising doses for their later trial participants.

(a)   (b)

(c) Objective 1, Scenario 2 — Mean true efficacy at predicted optimal dose

(d) Objective 1, Scenario 2 — Mean Cumulative Sum of Efficacy

(e) Objective 1, Scenario 3 — Mean true efficacy at predicted optimal dose

(f) Objective 1, Scenario 3 — Mean Cumulative Sum of Efficacy

(g) Objective 1, Scenario 4 — Mean true efficacy at predicted optimal dose

(h) Objective 1, Scenario 4 — Mean Cumulative Sum of Efficacy

266

(i)

(j)

(k)

(l)

(m)

(n)

Figure 10. Mean true efficacy at the predicted optimal dose (left) and mean cumulative sum of efficacy (right) against trial size for all seven objective 1 scenarios (top to bottom). These are the mean values and 95%CI values across 100 simulations. For the true efficacy plots (left), the brown lines show the minimum and maximum possible efficacy that could be achieved in that scenario. For the cumulative efficacy plots (right) the brown lines represent the maximum and minimum cumulative efficacy sum that could be expected for that scenario. For example, if the true

267

maximum efficacy for a scenario was 90%, no DOA could locate a dose with a true efficacy >90%, and no DOA could achieve a mean cumulative efficacy > 270 (=90% x 300) after testing 300 trial participants. A mean true efficacy/mean cumulative sum of efficacy curve being closer to the upper brown line reflects that DOA being the more effective for locating a maximally efficacious dose/maximising benefit to trial participants.

3.2. Objective 2. Evaluate the Correlated Beta Dose Optimisation Approach for optimising vaccine efficacy for a prime-dose/boost-dose administration.

3.2.1. True efficacy at predicted optimal dose

The DOA that best maximised true efficacy for a given scenario and trial size was considered to be the 'best' DOA for the aim of selecting an optimal dose for that scenario. The left-hand side of figure 11 shows the change in mean true efficacy with increasing numbers of trial participants for each of the four DOAs for each scenario, averaged across 100 simulated clinical trials. For each of these plots the upper and lower brown lines respectively show the maximal and minimal efficacy possible for that scenario. A mean true efficacy for a DOA being closer to the upper brown line relative to a second DOA indicates the first DOA being on average better at selecting a highly efficacious dose. Equivalently, a mean true efficacy being closer to the lower brown line would represent a DOA being of average worse at selecting a highly efficacious dose.

For all scenarios, the CoBe (Correlated Beta) DOA had similar or greater mean true efficacy relative to the other DOAs for all trial sizes [figure 11, LHS]. The mean true efficacy curve typically plateaued with between 60 and 90 trial participants.

The CoBe DOA had a lower/worse mean true efficacy than the Parametric DOA for scenarios 3 and 5 for small numbers of trial participants but had a higher mean true efficacy than the Parametric DOA for scenarios 2 and 7.

The CoBe DOA had a higher mean true efficacy than the Adaptive Naive DOA for scenarios 4 and 6 and for a small numbers of trial participants for scenario 7 and a large number of trial participants for scenario 1. Otherwise mean true efficacy was similar for these two DOAs.

Again, the Uniform Naive DOA had a lower/worse mean true efficacy than all other approaches for all scenarios, with the only exception being the Parametric DOA which had a lower mean true efficacy for scenarios 2 and 7.

3.2.2. Cumulative sum of efficacy

Cumulative sum of efficacy measures a DOA's capacity to maximise the benefit to trial participants. After a certain number of trial participants, a DOA with a higher cumulative efficacy would be considered 'more ethical' than a DOA with a lower cumulative efficacy, as it would reflect those trial participants having received on average more efficacious dosing. The right-hand side of figure 11 shows the change in mean cumulative sum of efficacy with increasing numbers of trial participants for each of the four DOAs for each scenario, averaged across 100 simulated clinical trials. For each of these plots the upper and lower brown lines respectively show the theoretical maximal and minimal mean cumulative efficacy possible for that scenario. A mean true efficacy for a DOA being closer to the upper brown line relative to a second DOA reflects that the trial participants for simulated clinical trials using the first DOA on average received more efficacious doses. If the relationship between number of trial participants and mean cumulative sum of efficacy for a DOA becomes parallel to the upper brown line after some number of trial participants, then that represents that the DOA gave a near optimal dose to all trial participants after that point. No DOA could exceed this upper brown line, as this upper brown line reflects a DOA for which every trial participant receives the dose that is truly most efficacious.

The same non-linearity in the relationship between number of trial participants and cumulative sum of efficacy for DOAs that used adaptive trial design (CoBe, Parametric, Adaptive Naive DOAs) that was observed in objective 1 was also observed in objective 2 [Figure 11, RHS].

The CoBe DOA typically had a similar or greater cumulative sum of efficacy relative to the other DOAs for all scenarios [figure 11, RHS]. The exception was for scenario 6, for which the Parametric DOA had a greater cumulative sum of efficacy, however the CoBe DOA had a greater cumulative sum of efficacy for scenarios 2 and 7. The Adaptive Naive DOA had a greater cumulative sum of efficacy than the CoBe DOA

269

for scenarios 2 and 6, however the CoBe DOA had a greater cumulative sum of efficacy for scenarios 4 and 7. The Uniform Naive DOA had a lower/worse cumulative sum of efficacy than the other DOAs for all scenarios, especially at large trial size.

(a)

(b)

(c)

(d)

(e)

Objective 2, Scenario 4

(f)

Objective 2, Scenario 4

(g)

Objective 2, Scenario 5

(h)

Objective 2, Scenario 5

(i)

Objective 2, Scenario 6

(j)
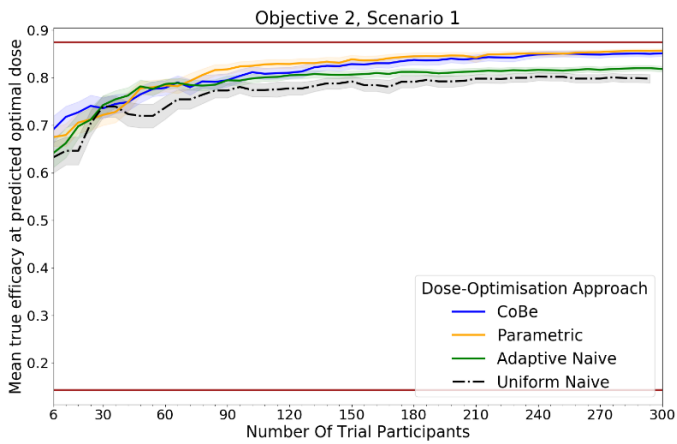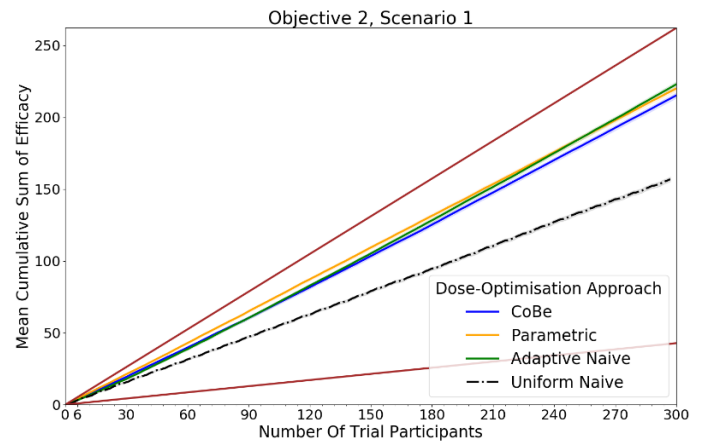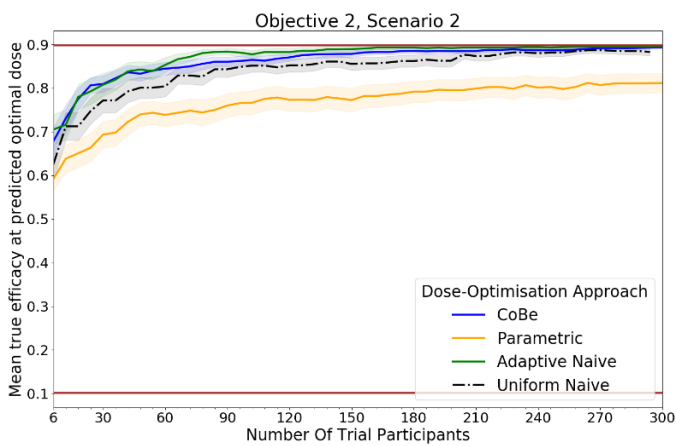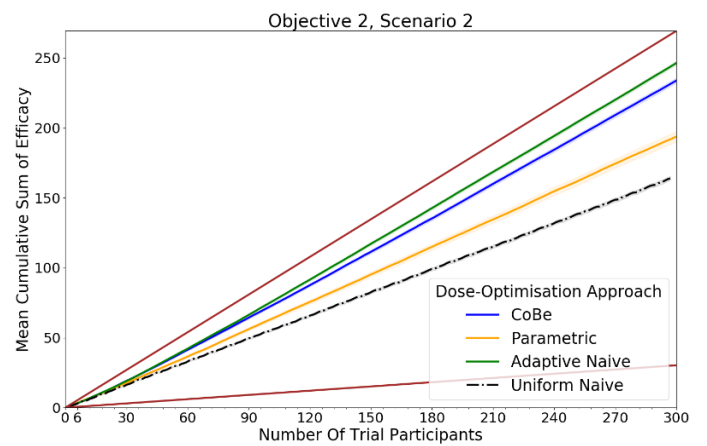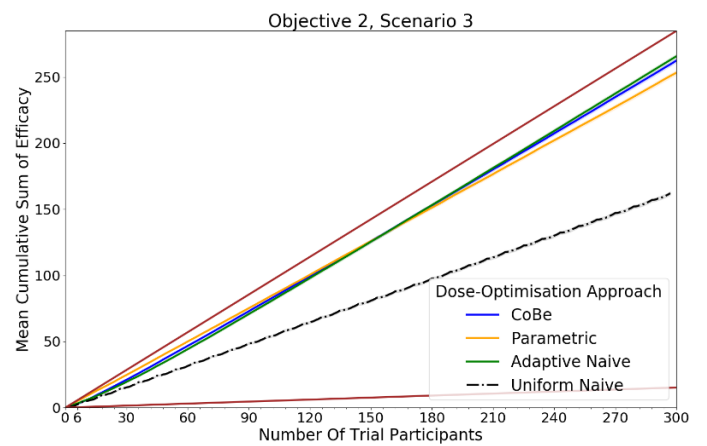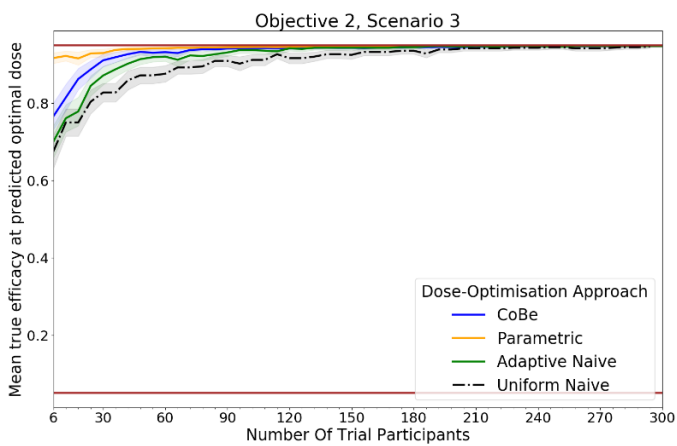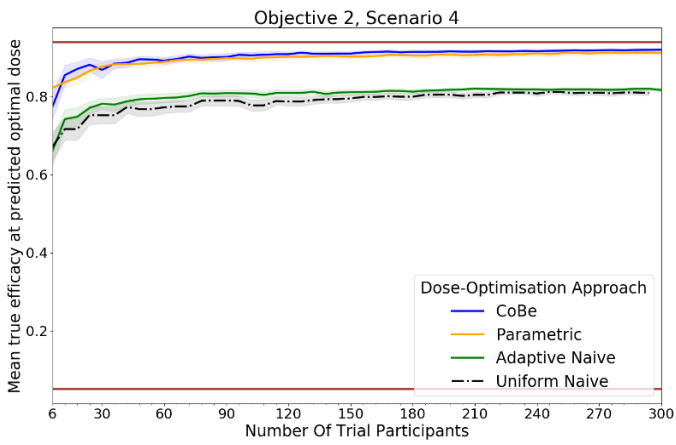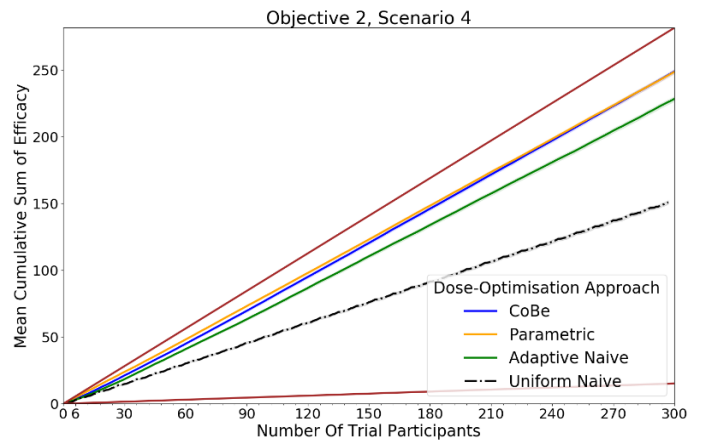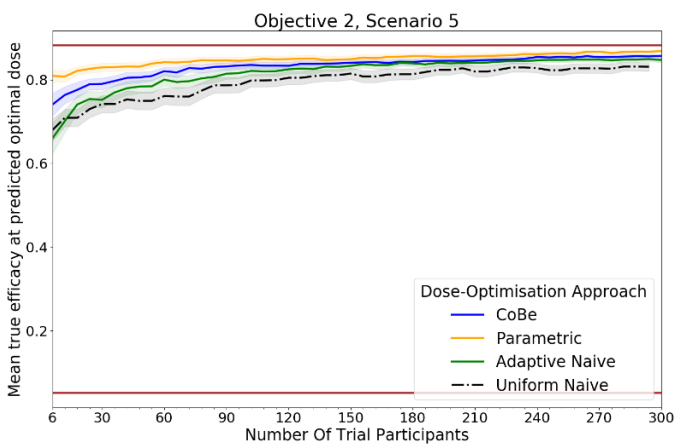
Objective 2, Scenario 6
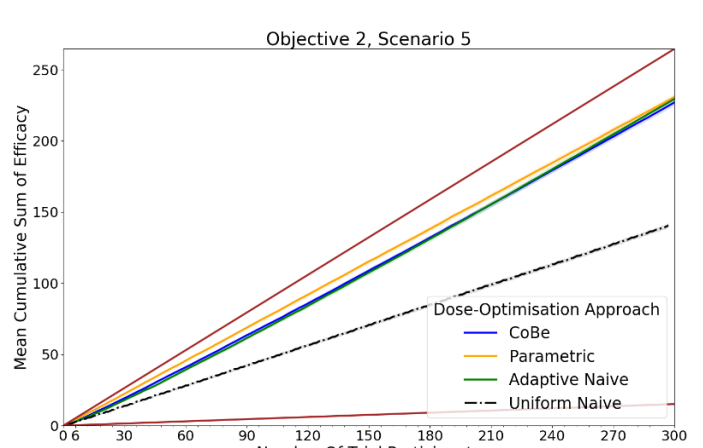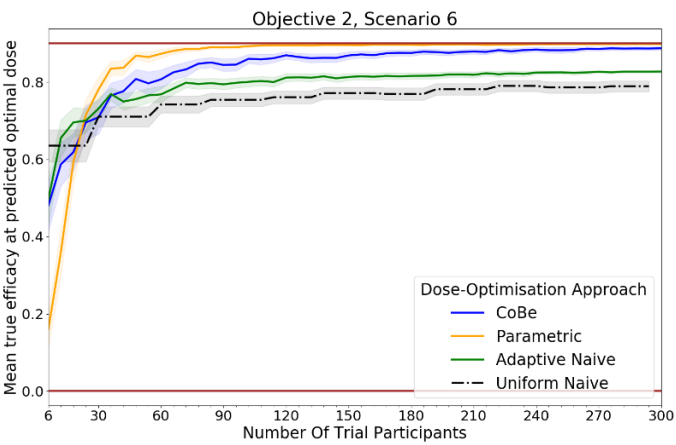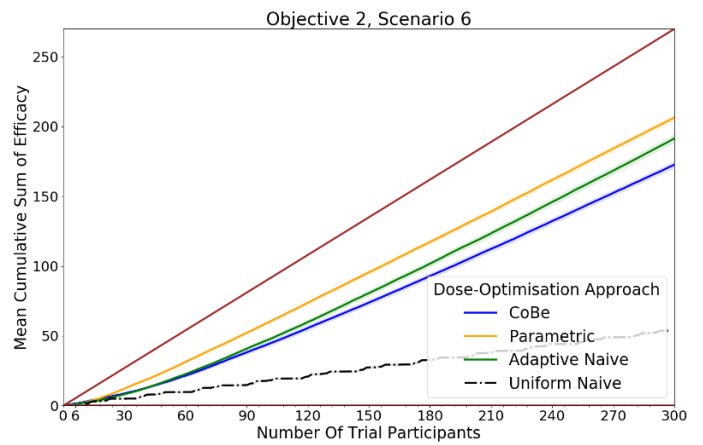
(k)

(l)
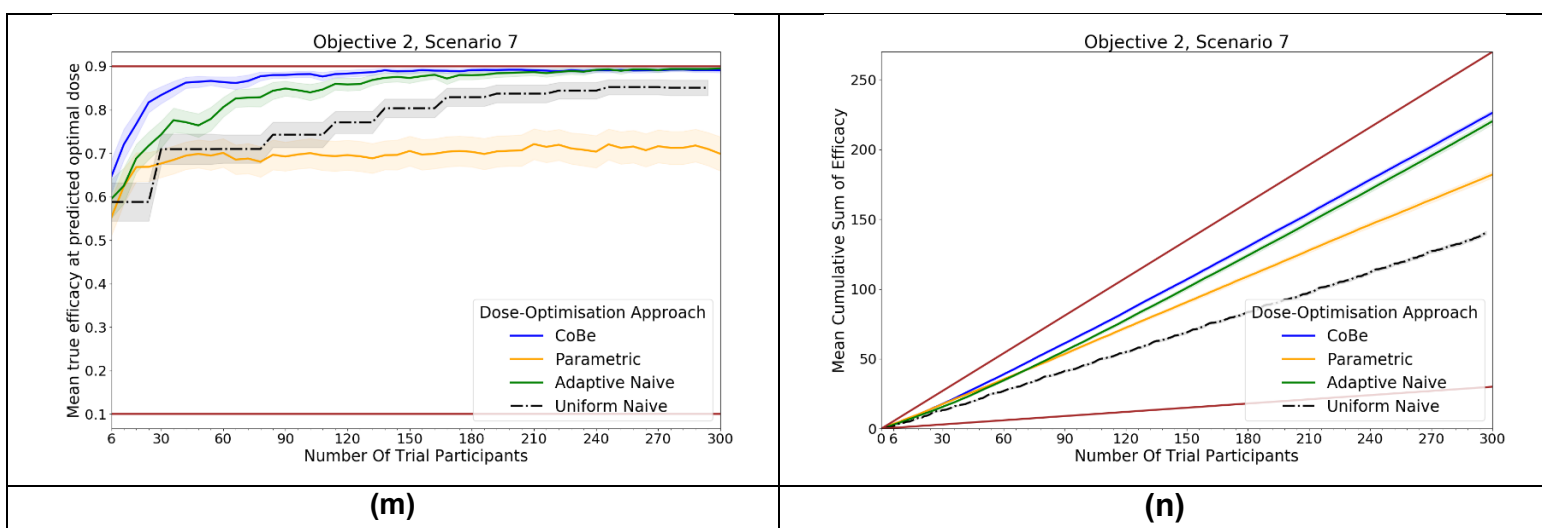
Figure 11. Mean true efficacy at the predicted optimal dose (left) and mean cumulative sum of efficacy (right) against trial size for all seven objective 2 scenarios (top to bottom). These are the mean values and 95%CI values across 100 simulations. For the true efficacy plots (left), the brown lines show the minimum and maximum possible efficacy that could be achieved in that scenario. For the cumulative efficacy plots (right) the brown lines represent the maximum and minimum cumulative efficacy sum that could be expected for that scenario.

3.3. Evaluate the Correlated Beta Dose Optimisation Approach for optimising vaccine utility, maximising efficacy, and minimising toxicity.

3.3.1. True utility at predicted optimal dose

The DOA that best maximised 'true utility at predicted optimal dose' (from here called 'true utility') for a given scenario and trial size was considered to be the 'best' DOA for the aim of selecting an optimal dose for that scenario. Utility was defined by the 'utility contour' function that increased with an increased probability of efficacy and increased with a decreased probability of toxicity. The left-hand side of figure 12 shows the change in mean true utility with increasing numbers of trial participants for each of the four DOAs for each scenario, averaged across 100 simulated clinical trials. For each of these plots the upper and lower brown lines respectively show the maximal and minimal utility possible for that scenario. A mean true utility for a DOA being closer to the upper brown line relative to a second DOA indicates the first DOA being on average better at selecting a high utility dose. Equivalently, a mean true efficacy being closer to the lower brown line would represent a DOA being on average worse at selecting a high utility dose.

272

For all scenarios, the CoBe (Correlated Beta) DOA had similar or greater mean true utility relative to the other DOAs for all trial sizes [figure 10, LHS]. The mean true utility curve typically plateaued with between 30 and 60 trial participants for the single dose administration scenarios (1-4), and between 60 and 90 trial participants for the prime/boost administration (5 and 6).

The CoBe DOA had similar mean true utility to the Parametric DOA for scenarios 1-4, however for scenarios 5 and 6 the CoBe DOA had a greater mean true utility relative to the Parametric DOA. The CoBe DOA had a greater mean true utility than either of the Adaptive Naive or Uniform Naive DOAs for all scenarios.

3.3.2. Cumulative sum of utility

Cumulative sum of utility measures a DOA's capacity to maximise the benefit to trial participants. After a certain number of trial participants, a DOA with a higher cumulative utility would be considered 'more ethical' than a DOA with a lower cumulative utility, as it would reflect those trial participants having received on average more efficacious/less toxic dosing. The right-hand side of figure 12 shows the change in mean cumulative utility with increasing numbers of trial participants for each of the four DOAs for each scenario, averaged across 100 simulated clinical trials. For each of these plots the upper and lower brown lines respectively show the theoretical maximal and minimal mean cumulative utility possible for that scenario. A mean true utility for a DOA being closer to the upper brown line relative to a second DOA reflects that the trial participants for simulated clinical trials using the first DOA on average received more efficacious/less toxic doses. If the relationship between number of trial participants and mean cumulative sum of utility for a DOA becomes parallel to the upper brown line after some number of trial participants, then that represents that the DOA gave a near optimal dose to all trial participants after that point. No DOA could exceed this upper brown line, as this upper brown line reflects a DOA for which every trial participant receives the dose that is truly optimal in regard to maximising efficacy whilst minimising toxicity according to the utility function.

The same non-linearity in the relationship between number of trial participants and cumulative utility for DOAs that used adaptive trial design (CoBe, Parametric,

273

Adaptive Naive DOAs) that was observed in objectives 1 and 2 was also observed in objective 3 [figure 11, RHS]. The CoBe DOA had a similar mean cumulative utility to the Parametric DOA for scenarios 1, 5 and 6, however the Parametric DOA had a greater mean cumulative utility for scenarios 2,3 and 4. The CoBe DOA had a similar or greater mean cumulative utility to the Adaptive Naive DOA for all scenarios other than scenario 2. The Uniform Naive DOA had a lower/worse cumulative utility than the other DOAs for all scenarios, especially at large trial size.
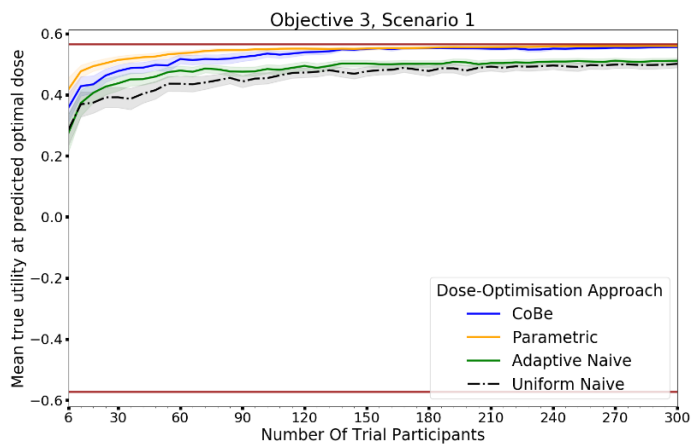
(a)
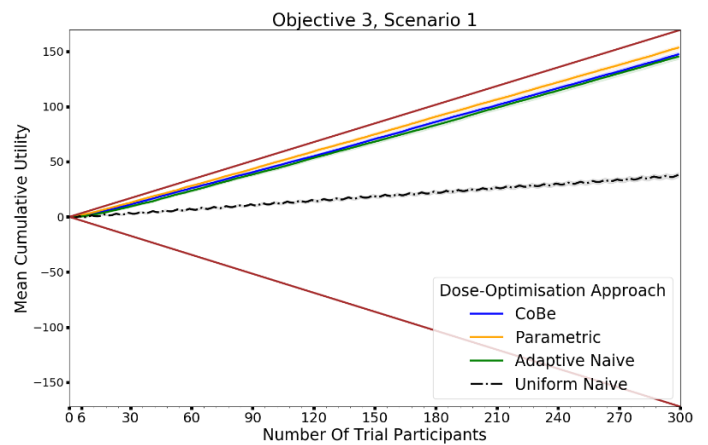
(b)
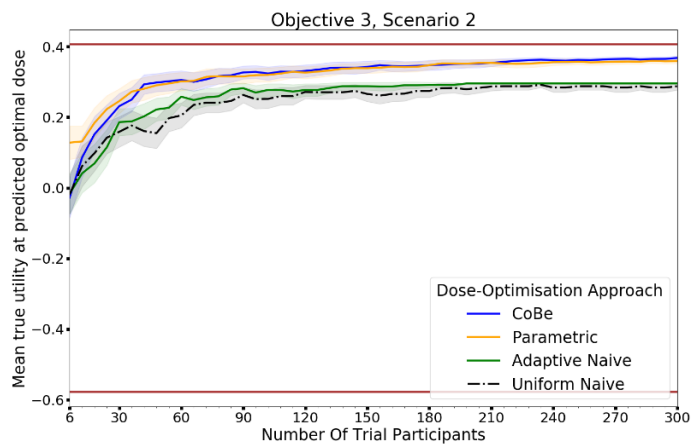
(c)
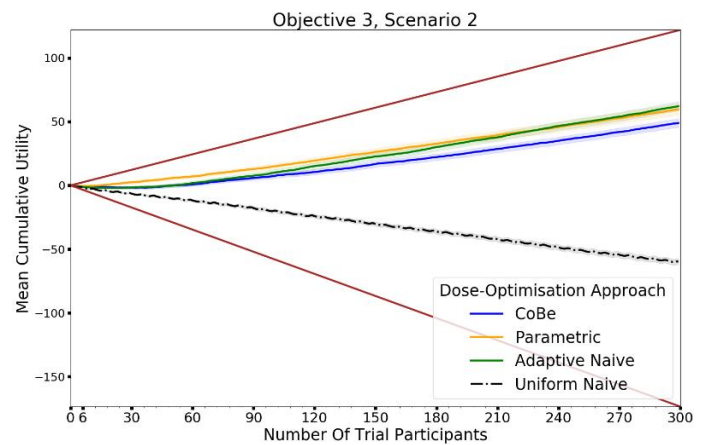
(d)
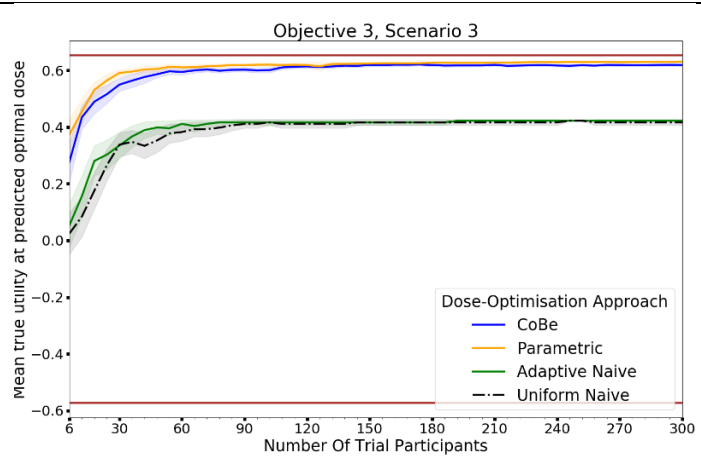
(e)
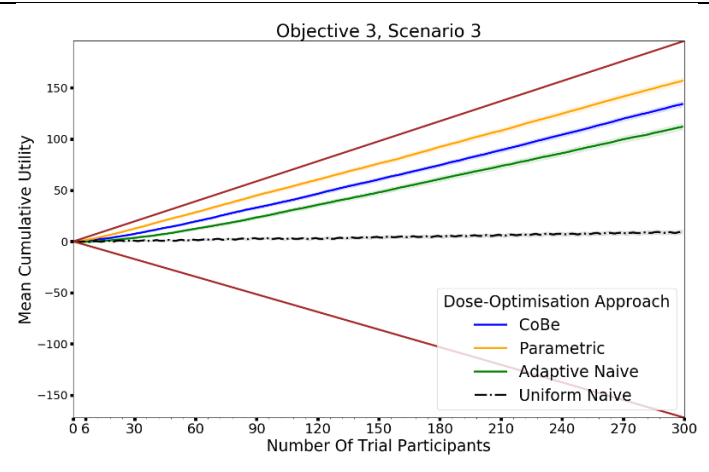
(f)
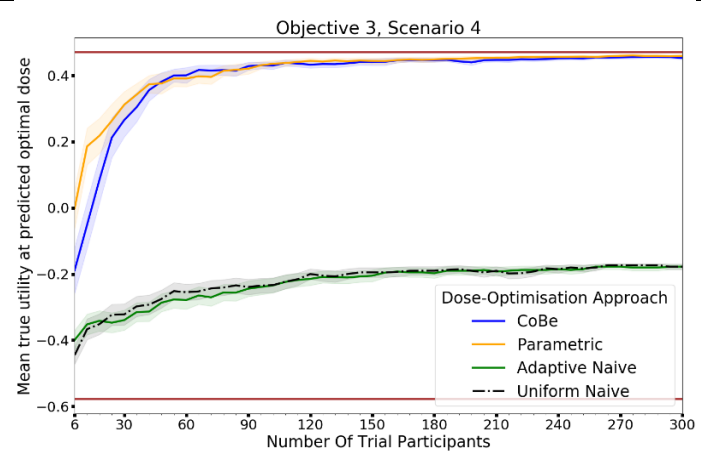
(g)
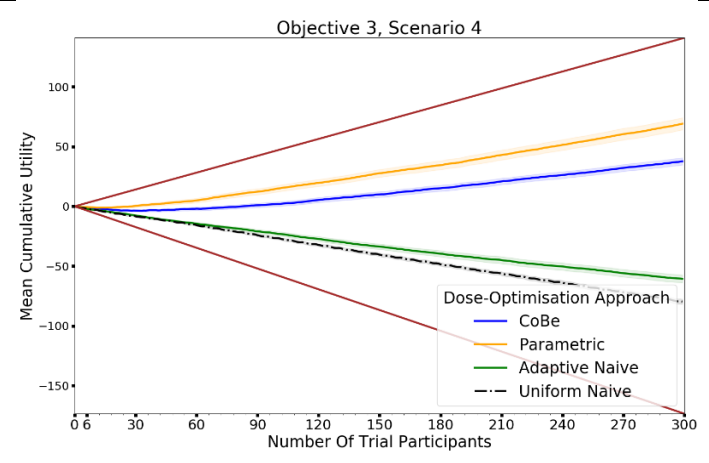
(h)
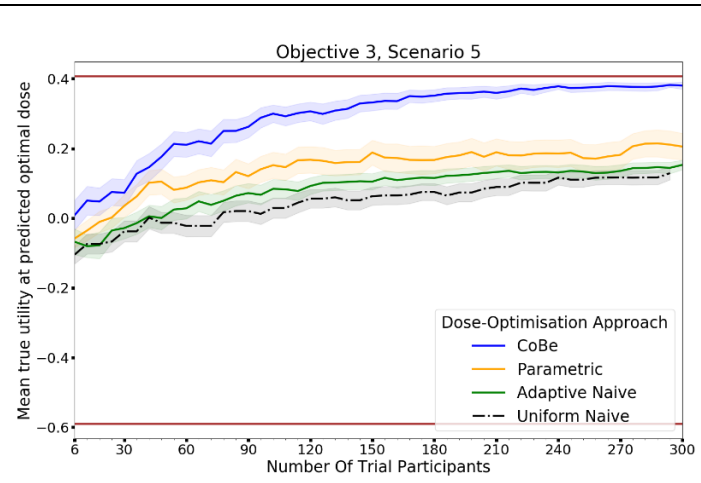
(i)
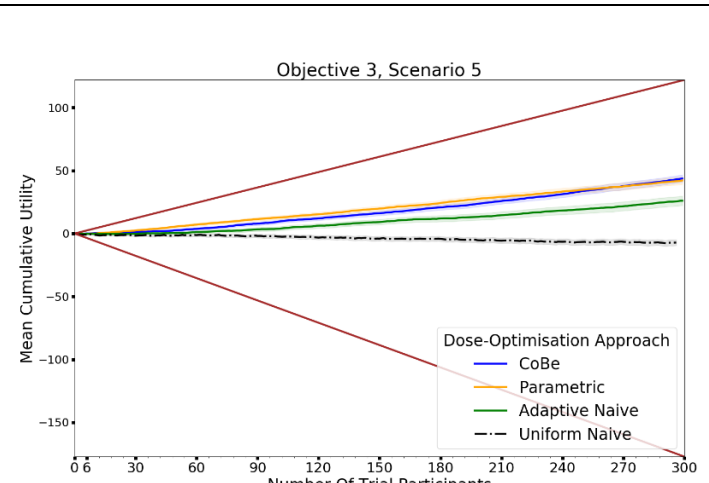
(j)
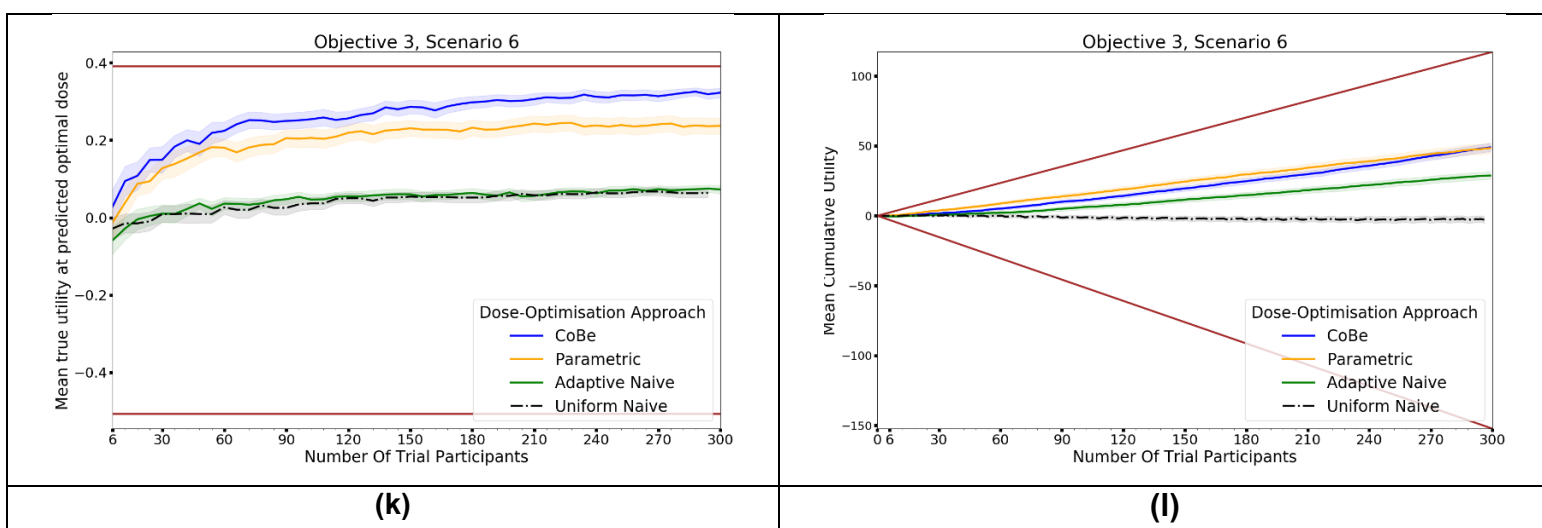
Figure 12. Mean true utility at the predicted optimal dose (left) and mean cumulative sum of utility (right) against trial size for all six objective 3 scenarios (top to bottom). These are the mean values and 95%CI values across 100 simulations. For the true utility at the predicted optimal dose plots (left), the brown lines show the minimum and maximum possible utility that could be achieved in that scenario. For the cumulative efficacy plots (right) the brown lines represent the maximum and minimum cumulative utility sum that could be expected for that scenario.

3.4. Objective 4. Evaluate the use of expert knowledge informed Continuous Correlated Beta Process priors for vaccine dose-optimisation

In this objective we assessed how the CoBe DOA that was discussed in objectives 1-3 would be impacted by the inclusion of expert priors as discussed in section 2.2.1.4. The black line in figure 13 ('No Prior') reflects the CoBe DOA as it was used in the previous objectives.

3.4.1. True efficacy/utility at predicted optimal dose

The DOA that best maximised 'true efficacy at predicted optimal dose' or 'true utility at predicted optimal dose' (from here called 'true efficacy/utility') for a given scenario and trial size was considered to be the 'best' DOA for the aim of selecting an optimal dose for that scenario. The left-hand side of figure 13 shows the change in mean true efficacy/utility with increasing numbers of trial participants for each of the four DOAs for each scenario, averaged across 100 simulated clinical trials. For each of these plots the upper and lower brown lines respectively show the maximal and minimal efficacy/utility possible for that scenario. A mean true efficacy/utility for a DOA being closer to the upper brown line relative to a second DOA indicates the first DOA being on average better at selecting a high efficacy/utility dose. Equivalently, a

276

mean true efficacy/utility being closer to the lower brown line would represent a DOA being on average worse at selecting a high efficacy/utility dose.

The CoBe DOAs that used 'Strong, Correct' and 'Very Strong, Correct' CCBP (continuous correlated beta process) priors had greater mean true efficacy/utility than the 'No Prior' CoBe DOA for all scenarios. The CoBe DOAs that used 'Strong, Incorrect' and 'Very Strong, Incorrect' CCBP priors had lower/worse mean true efficacy/utility than the 'No Prior' CoBe DOA for all scenarios. For all scenarios other than scenario 1, the CoBe DOA with the 'Very Strong, Incorrect' prior failed to predict a near optimal dose, even for large numbers of trial participants.

For the 'Very Strong, Correct' and 'Strong, Correct' CCBP prior DOAs, the mean true efficacy/utility decreased with the number of trial participants for early cohorts. This was expected as the expert prior is correct, therefore the initial predicted optimal dose is truly optimal and true efficacy/utility could not increase relative to this.

3.1.2. Cumulative sum of efficacy/utility

Cumulative sum of efficacy/utility measures a DOA's capacity to maximise the benefit to trial participants. After a certain number of trial participants, a DOA with a higher cumulative efficacy/utility would be considered 'more ethical' than a DOA with a lower cumulative efficacy/utility The right-hand side of figure 13 shows the change in mean cumulative utility with increasing numbers of trial participants for the CoBe DOAs with each of the five types of CCBP prior for each scenario, averaged across 100 simulated clinical trials. For each of these plots the upper and lower brown lines respectively show the theoretical maximal and minimal mean cumulative efficacy/utility possible for that scenario. A mean true efficacy/utility for a DOA being closer to the upper brown line relative to a second DOA reflects that the trial participants for simulated clinical trials using the first DOA on average received more efficacious or more efficacious/less toxic doses. No DOA could exceed this upper brown line, as this upper brown line reflects a DOA for which every trial participant receives the dose that is truly optimal.

The CoBe DOAs with 'Very Strong, Correct' and 'Strong, Correct' CCBP priors had greater cumulative efficacy/utility for all scenarios relative to the 'No Prior' CoBe

DOA [figure 13, RHS]. The CoBe DOAs with 'Very Strong, Incorrect' and 'Strong, Incorrect' CCBP priors had worse cumulative efficacy/utility for all scenarios relative to the 'No Prior' CoBe DOA.

(a) (b)

(c) (d)

(e) (f)

(g)

(h)

7.

(i)

(j)

(k)

(l)

279
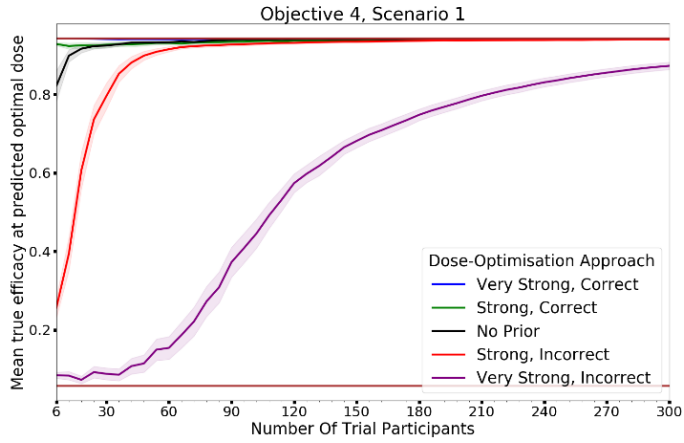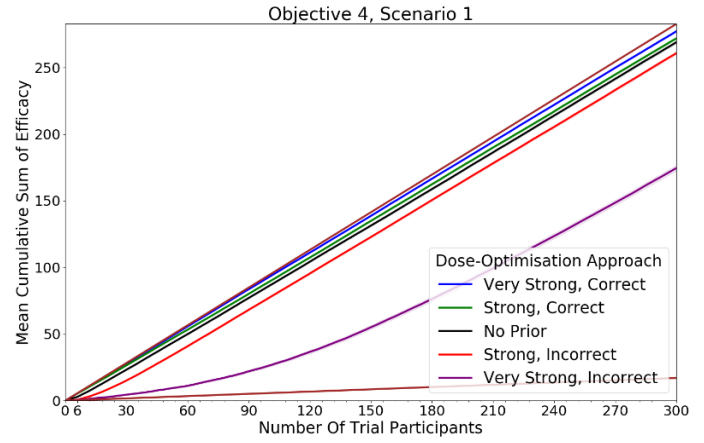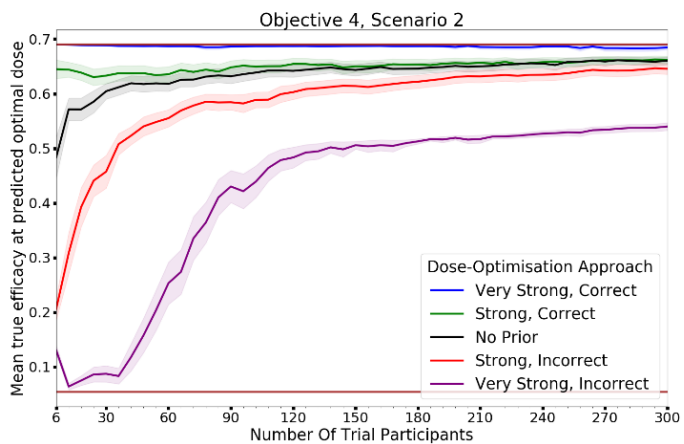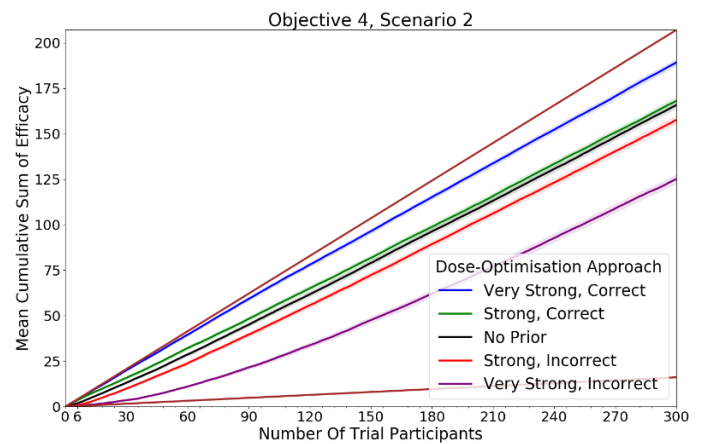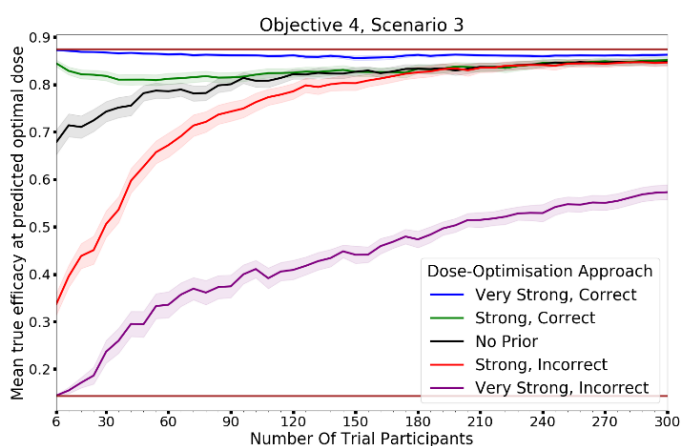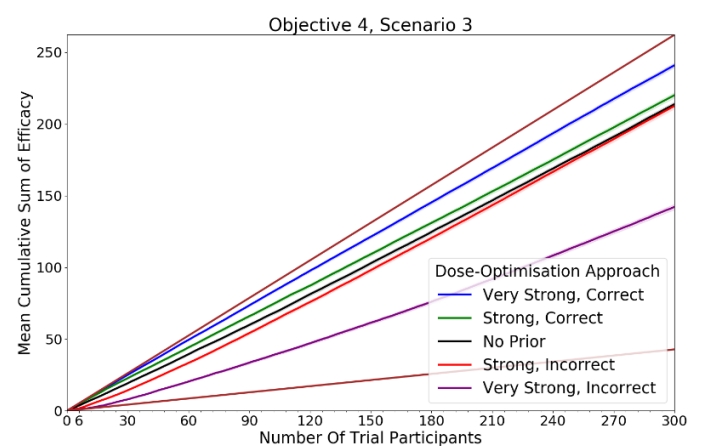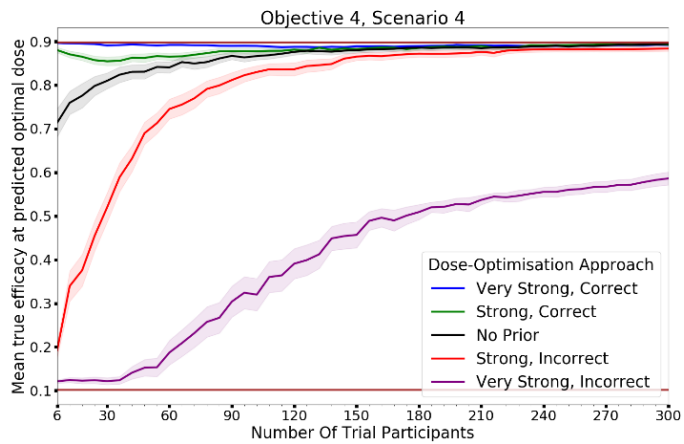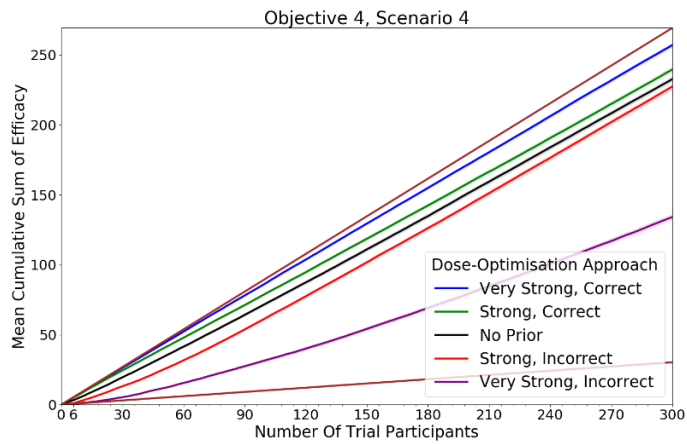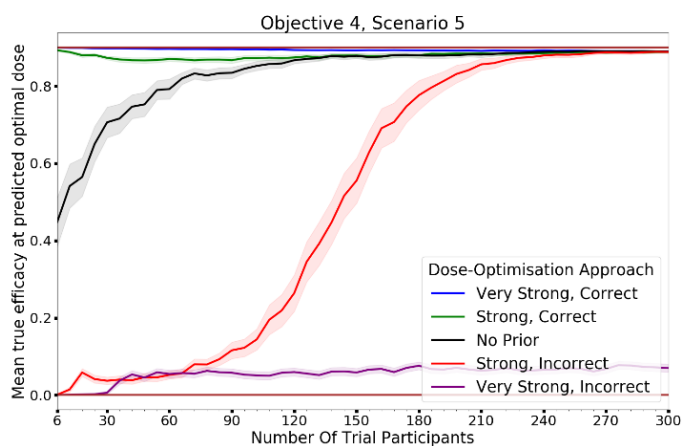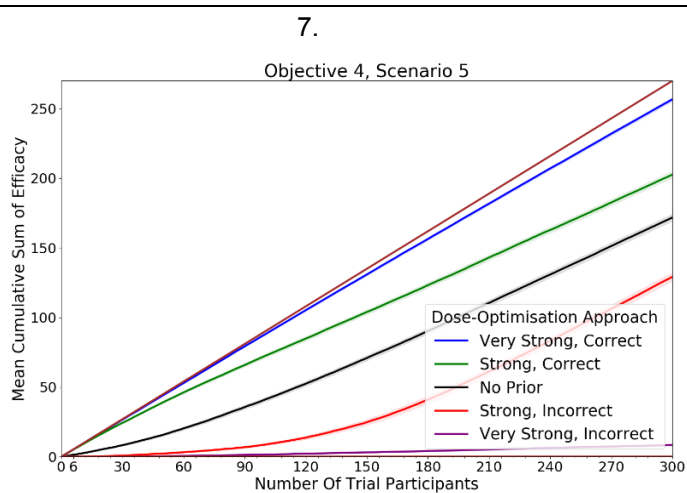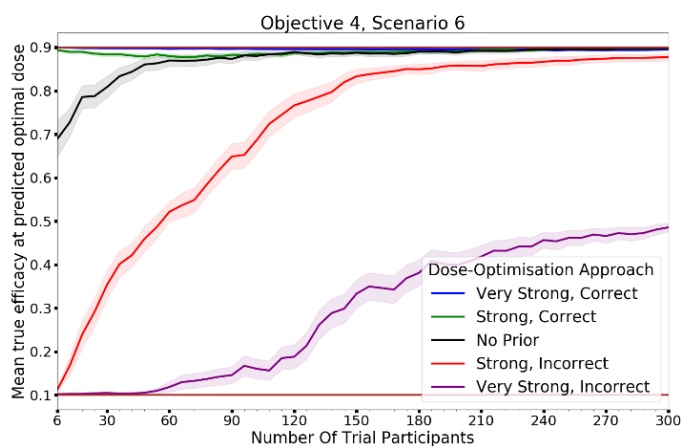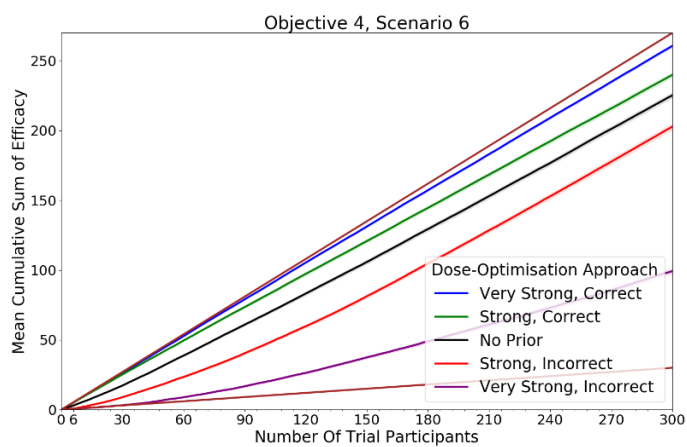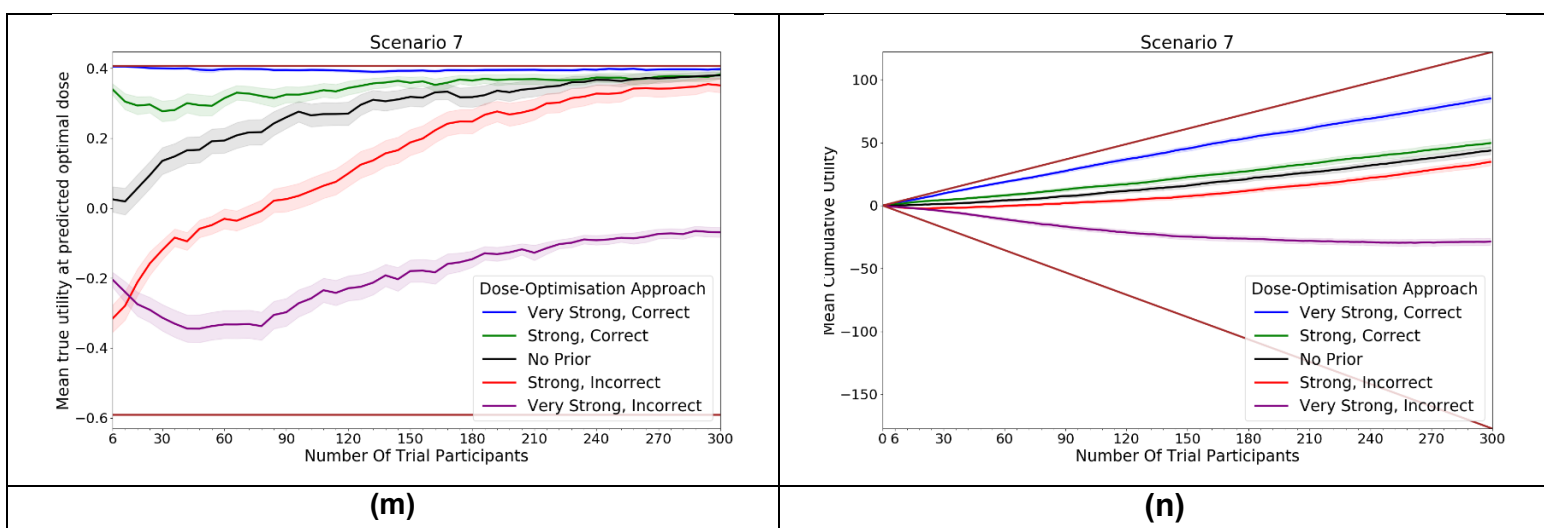
Figure 13. Mean true efficacy at the predicted optimal dose (S1-6, left), mean cumulative sum of efficacy (S1-6 6, right), mean true utility at the predicted optimal dose (S7, left), and mean cumulative sum of utility (S7, right) against trial size for all seven objective 4 scenarios (top to bottom). These are the mean values and 95%CI values across 100 simulations. For the true efficacy/utility at the predicted optimal dose plots (left), the brown lines show the minimum and maximum possible efficacy/utility that could be achieved in that scenario. For the cumulative efficacy/utility plots (right) the brown lines represent the maximum and minimum cumulative efficacy/utility sum that could be expected for that scenario.

4. Discussion

In this work we used simulation studies to evaluate the Correlated Beta (CoBe) dose optimisation approach (DOA), a novel mathematical-modelling methodology for selecting optimal vaccine dose that makes use of the non-parametric Continuous Correlated Beta Process (CCBP) model. We found that the CoBe DOA is effective and ethical for finding a vaccine dose which maximises efficacy and minimises toxicity for both single-administration and prime/boost administration regimens. The CoBe DOA typically had similar or preferable capacity to select optimal vaccine dose with maximal benefit to trial participants when compared to other DOAs. Additionally, this DOA can be further improved if there is correct and informative prior expert knowledge regarding vaccine dose-efficacy curves. This work suggests that mathematical modelling and adaptive trial design can lead to better vaccine dosing strategies. Further to this, non-parametric models and specifically the non-parametric CCBP model might be particularly useful if an appropriate parametric model is unknown. This may allow for more practical application of modelling in vaccine dose selection, accelerating vaccine development and saving lives.

This work is novel within the field of mathematical modelling-based vaccine dose selection, and the CCBP has also not previously been investigated for its potential to aid in optimising vaccine dose. The context in which we evaluated the methods was broad, as we conducted simulation studies that included single and prime/boost-administration vaccine dose-optimisation, both efficacy maximisation and toxicity minimisation, and the impact of incorporating expert opinion and knowledge into the modelling process.

The CoBe DOA has several strengths relative to the other DOAs discussed in this work. The CCBP models that were used to predict vaccine dose-response make simpler assumptions and are simpler to implement and interpret compared to other parametric/non-parametric mathematical models. Further, CCBP models can be extended to modelling prime/boost dose-response with only minor alterations and can be modified to benefit from expert knowledge with similar ease. The simplicity of the CCBP model did not hinder its ability to predict optimal vaccine dose in this simulation study, with CoBe DOA appearing equivalent or better at predicting optimal

dose relative to other common dose-optimisation approaches. As the CCBP model did not rely on any biologically-based assumptions, it could likely be generalised for the purposes of optimising vaccine dose, time between doses, adjuvant dose, or many other parameters related to vaccine administration that could impact vaccine efficacy and utility.

Beyond introducing and evaluating this new DOA, this work has several other strengths.

Firstly, we evaluated multiple other DOAs as part of this analysis, which highlights the potential strengths or weaknesses of these DOAs relative to each other. This analysis was conducted over many simulated scenarios, which showed that for different 'true' dose-efficacy/dose-toxicity relationships the performance of DOAs can vary. This work is also novel in highlighting that the potential efficacy/utility of a vaccine may be limited if dose is selected using a DOA that only considers a small number of potential doses, such as the Adaptive Naive and Uniform Naive DOAs. For example, for many scenarios (objective 1 scenario 3, objective 2 scenarios 1,4, and 6, and all objective 3 scenarios), the Adaptive Naive and Uniform Naive DOAs failed to find the true optimal dose as none of the small doses that were considered by these DOAs were optimal for those scenarios.

Finally, this work showed further evidence that using mathematical modelling and/or adaptive design may be both more effective for selecting optimal vaccine dose and more ethical than the 'Uniform Naive' DOA, which is equivalent to the standard approach in vaccine dose ranging trials.

There were weaknesses with both this work and the CoBe DOA that we proposed.

Firstly, this work and its findings are based on simulated clinical trial data, not empirical data. If none of the scenarios are accurate approximations to the true dose-efficacy of a vaccine, then the findings and recommendations of this work may not be relevant. We accounted for this weakness by investigating many qualitatively different scenarios.

Secondly, we included only binary measures of efficacy/toxicity. In practice, there may be non-binary outcomes of interest, for example CD4+ T Cell percentage or antibody titres. The CCBP model used for the CoBe DOA can only be used for binary responses, and further work would be needed to investigate similar DOAs for non-binary responses. Vaccine dose-optimisation based on non-binary responses would require more complicated utility functions and models, but this was beyond the scope of this work.

Thirdly we ignored some features that are commonly used in adaptive trial design and may be practically desirable. For example, stopping rules, which are criteria that allow for dose-finding trials to end early if there is sufficient evidence to suggest that one dose is optimal [59]. Additionally, we ignored escalation/de-escalation criteria, which are criteria that limit trial doses to a sub-range of the dosing domain until there is sufficient evidence to support escalating to larger and potentially more toxic doses [30,60]. Both of these trial design features would have added complexity to the implementation of the DOAs. Given that all evaluated DOAs did not include these features, we believe that our results were not biased by this weakness. Further work may need to be conducted to evaluate the effects of including stopping rules or escalation/de-escalation criteria. Similarly, to reduce the scope and complexity of this work, we also only compared the CoBe DOA against three other DOAs. Further work could conduct a comparison with other DOAs, for example rule-based designs [61], or the EffTox [34,35] or Bayesian Optimal Design algorithm [24].

We only investigated one parametric model each for single-administration dose-efficacy, single-administration dose-toxicity, prime/boost administration dose-efficacy, prime/boost administration dose-toxicity, and prime/boost/second-boost dose-efficacy. It is possible that the results for the parametric DOA for some scenarios would have been different if we had chosen different parametric models. For example, for scenario 6 of objective 1, the parametric DOA had a low mean true efficacy relative to the other DOAs after 300 trial participants. It is possible that this was due to that parametric model being misspecified for that scenario, and that a parametric DOA which used a parametric model of undulating dose-response would have been more effective in that scenario. Whilst this may have impacted our results, it also highlights why we believe non-parametric models may have potential for use

283

in vaccine dose-optimisation, as these may have reduced risk of choosing a model that may negatively impact selection of optimal vaccine dose compared to parametric models.

To reduce complexity, we did not evaluate in the main body of this work the effects of changes in cohort size, type of CCPB kernel, the CCBP kernel length hyperparameter values, the 'temperature' used for the SoftMax selection for the Parametric DOA, and the number of doses used for the Uniform Naive and Adaptive Naive DOAs. We do however provide an evaluation of the effects of changing in the supplementary materials [S7]. Additionally, whilst we showed that it is likely that mathematical modelling and adaptive trial design may lead to selection of more optimal vaccine doses and improve benefit to trial participants, these methods may increase the complexity and duration of conducting vaccine dose-ranging trials. Vaccine developers may need to consider whether this complexity is justified by the potential benefits of more ethical trials and improved clinical vaccine doses.

Comparing our work to previous findings, other DOAs that used non-parametric models were found to be as effective as using parametric modelling based DOAs for the selection of optimal dose [24,25], which is consistent with our findings. The use of mathematical modelling methods and adaptive design has previously been found to lead to more ethical clinical trials [17,62]. This is consistent with our finding that the Uniform Naive DOA, which represents a non-modelling approach, had the lowest cumulative sums of efficacy/utility across most scenarios. This work aligns well with other work in the field of Immunostimulation/Immunodynamic (IS/ID) modelling, which has proposed increased adoption of mathematical modelling for the purposes of optimal vaccine dosing.

There are several future possible areas for research. Firstly, while this work shows a theoretical validation of the CoBe DOA, clearly empirical validation of these methods is required. Secondly, further development and extension of the CoBe DOA would be beneficial to allow for uptake of these methodologies into clinical use. A quantitative method for selecting length hyperparameters for the CCBP model would be beneficial, and is discussed in [28,29]. Whilst inclusion of expert knowledge into CCBP models was evaluated in objective 4, a validated method of extrapolating from

animal vaccine dose-efficacy/dose-toxicity data to inform priors for these models would be highly beneficial considering that preclinical data are often used in the design of human dose-ranging trials. This may be important given the findings of objective 4, as for all scenarios the incorrect priors were detrimental to both selection of optimal dose and maximising benefit to trial participants. Further, the evaluation of CoBe DOA using more complicated utility functions of multiple immunological/toxicological responses is needed and may be more informative. Developing computational software such as an R or Python package may also be beneficial for allowing practical application of the CoBe DOA.

There is also potential for future research into other non-parametric models for the purpose of vaccine dose optimisation. Extensions of the CCBP or other non-parametric models should be investigated for modelling continuous or ordinal dose-response data. Additionally, we considered a homogenous trial population in this work. Developing methods of accounting for heterogeneity in the clinical trial population, either through trial participant randomisation or through augments to the models, could be important to ensuring maximal vaccine benefit.

5. Conclusions

Selection of optimal vaccine dose is an important but complicated endeavour. In this work we evaluated a novel approach for selecting optimal vaccine dose using the non-parametric CCBP model and adaptive trial design. Using mathematical models and/or adaptive design may lead to more effective and ethical vaccine dose-finding clinical trials, even if the shape of the dose-efficacy curve is unknown. These methods may also maximise benefit to vaccine clinical trial participants. This is the first novel investigation of modelling-based vaccine dose-optimisation approaches when compared to non-modelling vaccine dose-optimisation approaches. If developed further and implemented into vaccine clinical trials, mathematical modelling could accelerate vaccine development and save lives.

References

1. Kaur, R.J.; Dutta, S.; Bhardwaj, P.; Charan, J.; Dhingra, S.; Mitra, P.; Singh, K.; Yadav, D.; Sharma, P.; Misra, S. Adverse Events Reported From COVID-19 Vaccine Trials: A Systematic Review. *Indian J Clin Biochem* **2021**, *36*, 427–439, doi:10.1007/s12291-021-00968-z.

2. Benest, J.; Rhodes, S.; Quaife, M.; Evans, T.G.; White, R.G. Optimising Vaccine Dose in Inoculation against SARS-CoV-2, a Multi-Factor Optimisation Modelling Study to Maximise Vaccine Safety and Efficacy. *Vaccines* **2021**, *9*, 78, doi:10.3390/vaccines9020078.

3. Du, Z.; Wang, L.; Pandey, A.; Lim, W.W.; Chinazzi, M.; Piontti, A.P. y.; Lau, E.H.Y.; Wu, P.; Malani, A.; Cobey, S.; et al. Modeling Comparative Cost-Effectiveness of SARS-CoV-2 Vaccine Dose Fractionation in India. *Nat Med* **2022**, *28*, 934–938, doi:10.1038/s41591-022-01736-z.

4. Dolgin, E. Could Computer Models Be the Key to Better COVID Vaccines? *Nature* **2022**, *604*, 22–25, doi:10.1038/d41586-022-00924-8.

5. Aouni, J.; Bacro, J.N.; Toulemonde, G.; Colin, P.; Darchy, L. Utility-Based Dose Selection for Phase II Dose-Finding Studies. *Ther Innov Regul Sci* **2021**, *55*, 818–840, doi:10.1007/s43441-021-00273-0.

6. Bretz, F.; Dette, H.; Pinheiro, J. Practical Considerations for Optimal Designs in Clinical Dose Finding Studies. *Stat Med* **2010**, *29*, 10.1002/sim.3802, doi:10.1002/sim.3802.

7. Han, S. Clinical Vaccine Development. *Clin Exp Vaccine Res* **2015**, *4*, 46–53, doi:10.7774/cevr.2015.4.1.46.

8. David, S.; Kim, P.Y. Drug Trials. In *StatPearls*; StatPearls Publishing: Treasure Island (FL), 2022.

9. Commissioner, O. of the Step 3: Clinical Research. *FDA* **2019**.

10. The 5 Stages of COVID-19 Vaccine Development: What You Need to Know About How a Clinical Trial Works Available online: https://www.jnj.com/innovation/the-5-stages-of-covid-19-vaccine-development-what-you-need-to-know-about-how-a-clinical-trial-works (accessed on 31 July 2022).

11. Mhaskar, R.; Miladinovic, B.; Guterbock, T.M.; Djulbegovic, B. When Are Clinical Trials Beneficial for Study Patients and Future Patients? A Factorial Vignette-Based Survey of Institutional Review Board Members. *BMJ Open* **2016**, *6*, e011150, doi:10.1136/bmjopen-2016-011150.

12. Rhodes, S.J.; Knight, G.M.; Kirschner, D.E.; White, R.G.; Evans, T.G. Dose Finding for New Vaccines: The Role for Immunostimulation/Immunodynamic Modelling. *J Theor Biol* **2019**, *465*, 51–55, doi:10.1016/j.jtbi.2019.01.017.

13. Handel, A.; Li, Y.; McKay, B.; Pawelek, K.A.; Zarnitsyna, V.; Antia, R. Exploring the Impact of Inoculum Dose on Host Immunity and Morbidity to Inform Model-Based Vaccine Design. *PLOS Computational Biology* **2018**, *14*, e1006505, doi:10.1371/journal.pcbi.1006505.

14. Lalonde, R.L.; Kowalski, K.G.; Hutmacher, M.M.; Ewy, W.; Nichols, D.J.; Milligan, P.A.; Corrigan, B.W.; Lockwood, P.A.; Marshall, S.A.; Benincosa, L.J.; et al. Model-Based Drug Development. *Clinical Pharmacology & Therapeutics* **2007**, *82*, 21–32, doi:10.1038/sj.clpt.6100235.

15. Milligan, P.A.; Brown, M.J.; Marchant, B.; Martin, S.W.; van der Graaf, P.H.; Benson, N.; Nucci, G.; Nichols, D.J.; Boyd, R.A.; Mandema, J.W.; et al. Model-Based Drug Development: A Rational Approach to Efficiently Accelerate Drug Development. *Clinical Pharmacology & Therapeutics* **2013**, *93*, 502–514, doi:10.1038/clpt.2013.54.

16. Onar, A.; Kocak, M.; Boyett, J.M. Continual Reassessment Method vs. Traditional Empirically-Based Design: Modifications Motivated by Phase I Trials in Pediatric Oncology by the Pediatric Brain Tumor Consortium. *J Biopharm Stat* **2009**, *19*, 437–455, doi:10.1080/10543400902800486.

17. Pallmann, P.; Bedding, A.W.; Choodari-Oskooei, B.; Dimairo, M.; Flight, L.; Hampson, L.V.; Holmes, J.; Mander, A.P.; Odondi, L.; Sydes, M.R.; et al. Adaptive Designs in Clinical Trials: Why Use Them, and How to Run and Report Them. *BMC Medicine* **2018**, *16*, 29, doi:10.1186/s12916-018-1017-7.

18. Meurer, W.J.; Legocki, L.; Mawocha, S.; Frederiksen, S.M.; Guetterman, T.C.; Barsan, W.; Lewis, R.; Berry, D.; Fetters, M. Attitudes and Opinions Regarding Confirmatory Adaptive Clinical Trials: A Mixed Methods Analysis from the Adaptive Designs Accelerating Promising Trials into Treatments (ADAPT-IT) Project. *Trials* **2016**, *17*, 373, doi:10.1186/s13063-016-1493-z.

19. Rhodes, S.J.; Zelmer, A.; Knight, G.M.; Prabowo, S.A.; Stockdale, L.; Evans, T.G.; Lindenstrøm, T.; White, R.G.; Fletcher, H. The TB Vaccine H56+IC31 Dose-Response Curve Is Peaked Not Saturating: Data Generation for New Mathematical Modelling Methods to Inform Vaccine Dose Decisions. *Vaccine* **2016**, *34*, 6285–6291, doi:10.1016/j.vaccine.2016.10.060.

20. Benest, J.; Rhodes, S.; Afrough, S.; Evans, T.; White, R. Response Type and Host Species May Be Sufficient to Predict Dose-Response Curve Shape for Adenoviral Vector Vaccines. *Vaccines* **2020**, *8*, 155, doi:10.3390/vaccines8020155.

21. Benest, J.; Rhodes, S.; Evans, T.G.; White, R.G. Mathematical Modelling for Optimal Vaccine Dose Finding: Maximising Efficacy and Minimising Toxicity. *Vaccines (Basel)* **2022**, *10*, 756, doi:10.3390/vaccines10050756.

22. Schorning, K.; Bornkamp, B.; Bretz, F.; Dette, H. Model Selection versus Model Averaging in Dose Finding Studies. *Statistics in Medicine* **2016**, *35*, 4021–4040, doi:10.1002/sim.6991.
23. Bodin, E.; Kaiser, M.; Kazlauskaite, I.; Dai, Z.; Campbell, N.; Ek, C.H. Modulating Surrogates for Bayesian Optimization. In Proceedings of the Proceedings of the 37th International Conference on Machine Learning; PMLR, November 21 2020; pp. 970–979.
24. Takahashi, A.; Suzuki, T. Bayesian Optimization Design for Dose-Finding Based on Toxicity and Efficacy Outcomes in Phase I/II Clinical Trials. *Pharmaceutical Statistics* **2021**, *20*, 422–439, doi:10.1002/pst.2085.
25. Takahashi, A.; Suzuki, T. Bayesian Optimization for Estimating the Maximum Tolerated Dose in Phase I Clinical Trials. *Contemporary Clinical Trials Communications* **2021**, *21*, 100753, doi:10.1016/j.conctc.2021.100753.
26. Lin, R.; Yin, G. STEIN: A Simple Toxicity and Efficacy Interval Design for Seamless Phase I/II Clinical Trials. *Statistics in Medicine* **2017**, *36*, 4106–4120, doi:10.1002/sim.7428.
27. Goetschalckx, R.; Poupart, P.; Hoey, J. Continuous Correlated Beta Processes. In Proceedings of the Proceedings of the Twenty-Second international joint conference on Artificial Intelligence - Volume Volume Two; AAAI Press: Barcelona, Catalonia, Spain, July 16 2011; pp. 1269–1274.
28. Rolland, P.; Kavis, A.; Immer, A.; Singla, A.; Cevher, V. Efficient Learning of Smooth Probability Functions from Bernoulli Tests with Guarantees. In Proceedings of the Proceedings of the 36th International Conference on Machine Learning; PMLR, May 24 2019; pp. 5459–5467.
29. Gompert, Z. A Continuous Correlated Beta Process Model for Genetic Ancestry in Admixed Populations. *PLoS One* **2016**, *11*, e0151047, doi:10.1371/journal.pone.0151047.
30. Van Meter, E.M.; Garrett-Mayer, E.; Bandyopadhyay, D. Dose-Finding Clinical Trial Design for Ordinal Toxicity Grades Using the Continuation Ratio Model: An Extension of the Continual Reassessment Method. *Clin Trials* **2012**, *9*, 303–313, doi:10.1177/1740774512443593.
31. James, G.D.; Symeonides, S.; Marshall, J.; Young, J.; Clack, G. Assessment of Various Continual Reassessment Method Models for Dose-Escalation Phase 1 Oncology Clinical Trials: Using Real Clinical Data and Simulation Studies. *BMC Cancer* **2021**, *21*, 7, doi:10.1186/s12885-020-07703-6.
32. Morris, T.P.; White, I.R.; Crowther, M.J. Using Simulation Studies to Evaluate Statistical Methods. *Stat Med* **2019**, *38*, 2074–2102, doi:10.1002/sim.8086.
33. O'quigley, J. Non-Parametric Optimal Design in Dose Finding Studies. *Biostatistics* **2002**, *3*, 51–56, doi:10.1093/biostatistics/3.1.51.
34. Thall, P.F.; Cook, J.D. Dose-Finding Based on Efficacy-Toxicity Trade-Offs. *Biometrics* **2004**, *60*, 684–693, doi:10.1111/j.0006-341X.2004.00218.x.
35. Brock, K.; Billingham, L.; Copland, M.; Siddique, S.; Sirovica, M.; Yap, C. Implementing the EffTox Dose-Finding Design in the Matchpoint Trial. *BMC Medical Research Methodology* **2017**, *17*, 112, doi:10.1186/s12874-017-0381-x.
36. Diniz, M.A.; Tighiouart, M.; Rogatko, A. Comparison between Continuous and Discrete Doses for Model Based Designs in Cancer Dose Finding. *PLOS ONE* **2019**, *14*, e0210139, doi:10.1371/journal.pone.0210139.
37. Kardani, K.; Bolhassani, A.; Shahbazi, S. Prime-Boost Vaccine Strategy against Viral Infections: Mechanisms and Benefits. *Vaccine* **2016**, *34*, 413–423, doi:10.1016/j.vaccine.2015.11.062.
38. Lu, S. Heterologous Prime-Boost Vaccination. *Curr Opin Immunol* **2009**, *21*, 346–351, doi:10.1016/j.coi.2009.05.016.
39. Chen, J.; Liu, T. Statistical Considerations on Implementing the MCP-Mod Method for Binary Endpoints in Clinical Trials. *Contemp Clin Trials Commun* **2020**, *19*, 100641, doi:10.1016/j.conctc.2020.100641.
40. Raineri, E.; Dabad, M.; Heath, S. A Note on Exact Differences between Beta Distributions in Genomic (Methylation) Studies. *PLoS One* **2014**, *9*, e97349, doi:10.1371/journal.pone.0097349.
41. Lee, S.M.; Cheung, Y.K. Calibration of Prior Variance in the Bayesian Continual Reassessment Method. *Stat Med* **2011**, *30*, 2081–2089, doi:10.1002/sim.4139.
42. Thall, P.F.; Herrick, R.C.; Nguyen, H.Q.; Venier, J.J.; Norris, J.C. Effective Sample Size for Computing Prior Hyperaramaters in Bayesian Phase I-II Dose-Finding. *Clin Trials* **2014**, *11*, 657–666, doi:10.1177/1740774514547397.
43. Weisstein, E.W. Beta Distribution Available online: https://mathworld.wolfram.com/ (accessed on 5 August 2022).
44. Thompson, W.R. On the Likelihood That One Unknown Probability Exceeds Another in View of the Evidence of Two Samples. *Biometrika* **1933**, *25*, 285–294, doi:10.2307/2332286.

45. Aziz, M.; Kaufmann, E.; Riviere, M.-K. On Multi-Armed Bandit Designs for Dose-Finding Clinical Trials. *J. Mach. Learn. Res.* **2021**, *22*, 14:686-14:723.

46. Shen, C.; Wang, Z.; Villar, S.; Schaar, M.V.D. Learning for Dose Allocation in Adaptive Clinical Trials with Safety Constraints. In Proceedings of the Proceedings of the 37th International Conference on Machine Learning; PMLR, November 21 2020; pp. 8730–8740.

47. Agrawal, S.; Goyal, N. Analysis of Thompson Sampling for the Multi-Armed Bandit Problem. In Proceedings of the Proceedings of the 25th Annual Conference on Learning Theory; JMLR Workshop and Conference Proceedings, June 16 2012; p. 39.1-39.26.

48. O'Quigley, J.; Iasonos, A.; Bornkamp, B. Dose-Response Functions. In *Handbook of Methods for Designing, Monitoring, and Analyzing Dose-Finding Trials: Handbooks of Modern Statistical Methods*; 2017; p. 199 ISBN 978-1-315-15198-4.

49. Mandrekar, S.J.; Qin, R.; Sargent, D.J. Model-Based Phase I Designs Incorporating Toxicity and Efficacy for Single and Dual Agent Drug Combinations: Methods and Challenges. *Stat Med* **2010**, *29*, 1077–1083, doi:10.1002/sim.3706.

50. Zhang, W.; Sargent, D.J.; Mandrekar, S. An Adaptive Dose-Finding Design Incorporating Both Toxicity and Efficacy. *Statistics in Medicine* **2006**, *25*, 2365–2383, doi:10.1002/sim.2325.

51. Reverdy, P.; Leonard, N.E. Parameter Estimation in Softmax Decision-Making Models With Linear Objective Functions. *IEEE Transactions on Automation Science and Engineering* **2016**, *13*, 54–67, doi:10.1109/TASE.2015.2499244.

52. Vamplew, P.; Dazeley, R.; Foale, C. Softmax Exploration Strategies for Multiobjective Reinforcement Learning. *Neurocomputing* **2017**, *263*, 74–86, doi:10.1016/j.neucom.2016.09.141.

53. Cao, X.H.; Stojkovic, I.; Obradovic, Z. A Robust Data Scaling Algorithm to Improve Classification Accuracies in Biomedical Data. *BMC Bioinformatics* **2016**, *17*, 359, doi:10.1186/s12859-016-1236-x.

54. Linn, K.A.; Gaonkar, B.; Satterthwaite, T.D.; Doshi, J.; Davatzikos, C.; Shinohara, R.T. Control-Group Feature Normalization for Multivariate Pattern Analysis of Structural MRI Data Using the Support Vector Machine. *Neuroimage* **2016**, *132*, 157–166, doi:10.1016/j.neuroimage.2016.02.044.

55. van den Berg, R.A.; Hoefsloot, H.C.; Westerhuis, J.A.; Smilde, A.K.; van der Werf, M.J. Centering, Scaling, and Transformations: Improving the Biological Information Content of Metabolomics Data. *BMC Genomics* **2006**, *7*, 142, doi:10.1186/1471-2164-7-142.

56. Virtanen, P.; Gommers, R.; Oliphant, T.E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat Methods* **2020**, *17*, 261–272, doi:10.1038/s41592-019-0686-2.

57. Morita, S.; Thall, P.F.; Takeda, K. A Simulation Study of Methods for Selecting Subgroup-Specific Doses in Phase I Trials. *Pharm Stat* **2017**, *16*, 143–156, doi:10.1002/pst.1797.

58. Morita, S.; Thall, P.F.; Müller, P. Prior Effective Sample Size in Conditionally Independent Hierarchical Models. *Bayesian Anal* **2012**, *7*, 10.1214/12-BA720, doi:10.1214/12-BA720.

59. Iasonos, A.; Wilton, A.S.; Riedel, E.R.; Seshan, V.E.; Spriggs, D.R. A Comprehensive Comparison of the Continual Reassessment Method to the Standard 3 + 3 Dose Escalation Scheme in Phase I Dose-Finding Studies. *Clin Trials* **2008**, *5*, 465–477, doi:10.1177/1740774508096474.

60. Goodman, S.N.; Zahurak, M.L.; Piantadosi, S. Some Practical Improvements in the Continual Reassessment Method for Phase I Studies. *Statistics in Medicine* **1995**, *14*, 1149–1161, doi:10.1002/sim.4780141102.

61. Li, S.; Xie, X.-J.; Heitjan, D.F. Flexible, Rule-Based Dose Escalation: The Cohort-Sequence Design. *Contemporary Clinical Trials Communications* **2020**, *17*, 100541, doi:10.1016/j.conctc.2020.100541.

62. Millen, G.C.; Yap, C. Adaptive Trial Designs: What Is the Continual Reassessment Method? *Archives of Disease in Childhood - Education and Practice* **2021**, *106*, 175–177, doi:10.1136/archdischild-2019-316931.

Supplementary material for paper 5

The 'utility contour' utility function was used in the main body of the work to represent an example utility function that could be used to quantitatively define the trade-off between efficacy and toxicity. This utility function was first described by Thall and Cook in 2004[1]. This was further described by Brock in 2017 [2].

$$U(p_{eff}, p_{tox}) = 1 - ((\frac{1 - p_{eff}}{1 - anchor_{eff}})^{rho} - (\frac{1 - p_{tox}}{1 - anchor_{tox}})^{rho})^{\frac{1}{rho}}$$

With parameters $anchor_{eff}, anchor_{tox}, rho$. Clinicians are queried for the smallest value for the vaccine's efficacy probability such that, if the probability of toxicity for that vaccine was zero, the vaccine would see clinical use. This efficacy probability value is $anchor_{eff}$. Clinicians are queried for the maximum value for the vaccine's toxicity probability such that, if the probability of efficacy for that vaccine was one, the vaccine would see clinical use. This toxicity probability value is $anchor_{tox}$. Clinicians are then asked to suggest an efficacy probability and a toxicity probability p_{eff}^*, p_{tox}^* with $anchor_{eff} < p_{eff}^* < 1$ and $0 < p_{tox}^* anchor_{tox}$, where a vaccine with these efficacy and toxicity probabilities would also only be on the threshold of clinical viability.

These questions define three points in the efficacy-toxicity outcome space, $(anchor_{eff}, 0), (1, anchor_{tox}), (p_{eff}^*, p_{tox}^*)$

These three points then can be used to define the neutral-utility curve. That is to say, $U(anchor_{eff}, 0) = U(1, anchor_{tox}) = U(p_{eff}^*, p_{tox}^*) = 0$.

$U(anchor_{eff}, 0) = U(1, anchor_{tox}) = 0$ is guaranteed for all values of rho, so we solve numerically

$$U(p_{eff}^*, p_{tox}^*) = 1 - ((\frac{1 - p_{eff}^*}{1 - anchor_{eff}})^{rho} - (\frac{p_{tox}^*}{anchor_{tox}})^{rho})^{\frac{1}{rho}} = 0$$
to calculate rho.

In this work we used $anchor_{eff}, anchor_{tox}, rho$ parameters from [2]. These were $anchor_{eff} = 0.4, anchor_{tox} = 0.7, rho = 2.07$, based on the points $(0.4, 0), (1, 0.7), (0.5, 0.4)$. Whilst these were not elicited from clinicians in a vaccine context, we choose to use these as they had been previously used in a clinical setting. Figure S1 shows the utility contour that arises from these parameters.
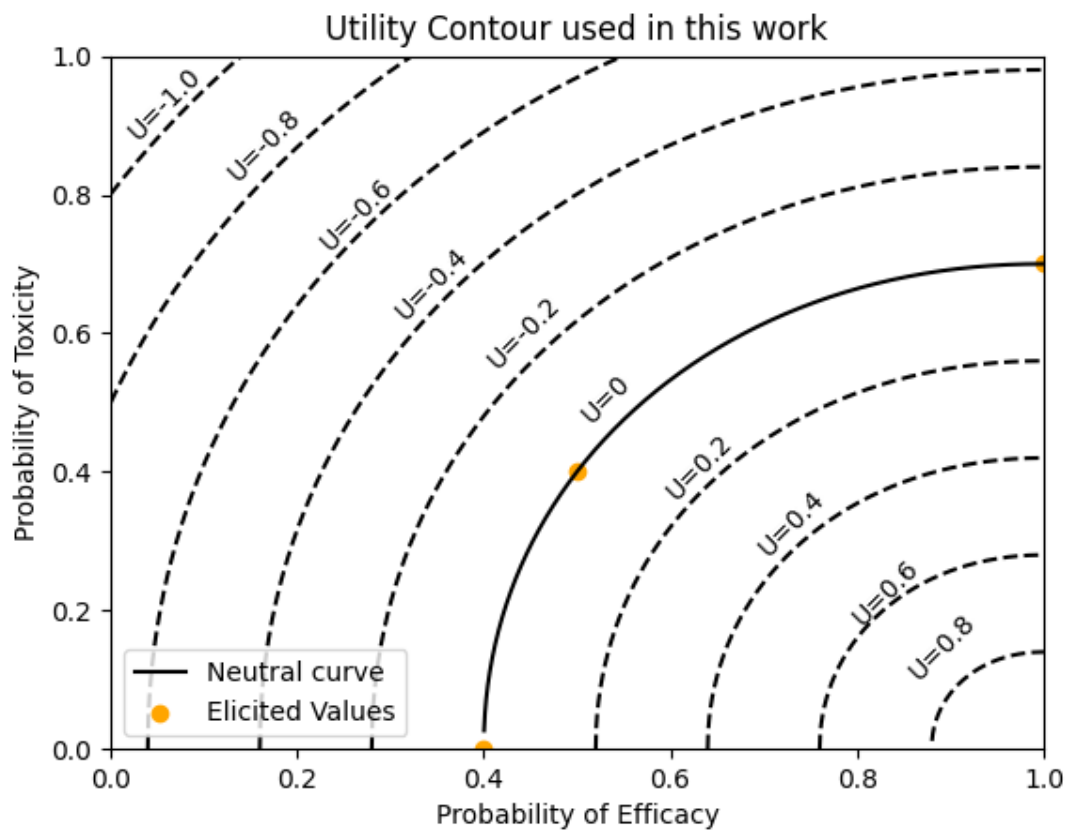
Figure S1. Utility contour used in this work. The thin black line shows the neutral utility curve. The orange dots show the three points used to define the neutral utility curve and function; (0.4, 0),(1,0.7), (0.5, 0.4).

S2 Length hyperparameters

In values were chosen using a geometric argument and validated empirically. See S7.1.2. For details.

S3 Pseudo-data

Pseudo-data or 'anchor points' were used to stabilise and inform models for which little real data are available. In short, we pretended in all Parametric DOAs that there

existed data which did not actually exist, but also down weighted these data during model calibration such that these data points were less important than real data.

S3.1. Efficacy Models

For single administration efficacy modelling, pseudo-data were of the form in table S.3.1.1. Thus, there were 30 pseudo individuals divided evenly over 3 doses. The weight of these datapoints was 0.05. Thus, the effective sample size of the pseudodata was 1.5 (=30x0.05), which was quickly minimal relative to the amount of real data. For example, after the first cohort the model would only be 20% influenced by these pseudo-data (1.5/(1.5+6) = 0.2) and hence 80% influenced by the 'real' (simulated) data.

Dose	Non-efficacy response	Efficacy Response
0.1	9	1
0.5	5	5
0.9	1	9

Table S.3.1.1. Efficacy pseudodata for single administration

For prime-boost administration efficacy modelling, pseudo-data were of the form in table S.3.1.2. Thus, there were 50 pseudo individuals divided evenly over 5 doses. The weight of these datapoints was 0.03. Thus, the effective sample size of the pseudodata was 1.5 (=50 x 0.03), which is quickly minimal relative to the amount of real data.

Dose	Non-efficacy response	Efficacy Response
0,0	9	1
0,1	5	5
1,0	5	5
1,1	1	9
0.5, 0.5	5	5

Table S.3.1.2. Efficacy pseudodata for prime/boost administration

For prime-boost administration efficacy modelling, pseudo-data were of the form in table S.3.1.3. Thus, there were 80 pseudo individuals divided evenly over 8 doses. The weight of these data points was 0.01875. Thus, the effective sample size of the pseudodata was 1.5 (=80x0.01875), which is quickly minimal relative to the amount of real data.

Dose	Non-efficacy response	Efficacy Response
0,0,0	1	9
0,0,1	5	5
0,1,0	5	5
0,1,1	5	5
1,0,0	5	5
1,0,1	5	5
1,1,0	5	5
1,1,1	9	1

Table S.3.1.3. Efficacy pseudodata for prime/boost administration

S3.2 Toxicity Model

For single administration toxicity modelling, pseudo-data were of the form in table S.3.2.1. Thus, there were 30 pseudo individuals divided evenly over 3 doses. The weight of these datapoints was 0.05. Thus, the effective sample size of the pseudodata was 1.5 (=30x0.05), which was quickly minimal relative to the amount of real data.

Dose	Non-toxic response	Toxic Response
0.1	9	1
0.5	5	5
0.9	1	9

Table S.3.2.1. Toxicity pseudodata for single administration

For prime-boost administration toxicity modelling, pseudo-data were of the form in table S.3.2.2. Thus, there were 50 pseudo individuals divided evenly over 5 doses. The weight of these datapoints was 0.03. Thus, the effective sample size of the pseudodata was 1.5 (=50x0.03), which was quickly minimal relative to the amount of real data.

Dose	Non-efficacy response	Efficacy Response
0,0	9	1
0,1	5	5
1,0	5	5
1,1	1	9
0.5, 0.5	5	5

Table S.3.2.2. Toxicity pseudodata for prime/boost administration

S4 SoftMax Selection Method

In this section of the supplementary we describe the SoftMax selection method of trial dose selection. We previously discussed this in [3] but replicate this description here.

Method

SoftMax selection is a method of action selection used commonly in both multi-armed bandit problems and reinforcement learning. We provide a description of action/dose selection under this method. Let A_1, A_2, ..., A_n be the n possible actions available to be taken, each with respective predicted utility U_1, U_2, ..., U_n. Then an action A_i is selected to evaluate (dose selected to trial) with probability

$$Probability\ of\ selecting\ action\ A_i = \frac{e^{inversetemperature \times U_i}}{\sum_{j=1}^{n} e^{inversetemperature \times U_j}}$$

where inverse_temperature is a hyperparameter which controls the degree of exploration. An increased inverse temperature leads to actions with lower predicted utility having lower probability of selection. For inverse_temperature = 0, which is the

lowest possible inverse_temperature, all actions are selected with equal probability (1/n). As inverse_temperature tends to infinity, this selection method tends to only select the action(s) with the maximum predicted utility. A random number generator is used to select an action with these probabilities.

inverse_temperature values

For the Parametric DOA, we used inverse_temperature = 6.9. This was chosen such that a predicted difference in utility of 0.1 would have a doubled probability of selection. This is shown

$$2e^{inversetemperature \times U_i} = e^{inversetemperature \times (U_i + 0.1)}$$

$$= e^{inversetemperature \times (U_i)} e^{inversetemperature \times (0.1)}$$

$$2 = e^{inversetemperature \times (0.1)}$$

$$\ln(2) = 0.69 = 0.1 \times inversetemperature$$

$$6.9 = inversetemperature$$

This may not have been optimal values, but optimal values are likely to vary depending on the scenario. This is further discussed in S7.4.

S5 Rescaled Dosing Domains

It may be reasonable to transform the 'raw' value of doses before modelling. This can help with stabilising computation [4]. These rescalings or transformations are of the form

$$x_i = transform(dose_i)$$

Where x_i is the transformed dose value that will be used for modelling and again $dose_i$ is the 'raw' dose value. In particular, using a log10 transform is common in drugs and in vaccines and would be given by

$$x_i = transform(dose_i) = log_{10}(dose_i)$$

In this work I assume that all prime/boost/second-boost doses had been transformed using a transform such that

$$0 = transform(dose_{min})$$

$$1 = transform(dose_{max})$$

Where $dose_{min}$ and $dose_{max}$ were respectively the smallest dose that could be given for the prime/boost/second-boost doses.

Two example transforms that would have this property would be the MinMax and Log10 MinMax transforms.

Minmax:

$$x_i = transform(dose_i) = \frac{dose_i - dose_{min}}{dose_{max} - dose_{min}}$$

Log10 Minmax:

$$x_i = transform(dose_i) = \frac{log_{10}(dose_i) - log_{10}(dose_{min})}{log_{10}(dose_{max}) - log_{10}(dose_{min})}$$

If $dose_{min}$=0, then a small constant may be added to ensure that the transform is well defined. One possibility is to choose this constant to be equal to the smallest positive (>0) dose in the dosing dimension.

$$x_i = transform(dose_i) = \frac{log_{10}(dose_i + dose_{leastnonzero}) - log_{10}(dose_{min} + dose_{leastnonzero})}{log_{10}(dose_{max} + dose_{leastnonzero}) - log_{10}(dose_{min} + dose_{leastnonzero})}$$

An alternative is to choose c = 1 - $dose_{min}$ and have

$$x_i = transform(dose_i) = \frac{log_{10}(dose_i + c) - log_{10}(dose_{min} + c)}{log_{10}(dose_{max} + c) - log_{10}(dose_{min} + c)}$$

Both would distort the data, but this may be justified by improved modelling stability.

S6 Scenario Creation

Simulation studies require researchers to define scenarios that are used to evaluate the approaches that the simulation study aims to investigate. In previous simulation

studies, which evaluated dose optimisation between only a small number of dosing groups, the efficacy probabilities and toxicity probabilities for each dosing group have been chosen by hand to create scenarios with specific qualitative features (e.g., peaking dose-efficacy curves). Little detail is typically given to how these are chosen, as with only a small number of dosing groups it is easy for readers to visualise.

In this work, scenarios could have 101, 441, or 1331 potential doses that needed to have defined efficacy and/or toxicity probabilities for single, prime/boost, and prime/boost/second-boost scenarios respectively. These were too many doses to choose by hand, so we had to use an algorithmic approach to setting true efficacy/toxicity probabilities for the doses in each scenario. We considered using parametric models or kernel-based algorithms to generate these probabilities, however we decided that using either parametric model or kernel based algorithms could bias our results. We choose to use a K-nearest neighbours style algorithm that involved iterative averaging, which produced smooth and continuous dose-response curves if the number of iterations was large enough,

This algorithm for took the following inputs:

- The prime doses that should be used in the iterative process
- Potentially the boost doses that should be used in the iterative process
- Potentially the second-boost doses that should be used in the iterative process
- Anchor Doses: Doses that we would like to set to a specific value in the scenario
- Anchor Probabilities: The Specific probabilities for the anchor doses.
- K, the number of neighbours for each dose that will be averaged. This increases the 'smoothness' of the resulting curve.
- The number of iterations of the algorithm to use.

For the algorithm, all doses (potential combinations of any prime and boost doses) were found. For each of these combinations, we set an initial probability of response of 0.5. Then the following iterative process was used:

296

For doses which were specified as anchor doses, set the probability of response equal to the respective anchor probability for that dose.

Set the probability of response for each dose equal to the mean value of its K-nearest neighbours, which was defined by Euclidean distance and included itself.

Repeat steps 1 and 2 as many times as specified.

By conducting a large number of iterations, the probability of response for each dose would be approximately the average of its surrounding doses, and there would be smooth dose-response curves between anchor points. Where we generated probabilities for doses that were outside of the [0,1] bounded dosing domain, these doses and respective probabilities were then excluded.

Rather than giving the probabilities for each dose for each scenario in tabular format, which we do not believe would be easy to parse, for each scenario we give the input parameters used to generate the dose-response curve. We hope that this show how we created scenarios with certain qualitative behaviours.

I only show the value for Objective 1 Scenario1 here for demonstration, the others are included in the appendix of this thesis [A.D.4]

Objective 1, Scenario 1 Efficacy

Inputs	Values
Prime doses	[-1.00, -0.99, …, 1.99, 2.00]
Anchor doses	[-0.5], [1.5]
Anchor probabilities	[0.05], [0.95]
K	21
Iterations	11

S7. Additional DOA Hyperparameters

S7.1. CoBe DOA: CCBP Kernel Type and Length Hyperparameters

The CoBe DOA uses a CCBP model of dose response, which relies on a kernel function to describe similarity between doses. In this work, we used the Squared Exponential kernel, as this was the kernel used by the original authors of the CCBP model. However, more recent work has suggested that other kernel functions should be considered, for example the 'SBC' kernel described by Rolland et al [5]. This was shown to outperform the squared exponential kernel for large amounts of data (>1000 data points). We chose to consider the squared exponential kernel due to considering typically <300 data points, but it was possible that the SBC kernel may have been superior. Additionally, both of these kernels require a length hyper parameter l to be specified. In the main body of the work, we used l = 0.2 for our squared exponential kernel, but again it is possible that this may have biased our results. We originally choose l = 0.2 as that would mean $K(d_i, d_i+0.15)0.5$, which seemed reasonable. We did not conduct a systematic optimisation of l prior to conducting the simulation studies, as this may have biased our findings in favour of the CoBe DOA. Additionally, we believed that it would be likely that the 'optimal' value for l would depend on the scenario.

We did however investigate of the effect that changing the kernel function or value of the length hyperparameter may have had after the main investigation was concluded. This was done by following the same methodology and scenarios as objective 1 of the main body of this work, but where all DOAs investigated were CoBe DOAs with one of two different kernel functions, each with 4 different potential length hyperparameters. We also investigated the 'uncorrelated' kernel described in the main body of the text. Results are shown in Figure S2.
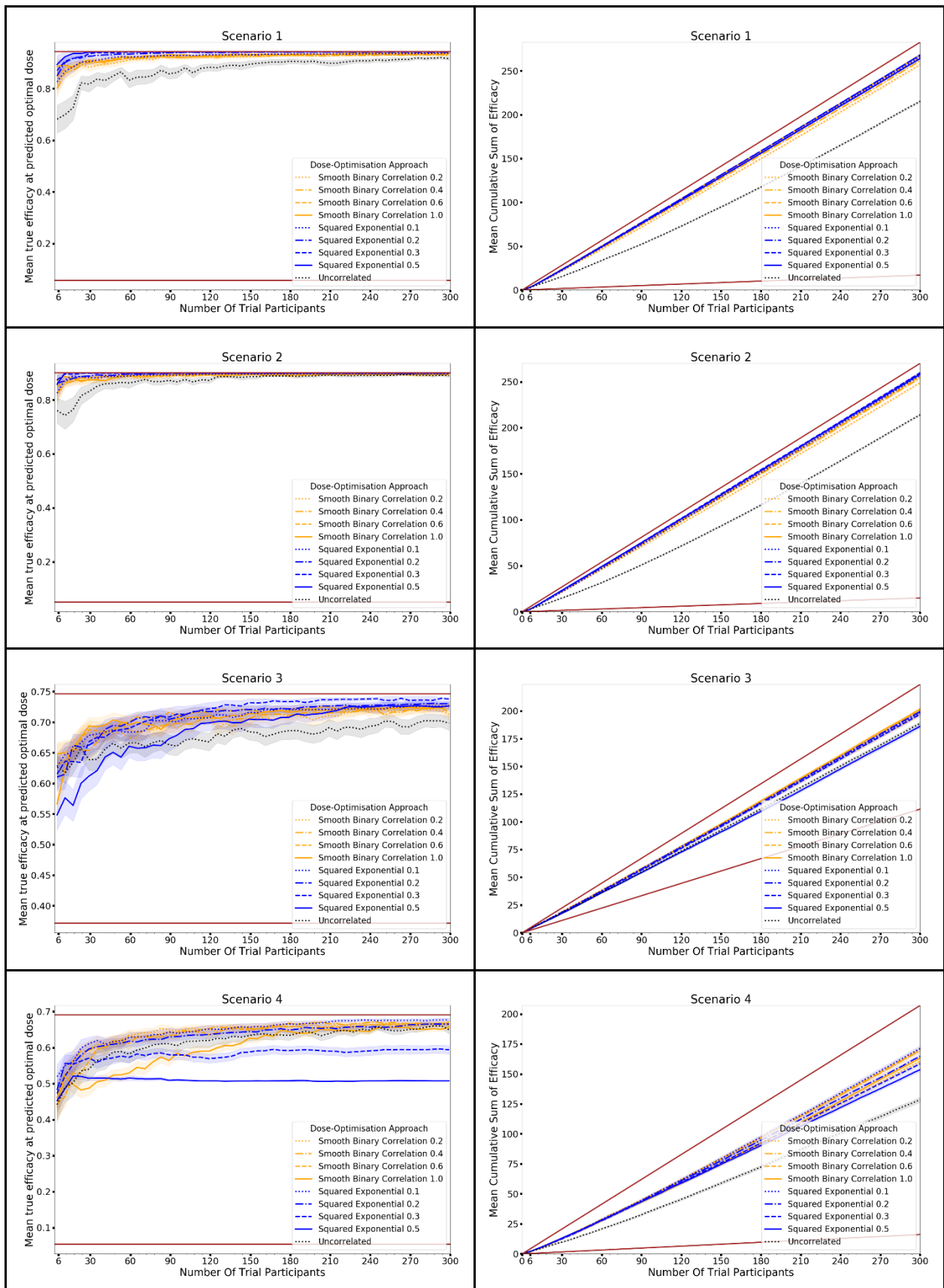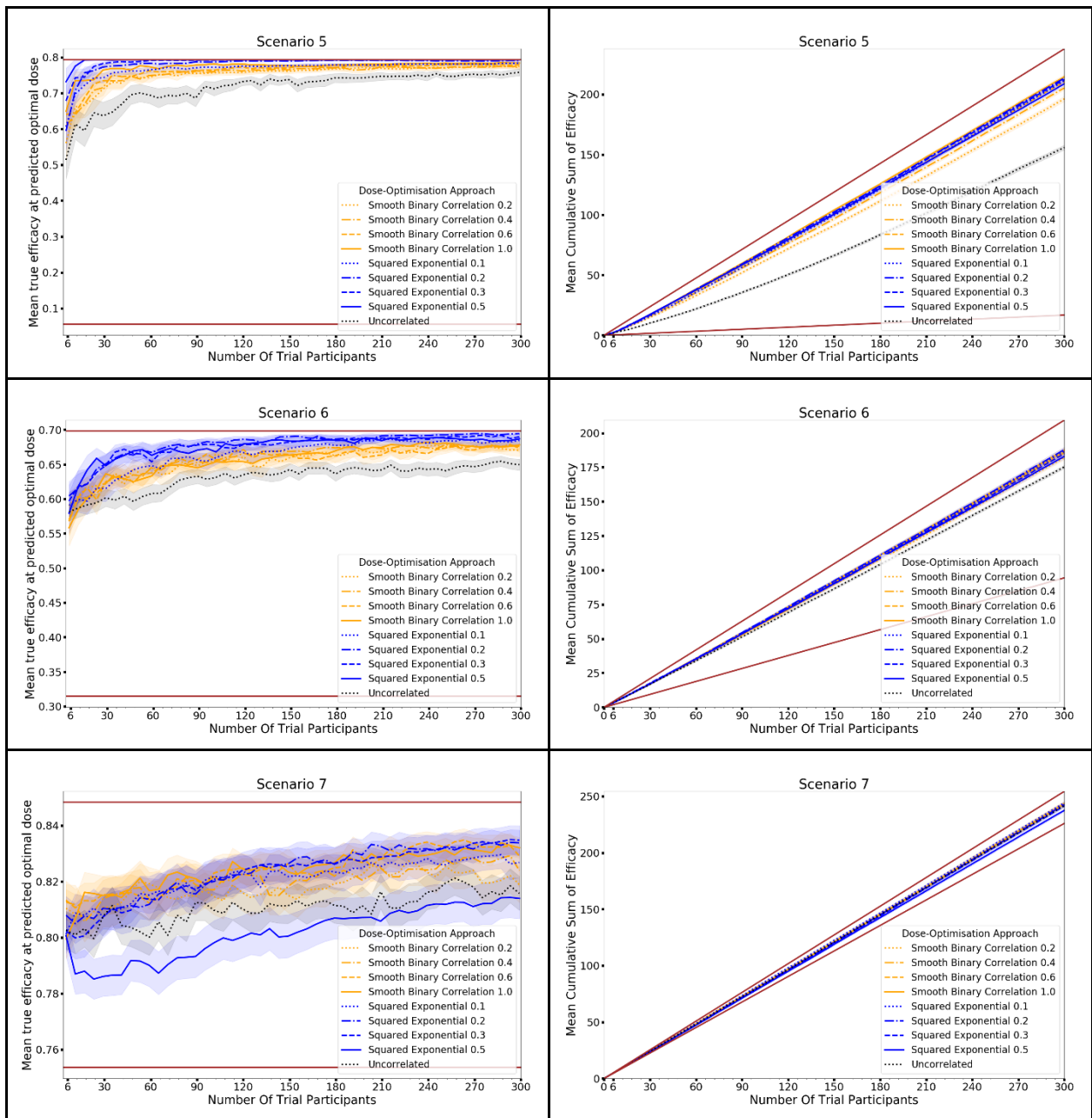
Figure S7.1. Mean true efficacy at the predicted optimal dose (left) and mean cumulative sum of efficacy (right) against trial size for all seven objective S7.1 scenarios (top to bottom). These are the mean values and 95%CI values across 100 simulations. For the true efficacy plots (left), the brown lines show the minimum and maximum possible efficacy that could be achieved in that scenario. For the cumulative efficacy plots (right) the brown lines represent the maximum and minimum cumulative efficacy sum that could be expected for that scenario.

These results suggest that using the SBC kernel rather than the squared exponential kernel would not have improved either metric for these scenarios. Additionally, using l=0.2 for the squared exponential appears to have been reasonable. l=0.5 performed well for some scenarios but was inconsistent and performed poorly for scenarios 3,4 and 7. As expected, the optimal value for l therefore depends on the true underlying

dose-response curve. That said, the CoBe DOA using the squared exponential kernel was often effective in each scenario with multiple different length hyperparameters. This may suggest that effective optimisation using the CoBe DOA does not require precision in the choice of length hyperparameter.

S7.2. CoBe DOA: CCBP Kernel Length Hyperparameters for Prime/Boost and Prime/Boost/Second-Boost administration

In this work the length parameters l=0.2 was used for modelling single-administration dose-response, $l_1=l_2=0.25$ were used for modelling prime/boost dose-response, and length parameters $l_1=l_2=l_3=0.4$ were used for modelling prime/boost/second-boost dose-response. We discussed the choice of l=0.2 above. Here we discuss why the values of the length hyperparameters for prime/boost and modelling prime/boost/second-boost were chosen. We note that these were not chosen based on systematic optimisation (to avoid advantaging the CoBe DOA relative to the other DOAs), but rather on geometrical principles.

Consider a 1-D dosing domain normalised such that the maximum dose is 1 and the minimum is 0. If a point is selected in the centre (0.5), then a 1-d sphere with radius = l = 0.2 around that point would have 40% of the dosing domain inside.

Consider a 2-D dosing domain normalised such that the maximum dose is 1 and the minimum is 0 for both dimensions. Let $l_1=l_2$. If the central point is selected in the centre ([0.5, 0.5]), then a 2-d sphere with radius l1 around that point would have $100\pi l_1^2\%$ of the dosing domain inside. For $l_1=0.2$, this would be 12.5%, so less of the dosing domain would be influenced per data point. Hence, we believed that it would be likely that we should choose $l_1=l_2>0.2$.

The choice of l1 such that 40% is inside is approximately .36. Hence choosing $l_1=l_2=0.36$ would approximately correspond to the same percentage of the dosing domain being in some sense 'strongly' influenced by a given data point in the centre of the dosing domain. However, using such a large value of l1 might limit optimisation of the prime or boost dose individually. We therefore choose to use some value between 0.2 and 0.36, so we choose 0.25.

301

A similar argument was used to choose $l_1=l_2=l_3=0.4$. Consider a 3-D dosing domain normalised such that the maximum dose is 1 and the minimum is 0 for both dimensions. If the central point is selected in the centre ([0.5, 0.5, 0.5]), then a 3-d sphere with radius l_1 around that point would have $100\frac{4\pi}{3}l_1^3\%$ of the dosing domain inside. Hence the choice of l_1 such that 40% is inside would be 0.66. Again, we choose to choose a value between 0.2 and 0.66 as a compromise between optimising within the 3-D dosing domain and being able to optimise along each dimension individually.

We investigated of the effect that this may have on selection of maximum efficacy dose and benefit to trial participants. This was done by following the same methodology as objective 2 of the main body of this work, however I only investigated scenarios 1 and 6. Scenario 1 was used to investigate the choice of $l_1=l_2=0.25$ and scenario 6 was used to investigate the choice of $l_1=l_2=l_3=0.4$. All DOAs investigated were CoBe DOAs but with varying length hyper parameters. I show these results in Figure S3.
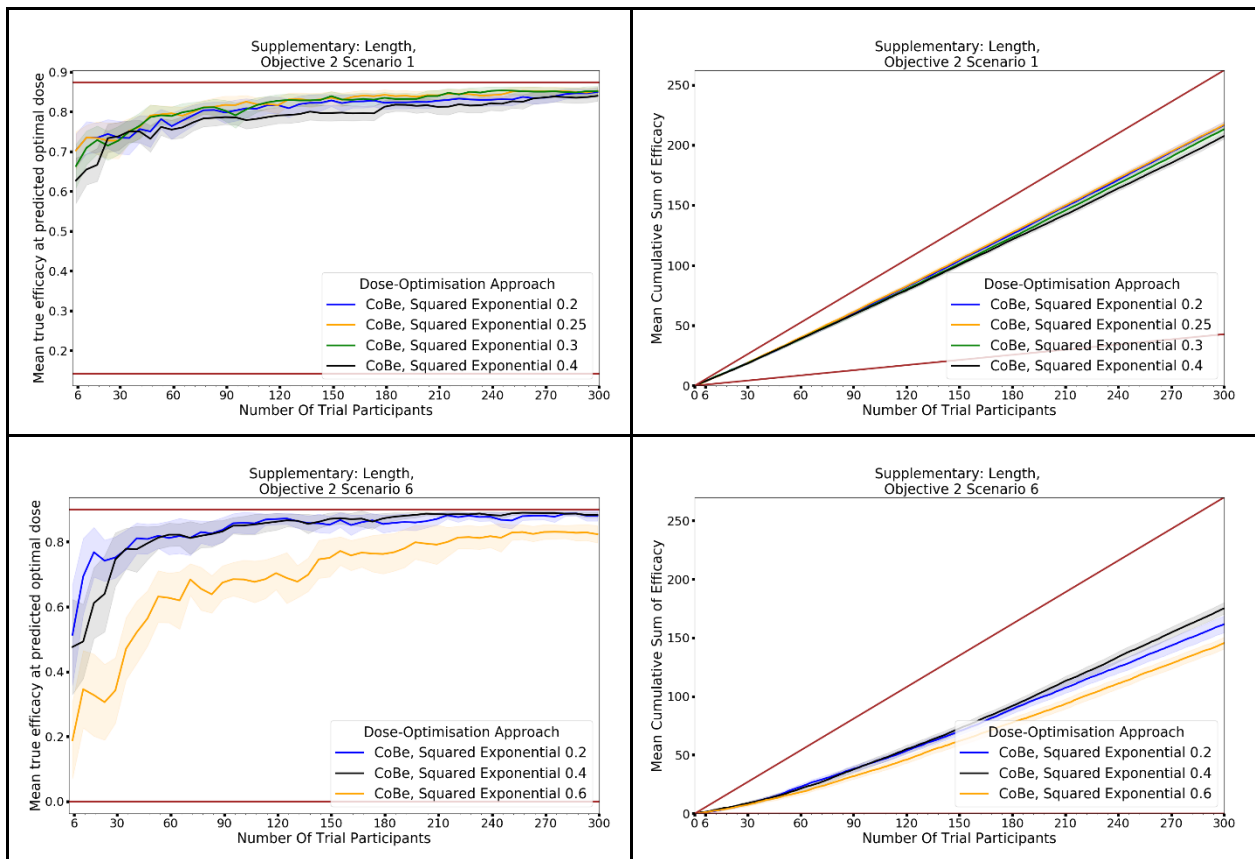
Figure S7.2. Mean true efficacy at the predicted optimal dose (left) and mean cumulative sum of efficacy (right) against trial size for all seven objective S7.2 scenarios (top to bottom). These are the mean values and 95%CI values across 100 simulations. For the true efficacy plots (left), the brown lines show the minimum and maximum possible efficacy that could be achieved in that scenario. For the cumulative efficacy plots (right) the brown lines represent the maximum and minimum cumulative efficacy sum that could be expected for that scenario.

These results suggest our choice of length hyperparameters were reasonable, though our initial concerns based on our geometric interpretation that $l_1=l_2=0.2$ or $l_1=l_2=l_3=0.2$ would be too small may not have been correct. $l_1=l_2=0.2$ and $l_1=l_2=l_3=0.2$ both performed well in this investigation, though this may not have been the case for other scenarios. It also suggests that the performance of the CoBe DOA may not be that sensitive to the choice of length hyperparameter.

S7.3. CoBe DOA: Cohort Size

For the CoBe DOA, and indeed all adaptive trial designs, clinicians must choose how to divide their total trial population into trial cohorts. In this work the N total trial participants were divided into C cohorts of cohort size b, with b = N/C. In the main body of the text the CoBe DOA used N = 300, C = 50, b = 6. There is clearly a trade-off to be made in terms of the size of b. As b decreases, C increases, meaning that a

larger number of cohorts would be required before the trial can be concluded. As b increases however, the DOA becomes less adaptive, with b = N (C=1) being not adaptive at all. Decreasing b increases the time taken to conduct a trial, increasing b may reduce the benefits of adaptive design that we have discussed in this work (increased benefit to trial participants and improved capacity to select optimal dose).

We investigated of the effect that this may have on selection of maximum efficacy dose and benefit to trial participants. This was done by following the same methodology and scenarios as objective 1 of the main body of this work, but where all DOAs investigated were CoBe DOAs with six different values for b. Results are shown in figure S4

Supplementary: Cohort Size, Objective 1 Scenario 3

Supplementary: Cohort Size, Objective 1 Scenario 4

Supplementary: Cohort Size, Objective 1 Scenario 5

Supplementary: Cohort Size, Objective 1 Scenario 6
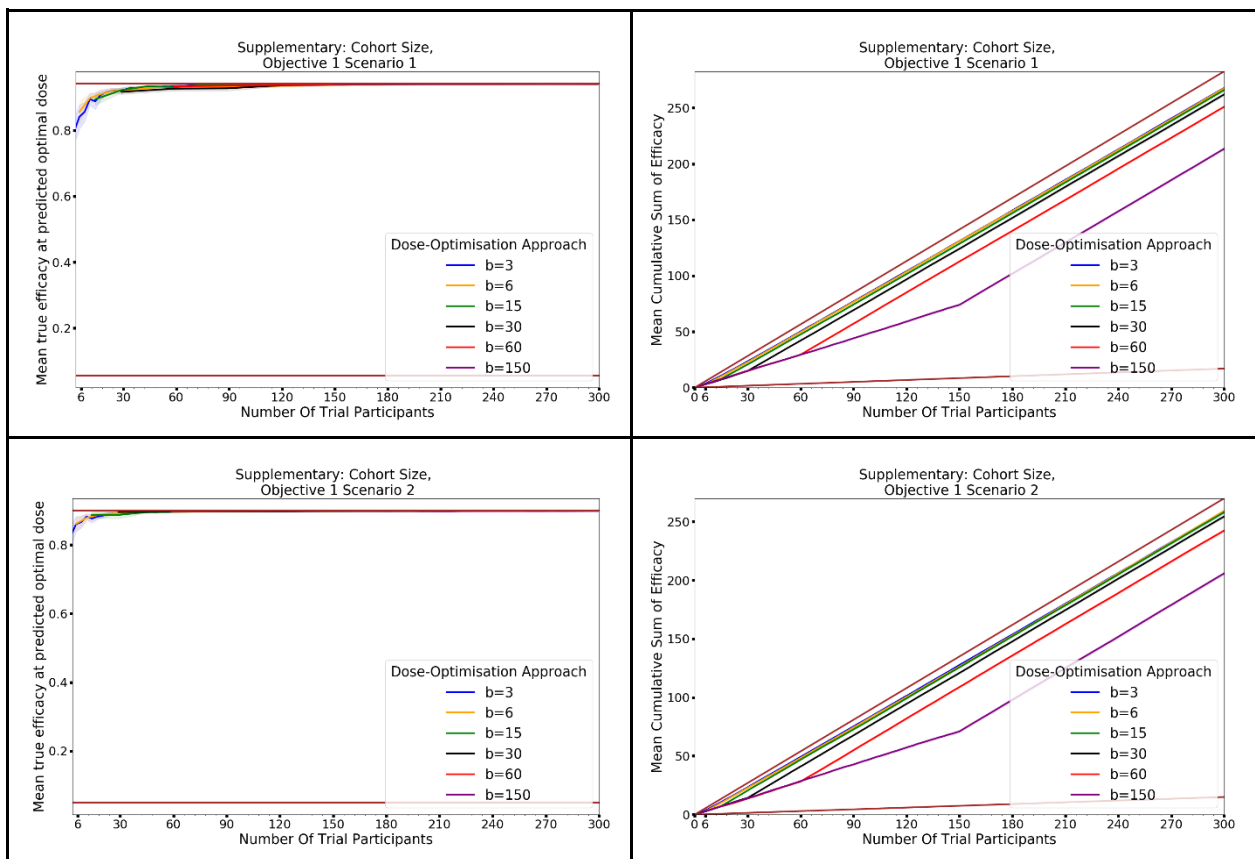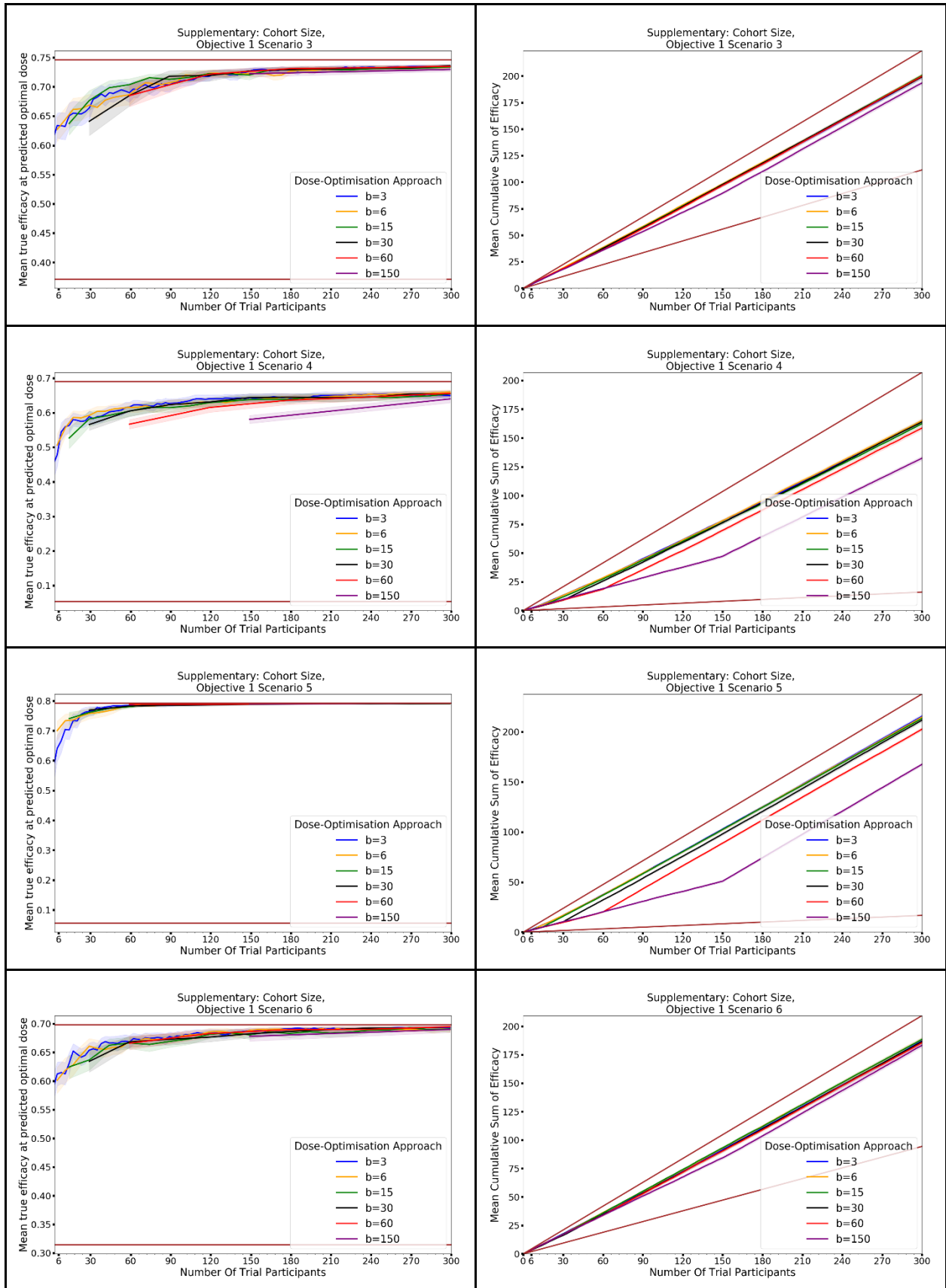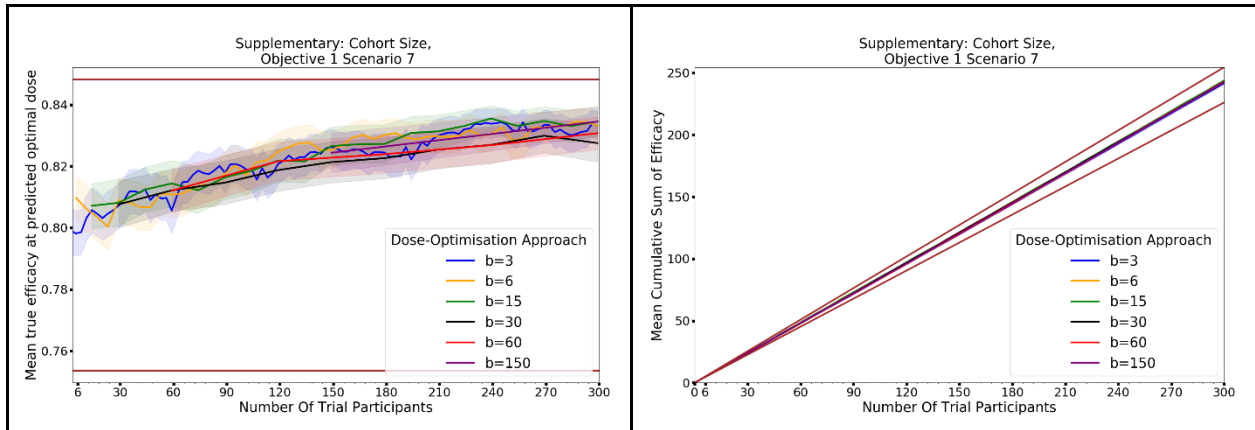
Figure S7.3. Mean true efficacy at the predicted optimal dose (left) and mean cumulative sum of efficacy (right) against trial size for all seven objective S7.3 scenarios (top to bottom). These are the mean values and 95%CI values across 100 simulations. For the true efficacy plots (left), the brown lines show the minimum and maximum possible efficacy that could be achieved in that scenario. For the cumulative efficacy plots (right) the brown lines represent the maximum and minimum cumulative efficacy sum that could be expected for that scenario.

The results of this investigation were as expected. Using fewer, larger cohorts led to a reduction in cumulative sum of efficacy, though the gradient of the cumulative sum of efficacy for all DOAs was typically equal for each of the DOAs after their first complete cohort. For some scenarios we observed that there was also a reduction in mean true utility at the predicted optimal. For example, for scenario 4 after 60 trial participants, the CoBe DOAs with b = 3, 6, 15, and 30 all outperformed the CoBe DOA that had conducted 1 cohort of b = 60. From a qualitative inspection, the CoBe should use at least three cohorts (C>2, b<N/2) to maximise vaccine efficacy, and that a large number of cohorts should be used to maximise benefit to trial participants. This may however vary by scenario. The trade-off between this and reducing the time requirements of vaccine clinical trials would require consideration with vaccine developers, but we hope that this section has shown that our results were unlikely to be biased by our choice of b=6.

S7.4. Parametric DOA: SoftMax temperature hyperparameter

The Parametric DOA used a SoftMax selection function as the method of trial selection, which we have previously shown to be effective for addressing potential concerns relating to the exploration/exploitation trade-off [3]. As discussed, this selection method depends on an 'inverse-temperature' hyperparameter t, which in the main body of this work we set as inverse_temperature t = 6.9. Increasing the

value of inverse-temperature would make the parametric DOA more exploitative. Decreasing the value of inverse-temperature would make the parametric DOA more explorative.

We investigated the effect that this may have on selection of maximum efficacy dose and benefit to trial participants. This was done by following the same methodology and scenarios as objective 1 of the main body of this work, but where all DOAs investigated were parametric DOAs with five different inverse-temperature hyper parameters. Results are shown in Figure S5.
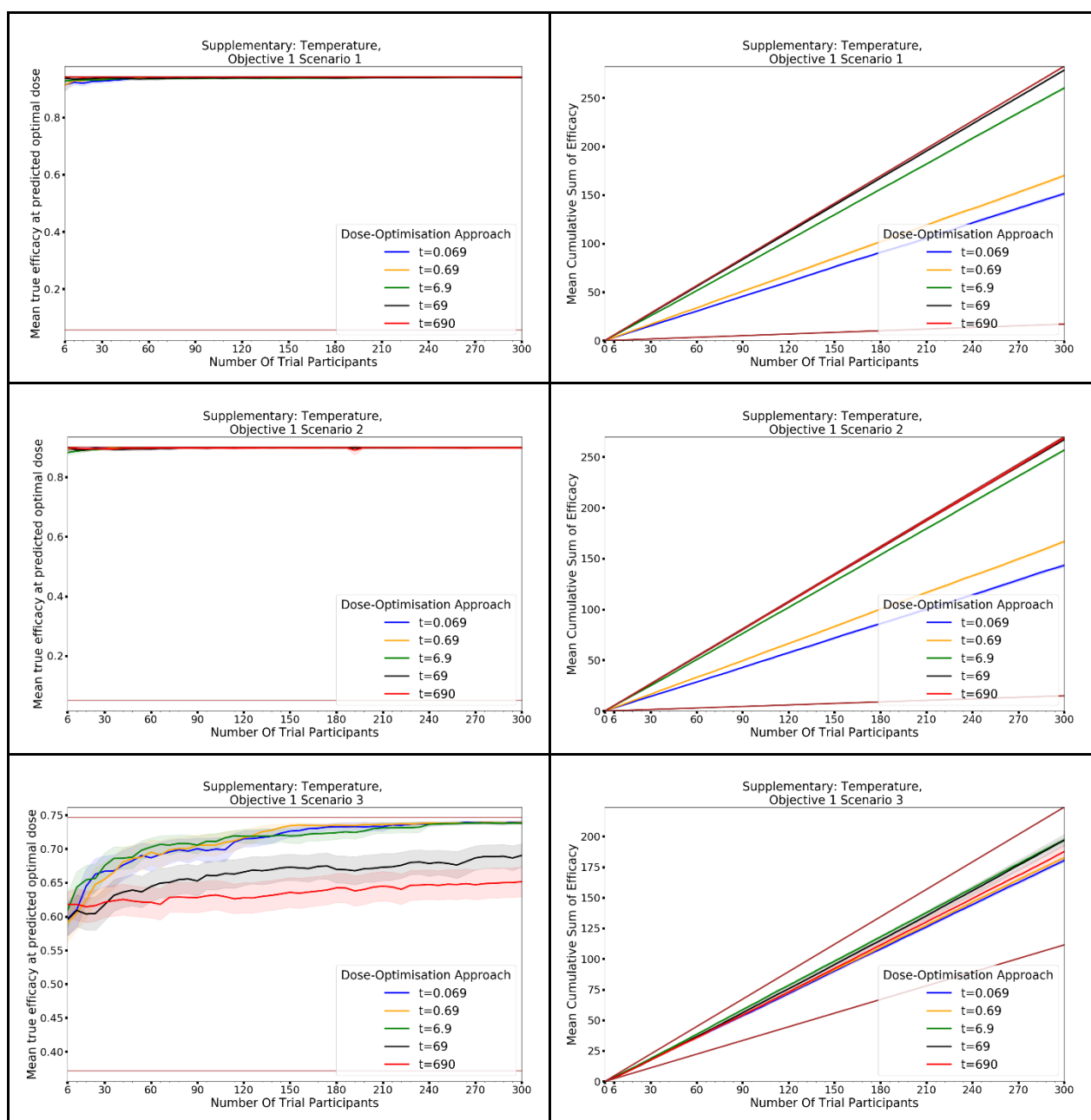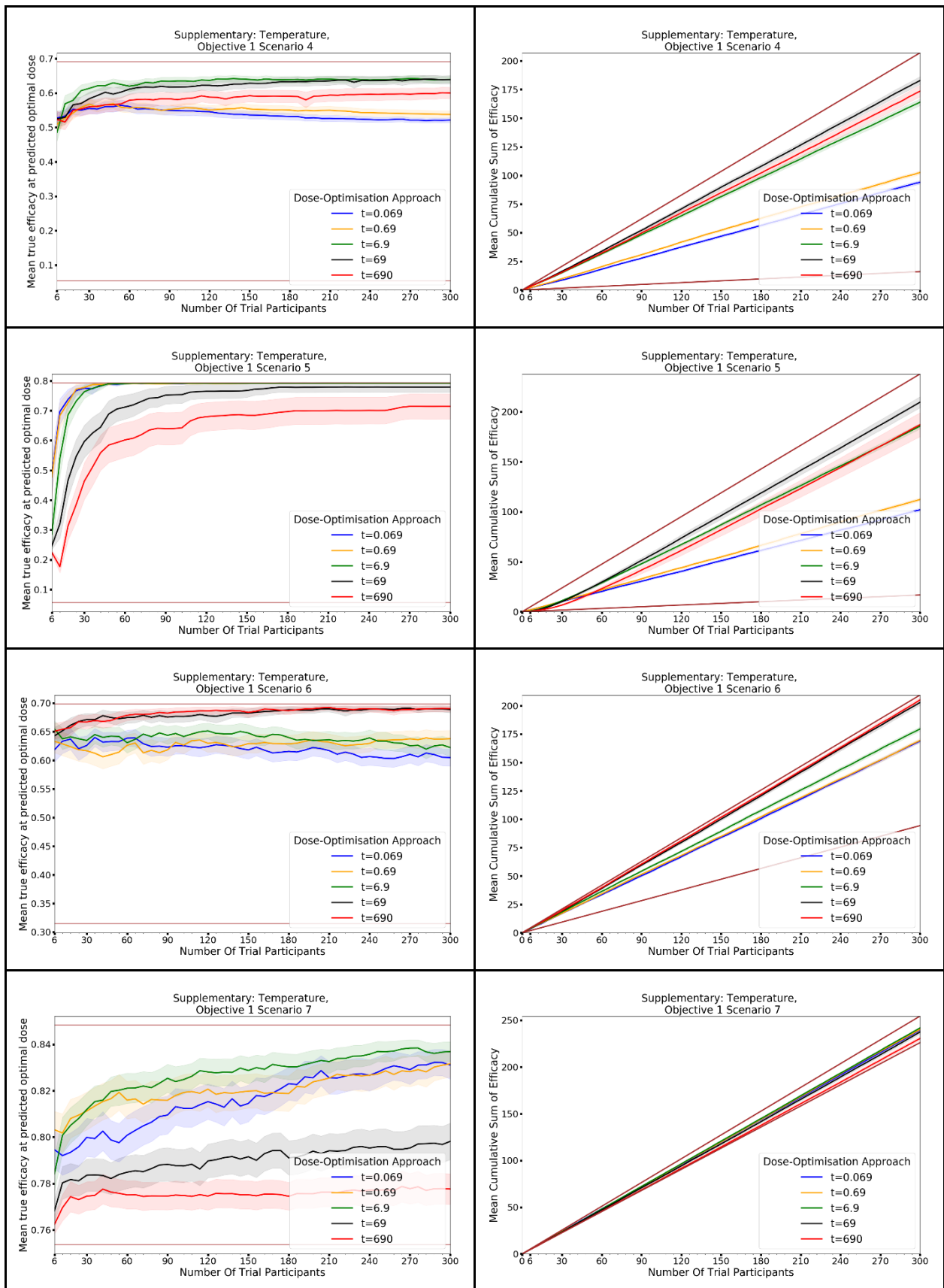
Figure S7.4. Mean true efficacy at the predicted optimal dose (left) and mean cumulative sum of efficacy (right) against trial size for all seven objective S7.4 scenarios (top to bottom). These are the mean values and 95%CI values across 100 simulations. For the true efficacy plots (left), the

brown lines show the minimum and maximum possible efficacy that could be achieved in that scenario. For the cumulative efficacy plots (right) the brown lines represent the maximum and minimum cumulative efficacy sum that could be expected for that scenario.

As expected, for different scenarios the 'optimal' value for inverse temperature varied. For example, t = 690 was best able to locate a maximally efficacious dose for scenario 6 but was worst for scenario 7. For t = 0.069 and t = 0.69, the mean cumulative sum of efficacy was typically far lower for all numbers of trial participants compared to when larger values of t were used. We believe that the results for t=6.9 show that it was typically optimal or near optimal for all scenarios other than scenario 6. Given this finding, we believe that this choice of t was reasonable in the main body of work, and that the effectiveness of the parametric DOA was not underestimated due to our choice of t=6.9. This supplementary section agrees with previous finding that parametric modelling-based adaptive design should not typically use a highly exploitative method of trial dose selection.

S7.5. Uniform Naive DOA: Number of Doses

For the uniform naive DOA, we had the number of dosing groups that would be investigated by the DOA, b. We choose b = 6 for single dose administration, b = 9 for prime/boost administration and b=27 for prime/boost/second-boost administration. Investigating only a small number of dosing groups could limit vaccine efficacy/utility, as if none of the b doses are optimal then the true optimal dose could not be selected. On the other hand, if b is too large then there is a reduction in the number of individuals available per dosing group. For example, if 30 dosing groups were investigated with a total number of trial participants N = 120, then only 4 individuals would be tested per group, which may not be sufficient to determine which dose is optimal.

We investigated of the effect that changing b might have on selection of maximum efficacy dose and benefit to trial participants. This was done by following the same methodology and scenarios as objective 1 of the main body of this work, but where all DOAs investigated were CoBe DOAs with six different values for b. The doses investigated for each value of b are given in table S1, with doses being as evenly distributed across the dosing domain as possible. Results are shown in figure S6.

Number of dosing groups (b)	Doses investigated
1	0.5
2	0.33, 0.66
3	0.0, 0.5, 1.0
4	0.0, 0.33, 0.66, 1.0
5	0.0, 0.25, 0.5, 0.75, 1.0
6	0.0, 0.2, 0.4, 0.6, 0.8, 1.0
10	0.0, 0.11, 0.22, 0.33, 0.44, 0.55, 0.66, 0.77, 0.88, 1.0

Table S7.5. Doses investigated depending on b.

310

Supplementary: Uniform Naive number of dose, Objective 1 Scenario 3

Supplementary: Uniform Naive number of dose, Objective 1 Scenario 3

Supplementary: Uniform Naive number of dose, Objective 1 Scenario 4

Supplementary: Uniform Naive number of dose, Objective 1 Scenario 4

Supplementary: Uniform Naive number of dose, Objective 1 Scenario 5

Supplementary: Uniform Naive number of dose, Objective 1 Scenario 5

Supplementary: Uniform Naive number of dose, Objective 1 Scenario 6

Supplementary: Uniform Naive number of dose, Objective 1 Scenario 6

Figure S7.5. Mean true efficacy at the predicted optimal dose (left) and mean cumulative sum of efficacy (right) against trial size for all seven objective S7.5 scenarios (top to bottom). These are the mean values and 95%CI values across 100 simulations. For the true efficacy plots (left), the brown lines show the minimum and maximum possible efficacy that could be achieved in that scenario. For the cumulative efficacy plots (right) the brown lines represent the maximum and minimum cumulative efficacy sum that could be expected for that scenario.

The results of this investigation were clearly biased in a way that could not easily be controlled for. How well the Uniform Naive DOA performed for locating maximum efficacy dose depended on the scenario. For example, the Uniform Naive DOA with b=1 'locates' the optimal dose immediately for scenario 3, as the only dose investigated is 0.5 and this is indeed the true optimal dose. b = 3 also locates the optimal dose very well for this scenario, whereas b = 2 performs very poorly and neither 0.33 nor 0.66 are near the true optimal dose. This is not a reflection that using two dosing groups is in general worse than using one dosing group or three dosing groups, but a reflection that the performance of the uniform naive DOA is dependent on how optimal the doses it investigates are. This sensitivity is particularly true for small b. We choose b=6 as that is the number of dosing groups that was used in [6], and because investigation of at least 5 or 6 dosing groups was suggested by [7].

There was trivial difference in cumulative sum of efficacy, as none of these DOAs used adaptive design.

S7.6. Adaptive Naive DOA: Number of Doses

This section is exactly the same as discussed for the Uniform Naive DOA, except using the Adaptive Naive DOA. Results are shown in figure S7.

Supplementary: Adaptive Naive number of dose, Objective 1 Scenario 1

Supplementary: Adaptive Naive number of dose, Objective 1 Scenario 1

Supplementary: Adaptive Naive number of dose, Objective 1 Scenario 2

Supplementary: Adaptive Naive number of dose, Objective 1 Scenario 2

Supplementary: Adaptive Naive number of dose, Objective 1 Scenario 3

Supplementary: Adaptive Naive number of dose, Objective 1 Scenario 3

Supplementary: Adaptive Naive number of dose, Objective 1 Scenario 4

Supplementary: Adaptive Naive number of dose, Objective 1 Scenario 4

Figure S7.6. Mean true efficacy at the predicted optimal dose (left) and mean cumulative sum of efficacy (right) against trial size for all seven objective S7.6 scenarios (top to bottom). These are the mean values and 95%CI values across 100 simulations. For the true efficacy plots (left), the brown lines show the minimum and maximum possible efficacy that could be achieved in that scenario. For the cumulative efficacy plots (right) the brown lines represent the maximum and minimum cumulative efficacy sum that could be expected for that scenario.

Findings were similar to those for the Uniform Naive DOA, with the performance of each DOA being dependent largely on whether any of the doses that were investigated were indeed optimal. There was a difference between the DOAs with regards to cumulative sum of efficacy, however again this was dependent on the scenario. For example, b=1 and b=3 had large cumulative sums of efficacy for

scenario 3, with the Uniform Naive DOA that had b=1 giving the optimal dose of 0.5 to all individuals. We would argue that this does not reflect consistent capacity for dose optimisation and ethical trial design, especially given the inferior performance of this DOA for all environments where the one investigated dose was not optimal. b=6 was more consistent across different scenarios, justifying our choice.

Data Availability Statement: Data and code for this work are available through https://github.com/ISIDLSHTM/CoBeDOA_Data and https://github.com/ISIDLSHTM/CoBeDOA

Paper 5 Supplementary References

1. Thall, P.F.; Cook, J.D. Dose-Finding Based on Efficacy-Toxicity Trade-Offs. *Biometrics* **2004**, *60*, 684–693, doi:10.1111/j.0006-341X.2004.00218.x.
2. Brock, K.; Billingham, L.; Copland, M.; Siddique, S.; Sirovica, M.; Yap, C. Implementing the EffTox Dose-Finding Design in the Matchpoint Trial. *BMC Medical Research Methodology* **2017**, *17*, 112, doi:10.1186/s12874-017-0381-x.
3. Benest, J.; Rhodes, S.; Evans, T.G.; White, R.G. Mathematical Modelling for Optimal Vaccine Dose Finding: Maximising Efficacy and Minimising Toxicity. *Vaccines (Basel)* **2022**, *10*, 756, doi:10.3390/vaccines10050756.
4. Thall, P.F. BAYESIAN ADAPTIVE DOSE-FINDING BASED ON EFFICACY AND TOXICITY. *Journal of Statistical Research* **2012**, *46*, 187–202.
5. Rolland, P.; Kavis, A.; Immer, A.; Singla, A.; Cevher, V. Efficient Learning of Smooth Probability Functions from Bernoulli Tests with Guarantees. In Proceedings of the Proceedings of the 36th International Conference on Machine Learning; PMLR, May 24 2019; pp. 5459–5467.
6. Takahashi, A.; Suzuki, T. Bayesian Optimization Design for Dose-Finding Based on Toxicity and Efficacy Outcomes in Phase I/II Clinical Trials. *Pharmaceutical Statistics* **2021**, *20*, 422–439, doi:10.1002/pst.2085.
7. Diniz, M.A.; Tighiouart, M.; Rogatko, A. Comparison between Continuous and Discrete Doses for Model Based Designs in Cancer Dose Finding. *PLOS ONE* **2019**, *14*, e0210139, doi:10.1371/journal.pone.0210139.

Chapter 7. Discussion & Conclusion

The current methods of selecting vaccine dose often rely on empirical comparison of dose-immunogenicity and dose-toxicity data. This is in contrast to the modelling-based methodologies used in selecting optimal drug doses, which have become prevalent due to their capacity to increase drug effectiveness and safety whilst reducing the size of clinical trial required to locate optimal dose. The field of Immunostimulation/Immunodynamic (IS/ID)/ mathematical modelling aims to use mathematical modelling of vaccine dose-response data to select optimal vaccine dose. Modelling for the purpose of selecting dose is already common in model-based drug development. Prior IS/ID modelling has found that the vaccine dose-efficacy relationship may be peaking rather than saturating, which suggests that model-based drug development methodologies which assume a saturating dose-efficacy may not be reasonable and could be selecting sub-optimal doses than may be detrimental to efficacy. Additionally, prior work has primarily only considered optimal dose as that which maximises vaccine immunogenicity/efficacy.

In this thesis I have expanded the field of IS/ID modelling for vaccine dose optimisation. This was done by; gathering adenoviral vector vaccine dose-response data through a systematic literature review, investigating the prevalence of predicted peaking versus saturating dose-response for this adenoviral vector vaccine dose-response data, extending IS/ID vaccine dose optimisation as a multi-factor optimisation problem, evaluating the potential impact of correctly or incorrectly assuming a peaking/saturating dose-efficacy response, evaluating the impact of adaptive trial design on optimal dose selection using IS/ID methodologies, and evaluating the novel non-parametric modelling based 'Correlated Beta' dose optimisation approach with emphasis on multi-dimensional vaccine dose-optimisation. I have then discussed and given implications regarding the future application of IS/ID modelling in vaccine dose-ranging trial design and vaccine dose selection.

Each of the publications included within this thesis outline the associated strengths, weaknesses, and implications of those chapters. Here I outline the overall strengths,

weaknesses, and implications of these works in the context of this thesis, along with the future work needed to further develop the fields of mathematical and IS/ID modelling.

Summary of findings

I aimed to gather vaccine dose-immunological response data through systematic review and give a descriptive analysis of these data (chapter 2/paper 1). I hence aimed to establish the prevalence of predicted peaking versus predicted saturating dose response curves for a specific vaccine class (adenoviral vector vaccines) (chapter 3/paper 2). I then aimed to use mathematical modelling of dose-efficacy, dose-toxicity, and dose-cost using phase 1 clinical trial data to extend IS/ID into multi-objective dose optimisation (chapter 4/paper 3). This involved proposing three utility functions that included these three models and predicting optimal dose as defined by these utility functions for the Ad5-nCoV vaccine.

I then aimed to use simulation study methodologies to conduct a theoretical analysis of mathematical modelling for vaccine dose optimisation. I first investigated the impact of correctly or incorrectly assuming a peaking/saturating dose-efficacy curve shape, along with the potential inclusion of mathematical model informed adaptive trial design, on selection of 'optimal' dose and ethical trial design (chapter 5/paper 4). I then aimed to evaluate whether the use of the non-parametric CCBP model for modelling vaccine dose-response, which did not require the assumption of an underlying dose-response curve shape, could be beneficial in locating optimal vaccine dose (chapter 6/paper 5).

My findings were as follows. I extracted data from 35 studies and found that 94% of these studies investigated less than 6 different doses (chapter 2/paper 1). These data were best described as peaking for 22% of the data, best described as saturating for 4.7% of these data, and there was no significant evidence for 73.3% of the data (chapter 3/paper 2). I found that the model predicted optimal vaccine dose for the Ad5-nCoV vaccine was between 1.1-1.5 x 10^{11} viral particles, depending on the utility function used to define 'optimal' (chapter 4/paper 3).

I showed that assuming a peaking dose-efficacy curve was typically preferable to assuming a saturating dose-efficacy curve for the purposes of selecting optimal dose as defined by a utility function maximising efficacy and minimising toxicity. Weighted model averaging also was effective for selecting optimal vaccine dose (chapter 5/paper 4). I found that combining the non-parametric CCBP model with a Thompson Sampling method of adaptive trial design was an effective approach for optimising vaccine dose and was consistently better than approaches that used neither mathematical modelling nor adaptive design (chapter 6/paper 5). Mathematical modelling based adaptive trial design also led to more beneficial vaccine response for simulated clinical trial participants.

Strengths

I here describe the strengths of each paper within the context of the aims of this thesis, then describe the overall strengths of this thesis.

Strengths of chapter 2/paper 1

In chapter 2, I collected data from published adenoviral dose-immunogenicity studies. I listed the studies that were available in the literature by host species, number of dosing groups, adenovirus serotype and response type, and also showed the frequency distribution of doses which had previously been trialled in human and murine studies. This data and collation of studies relevant to the investigation of adenoviral dose-immunogenicity enabled the work in chapter 3, and I believe is of potential use for future vaccine dose-response modellers. Additionally, this work highlighted a potential issue for vaccine dose-response modelling. 94% of studies investigated five or fewer dosing groups, whereas previous work has highlighted that modelling is benefitted by considering a larger number of dosing groups [196]. This previous work found that using more than five dosing groups was beneficial when using modelling-based dose selection. Whilst that work investigated MTD identification with model informed adaptive design, I believe that in combination with the findings of paper 1 it is reasonable to expect that IS/ID modelling of adenoviral

vector vaccines may be limited by the small number of dosing groups typically investigated.

Strengths of chapter 3/paper 2

In chapter 3, I found evidence to support the hypothesis that some adenoviral vaccine dose-immunogenicity may be better described by a peaking curve than a saturating curve. Whilst for many of the data there was no evidence to support either curve shape, and hence true prevalence of peaking curve shape relative to saturating curve shape could not be determined, I believe that this supports the findings of previous IS/ID studies that vaccine dose-immunogenicity (and hence dose-efficacy) should not be assumed to always be best described as saturating.

As noted above, it was not surprising that curve shape could not be identified for the majority of data given the typically small number of dosing groups. Chen and Lui conducted a simulation study which was published after paper 2, showing that true curve shape could not consistently be accurately determined with only 4 small (n<10) dosing groups if conducting model selection when using ΔAIC>0 as the model selection criteria [138]. Again, the findings of paper 2 highlight that present adenoviral dose-ranging trials may not be designed in a way that allows for identification of dose-response curve shape. Using a larger number of smaller dosing groups may improve mathematical modelling for vaccine dose selection.

Strengths of chapter 4/paper 3

In chapter 4, I considered multiple potential utility functions that could be used to define optimal vaccine dose, including efficacy, toxicity, and cost in these utility functions. Whilst the dose-cost model and the integration of cost into the utility function were both simple, this represented an advance in complexity and scope beyond previous IS/ID studies. Additionally, this work highlighted that different doses could be considered optimal depending on which utility function was used to define optimality. Whilst this previous statement is trivial [96], I believe it does provide an

implication that the concept of an 'optimal' dose is dependent on the utility function, and thus utility functions must be defined to enable 'dose-optimisation'.

Within the scope of conducting IS/ID modelling, this work used clinical trial data, herd immunity thresholds from epidemiological studies, and estimates of costs from industry reports. This highlighted the capacity for IS/ID modelling to include data and findings from a variety of sources in order to aid in development of models and utility functions. This then presents the capacity for IS/ID modelling to be considerate of the context that the vaccine will be used when predicting 'optimal' dose. In particular, this work did not require discussion with the clinicians and developers that conducted the dose-ranging trial I gathered data from. This suggests that, whilst I have suggested that designing clinical trials to include a larger number of dosing groups would improve modelling, IS/ID methods to determine optimal dose can be used even when modellers are not involved in the design of the dose ranging trials.

In the supplementary materials, I conducted a sensitivity analysis of the parameters on optimal dose for each of the utility functions. This was beneficial in showing which of the model parameters would most change the predicted optimal dose if the parameter were misspecified. For example, S2.2.1 showed that the 'costless' utility function was insensitive to changes in the 'max' parameter of the dose-seroconversion model. This could imply that, for this utility function, minimisation of the variance of the 'max' parameter could be ignored if future dosing-ranging studies were conducted for this vaccine using d-optimal [appendix A.E.] design, allowing more efficient trial dose selection to better identify the parameters that would affect which dose was predicted optimal.

Strengths of chapter 5/paper 4

In chapter 5, I conducted a simulation study to investigate the effects of trial size, assumed efficacy model, and method of trial dose selection on the capacity of modelling to predict optimal vaccine dose, maximise benefit to trial participants, and improve accuracy of model predictions of vaccine utility. This represented the first IS/ID simulation study to investigate the theoretical capacity of IS/ID modelling

methods in locating maximum utility dose. The inclusion of ordinal toxicity modelling had also not previously been considered in IS/ID modelling.

Here I showed that trial size was a significant factor in how effective modelling methods may be in determining optimal vaccine dose. Increasing trial size can improve dose selection, but there were decreasing benefits beyond certain trial sizes (60 trial participants). The 'uniform' method of trial dose selection was also shown to be potentially as effective for locating maximum utility dose as a continual reassessment modelling method of trial dose selection method, which had previously been shown to be effective in locating both maximum tolerated dose [160] and maximum utility dose [197]. The 'uniform' method of trial dose selection can be considered as the most extreme version of using a 'large number of small dosing groups', further implying that such methods of trial dose selection seem beneficial for improving the effectiveness of modelling-based approaches of vaccine dose selection.

Finally, the findings of this work suggest that assuming a saturating dose-efficacy curve may be less useful than assuming a peaking dose-efficacy curve if it is possible that the dose-efficacy curve is truly peaking. It also suggests that IS/ID modellers, vaccine developers or clinicians should consider weighted model averaging. The effectiveness of model averaging was not unexpected given previous findings in dose-response modelling of drugs and toxins [138,143,198,199]. However, this work highlighted the potential for model-averaging in vaccines IS/ID and mathematical modelling, where I and others have shown there is potential for model uncertainty.

Strengths of chapter 6/paper 5

In chapter 6, I described the application of CCBP models in vaccine dose-response modelling and introduced the CoBe dose optimisation approach for designing and conducting vaccine dose-finding clinical trials. I then compared this to three other representative dose optimisation approaches. This included comparison with model-based and model-free dose optimisation approaches. The comparison dose-optimisation approaches also reflected both adaptive and non-adaptive trial design.

This represented a development in the modelling of dose-response for prime/boost and prime/boost/second-boost paradigm vaccines beyond what had previously been investigated in IS/ID modelling. Additionally, whilst this work described modelling of prime/boost and prime/boost/second-boost vaccines, I believe that it also shows that the CoBe dose optimisation approach could be applied for optimising time between prime/boost doses as well as antigen/adjuvant doses.

The overall strengths of the thesis are as follows.

Showed that the assumption of a saturating dose-immunogenicity/dose-efficacy curve shape may not be reasonable or necessary in order to identify optimal vaccine dose

My work highlights that, at least for single-administration replication-deficient adenoviral vector vaccines, the assumption of a saturating dose-immunogenicity and dose-efficacy curve is not justified. Further, dose-efficacy does not have to be assumed to be saturating in order for statistical IS/ID modelling of vaccine to be used effectively. For IS/ID modelling of vaccines in general, there are likely to be more effective dose-response models than the sigmoid saturating model that is common in modelling drug dose-response, using for example peaking models, weighted model averaging, or non-parametric modelling.

The modelling-based meta-analysis in paper 2 of the adenoviral vector vaccine dose-immunogenicity data collected in paper 1 showed that there was evidence to support 22.0% of the available adenoviral vector vaccine dose-immunogenicity data being best described by a peaking dose-response curve. This showed that, at least for adenoviral vector vaccines, dose-response should not be assumed to be saturating and supports previous findings of Handel et al. and Rhodes et al. that vaccine dose-immunogenicity may be peaking. In paper 4 I found that modelling dose-efficacy using a peaking latent-quadratic model may be more effective than assuming a saturating model. The dose-optimisation approaches that assumed a saturating dose-efficacy curve performed poorly relative to those that assumed a peaking efficacy curve when the true dose-efficacy was peaking. However, the dose-optimisation approaches that assumed a peaking dose-efficacy curve often performed comparatively well relative to those that assumed a saturating efficacy

curve when the true dose-efficacy was saturating. Thus, I believe that it is not reasonable to assume that all vaccine dose-immunogenicity is saturating, and that assuming a peaking dose-efficacy curve is a more reasonable 'default' assumption for vaccine IS/ID modelling for clinical trials that investigated a large number of dosing groups or use a continual-modelling based trial design.

I also presented other methods for modelling vaccine dose-efficacy that do not require the assumption of a saturating dose-response curve. Beyond assuming that dose-efficacy is peaking, weighted model averaging of feasible models of vaccine dose response allows modellers to account for model uncertainty. In addition, weighted model averaging seemed effective for both the retrospective modelling of data from a large number of dosing groups (paper 4 objective 1) or for conducting vaccine clinical trials in an adaptive design setting similar to the continual reassessment method (paper 4 objective 2). The CCBP model, which only assumes that the true vaccine dose-response curve-shape is smooth and continuous, was also shown to be effective for locating optimal vaccine dose in the adaptive design setting (paper 5) relative to other dose-optimisation approaches. This follows on from previous work that showed that flexible non-parametric models can be as effective for locating optimal drug dose as parametric models [75,76], even when those parametric models are reasonable approximations of true dose-response. Thus, I believe that this thesis shows that it is likely not reasonable to assume that the maximum efficacy dose is always the largest dose (paper 2), and that for binary outcome measures of efficacy assuming a saturating dose-efficacy curve is not needed nor preferable to using a model that can capture peaking behaviour (papers 4 and 5).

Modelling and adaptive design are likely to be effective for selecting optimal vaccine dose, and may be preferable to dose-optimisation approaches that do not use modelling/adaptive design

This work further demonstrates the theoretical potential for IS/ID and mathematical modelling to improve vaccine dosing. In paper 3 I showed that mathematical modelling could be used on real world data to predict optimal dose as defined by a number of different utility functions, though I could not validate these findings. In

papers 4 and 5 I showed that mathematical modelling could be used to predict optimal dose and, whilst these did not use real data, I was able to compare the true utilities of model predicted optimal doses against the true utilities of the true optimal doses for the simulation study scenarios. I believe that these findings in papers 4 and 5 suggest that mathematical modelling is likely to be effective as a method of optimising vaccine dose, as hypothesised given the prevalence of mathematical modelling in drug dose selection. Whilst similar statistical models of drug dose-response have already been validated as effectives tools for dose-selection in model-based drug development, with papers 4 and 5 I showed that they may be repurposed to be effective for vaccine-specific dose optimisation problems. For example, dose-optimisation when concurrently considering a peaking dose-efficacy curve and ordinal toxicity gradings (chapter 5/ paper 4), and dose-optimisation when maximising a utility function of efficacy/toxicity for prime/boost paradigm vaccines (chapter 6/paper 5, objectives 3 and 4).

I also showed how mathematical modelling could be used, not only as a retrospective method of predicted optimal vaccine dose from clinical trial data, but also to guide vaccine dose-ranging clinical trial design. Previous work in model-based drug development has suggested that the use of mathematical modelling in dose selection can result in accelerated clinical trials with smaller sample sizes [156,200,201]. It was also to potentially increase benefit to participants in drug trials where a saturating dose-efficacy can be assumed [202], and our findings support that improving expected benefit to trial participants is also possible when a saturating dose-efficacy cannot be assumed (papers 4 and 5). This could lead to more ethical vaccine clinical trials. This may be more important in vaccine development than for the selection of dose for some drugs, for example the oncological drugs that CRM was developed for. This is because, as noted in chapter 1, individuals that are enrolled in dose-ranging trials for prophylactic vaccines are typically healthy unlike trials in oncological drugs. I also noted in chapter 1 that public perception of vaccine safety and efficacy may influence vaccine uptake, and so it may be possible that maximising benefit to trial participants may improve vaccine uptake. These factors combine to suggest that maximising benefit to vaccine dose-ranging trial participants is important, and this thesis I believe supports the hypothesis that mathematical

modelling-informed adaptive trial design can improve expected benefit to trial participants.

In paper 5 I also compared two modelling-informed adaptive trial design vaccine dose-optimisation approaches to two dose-optimisation approaches that did not use modelling. The 'uniform naive' dose-optimisation approach, which used neither modelling nor adaptive design, was typically less effective at locating optimal vaccine dose and less beneficial to trial participants than the dose-optimisation approaches that used modelling and/or adaptive design. The 'adaptive naive' dose-optimisation approach better maximised benefit to trial participants than the 'uniform naive' dose-optimisation approach. However, this dose-optimisation approach considered only a small number of potential trial doses, which may be suboptimal when attempting to select optimal vaccine dose. Where none of the small number of doses were near optimal, neither the 'uniform naive' nor 'adaptive naive' dose-optimisation approaches were able to maximise vaccine utility. This might suggest that one of the benefits of modelling may be that it allows for selection between larger numbers of potential doses than non-modelling-based approaches.

This finding agrees with the findings of Diniz et al. [196], which suggested that a large number of dosing groups should be considered in model-based dose-ranging trials. My findings appear to contradict the works of Chen and Lui[138], which found that modelling-based identification of the maximally efficacious dose was improved by considering only 4 dosing groups of 30 individuals relative to considering 6 dosing groups of 20 individuals. However, in that work at least one of the dosing groups was at the true maximally efficacious dose for both the 4 groups and 6 group dose-optimisation approaches, therefore I do not believe that there is a contradiction between our works. I discuss further the effects of changing the number of dosing groups for the 'uniform naive' dose-optimisation approach in Appendix A.F.

Taken as a whole, this work highlights a number of reasons that mathematical modelling is likely to be effective for the purposes of selecting optimal vaccine dose and guiding vaccine clinical trial design. The mathematical models and methods that I discussed in this work were comparatively simple relative to the mechanistic "Quantitative Systems Pharmacology" type models that had previously been

discussed in IS/ID modelling and have seen success for drug dose optimisation. If simple IS/ID modelling is able to be effectively used to select optimal vaccine dose, then there may be increased potential for practical application of IS/ID modelling. Specifically, I believe this work highlights that conducting dose-finding studies based on direct comparison of only a small number of dose ranging groups without modelling or adaptive design are likely to be less useful than vaccine dose-optimisation that uses mathematical IS/ID modelling and/or adaptive design.

These findings have particular relevance given the vaccination response to the COVID-19 pandemic, with the government of the United States of America having to implement 'operation warp speed' in order to allow for rapid vaccine development relative to the traditional development pathway [203]. In response to this computational and mathematical modelling have been suggested as alternatives that may have better enabled safe and effective dosing in the accelerated timeframe of pandemic vaccine development [191].

Consideration of vaccine dose-optimisation as a potentially multi-objective optimisation problem, and discussion of multiple utility functions

Previous IS/ID work has predominantly focused on modelling and maximisation of immunogenicity, though the importance of safety and dose-sparing have been highlighted [23,49]. Handel et al. [49] considered defining optimal dose through a simple utility function that aimed to maximise antibody level whilst minimising morbidity, however this thesis represents the first piece of work in the IS/ID field to emphasise vaccine-dose selection as a multi-objective optimisation problem. Throughout the chapters of this thesis, I have used six different utility functions: the herd immunity threshold in paper 3, the costless utility function in paper 3, the costed utility function in paper 3, the weighted efficacy/ordinal toxicity utility function in paper 4, maximum efficacy in paper 5, and the utility contour utility function in paper 5. These show a breadth of potential ways that 'optimal dose' can be defined. For further description of potential utility functions see Appendix A.C.

By conducting simulation studies in papers 4 and 5 I showed that IS/ID modelling can be used when considering vaccine dose-optimisation as a multi-objective

optimisation problem. This extends the work of Handel et al. [49] , which did not conduct simulation studies to address whether the dose that maximised their presented utility function was likely to truly be optimal. The simulation studies conducted here show that with sufficient data modelling may be effective for prediction of vaccine dose that maximises a multi-objective utility function.

The consideration of vaccine dose optimisation as a multi-objective optimisation problem however also identified that multi-objective dose optimisation using IS/ID modelling may be non-trivial. In paper 5, objective 3, I investigated a continual modelling-based dose optimisation approach based on parametric models of dose-efficacy and dose-toxicity for prime/boost administration. For objective 3 scenario 5, the doses predicted to have maximal utility by this parametric modelling approach were typically sub-optimal. I had initially believed that both the efficacy and toxicity models were appropriate models of dose-efficacy and dose-toxicity, particularly as the same dose-efficacy model was shown to be able to locate the maximally efficacious dose in objective 2 for that same dose-efficacy curve. This might suggest that we should not assume that models that are appropriate in single-objective optimisation will also be effective for determining optimal dose as defined by a multi-objective utility function. Again, this highlights the importance of conducting simulation studies.

Highlighted potential pitfalls in IS/ID modelling and vaccine dose optimisation

In this thesis I discovered, investigated, and highlighted some potential pitfalls that I believe have implications for the future development of the field of IS/ID modelling and for optimal vaccine dose selection. Specifically, these were the potential need to consider the exploration/exploitation trade-off, the potential for 'optimistic biases', and the potential that vaccine dose-ranging trials may not presently provide sufficient data to consistently determine dose-response curve shape.

The concept of an 'exploration/exploitation' trade-off within the context of 'multi-armed bandit' problems has been well discussed. 'Exploitation' in the context of adaptive-design vaccine dose-ranging trials would mean that only doses that are predicted to be optimal are investigated, whereas 'exploration' involves investigating

327

doses that are predicted to be sub-optimal. Being too 'exploitative' could prevent optimal doses from being found [204]. This may be because, if only the dose that is predicted optimal is tested, and the model is accurate regarding its predictions of efficacy and toxicity for that dose, then new trial data would be unlikely to change which dose the model predicts as optimal, even if that predicted optimal dose is in-fact sub-optimal.

This trade-off is typically discussed in the context of discrete doses [150,205,206] ('arms' in the parlance of that literature) without modelling, but in paper 4 I showed that including exploration in model-informed adaptive design dose-optimisation approaches may improve mathematical modelling informed vaccine dose selection. I found through further simulation that exploration became even more important as the number of trial participants increased [appendix A.G.]. I have highlighted that conducting mathematical model-based dose-ranging trials without considering this trade-off, as in the style of traditional continual reassessment modelling, may not be beneficial to selecting optimal vaccine dose.

I highlighted the potential for model predictions of vaccine utility to be 'optimistically biased', where the prediction of utility at the predicted optimal dose is more likely to be an overestimation than an underestimation compared to the true utility at that dose. Whilst the potential for models to overestimate utility at predicted maxima is not novel [207–209], it has not been discussed in the context of IS/ID modelling or model-based drug development to my knowledge. This represents a pitfall in present IS/ID modelling. This is also a potential pitfall in vaccine dose-optimisation approaches that do not use modelling, as discussed in the supplementary of paper 4, and direct comparison-based dose-optimisation approaches may also be vulnerable to optimistic bias. A further discussion of this is included in Appendix A.F.2. This pitfall and the one above would not have been noticed if I had not conducted simulation studies, again highlighting that theoretical analysis of IS/ID methodologies through simulation studies is important to the development of the field.

The final pitfall of IS/ID modelling this thesis suggests is that data from vaccine dose-ranging trials may not be presently sufficient to determine dose-

immunogenicity/dose-efficacy curve shape. Whilst I was able to predict optimal vaccine dose using published data in chapter 4, having sufficient data to identify the dose-response relationship and validate any modelling assumptions regarding dose-response curve shape may be important to increasing trust in model predictions. IS/ID practitioners should be aware that curve shape may not be possible to determine using data from vaccine dose-ranging trials as these are presently designed and must be prepared to account for this uncertainty.

Weaknesses & Challenges

Weaknesses of chapter 2/paper 1

In chapter 2, I only collected data on adenoviral vector vaccines, and additionally this data only consisted of immunological response data to a single administration of vaccine. Whilst this was the data that would be required to answer the questions posed in chapter 3, it did mean that analysis could not be done for other classes of vaccines. Secondly, it also meant that I could only investigate modelling prime-boost dose-immunogenicity through the simulation studies in chapter 6 and without real world data. Thirdly, I did not collect data on the prevalence of adverse events by dosing groups. This was again as the main of my work in chapter 2 involved the dose-response curve shape of immunogenicity responses. Collecting toxicity data may have been beneficial to informing safe adenoviral vector vaccine dose ranging trials.

Weaknesses of chapter 3/paper 2

In chapter 3, I only included a single representative model for each of the peaking and saturating curves. The results may have been different if I had chosen different models, but the two chosen models were reasonable given that they were models that had previously been used for IS/ID modelling [50]. We also did not include models that represented simpler or more complicated dose-response dynamics. An exponential model [210] may have been a good description for some of the data despite being a simpler model but may not be biologically justifiable to extrapolate from. Biphasic and triphasic dose-response models have been suggested [211] but use 6 or 9 model parameters respectively, so for most of the available dose-ranging

studies I would not have been able to be used to calibrate these models. Additionally, my interest was in the prevalence of vaccine dose-response data that was best described as peaking or saturating, not on finding the best possible model for each of the dose-response datasets. Therefore, I considered it reasonable to only include representatives of these two behaviours

A further weakness was the method of calibration that was used. The parameters for the representative models were found by minimising the root mean squared error for each of the dose-response datasets. Technically such calibration would only be statistically valid if the data for each of the dosing groups could be assumed to be distributed around the true response value for that dose with homoscedastic, uncorrelated error with expectation zero (normally distributed with mean zero for example) [212]. Such assumption may not be valid for immunologic data [213]. Despite this, mean squared error and root mean squared error are commonly used for dose-response model fitting and model selection purposes [135], and I do not believe that using a different metric would have impacted the results. Further to this, I used AIC as the metric to determine which model best described data. I considered using the 'corrected AIC' or BIC metrics, but this would have not changed the results as both the peaking and saturating models had three parameters, so the regularisation term for both models can be ignored.

Whilst I do not believe that it was due to my implementation, the greatest weakness of this chapter was that curve shape could not be determined to be best described as peaking or saturating for 73.3% of the data. I believe that this was due to the dose-response data often having too few dosing groups to determine the curve shape, or not investigating sufficiently large doses to observe saturating or peaking behaviour. For these vaccines, the optimal dose may have been limited by toxicity, cost, or practical implementation reasons, but that still meant that curve shape could not be determined, which limited statistical capacity to determine the prevalence of adenoviral vector vaccine dose-response data being truly peaking or saturating.

Finally, one of the challenges of implementing this work was considering whether I should use a pooling approach. It is suggested that for meta-analyses which aim to determine curve shape, a pooling approach can be used [137,214]. This involves

hierarchical modelling to estimate model parameters that fit all of the available data from different datasets. Whilst these pooling techniques are potentially useful tools in determination of dose-response shape, it requires the assumption that the underlying dose-response curve shape (up to reparameterisation) is the same for all data [215]. From discussion with my supervisors and consideration of Handel et al. [49] it seemed more reasonable that for some vaccines dose-response is peaking and for others it is saturating. Therefore, this assumption did not seem reasonable. Alternatively, I could have assumed that all data that had the same host-species, response type, adenoviral vector species, and route of administration would have the same dose response curve shape, and then pooled data based on those factors. This would have prevented the analysis in objective 2, which assessed whether that assumption was reasonable.

Weaknesses of chapter 4/paper 3

In chapter 4, there were weaknesses that could not be addressed with the available data. Firstly, I did not account for covariates that may have influenced dose response, such as weight, age, or gender. Instead, I assumed that dose was the only factor that impacted the probability of seroconversion or grade 1 / 3 adverse event occurrence between the dosing groups. Given that there was some heterogeneity in age between the three dosing groups, this may not have been a reasonable assumption. I also decided to derive parameters for the dose-cost model from literature rather than through querying the vaccine developers. I collaborated with an economist (Matthew Quaife) to ensure that my cost model was reasonable and conducted a sensitivity analysis to investigate the potential effects of the cost model parameters being poorly estimated. I also assumed that the probabilities of individuals experiencing seroconversion and toxicity were mutually independent, rather than using a copula method to derive a joint probability from marginals [216]. Again, this would have required a more detailed dataset than was publicly available.

To counteract these weaknesses, I considered contacting the vaccine developers to inquire after a more detailed dataset. I decided against this as I did not want to have any potential to be biased in my findings through contact with the developers. The main purpose of this work was to demonstrate the potential for mathematical

modelling in multi-objective vaccine dose optimisation, and for showing that 'optimal dose' was dependent on the utility function used to define vaccine utility. This work was not used to actually choose a clinical dose for this vaccine, and therefore my simplified modelling was appropriate to my aims.

The final two challenges I faced in this chapter were on validation and scope. I was unable to validate whether the doses that I predicted to be optimal were actually optimal. This would either have required further extensive dose-ranging trials or perfect knowledge of the true underlying dose response curves. The first was not possible as the developers did not investigate neutralising antibody titres in later dose-response trials and this PhD did not have scope to conduct such a trial. The latter was not possible practically, and is only possible in simulation studies, so this weakness of this paper was accounted for in papers 4 and 5. With regards to scope, I only included three simple utility functions. Including a greater quantity and complexity of utility functions could have been interesting, but I decided that explaining a smaller number of simple utility functions thoroughly would be more demonstrative of the potential benefits and use of IS/ID and mathematical modelling in vaccine dose selection.

Weaknesses of chapter 5/paper 4

In chapter 5, there were simplifications made which may have impacted ecological validity of the findings. I did not include stopping rules in the continual modelling-based approaches, which are often used in continual modelling [217]. I did not believe that including these would have improved this work, as doing so would have increased the complexity of the work and meant that 'trial size' was not well defined. I also did not include any dose-optimisation approaches that did not use mathematical modelling, or that modelled dose response using data from a small number of dosing groups (as done in chapter 4). Model-free dose-optimisation approaches were investigated in chapter 6. Modelling based on data from only a small number of dosing groups would have required making assumptions on how close to the true-optimal dose that these dosing groups should be, and accounting for or investigating these assumptions would have greatly increased the complexity of this work.

Other weaknesses and challenges of this work relate to it being a simulation study. The absence of real data meant that there was no way to show that these findings are true for actual vaccine dose-optimisation. Simulation studies are only useful if the scenarios they investigate are representative of real-world vaccine dose-response. To account for this, I used many scenarios with qualitatively different dose-efficacy, dose-toxicity, and dose-utility curves.

Whilst using a large number of scenarios was beneficial, it did present a challenge in the presentation of findings. Presenting plots of simple regret, inaccuracy, and average regret for each of the objectives and scenarios would have increased the size of the published paper and potentially led to confusion in parsing of the results. The alternative was to present a summary of these metrics across all scenarios, which I did, and I believe improved the clarity of the findings. Technically this is only valid if all real-world vaccines can be described accurately by one of the 14 scenarios with each scenario being equally probable. This was unlikely to be true. However, as the findings of the effects of trial size/assumed efficacy model/method of dose selection for the summary plot were similar to those that observed for most of the scenarios I do not believe that this impacted the results. Additionally, I included the results for each scenario in the supplementary and advised readers to consider these.

Despite the weaknesses of simulation studies, the benefits of simulation studies are that they allow for validation of whether the model predicted optimal doses are indeed optimal. I therefore believe that the combinations of chapter 4 and 5 together represent both theoretical and practical validation of mathematical IS/ID modelling for vaccine dose-optimisation.

Weaknesses of chapter 6/paper 5

Chapter 6 had the same weaknesses as chapter 5 with regards to being a simulation study. It also shares the same weakness of not including stopping rules for the same reasons. The other weakness of this work that was not mentioned in the body of the work is that CCBP model and the CoBe dose-optimisation approach are not presently available as part of any common statistical software. However,

investigation and theoretical validation of this model and approach was sensible prior to development of software or tools.

Additionally, a challenge in this work was in choosing the parametric models of dose-efficacy and dose-toxicity for the objectives discussing prime/boost and prime/boost/second-boost paradigm vaccine dose optimisation. I could not find any models of such that had been well validated. I therefore choose to use a model that was similar to that which I had found effective for single-administration vaccine dose optimisation in chapter 5. It is possible that other better models may exist, but the optimal parametric model would likely depend on the scenario, and I believe that the models used were reasonable.

There were some overall limitations and challenges to this thesis as a whole.

Quantitative allometric scaling and population covariates were not accounted for

Previous work has highlighted that a topic of interest in IS/ID modelling is in developing methods of allometric scaling, allowing prediction of human efficacy/toxicity for given vaccine doses based on preclinical animal studies [23,51]. A weakness of this thesis was that I did not further develop allometric scaling tools for IS/ID modelling. Further, the finding of paper 2 that dose-immunogenicity curve shape may depend on the species being vaccinated may prevent or limit vaccine allometric scaling between species. For example, if mouse dose-immunogenicity was peaking and human dose-immunogenicity was saturating for some vaccine, allometric scaling factors of model parameters may not be possible. Despite this, there is still a practical need to predict clinical immunogenicity/efficacy from preclinical data which this work has failed to address. Given the weaknesses of paper 2, namely that dose-response curve shape could only be determined for 26.7% percentage of dose-response data and that I only investigated adenovirus vectored vaccines, further investigation into the effect of host on dose-response curve shape is needed. Also, whilst the results of that paper were deemed statistically significant, the small number of datasets for which dose-response curve

shape could be determined may mean that more data are required to improve confidence in these findings.

Allometric scaling of dose-toxicity curves may also be reasonable if we make the common assumption that vaccine dose-toxicity is likely to be saturating for all species. This was a topic that was unexplored in this work despite it being likely to be feasible, and therefore represents a weakness of this work.

A similar idea to allometric scaling between species is to include population covariates within the statistical analysis and modelling. This allows for prediction of dose-immunogenicity/efficacy/toxicity in an untested/minimally tested population based on data from a well-tested population, and also allows for individualised dosing. Example covariates are age and gender, which have been shown to affect vaccine efficacy [218], or the haplotype of a vaccinee's major histocompatibility complexes [219]. None of our models were built to allow for response to depend on covariates and assumed that dose was the only predictive variable. I note that in my experience much of the dose-response modelling and continual reassessment modelling literature also does not include covariate analysis, so this is not unique to my work. Additionally, other than for paediatric populations, vaccines are typically not dosed differently for different populations based on these covariates [220]. I believe that further work into including covariates into dose-response models may be reasonable, and the absence of such represents a weakness of this work. However, unless there is the potential for clinical vaccine dose to vary depending on the covariates of vaccinated populations/individuals, the benefit of such modelling may be reduced relative to the benefits of covariate modelling in model-based drug dosing for individuals-based dosing.

Only binary efficacy outcomes were considered, and predictive immunogenicity-efficacy modelling was not investigated

Two further weaknesses of this work involve the concept of vaccine efficacy and how I defined this. In the background of this work, I discussed the concept of correlates and surrogates of protection, and the issues involved with these that mean that accurate binary measures of efficacy may not be available. The first weakness then

was that I assumed that vaccine efficacy in dose-ranging trials was binary and measurable. Secondly, I did not discuss or contribute to the development of models that may aid in predicting vaccine efficacy from immunogenicity data through these chapters.

With regards to the first weakness, in chapters 2 and 3 I considered modelling of dose-immunogenicity. For chapters 4, 5 and 6 I assumed that vaccine efficacy was an outcome that was binary (individuals were protected or not protected) and measurable. In chapters 5 and 6, I showed that modelling can be effectively used to maximise the probability of a dose-dependent binary efficacy outcome and minimise probability of dose-dependent ordinal/binary toxicity. For vaccines where such a measure is not available these methods would not be applicable.

Further, the use of binary thresholds of protection have previously been criticised [221], as it is unlikely that there is some true immunological threshold where a response in excess of this threshold offers 100% protection against disease and immunogenicity below this offers 0% protection. Instead it has been suggested that any increase in response offers some increasing level of protection [36]. Therefore, as IS/ID modelling develops it may be preferable to combine dose-immunogenicity and immunogenicity-efficacy models rather than modelling dose-efficacy as was done in chapters 4, 5 and 6.

This leads to the second of these weaknesses, that I did not contribute to developing immunogenicity-efficacy models despite their potential relevance in the long-term goals of IS/ID modelling. Whilst I did not do this, this is an active area of research. Khoubury et al. found that vaccine efficacy against severe COVID-19 disease could be predicted based on neutralising antibodies titre using mathematical modelling [38]. This work required a meta-analysis of many phase 2 and 3 immunogenicity and efficacy data from seven COVID-19 vaccines. Such data may not always be available, with far more research on COVID-19 being conducted that should be expected for most diseases [222].

Therefore, whilst these represent weaknesses in the work and potential limitations in the present field of IS/ID, I believe that assuming that binary efficacy outcomes were

available was reasonable. This was partially because binary thresholds of protection such as seroconversion are commonly used to choose optimal vaccine dose in practise. This was also given the novelty of the field: it appeared relevant to investigate simplistic models and definitions of efficacy/utility that could be more generally applied in practical vaccine dose selection.

Time-response dynamics were not considered

I did not use any models within this work that could be used to describe the dynamics of immunogenicity with respect to time. Models can be used to describe how immune response increases, peaks, and then decays following vaccination. In combination with immunogenicity models these temporal models can be used to predict the decline of vaccine efficacy over time [223]. This could be important for deciding optimal dose, and therefore there exist a benefit to investigating and using models that can describe longitudinal dose-response. Previous IS/ID models of vaccine dose-immunogenicity described by Rhodes et al. and Handel et al. have been used to describe the increase and then decline of cellular and humoral immune response [49,50]. Time-response dynamics were also considered by Kumbhari et al. and used to suggest optimal interval between vaccine doses[130].

Whilst this represents a weakness of the modelling methods I discussed in my chapters; I believe that the modelling I described was reasonable. Firstly, these mechanistic models that can describe longitudinal dose response may be more complicated, may require more data, and require that immunogenicity data be available at multiple time points. Secondly, the primary aim of IS/ID modelling is in aiding selection of optimal vaccine dose. The benefit of longitudinal dose-response modelling may not be relevant unless the utility functions used to define optimal dose either depend on immunogenicity at multiple time points or on model-predicted immunogenicity at time point(s) for which immunogenicity was not measured. For example, the utility function in Handel et al. uses the area-under-curve of the vaccine time-morbidity curve as part of their proposed utility function [49]. This again represents a weakness of this work that may be more relevant as the field of IS/ID develops.

337

A weakness of the chapters that provided a simulation study of clinical trial design and dose-optimisation approaches (5 and 6) was that I did not discuss the inclusion of a placebo group in vaccine dose ranging trials. Placebo arms are commonly used in vaccine trials and may be particularly useful for determining whether adverse events are vaccine related [20]. These can also be used to provide evidence that the vaccine does cause immunogenicity/efficacy.

I decided against including placebo arms in my simulation studies as it added additional complexity without providing meaningful extensions of the field. For CRM-style dose optimisation, previous work has already investigated how modelling-based dose-optimisation can be used when there is a placebo group [224]. One suggested method is to conduct a CRM style dose-optimisation approach using half of the trial participants and to assign the other half to the placebo arm. Inclusion of a placebo arm for the adaptive naive dose-optimisation approach in paper 5 could be done by simply including another arm for each of the scenarios to represent the placebo dosing-group.

Whilst I believe that not discussing the inclusion of placebo dosing groups was reasonable, there is some potential for future investigation. Inclusion of a placebo dosing group could be beneficial to implementing stopping rules. Investigating the inclusion of placebo groups may be needed if the ethical and regulatory bodies require placebo arms in dose-ranging trials, though I note that recently the European Medicines Agency described the inclusion of a control group a choice not a requirement, and the Food and Drug Administration noted that there are both advantages and disadvantages to including placebo arms in a dose-ranging trial [225,226].

Implications

Each chapter has specific implications for the development of the field of IS/ID modelling and vaccine dose optimisation.

In paper 1 and 2, I showed that there is evidence to suggest dose-immunogenicity for some adenoviral vector vaccines may be best described by a peaking dose-response curve. This furthers the evidence that future mathematical-modelling and vaccine dose-response optimisation should not assume that increasing dose will increase vaccine efficacy. If vaccine dose-efficacy curve shape is incorrectly assumed to be saturating this could lead to vaccine dose being selected with decreased efficacy and/or increased adverse event risk. This implies that vaccine development which assumes a saturating dose-efficacy relationship may limit the benefit of the vaccine to public health. Additionally, I found evidence to suggest that the dose-immunogenicity curve shape may be dependent on the host species being vaccinated, which may imply that future attempts to develop allometric scaling methodologies for translating dose between species may be limited.

In paper 3, I showed that mathematical modelling can be applied to select a vaccine dose that is optimal when considering vaccine dose selection as a multi-objective optimisation problem. The doses that I predicted were optimal were dependent on the utility function that was used. This implies that doses may or may not be considered 'optimal' depending on the utility function chosen. Therefore, where vaccine doses are selected using vaccine dose-finding trials with no specified utility function, 'optimality' of these vaccine doses is not well defined.

In paper 4, I predicted through simulation studies that model-informed adaptive trial designs may improve the expected benefit to trial participants for vaccine dose-ranging trials, as later trials participants in adaptive design clinical trials are likely to receive doses that are predicted to be effective given the data from early trials participants. Therefore, vaccine developers should consider these dose-optimisation approaches. I also found that model predictions may overestimate the utility of the predicted optimal dose ('optimistic bias'), which could lead to suboptimal vaccine policy and practical vaccine benefit being less than predicted. I also found that weighted average models of dose-efficacy may be beneficial for the selection of optimal vaccine dose, which may reduce the potential for model misspecification to cause predictions of sub-optimal vaccine dosing. The final implication was that 'exploitative'/'greedy' selection of trial doses in adaptive design (where only the predicted optimal dose after each cohort is investigated in the next cohort) may

prevent optimal vaccine dose from being found, as this selection method may provide insufficient data to properly estimate the entire dose-response curve. Therefore, consideration of the exploration/exploitation trade-off through testing doses that are predicted to be suboptimal may improve IS/ID mathematical modelling-based dose-optimisation. The recommendations of this work for the field of IS/ID modelling are that weighted model averaging should be considered when modelling dose efficacy, vaccine dose-ranging trials that use adaptive design must balance exploration and exploitation, and that vaccine developers should be cautious when reporting model-based predictions of the utility of the predicted optimal dose. In paper 5, I found through simulation that the CoBe dose-optimisation approach, which combined the Continuously Correlated Beta Process model with Thompson sampling, was an effective and ethical approach to conducting vaccine dose-finding studies for the scenarios that I evaluated it on. This was particularly true for prime-boost vaccine dose selection which aimed to maximise a utility function of efficacy and toxicity. This implies that the CoBe dose-optimisation approach may be an effective approach for selecting optimal vaccine dose that reduces the number of individuals required to locate optimal vaccine dose and may therefore reduce cost of clinical trials. It also furthers evidence that non-parametric models are useful for selecting optimal dose when dose-efficacy curve shape is unknown.

Overall implications of the work below

The overall findings of this work suggest that whilst it may not be reasonable to assume that vaccine dose-efficacy curve shape is always best described as saturating, as is possible for model-based drug development, IS/ID mathematical modelling has the potential to be a useful tool in ensuring that vaccines are dosed to provide maximal benefit.

Relative to the mechanistic models highlighted in previous IS/ID modelling research, the models and methodologies described in this thesis were simple. The parametric and non-parametric models used in chapters 4-6 required fewer assumptions regarding underlying immunodynamics than these previous mechanistic models and could be used to model the dose-response curve for any binary outcome response. Whilst mechanistic modelling of dose-immunogenicity has shown potential for the

purposes of vaccine dose optimisation, this thesis implies that even simple models of dose-response can be used to successfully optimise vaccine dose. Use of a simpler model reduces uncertainty in predictive ability and generalisability of the model, which is sometimes the case for more complex models, and improves accessibility.

This thesis suggests that there are modelling based possibilities for vaccine dose selection that may not have been previously available or validated. The simulation studies in this work showed that mathematical modelling can still be effective when dose-efficacy is non-saturating, which validates the use of IS/ID modelling and implies that vaccine doses can be selected with more quantitative justification. This implication is important, as model-based dose selection is already considered preferable in drugs. This work also implies that vaccine clinical trials could be conducted more efficiently and ethically using modelling and adaptive design, which may reduce trial size and improve willingness for participants to enrol and reduce loss to follow-up.

Clinical and vaccine development implications of the overall thesis

I believe that combined findings of papers 2 and 4 imply that the assumption of a saturating dose-efficacy curve may neither be justified by historical data nor beneficial to optimal dose selection when compared to assuming a peaking dose-efficacy curve or using a weighted average of peaking and saturating models. Non-parametric models of dose-efficacy and dose-toxicity such as the CCBP model should also be considered. My findings agree with previous works that have shown that non-parametric models are able to locate optimal drug dose with similar capacity to parametric models which assume a dose-response curve shape. This implies that vaccine developers may not need to assume that a specific curve shape will best describe the true vaccine dose-efficacy curve in order to effectively use mathematical modelling. These findings also imply that clinical trial designs which assume a saturating dose-efficacy curve, such as the 3+3 or other maximum-tolerated dose designs commonly used in drug development, may not be appropriate or effective for locating optimal vaccine dose.

In papers 4 and 5 I found that incorporating adaptive trial design in simulated vaccine dose-finding trials, with or without mathematical modelling, led to increased benefit to trial participants. This implies that vaccine dose-ranging trials which do not use these methods may not be as ethical with regards to maximising benefit to trial participants.

A finding of chapter 6/paper 5 was that considering only a small number of doses within a dose-finding trial may limit vaccine utility, in agreement with Diniz et al [196]. This may imply that methodologies of vaccine dose-finding that involve a small number of doses may reduce vaccine benefit. In combinations with the findings of chapter 2/paper 1, which showed that typically less than six dosing groups were used in adenoviral dose-ranging trials, this implies that adenoviral vector vaccines may be sub-optimally dosed.

This work has also implied that, given that vaccine dose-optimisation should potentially be viewed as a multi-objective optimisation problem, it is important to consider how 'optimal' dose is defined. This work also implies that selection of the utility function used to define 'optimal' dose a-priori of dose-ranging trials being conducted is important. This follows from my finding that adaptive trial design leads to more ethical clinical trials and reduces the number of trial participants required to locate optimal dose, and the knowledge that these adaptive trial designs (both SoftMax selection and Thompson sampling) require a-priori definition of a utility function.

Future research and IS/ID implications of the overall thesis

This work had two major implications regarding future research and IS/ID modelling. Firstly, the findings of chapter 5 showed that mathematical modelling based adaptive design which only conducts trials using the predicted optimal dose are likely to be sub-optimal for the purposes of optimising vaccine dose. This may imply that clinical trials conducted using such a method of trial dose selection may have had a reduced probability to locate the true optimal dose. I recommend that the

exploration/exploitation trade-off be considered for all adaptive design vaccine clinical trials.

Secondly, the finding that 'overestimation bias' is possible when predicting optimal vaccine dose using mathematical modelling. This implies that policy decisions made using predictions of utility for the predicted optimal dose may be sub-optimal. This finding may also be the case when vaccine dose is selected without using mathematical modelling. This implication is therefore also relevant to vaccine developers and statisticians conducting direct-comparison vaccine dose-ranging studies.

Social Implications of the overall thesis

In chapter 6/paper 5 I showed that dose-optimisation approaches which used no adaptive design or modelling, and instead choose 'optimal dose' via direct comparison of a small number of dosing groups, required a larger number of individuals to locate optimal vaccine dose. These dose-optimisation approaches that used no adaptive design or modelling were also less beneficial to the simulated trial participants relative to dose-optimisation approaches which did use adaptive trial design. 3+3 designs or other such dose-optimisation approaches that assume that efficacy increases with dose may not be justified either due to the findings of paper 2. These findings imply, given that such dose-optimisation approaches are commonly used for conducting vaccine dose-ranging trials, that many modern vaccines may be sub-optimally dosed.

I note that the statement 'vaccines may be sub-optimally dosed' may be difficult to prove or refute for vaccines where researchers or vaccine developers did not define a quantitative utility function. In these cases, it is possible that a utility function could be retro-actively chosen to justify or criticise the choice of dose.

This work implies that a combination of considering only a small number of doses, not using modelling, not using adaptive trial design, assuming a saturating dose-efficacy curve, and not defining a quantitative utility function could lead to suboptimal vaccine doses and decrease the potential societal benefits that vaccines could provide. Conducting vaccine dose-finding trials using better methods could lead to

better vaccine dosing, which could improve the societal benefits of vaccines and save lives.

Future work

Validation of theoretical methods

The most important piece of future work is validating the predictions made in my work. I showed in both paper 4 and 5 that parametric models may not need completely accurate descriptions of the true dose-response in order to enable effective dose-optimisation. Despite this, vaccine developers and regulatory bodies need to be confident that the predictions of these models can be trusted before these models can be used in practice.

Bonate [227] makes the distinction that dose-response models should be both 'validated' and 'credible'. 'Validated' models are those for which the model's predictions of past and future data are dependable. 'Credible' models are those which are accepted by users or decision makers. Validation and credibility are not necessarily the same; a model can be validated without being credible or credible without being validated. However, validation tends to improve credibility, as does model interpretability. I hope that in providing simulation studies of these models and primarily using simple statistical models I have improved the credibility of IS/ID modelling, and that with validation these models become credible for practical use. A validation study could aid in this goal.

A simple trial validation study of these models could be to gather data from a published dose-ranging trial with at least three dosing groups and calibrate a dose-response model to this data using the techniques discussed in this work. This model would then be used to predict response for a dose that was not used in the original dose-ranging trial, then a clinical trial conducted of this 'validation' dose using equivalent populations and data-collection methods as the original published trial. A hypothesis testing approach could then be used to evaluate whether the models' predictions matched the empirical results. The validation dose should be chosen to be between the largest and smallest doses that were evaluated in the original dose-ranging trial (as statistical models are primarily useful for interpolation rather than

extrapolation). Additionally, given the findings to an 'optimistic bias' in chapter 5, the validation dose should not be at the predicted optimal dose.

If funding were available, then a specific validation study could be conducted to investigate this, following the same steps as described above except replacing the previously published data with data from a dose-ranging trial specifically conducted for this purpose. In this case, for a single-administration vaccine, I would propose that the initial dose-ranging trial preferentially investigate a large number of small dosing groups rather than a small number of large dosing groups. This data could then be used for modelling, and prediction/validation conducted as above.

Alternatively, if it were not possible to conduct further clinical trials, a similar validation study could also be done using only published data. This could be conducted by calibrating models to data from all but one dosing group of a published dose-ranging trial and predicting the responses for the excluded dosing group. However, this may be less trusted, as an unscrupulous researcher could bias such a study by choosing data that is very 'predictable', or by leveraging their knowledge of the validation data when building the model to improve the apparent predictive validity.

I note in this recommendation of future work that similar validation of mechanistic modelling should also be conducted to further the field of IS/ID modelling, and that for these models a validation of extrapolated predictions may also be reasonable to investigate. I also note that, according to Bonate, models 'cannot be 100% validated', therefore validation may be a question of how accurate a model must be to be acceptable to modellers and developers [227].

Practical application of mathematical modelling in vaccine dose optimisation

Once these modelling techniques are both validated and credible, the next step in furthering the field of IS/ID and mathematical modelling for vaccine dose optimisation would be practical application of these methods for the purposes of informing real-world vaccine dosing and dose-ranging clinical trials. The 17DD yellow-fever vaccine was found to have been initially overdosed, wasting vaccine supply in a limited vaccine setting. It has also been suggested that the clinical doses for the MRNA-

1273 and BNT162b COVID-19 vaccines may be in excess of the optimal doses, with a decreased dose potentially allowing for increased global benefit [228,229]. Thus, there is already discussion on whether the present empirical methods of dose selection are leading to optimal vaccine dosing [191].

In this work I investigated the theoretical application of mathematical modelling methodologies in optimisation vaccine dose, but whether these methods outperform current methodologies still requires validation. This would require vaccine developers to believe that there is potential benefit to using such methods, along with approval from regulatory bodies. I note that model-based drug development has been shown to be practically effective in choosing optimal drug dose, and so there is evidence to suggest that modelling is theoretically effective. I do however believe that the terminal goal of the development of the field of IS/ID modelling is improving vaccine dosing and saving lives. Thus, theoretical work is only beneficial if it enables or inspires practical use of these tools, or if it highlights pitfalls that could lead to suboptimal dosing or policy decisions (such as 'optimistic bias').

These mathematical modelling methods and dose-optimisation approaches are likely to be of interest to vaccine developers/pharmaceutical companies and regulatory bodies, and it is likely that these can feasibly be incorporated into and improve the current vaccine development pathway.

These findings and methods are likely to be of interest to vaccine developers/pharmaceutical companies. If modelling can lead to more efficacious and less toxic vaccines, then there may be increased vaccine uptake. If similar efficacy can be achieved with smaller doses, then the cost of vaccine manufacture may be reduced. If optimal dose can be located using a smaller number of trial participants, this could reduce the cost of clinical trials in vaccine development. Further, it is possible that these methods could reduce the number of vaccine candidates that fail

to proceed past early phase clinical trials[9], reducing the cost of vaccine development. All of these factors could make vaccine development more profitable/less expensive.

Regulators have an interest in ensuring that marketed vaccines are maximally safe and effective. Regulators and ethics board also have an interest in ensuring that clinical trials are conducted ethically in a way that minimises harm or maximises benefit to trial participants. For disease control during a pandemic, regulators may desire accelerated vaccine trials and dose selection whilst ensuring that the above two interests (safe and effective vaccines, ethical trials) are maintained. Based on this thesis and on previous findings I believe that regulators therefore would have an interest in conducting future vaccine development using IS/ID mathematical modelling and adaptive trial design.

If these models and methods had been available to and accepted by developers and regulators during the COVID-19 pandemic, vaccine development could have been further accelerated, fewer vaccine candidates may have been needed, and there might be improved vaccine dosing and reduced cost of development. Given the simplicity of the statistical models discussed in this thesis relative to mechanistic modelling, I believe that the complexity of future work incorporating these models into vaccine development would be reduced. Given that such models have historically been used and recommended for use in phase I/II drug trials, I believe that these are the trial phases that developers and modellers should first look to practically apply these methods in.

These methods could also be used pre-clinically to determine optimal vaccine dose in animals, which may be useful if allometric scaling methods can be further developed. Additionally, initial practical application of these models and dose-

---

[9] Consider, for example, chapter 6/paper 5, objective 3, scenario 4. None of the six doses investigated using the 'Uniform Naive' dose-optimisation approach (representing a traditional method of selecting vaccine doses) had a utility greater than zero. Therefore, this DOA could not locate a dose with a utility greater than zero. The dose-optimisation approaches that used modelling on average located a dose with utility greater than zero with fewer than 30 trial participants. If utility greater than zero was a go/no-go criterion for a vaccine with these dose-efficacy and dose-toxicity curves, then modelling could be the difference between a marketable and failed vaccine candidate.

optimisation approaches in preclinical trials may be reasonable for improving the credibility of these methods for later application in human clinical trials. I also believe that these methods could be practically applied for smaller phase IV trials for vaccines that have already been marketed. In this way mathematical modelling could be used to improve dosing of vaccines that are known to be safe and effective. As these vaccines are already known to be safe, regulators may be less hesitant about applying these methods.

The additional importance in practical application of these tools is that it may highlight areas where the field of IS/ID modelling needs additional further work. Conducting clinical trials using these methodologies may present issues of practicality, for example how to incorporate adaptive trial design with potential blinding or double-blinding requirements. Only through interaction and discussion between modellers and vaccine developers can tools be developed that are practical, ethical, and effective.

Model extensions

In addition to these methods seeing further practical development, there are a number of areas of future work regarding improvements and extensions to IS/ID models.

Including population covariates

As noted, a weakness of this thesis was in not discussing the inclusion of covariates into the dose-response modelling process. In model-based drug development, inclusion of the covariates of age, weight, or disease progression/severity are may be used [177] and have previously been used for IS/ID modelling to account for within population variation [51]. These can be incorporated in several ways. Hierarchical and mixed effects modelling have been suggested as useful for incorporating covariates into both mechanistic and statistical models [230,231]. Wijesinha and Piantadosi [232] suggested a parametric method for statistical models that involved estimating separate model parameters, one each for data from female and male trial participants. They also suggested a semi-parametric approach, where covariates can increase response probability in either a dose-dependent or dose-

independent method. All of these methods were highlighted as greatly increasing the number of parameters that must be estimated from the data, meaning a greater trial size might be needed to estimate parameters of the dose response curve. The CCBP model has also been extended to include contextual factors [233] and investigating whether this could allow for covariates to be adjusted for in the CoBe DOA would be an interesting area of future work.

Developing parametric models for prime/boost administration vaccines

In paper 5 I investigated the use of vaccine dose-optimisation approaches that made use of parametric models. The findings of that paper showed that such methods are theoretically effective for maximising efficacy or utility for single-administration paradigm vaccines. Maximising efficacy dose for prime/boost paradigm vaccines was also typically reasonable, however the parametric dose-optimisation approach performed poorly for selecting a prime/boost paradigm dose that maximised the utility function of efficacy and toxicity (paper 5, objective 3, scenario 5 and 6).

The findings of that paper showed that the non-parametric CoBe dose optimisation approach was likely to be effective in that setting. However, given that parametric models are more common than non-parametric models for the purposes of drug dose optimisation, it might be useful to investigate other potential parametric dose-efficacy and dose-toxicity models in that setting. Simulation studies could be used to determine whether mechanistic compartmental or agent-based models may be effective. Alternatively, statistical models other than the simple models included in paper 5 should be considered.

Whilst I discuss prime/boost dose optimisation in paper 5, the multi-dimensional dosing scenarios in objectives 2, 3 and 4 for that paper could also be considered to represent antigen/adjuvant dose optimisation. I believe that the findings or paper 5 suggest that it may be beneficial to conduct further simulation studies to investigate IS/ID modelling when locating the optimal doses of single-administration antigen-adjuvant paradigm vaccine. Modelling has been suggested as potentially useful for finding optimal antigen-adjuvant dose [234], but not applied or investigated practically to my knowledge. I believe that CCBP models and the CoBe dose-

optimisation approach would be likely to be effective for locating such a dose, given the findings of paper 5. Parametric models of antigen/adjuvant dose-response could be investigated as well.

Developing models for optimising time between doses

In the background I highlighted that magnitude of vaccine administered ('dose') is not the only decision that is faced by vaccine developers and that can be decided using dosing running trials. The prime-boost interval may also affect vaccine efficacy and toxicity. Studies have found that changing this interval can affect immunogenicity [235]. Findings for two COVID-19 vaccines found no effect from the prime-boost interval on toxicity [236], but I do not consider this to be conclusive evidence for all vaccines. There is future work in developing models that may allow for the optimal prime-boost interval to be determined. Agent-based modelling of the immune system has been used previously to describe how optimal prime-boost interval could be determined [237]. However, this was a simulation study and thus suffers from the same weaknesses I have previously noted. They also only described optimising the prime-boost interval, with prime and boost doses fixed. Thus, future work would be required before modelling can be used to concurrently optimise prime dose, boost dose and prime-boost interval.

Again, I believe that the findings of paper 5 could be generalised to suggest that the CoBe dose-optimisation approach may be effective here, as no parametric form for the relationship between interval-response would need to be assumed. Further simulation studies using scenarios which are specifically believed to reflect the interval-response relationship that may be expected for real life vaccines could be conducted to validate this.

Investigating parametric dose-response modelling that uses Bayesian inference

Previously in model-based drug decision making and literature regarding the continual reassessment method, both frequentist and Bayesian methodologies have been used and there is no strong guidance to suggest either is more preferable [200,238].

Frequentist modelling typically implies that no prior is placed over the model parameters, and that point estimates (such as the maximum likelihood estimate or least squares estimate) are calculated for model parameters and model predictions. Using point estimates is sometimes called 'likelihood-based' inference [239]. Bayesian modelling typically implies that there are priors placed over model parameters and that posterior distribution will be determined for model parameters and predictions. Determining these posterior distributions is sometimes called 'Bayesian' inference [160].

With the exception of the CCBP models in paper 5, all modelling in this work was conducted using 'likelihood-based' inference. Whilst I believe that I showed that such models could be effective for optimising vaccine dose, there may be merit in exploring Bayesian inference.

Bayesian inference for estimation of posteriors for model parameters and predictions could enable the use of techniques that have previously been described in the literature. Some clinical trials conducted in the CRM style have used these posteriors to define their stopping rules [217]. With regards to my work, Bayesian inference of model parameters in papers 4 may have allowed for a Thompson Sampling method of incorporating exploration into the trial dose selection rather than the SoftMax selection method described. This would have also been possible for the parametric dose-optimisation approach in paper 5. Whilst I showed in paper 4 that the SoftMax method of trial dose selection was effective, Thompson Sampling may be more parsimonious and intrinsically balance exploration and exploitation.

I therefore believe that it would be worth conducting simulation studies to compare the effectiveness of Bayesian inference to that of likelihood-based inference for conducting clinical trials to determine optimal vaccine dose. However, likelihood-based inference is still likely to be effective in practise, and one of the pioneers of Bayesian-modelling for dose-ranging trials Peter Thall suggested that Bayesian modelling has had limited use in practise despite a strong theoretical grounding, so such future research may be of less importance than the other topics mentioned here [107,240].

In paper 4 I showed that it may be important to include exploration in dose-ranging vaccine trials that used modelling and adaptive design. I investigated this further and found that this became even more important as trial size increased, see [appendix A.G.]. Within that paper and in paper 5 I used SoftMax selection, and whilst this appeared to improve dose selection over the fully continual standard method, it was a simple heuristic for incorporating exploration. The uniform dose selection method also appeared to be reasonably effective but led to high average regret.

As noted above, Bayesian inference may allow for the use of Thompson Sampling based methods of trial dose selection. Augmentations of the Thompson Sampling method, such as the acceleratedTS or TS-UCB algorithms, could be investigated [241,242]. There may be other methods of trial dose selection and trial dose selection that could warrant investigation, in particular D-optimal design.

If parametric models of dose-response are known, then D-optimal designs can be used [161,162,243]. Here, the method of trial dose selection is to distribute trial doses in a way that maximises the 'expected information gain' about model parameters. This can allow for 'efficient' use of trial participants and hence smaller clinical trials. D-optimal design has also been extended to 'penalised D-optimal design', which can balance the efficiency of the trial with the benefit to trial participants, theoretically leading to more ethical trial design [244]. D-optimal design can be used both prospectively to design a clinical trial or used to guide adaptive trial design [163].

Whilst D-optimal design has been suggested to be effective for optimising drug dose, it is a significantly more complicated methodology than any of those described in the papers contained in this work and has limited use in practical dose-finding [245]. Additionally, it is only feasible if a parametric model or models are known a-priori, including parameterisation, and if clinicians and modellers are confident that these models are accurate to the truth. Such assumptions may not be valid for all vaccine development, as discussed throughout this work. Regardless, such methods warrant

investigation. A brief description of D-optimal design and its application to one of the scenarios in paper 4 is given in Appendix A.E.

Further development of mechanistic models and immunogenicity-efficacy models

In this work, I chose to use statistical modelling methodologies to describe vaccine dose-response and to predict optimal vaccine dose. Whilst I believe that this was appropriate due to their simplicity and being comparatively more applicable to general vaccine development than mechanistic models of specific immunogenicity response, I also believe there is merit to additional research into mechanistic dose-immunogenicity models. As described, these have already been discussed and appear to have been effective where they have been used.

Additionally, for some immune responses there may not be a consensus on the underlying biology and hence the correct mechanistic dose-immunogenicity models. The works of Moore et al. [67] and Mayer et al. [69] propose a combined three different models of T cell proliferation. Moore in particular highlights that the antigen dependent model for T cell proliferation that was previously accepted may not well describe observed T cell response to a yellow fever vaccine. Kumbhari et al.'s proposed models [82,130] of murine CD8+ to an adjuvanted melanoma peptide vaccine [246] were again qualitatively different to the three described above. More works similar to that of Rhodes et al. [50] and Moore et al. [67], where multiple models are compared, would be beneficial to developing the field. Papers performing predictive validation of mechanistic dose-immunogenicity models would also be beneficial for demonstrating that such modelling may be effective for the purposes of dose-optimisation. However, these models are data dependent and the kind of data needed to parameterise a 'useful' mechanistic model is not regularly generated by traditional vaccine clinical trials.

Relatedly, I believe again that immunogenicity-efficacy model development would also benefit IS/ID modelling, where such work is feasible, and that this is needed to maximise the potential of mechanistic dose-immunogenicity models for selecting optimal vaccine dose.

IS/ID modelling and mathematical modelling for optimal vaccine dose selection appears theoretically promising based on the findings of this thesis and previous research. Use of these methodologies in practice would require the collaboration of vaccine developers with modellers, which may require an additional expense beyond what is typically required. Additionally, the models and methods previously investigated in drugs which require the assumption that efficacy increases with dose may not be justified, and so statistical analysis and trial design may have to be bespoke. This may be expensive, and vaccine developers may not be confident that the potential benefits of modelling justify these additional costs, complexities, and deviations from the standards of vaccine trial design.

I therefore believe that an important piece of future work would be developing and validating statistical software that encompasses all elements of IS/ID modelling. This would be for the purpose of vaccine dose-response modelling and conducting vaccine dose finding studies. This software would aim to allow developers to both trust the modelling more through validation and to reduce the expense of including such methods. Ideally, such a piece of software would therefore be available under an open-source licence, simple to use an intuitive graphical user interface, and flexible enough to incorporate models for various paradigms of vaccine administration. Uptake of such software would also require promotion of the software and methods through publications and conferences to maximise the proportion of vaccine developers that would be aware that these tools were available.

There are some examples of similar software that has been developed for MBDD. Software such as the Monolix, NONMEM, and Phoenix suite of tools are examples of such software for model-based drug development that provide simple graphical interfaces [247–249]. These however are not open-source. The 'trialr' and 'crm' R packages are open-source but require some familiarity with the R language to use[250,251].

354

Additionally, in paper 5 I highlighted that CCBPs model and CoBe dose-optimisation approach may be effective for locating optimal vaccine dose. Neither of these are available as part of an available piece of statistical software that is designed for practical use. Until a piece of software is designed that would allow for practical implementation of this model and dose-optimisation approach, use of these may be limited.

Further, there is also the need for practical frameworks to be developed to enable the use of these methodologies in practical vaccine dose selection. Model-based drug development is well validated, and organisations like the European Medicines Agency (EMA) have guidance for how pharmacokinetic/pharmacodynamic studies should be conducted [252]. However, the EMA guidelines for clinical evaluation of new vaccine discusses mathematical modelling only in the context of predicting the need for potential boosters after the primary series and provides no explicit guidance for the use of mathematical modelling in dose selection [253]. This may increase hesitancy to use these methods.

Development of clinically relevant utility functions

I argue in this work that vaccine dose optimisation should be considered as a multiple-objective optimisation problem, and that vaccine developers must specify quantitative utility functions in order for the 'optimal' dose to be meaningfully defined. Frameworks for developing utility functions that reflect clinician's belief about the relative importance of vaccine related immunogenicity, toxicity or other objectives would likely be beneficial.

Additionally, in this work I mainly discussed 'a-priori' multi-objective optimisation, where a utility function is chosen prior to data collection/independently of the observed data and modelling predictions. Multi-objective optimisation literature outside of dose selection has discussed 'a-posteriori' multi-objective optimisation [96]. Vaccine multi-objective dose optimisation using this method would not define a utility function prior to experimentation (or ever), after data collection and modelling key stakeholders would be told which doses are on the Pareto front along with the predictions of efficacy/toxicity/etc. for these doses, and then stakeholders would

state their preference between the doses based on these predictions. As these 'a-posteriori' methods do not define a utility function prior to experimentation, the continual modelling or adaptive design methodologies that I have suggested to be useful in ensuring ethical clinical trials could not be used.

Alternatively, 'interactive' multi-objective optimisation has also been discussed. Vaccine multi-objective dose optimisation using these methods would follow an adaptive trial design where the utility function is specified 'a-priori' but may be altered at interim stages of a trial. This allows for adaptive trial design whilst also allowing for changes in the utility function if clinicians believe that it no longer reflects their preferences. Such methodologies are typically complex to implement. Branke et al. [101] suggested a method of 'interactive' multi-objective optimisation, under which key stakeholders would be asked their order of preference between the doses based on current model prediction at interim time points. An ordinal regression model is calibrated using this preference data, and then the ordinal regression model is used as the utility function for choosing the doses to investigate in the next cohort. Such a method would likely complicate the clinical trial, and also means that the utility function is less explainable.

The 'a-priori' approach to multi-objective optimisation is typically the most supported by classical multi-objective optimisation literature. 'A-posteriori' methods may be vulnerable to retrofitting of utility functions to justify the selection of a specific dose, are less well researched, and may not allow for adaptive design [96]. 'Interactive' methods are more complicated but may be beneficial for ensuring that utility functions are clinically meaningful. I therefore believe that IS/ID modelling would be most benefited by 'a-priori' consideration of multi-objective optimisation, and that future work in developing this is needed. Development of 'interactive' methods of adaptive trial design may also warrant future investigation.

It would also be beneficial for any software that is designed to enable vaccine dose optimisation should include a variety of utility functions. Additionally, it would be preferable that such software should be able to guide developers towards selecting a clinically relevant utility function [Appendix A.C.]. I have created a simple Python project that does this for the flexible 'weighted overall desirability' utility function and

discussed by [244,254] though this does not presently have a graphical user interface and requires some familiarity with the Python programming language to use [255].

I believe that defining a utility function, particularly a-priori of conducting dose ranging trials, is beneficial, and that future work is needed in developing clinically relevant utility functions. This is because it allows continual-modelling/adaptive trial design to be conducted in such a way as to find the dose and shows quantitative consideration of the potential effect that vaccine dose may have.

Accounting for overestimation of utility

Within chapter 5 I found that the model predicted utility at the predicted optimal dose may be 'optimistically biased', meaning that this predicted utility was typically higher than the true utility of that dose. Prior to conducting this study I had assumed, if the models used were appropriate, that the prediction of utility around the predicted optimal dose would be normally distributed around the true value for utility at that dose. However, these findings were consistent with other modelling studies [207–209]. Further, as stated in that chapter, I do not believe that this effect is limited only to modelling based methods of optimisation and believe that it is also prevalent in direct comparison dose-optimisation approaches [Supplementary paper 4]. Clearly if estimates of vaccine efficacy and utility are biased by the process of selecting the predicted optimal dose, then these estimates should be used cautiously if being used to make policy decisions. In particular, epidemiological models have used and may continue to use clinical trial data to predict the societal benefit of a vaccination programme. If the estimates of vaccine utility are overestimated, then this could lead to incorrect estimations and sub-optimal disease prevention.

Future work therefore needs to be done to ensure that vaccine efficacy and utility estimates are not biased. This needs to be investigated for both modelling and non-modelling-based methods of vaccine dose optimisation. For non-modelling-based dose selection, if there are more than two dosing groups and the outcomes of interest are either binomially distributed or normally distributed with homoscedastic variance, then a Stein-Type estimator could be used [256]. However, a PubMed

search for [vaccin* AND Stein AND dos*] on the 13th of June 2022 yielded no results, implying that such methods may not be commonly used in vaccine dose-ranging trials.

For modelling-based dose selection, data could be split in two, half used to calibrate the model used to predict the optimal dose and half used to calibrate the model used to predict the utility at that dose [209]. This halving method means that less data are available to choose dose, which may not be desirable given the typically small size of vaccine dose ranging trials. Alternatively, cross-validation or parameter perturbation-based approaches have been suggested [208]. Cross validation resulted in an underestimation bias. The parameter perturbation method was shown to be unbiased and accurate in a simulation study, but the variances of these estimates were large and so estimates were less precise. Therefore, if it is desirable that vaccine dose-ranging trials provide accurate and precise estimates of vaccine efficacy/utility, then future work is required.

Conclusion

This thesis has expanded the field of IS/ID modelling for the optimisation of vaccine dose. I expanded this field to discuss vaccine dose optimisation as a multi-objective optimisation problem and showed that the assumption of a saturating dose-efficacy curve shape is not needed for mathematical modelling methods to be effective in predicting optimal vaccine dose. I developed and highlighted methods for model-based adaptive clinical trial design which are novel in vaccine development, and, drawing on experience in drug development, are practically applicable. I also showed the potential for conducting more efficient, effective, and ethical vaccine dose-finding trials which is an essential consideration in a cost-heavy, high-pressure field of research. Further, I showed that model-averaging and non-parametric modelling methods could be used to locate optimal vaccine dose even if the true dose-efficacy curve shape is not known a-priori. The Correlated Beta dose optimisation approach that I developed is an innovative and simple approach for vaccine dose optimisation. This leverages the benefits of non-parametric modelling and adaptive design to optimise vaccine dose finding, and I believe has the potential for broad application for both single-administration and prime-boost administration. I believe that future

investigation and practical implementation of mathematical modelling is warranted and has the potential to not only improve vaccine dose selection but also allow for more ethical trial design. Mathematical IS/ID modelling may become an increasingly relevant tool in ensuring safe and effective vaccination campaigns.

References

1. Riedel, S. Edward Jenner and the History of Smallpox and Vaccination. *Proc (Bayl Univ Med Cent)* **2005**, *18*, 21–25.
2. UK Immunisation Schedule: The Green Book, Chapter 11 Available online: https://www.gov.uk/government/publications/immunisation-schedule-the-green-book-chapter-11 (accessed on 13 August 2022).
3. van Riel, D.; de Wit, E. Next-Generation Vaccine Platforms for COVID-19. *Nat. Mater.* **2020**, *19*, 810–812, doi:10.1038/s41563-020-0746-0.
4. Ghattas, M.; Dwivedi, G.; Lavertu, M.; Alameh, M.-G. Vaccine Technologies and Platforms for Infectious Diseases: Current Progress, Challenges, and Opportunities. *Vaccines (Basel)* **2021**, *9*, 1490, doi:10.3390/vaccines9121490.
5. Sela, M.; Hilleman, M.R. Therapeutic Vaccines: Realities of Today and Hopes for Tomorrow. *Proc Natl Acad Sci U S A* **2004**, *101*, 14559, doi:10.1073/pnas.0405924101.
6. Gilbert, S.C. T-Cell-Inducing Vaccines – What's the Future. *Immunology* **2012**, *135*, 19–26, doi:10.1111/j.1365-2567.2011.03517.x.
7. Greenwood, B. The Contribution of Vaccination to Global Health: Past, Present and Future. *Philos Trans R Soc Lond B Biol Sci* **2014**, *369*, 20130433, doi:10.1098/rstb.2013.0433.
8. Wouters, O.J.; Shadlen, K.C.; Salcher-Konrad, M.; Pollard, A.J.; Larson, H.J.; Teerawattananon, Y.; Jit, M. Challenges in Ensuring Global Access to COVID-19 Vaccines: Production, Affordability, Allocation, and Deployment. *Lancet* **2021**, *397*, 1023–1034, doi:10.1016/S0140-6736(21)00306-8.
9. Vaccine Testing and Approval Process | CDC Available online: https://www.cdc.gov/vaccines/basics/test-approve.html (accessed on 13 August 2022).
10. Vaccine Development, Testing, and Regulation Available online: https://historyofvaccines.org/vaccines-101/how-are-vaccines-made/vaccine-development-testing-and-regulation (accessed on 13 August 2022).
11. Strategies, I. of M. (US) C. on the C.V.I.P.A.; Mitchell, V.S.; Philipose, N.M.; Sanford, J.P. *Stages of Vaccine Development*; National Academies Press (US), 1993;
12. The Complex Journey of a Vaccine – The Steps Behind Developing a New Vaccine Available online: https://www.ifpma.org/resource-centre/the-complex-journey-of-a-vaccine-final/ (accessed on 13 August 2022).
13. Han, S. Clinical Vaccine Development. *Clin Exp Vaccine Res* **2015**, *4*, 46–53, doi:10.7774/cevr.2015.4.1.46.
14. Singh, K.; Mehta, S. The Clinical Development Process for a Novel Preventive Vaccine: An Overview. *J Postgrad Med* **2016**, *62*, 4–11, doi:10.4103/0022-3859.173187.
15. Suvarna, V. Phase IV of Drug Development. *Perspect Clin Res* **2010**, *1*, 57–60.
16. Gouglas, D.; Le, T.T.; Henderson, K.; Kaloudis, A.; Danielsen, T.; Hammersland, N.C.; Robinson, J.M.; Heaton, P.M.; Røttingen, J.-A. Estimating the Cost of Vaccine Development against Epidemic Infectious Diseases: A Cost Minimisation Study. *The Lancet Global Health* **2018**, *6*, e1386–e1396, doi:10.1016/S2214-109X(18)30346-2.
17. Pronker, E.S.; Weenen, T.C.; Commandeur, H.; Claassen, E.H.J.H.M.; Osterhaus, A.D.M.E. Risk in Vaccine Research and Development Quantified. *PLoS One* **2013**, *8*, e57755, doi:10.1371/journal.pone.0057755.
18. Bok, K.; Sitar, S.; Graham, B.S.; Mascola, J.R. Accelerated COVID-19 Vaccine Development: Milestones, Lessons, and Prospects. *Immunity* **2021**, *54*, 1636–1651, doi:10.1016/j.immuni.2021.07.017.
19. Research, C. for B.E. and Clinical Considerations for Therapeutic Cancer Vaccines Available online: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/clinical-considerations-therapeutic-cancer-vaccines (accessed on 13 August 2022).

20. Haas, J.W.; Bender, F.L.; Ballou, S.; Kelley, J.M.; Wilhelm, M.; Miller, F.G.; Rief, W.; Kaptchuk, T.J. Frequency of Adverse Events in the Placebo Arms of COVID-19 Vaccine Trials: A Systematic Review and Meta-Analysis. *JAMA Network Open* **2022**, *5*, e2143955, doi:10.1001/jamanetworkopen.2021.43955.

21. Campi-Azevedo, A.C.; de Almeida Estevam, P.; Coelho-dos-Reis, J.G.; Peruhype-Magalhães, V.; Villela-Rezende, G.; Quaresma, P.F.; Maia, M. de L.S.; Farias, R.H.G.; Camacho, L.A.B.; Freire, M. da S.; et al. Subdoses of 17DD Yellow Fever Vaccine Elicit Equivalent Virological/Immunological Kinetics Timeline. *BMC Infect Dis* **2014**, *14*, 391, doi:10.1186/1471-2334-14-391.

22. Research, C. for B.E. and Development and Licensure of Vaccines to Prevent COVID-19 Available online: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/development-and-licensure-vaccines-prevent-covid-19 (accessed on 13 August 2022).

23. Rhodes, S.J.; Knight, G.M.; Kirschner, D.E.; White, R.G.; Evans, T.G. Dose Finding for New Vaccines: The Role for Immunostimulation/Immunodynamic Modelling. *J Theor Biol* **2019**, *465*, 51–55, doi:10.1016/j.jtbi.2019.01.017.

24. Lalonde, R.L.; Kowalski, K.G.; Hutmacher, M.M.; Ewy, W.; Nichols, D.J.; Milligan, P.A.; Corrigan, B.W.; Lockwood, P.A.; Marshall, S.A.; Benincosa, L.J.; et al. Model-Based Drug Development. *Clinical Pharmacology & Therapeutics* **2007**, *82*, 21–32, doi:10.1038/sj.clpt.6100235.

25. Kimko, H.; Pinheiro, J. Model-Based Clinical Drug Development in the Past, Present and Future: A Commentary. *Br J Clin Pharmacol* **2015**, *79*, 108–116, doi:10.1111/bcp.12341.

26. Lu, S. Heterologous Prime-Boost Vaccination. *Curr Opin Immunol* **2009**, *21*, 346–351, doi:10.1016/j.coi.2009.05.016.

27. Siddiqui, A.; Adnan, A.; Abbas, M.; Taseen, S.; Ochani, S.; Essar, M.Y. Revival of the Heterologous Prime-boost Technique in COVID-19: An Outlook from the History of Outbreaks. *Health Sci Rep* **2022**, *5*, e531, doi:10.1002/hsr2.531.

28. Hall, V.G.; Ferreira, V.H.; Wood, H.; Ierullo, M.; Majchrzak-Kita, B.; Manguiat, K.; Robinson, A.; Kulasingam, V.; Humar, A.; Kumar, D. Delayed-Interval BNT162b2 MRNA COVID-19 Vaccination Enhances Humoral Immunity and Induces Robust T Cell Responses. *Nat Immunol* **2022**, *23*, 380–385, doi:10.1038/s41590-021-01126-6.

29. Payne, R.P.; Longet, S.; Austin, J.A.; Skelly, D.T.; Dejnirattisai, W.; Adele, S.; Meardon, N.; Faustini, S.; Al-Taei, S.; Moore, S.C.; et al. Immunogenicity of Standard and Extended Dosing Intervals of BNT162b2 MRNA Vaccine. *Cell* **2021**, *184*, 5699-5714.e11, doi:10.1016/j.cell.2021.10.011.

30. Fox, C.B.; Kramer, R.M.; Barnes V, L.; Dowling, Q.M.; Vedvick, T.S. Working Together: Interactions between Vaccine Antigens and Adjuvants. *Ther Adv Vaccines* **2013**, *1*, 7–20, doi:10.1177/2051013613480144.

31. Pulendran, B.; S. Arunachalam, P.; O'Hagan, D.T. Emerging Concepts in the Science of Vaccine Adjuvants. *Nat Rev Drug Discov* **2021**, *20*, 454–475, doi:10.1038/s41573-021-00163-y.

32. Nauta, J. Correlates of Protection. In *Statistics in clinical vaccine trials*; Springer: Heidelberg, 2010; p. p107 ISBN 978-3-642-14690-9.

33. Jamrozik, E.; Selgelid, M.J. COVID-19 Human Challenge Studies: Ethical Issues. *The Lancet Infectious Diseases* **2020**, *20*, e198–e203, doi:10.1016/S1473-3099(20)30438-2.

34. Black, S. The Costs and Effectiveness of Large Phase III Pre-Licensure Vaccine Clinical Trials. *Expert Rev Vaccines* **2015**, *14*, 1543–1548, doi:10.1586/14760584.2015.1091733.

35. Sadoff, J.C.; Wittes, J. Correlates, Surrogates, and Vaccines. *The Journal of Infectious Diseases* **2007**, *196*, 1279–1281, doi:10.1086/522432.

36. Dudášová, J.; Laube, R.; Valiathan, C.; Wiener, M.C.; Gheyas, F.; Fišer, P.; Ivanauskaite, J.; Liu, F.; Sachs, J.R. A Method to Estimate Probability of Disease and Vaccine Efficacy from Clinical Trial Immunogenicity Data. *NPJ Vaccines* **2021**, *6*, 133, doi:10.1038/s41541-021-00377-6.

37. Nauta, J. Two Seroresponse Rates. In *Statistics in clinical vaccine trials*; Springer: Heidelberg, 2010; p. p28 ISBN 978-3-642-14690-9.

38. Khoury, D.S.; Cromer, D.; Reynaldi, A.; Schlub, T.E.; Wheatley, A.K.; Juno, J.A.; Subbarao, K.; Kent, S.J.; Triccas, J.A.; Davenport, M.P. Neutralizing Antibody Levels Are Highly Predictive of Immune Protection from Symptomatic SARS-CoV-2 Infection. *Nat Med* **2021**, *27*, 1205–1211, doi:10.1038/s41591-021-01377-8.

39. Gilbert, P.B.; Gabriel, E.E.; Miao, X.; Li, X.; Su, S.-C.; Parrino, J.; Chan, I.S.F. Fold Rise in Antibody Titers by Measured by Glycoprotein-Based Enzyme-Linked Immunosorbent Assay Is an Excellent Correlate of Protection for a Herpes Zoster Vaccine, Demonstrated via the Vaccine Efficacy Curve. *J Infect Dis* **2014**, *210*, 1573–1581, doi:10.1093/infdis/jiu279.

40. Gilbert, P.B.; Luedtke, A.R. Statistical Learning Methods to Determine Immune Correlates of Herpes Zoster in Vaccine Efficacy Trials. *The Journal of Infectious Diseases* **2018**, *218*, S99–S101, doi:10.1093/infdis/jiy421.

41. Kaur, R.J.; Dutta, S.; Bhardwaj, P.; Charan, J.; Dhingra, S.; Mitra, P.; Singh, K.; Yadav, D.; Sharma, P.; Misra, S. Adverse Events Reported From COVID-19 Vaccine Trials: A Systematic Review. *Indian J Clin Biochem* **2021**, *36*, 427–439, doi:10.1007/s12291-021-00968-z.

42. Mutsch, M.; Zhou, W.; Rhodes, P.; Bopp, M.; Chen, R.T.; Linder, T.; Spyr, C.; Steffen, R. Use of the Inactivated Intranasal Influenza Vaccine and the Risk of Bell's Palsy in Switzerland. *N Engl J Med* **2004**, *350*, 896–903, doi:10.1056/NEJMoa030595.

43. Mahase, E. Covid-19: Oxford Researchers Halt Vaccine Trial While Adverse Reaction Is Investigated. *BMJ* **2020**, *370*, m3525, doi:10.1136/bmj.m3525.

44. Bonate, P.L. *Pharmacokinetic-Pharmacodynamic Modeling and Simulation*; 2. ed.; Springer: New York, 2011; ISBN 978-1-4419-9484-4.

45. O'Quigley, J.; Iasonos, A.; Bornkamp, B. *Handbook of Methods for Designing, Monitoring, and Analyzing Dose-Finding Trials: Handbooks of Modern Statistical Methods*; 2017; ISBN 978-1-315-15198-4.

46. Xiong, Y.; Fan, J.; Kitabi, E.; Zhang, X.; Bi, Y.; Grimstein, M.; Yang, Y.; Earp, J.C.; Zheng, N.; Liu, J.; et al. Model-Informed Drug Development Approaches to Assist New Drug Development in the COVID-19 Pandemic. *Clin Pharmacol Ther* **2022**, *111*, 572–578, doi:10.1002/cpt.2491.

47. Milligan, P.A.; Brown, M.J.; Marchant, B.; Martin, S.W.; van der Graaf, P.H.; Benson, N.; Nucci, G.; Nichols, D.J.; Boyd, R.A.; Mandema, J.W.; et al. Model-Based Drug Development: A Rational Approach to Efficiently Accelerate Drug Development. *Clin Pharmacol Ther* **2013**, *93*, 502–514, doi:10.1038/clpt.2013.54.

48. Rhodes, S.J.; Guedj, J.; Fletcher, H.A.; Lindenstrøm, T.; Scriba, T.J.; Evans, T.G.; Knight, G.M.; White, R.G. Using Vaccine Immunostimulation/Immunodynamic Modelling Methods to Inform Vaccine Dose Decision-Making. *NPJ Vaccines* **2018**, *3*, 36, doi:10.1038/s41541-018-0075-3.

49. Handel, A.; Li, Y.; McKay, B.; Pawelek, K.A.; Zarnitsyna, V.; Antia, R. Exploring the Impact of Inoculum Dose on Host Immunity and Morbidity to Inform Model-Based Vaccine Design. *PLOS Computational Biology* **2018**, *14*, e1006505, doi:10.1371/journal.pcbi.1006505.

50. Rhodes, S.J.; Zelmer, A.; Knight, G.M.; Prabowo, S.A.; Stockdale, L.; Evans, T.G.; Lindenstrøm, T.; White, R.G.; Fletcher, H. The TB Vaccine H56+IC31 Dose-Response Curve Is Peaked Not Saturating: Data Generation for New Mathematical Modelling Methods to Inform Vaccine Dose Decisions. *Vaccine* **2016**, *34*, 6285–6291, doi:10.1016/j.vaccine.2016.10.060.

51. Rhodes, S.J.; Sarfas, C.; Knight, G.M.; White, A.; Pathan, A.A.; McShane, H.; Evans, T.G.; Fletcher, H.; Sharpe, S.; White, R.G. Using Data from Macaques To Predict Gamma Interferon Responses after Mycobacterium Bovis BCG Vaccination in Humans: A Proof-of-Concept Study of Immunostimulation/Immunodynamic Modeling Methods. *Clin Vaccine Immunol* **2017**, *24*, e00525-16, doi:10.1128/CVI.00525-16.

52. Rhodes, S.J.; Knight, G.M.; Fielding, K.; Scriba, T.J.; Pathan, A.A.; McShane, H.; Fletcher, H.; White, R.G. Individual-Level Factors Associated with Variation in Mycobacterial-Specific Immune Response: Gender and Previous BCG Vaccination Status. *Tuberculosis (Edinb)* **2016**, *96*, 37–43, doi:10.1016/j.tube.2015.10.002.

53. Aronson, J.K.; Ferner, R.E. The Law of Mass Action and the Pharmacological Concentration–Effect Curve: Resolving the Paradox of Apparently Non-Dose-Related Adverse Drug Reactions. *British Journal of Clinical Pharmacology* **2016**, *81*, 56–61, doi:10.1111/bcp.12706.

54. Guo, B.; Li, Y. Bayesian Designs of Phase II Oncology Trials to Select Maximum Effective Dose Assuming Monotonic Dose-Response Relationship. *BMC Med Res Methodol* **2014**, *14*, 95, doi:10.1186/1471-2288-14-95.

55. Bonate, P.L. Types of Model. In *Pharmacokinetic-pharmacodynamic modeling and simulation*; Springer: New York, 2011; pp. p6-8 ISBN 978-1-4419-9484-4.

56. Giorgi, M.; Desikan, R.; van der Graaf, P.H.; Kierzek, A.M. Application of Quantitative Systems Pharmacology to Guide the Optimal Dosing of COVID-19 Vaccines. *CPT Pharmacometrics Syst Pharmacol* **2021**, *10*, 1130–1133, doi:10.1002/psp4.12700.

57. Thiel, C.; Smit, I.; Baier, V.; Cordes, H.; Fabry, B.; Blank, L.M.; Kuepfer, L. Using Quantitative Systems Pharmacology to Evaluate the Drug Efficacy of COX-2 and 5-LOX Inhibitors in Therapeutic Situations. *npj Syst Biol Appl* **2018**, *4*, 1–12, doi:10.1038/s41540-018-0062-3.

58. Hartmann, S.; Biliouris, K.; Lesko, L.J.; Nowak-Göttl, U.; Trame, M.N. Quantitative Systems Pharmacology Model-Based Predictions of Clinical Endpoints to Optimize Warfarin and Rivaroxaban Anti-Thrombosis Therapy. *Frontiers in Pharmacology* **2020**, *11*.

59. Bai, J.P.F.; Earp, J.C.; Florian, J.; Madabushi, R.; Strauss, D.G.; Wang, Y.; Zhu, H. Quantitative Systems Pharmacology: Landscape Analysis of Regulatory Submissions to the US Food and Drug Administration. *CPT Pharmacometrics Syst Pharmacol* **2021**, *10*, 1479–1484, doi:10.1002/psp4.12709.

60. Menezes, B.; Cilliers, C.; Wessler, T.; Thurber, G.M.; Linderman, J.J. An Agent-Based Systems Pharmacology Model of the Antibody-Drug Conjugate Kadcyla to Predict Efficacy of Different Dosing Regimens. *AAPS J* **2020**, *22*, 29, doi:10.1208/s12248-019-0391-1.

61. Joslyn, L.R.; Linderman, J.J.; Kirschner, D.E. A Virtual Host Model of Mycobacterium Tuberculosis Infection Identifies Early Immune Events as Predictive of Infection Outcomes. *Journal of Theoretical Biology* **2022**, *539*, 111042, doi:10.1016/j.jtbi.2022.111042.

62. Cilfone, N.A.; Kirschner, D.E.; Linderman, J.J. Strategies for Efficient Numerical Implementation of Hybrid Multi-Scale Agent-Based Models to Describe Biological Systems. *Cell Mol Bioeng* **2015**, *8*, 119–136, doi:10.1007/s12195-014-0363-6.

63. Chelliah, V.; Lazarou, G.; Bhatnagar, S.; Gibbs, J.P.; Nijsen, M.; Ray, A.; Stoll, B.; Thompson, R.A.; Gulati, A.; Soukharev, S.; et al. Quantitative Systems Pharmacology Approaches for Immuno-Oncology: Adding Virtual Patients to the Development Paradigm. *Clin Pharmacol Ther* **2021**, *109*, 605–618, doi:10.1002/cpt.1987.

64. Duffull, S.B.; Wright, D.F.B.; Winter, H.R. Interpreting Population Pharmacokinetic-Pharmacodynamic Analyses – a Clinical Viewpoint. *Br J Clin Pharmacol* **2011**, *71*, 807–814, doi:10.1111/j.1365-2125.2010.03891.x.

65. Cheng, Y.; Straube, R.; Alnaif, A.E.; Huang, L.; Leil, T.A.; Schmidt, B.J. Virtual Populations for Quantitative Systems Pharmacology Models. In *Systems Medicine*; Bai, J.P.F., Hur, J., Eds.; Methods in Molecular Biology; Springer US: New York, NY, 2022; pp. 129–179 ISBN 978-1-07-162265-0.

66. Gaweda, A.E.; McBride, D.E.; Lederer, E.D.; Brier, M.E. Development of a Quantitative Systems Pharmacology Model of Chronic Kidney Disease: Metabolic Bone Disorder. *American Journal of Physiology-Renal Physiology* **2021**, *320*, F203–F211, doi:10.1152/ajprenal.00159.2020.

67. Moore, J.R.; Ahmed, H.; McGuire, D.; Akondy, R.; Ahmed, R.; Antia, R. Dependence of CD8 T Cell Response upon Antigen Load During Primary Infection : Analysis of Data from Yellow Fever Vaccination. *Bull Math Biol* **2019**, *81*, 2553–2568, doi:10.1007/s11538-019-00618-9.

68. Zarnitsyna, V.I.; Handel, A.; McMaster, S.R.; Hayward, S.L.; Kohlmeier, J.E.; Antia, R. Mathematical Model Reveals the Role of Memory CD8 T Cell Populations in Recall Responses to Influenza. *Front Immunol* **2016**, *7*, 165, doi:10.3389/fimmu.2016.00165.

69. Mayer, A.; Zhang, Y.; Perelson, A.S.; Wingreen, N.S. Regulation of T Cell Expansion by Antigen Presentation Dynamics. *Proceedings of the National Academy of Sciences* **2019**, *116*, 5914–5919, doi:10.1073/pnas.1812800116.

70. Seder, R.A.; Ahmed, R. Similarities and Differences in CD4+ and CD8+ Effector and Memory T Cell Generation. *Nat Immunol* **2003**, *4*, 835–842, doi:10.1038/ni969.

71. Chen, C.; Ortega, F.; Alameda, L.; Ferrer, S.; Simonsson, U.S.H. Population Pharmacokinetics, Optimised Design and Sample Size Determination for Rifampicin, Isoniazid, Ethambutol and Pyrazinamide in the Mouse. *European Journal of Pharmaceutical Sciences* **2016**, *93*, 319–333, doi:10.1016/j.ejps.2016.07.017.

72. González-Sales, M.; Nekka, F.; Tanguay, M.; Tremblay, P.; Li, J. Modelling the Dose–Response Relationship: The Fair Share of Pharmacokinetic and Pharmacodynamic Information. *Br J Clin Pharmacol* **2017**, *83*, 1240–1251, doi:10.1111/bcp.13225.

73. Baker, R.E.; Peña, J.-M.; Jayamohan, J.; Jérusalem, A. Mechanistic Models versus Machine Learning, a Fight Worth Fighting for the Biological Community? *Biol Lett* **2018**, *14*, 20170660, doi:10.1098/rsbl.2017.0660.

74. Brown, L.V.; Gaffney, E.A.; Wagg, J.; Coles, M.C. Applications of Mechanistic Modelling to Clinical and Experimental Immunology: An Emerging Technology to Accelerate Immunotherapeutic Discovery and Development. *Clin Exp Immunol* **2018**, *193*, 284–292, doi:10.1111/cei.13182.

75. Takahashi, A.; Suzuki, T. Bayesian Optimization Design for Dose-Finding Based on Toxicity and Efficacy Outcomes in Phase I/II Clinical Trials. *Pharmaceutical Statistics* **2021**, *20*, 422–439, doi:10.1002/pst.2085.

76. Takahashi, A.; Suzuki, T. Bayesian Optimization for Estimating the Maximum Tolerated Dose in Phase I Clinical Trials. *Contemporary Clinical Trials Communications* **2021**, *21*, 100753, doi:10.1016/j.conctc.2021.100753.

77. Lin, R.; Yin, G. STEIN: A Simple Toxicity and Efficacy Interval Design for Seamless Phase I/II Clinical Trials. *Statistics in Medicine* **2017**, *36*, 4106–4120, doi:10.1002/sim.7428.

78. Li, W.; Fu, H. Bayesian Isotonic Regression Dose–Response Model. *Journal of Biopharmaceutical Statistics* **2017**, *27*, 824–833, doi:10.1080/10543406.2016.1265535.

79. Mandrekar, S.J.; Qin, R.; Sargent, D.J. Model-Based Phase I Designs Incorporating Toxicity and Efficacy for Single and Dual Agent Drug Combinations: Methods and Challenges. *Stat Med* **2010**, *29*, 1077–1083, doi:10.1002/sim.3706.

80. Le Tourneau, C.; Lee, J.J.; Siu, L.L. Dose Escalation Methods in Phase I Cancer Clinical Trials. *J Natl Cancer Inst* **2009**, *101*, 708–720, doi:10.1093/jnci/djp079.

81. Farhang-Sardroodi, S.; Korosec, C.S.; Gholami, S.; Craig, M.; Moyles, I.R.; Ghaemi, M.S.; Ooi, H.K.; Heffernan, J.M. Analysis of Host Immunological Response of Adenovirus-Based COVID-19 Vaccines. *Vaccines* **2021**, *9*, 861, doi:10.3390/vaccines9080861.

82. Kumbhari, A.; Kim, P.S.; Lee, P.P. Optimisation of Anti-Cancer Peptide Vaccines to Preferentially Elicit High-Avidity T Cells. *Journal of Theoretical Biology* **2020**, *486*, 110067, doi:10.1016/j.jtbi.2019.110067.

83. Zhang, H.; Chiang, A.Y.; Wang, J. Improving the Performance of Bayesian Logistic Regression Model with Overdose Control in Oncology Dose-Finding Studies. *Statistics in Medicine n/a*, doi:10.1002/sim.9402.

84. Research, C. for B.E. and Toxicity Grading Scale for Healthy Adult and Adolescent Volunteers Enrolled in Preventive Vaccine Clinical Trials Available online: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/toxicity-grading-scale-healthy-adult-and-adolescent-volunteers-enrolled-preventive-vaccine-clinical (accessed on 13 August 2022).

85. Van Meter, E.M.; Garrett-Mayer, E.; Bandyopadhyay, D. Dose-Finding Clinical Trial Design for Ordinal Toxicity Grades Using the Continuation Ratio Model: An Extension of the Continual Reassessment Method. *Clin Trials* **2012**, *9*, 303–313, doi:10.1177/1740774512443593.

86. Stampfer, H.G.; Gabb, G.M.; Dimmitt, S.B. Why Maximum Tolerated Dose? *Br J Clin Pharmacol* **2019**, *85*, 2213–2217, doi:10.1111/bcp.14032.

87. Yuan, Y.; Yin, G. Bayesian Hybrid Dose-Finding Design in Phase I Oncology Clinical Trials. *Stat Med* **2011**, *30*, 2098–2108, doi:10.1002/sim.4164.

88. Love, S.B.; Brown, S.; Weir, C.J.; Harbron, C.; Yap, C.; Gaschler-Markefski, B.; Matcham, J.; Caffrey, L.; McKevitt, C.; Clive, S.; et al. Embracing Model-Based Designs for Dose-Finding Trials. *Br J Cancer* **2017**, *117*, 332–339, doi:10.1038/bjc.2017.186.

89. Fraisse, J.; Dinart, D.; Tosi, D.; Bellera, C.; Mollevi, C. Optimal Biological Dose: A Systematic Review in Cancer Phase I Clinical Trials. *BMC Cancer* **2021**, *21*, 60, doi:10.1186/s12885-021-07782-z.

90. O'Quigley, J.; Iasonos, A.; Bornkamp, B. Choice of Endpoints. In *Handbook of Methods for Designing, Monitoring, and Analyzing Dose-Finding Trials: Handbooks of Modern Statistical Methods*; 2017; pp. p15-16 ISBN 978-1-315-15198-4.

91. Strohbehn, G.W.; Parker, W.F.; Reid, P.D.; Gellad, W.F. Socially Optimal Pandemic Drug Dosing. *The Lancet Global Health* **2021**, *9*, e1049–e1050, doi:10.1016/S2214-109X(21)00251-5.

92. Liu, K.; Lou, Y. Optimizing COVID-19 Vaccination Programs during Vaccine Shortages. *Infect Dis Model* **2022**, *7*, 286–298, doi:10.1016/j.idm.2022.02.002.

93. Du, Z.; Wang, L.; Pandey, A.; Lim, W.W.; Chinazzi, M.; Piontti, A.P. y.; Lau, E.H.Y.; Wu, P.; Malani, A.; Cobey, S.; et al. Modeling Comparative Cost-Effectiveness of SARS-CoV-2 Vaccine Dose Fractionation in India. *Nat Med* **2022**, *28*, 934–938, doi:10.1038/s41591-022-01736-z.

94. Martonosi, S.E.; Behzad, B.; Cummings, K. Pricing the COVID-19 Vaccine: A Mathematical Approach. *Omega* **2021**, *103*, 102451, doi:10.1016/j.omega.2021.102451.

95. Allmendinger, R.; Emmerich, M.T.M.; Hakanen, J.; Jin, Y.; Rigoni, E. Surrogate-Assisted Multicriteria Optimization: Complexities, Prospective Solutions, and Business Case. *Journal of Multi-Criteria Decision Analysis* **2017**, *24*, 5–24, doi:10.1002/mcda.1605.

96. Emmerich, M.T.M.; Deutz, A.H. A Tutorial on Multiobjective Optimization: Fundamentals and Evolutionary Methods. *Nat Comput* **2018**, *17*, 585–609, doi:10.1007/s11047-018-9685-y.

97. Wang, H.; Olhofer, M.; Jin, Y. A Mini-Review on Preference Modeling and Articulation in Multi-Objective Optimization: Current Status and Challenges. *Complex Intell. Syst.* **2017**, *3*, 233–245, doi:10.1007/s40747-017-0053-9.

98. Zheng, Y.; Pu, A. Utility-Based Multi-Stakeholder Recommendations by Multi-Objective Optimization. In Proceedings of the 2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI); December 2018; pp. 128–135.

99. Aouni, J.; Bacro, J.N.; Toulemonde, G.; Colin, P.; Darchy, L. Utility-Based Dose Selection for Phase II Dose-Finding Studies. *Ther Innov Regul Sci* **2021**, *55*, 818–840, doi:10.1007/s43441-021-00273-0.

100. Thall, P.F.; Cook, J.D. Dose-Finding Based on Efficacy-Toxicity Trade-Offs. *Biometrics* **2004**, *60*, 684–693, doi:10.1111/j.0006-341X.2004.00218.x.

101. Branke, J.; Greco, S.; Słowiński, R.; Zielniewicz, P. Learning Value Functions in Interactive Evolutionary Multiobjective Optimization. *IEEE Transactions on Evolutionary Computation* **2015**, *19*, 88–102, doi:10.1109/TEVC.2014.2303783.

102. Peterson, M. Von Neumman and Morgenstern's Interval Scale. In *An introduction to decision theory*; Cambridge introductions to philosophy; Cambridge University Press: Cambridge, United States ; New York, NY, USA, 2017; pp. p101-112 ISBN 978-1-107-15159-8.

103. Thall, P.F.; Russell, K.E. A Strategy for Dose-Finding and Safety Monitoring Based on Efficacy and Adverse Outcomes in Phase I/II Clinical Trials. *Biometrics* **1998**, *54*, 251–264, doi:10.2307/2534012.

104. Zhang, W.; Sargent, D.J.; Mandrekar, S. An Adaptive Dose-Finding Design Incorporating Both Toxicity and Efficacy. *Statistics in Medicine* **2006**, *25*, 2365–2383, doi:10.1002/sim.2325.

105. Yin, G.; Li, Y.; Ji, Y. Bayesian Dose-Finding in Phase I/II Clinical Trials Using Toxicity and Efficacy Odds Ratios. *Biometrics* **2006**, *62*, 777–787, doi:10.1111/j.1541-0420.2006.00534.x.

106. O'Quigley, J.; Iasonos, A.; Bornkamp, B. A Bayesian Formulation. In *Handbook of Methods for Designing, Monitoring, and Analyzing Dose-Finding Trials: Handbooks of Modern Statistical Methods*; 2017; pp. p84-85 ISBN 978-1-315-15198-4.

107. Brock, K.; Billingham, L.; Copland, M.; Siddique, S.; Sirovica, M.; Yap, C. Implementing the EffTox Dose-Finding Design in the Matchpoint Trial. *BMC Medical Research Methodology* **2017**, *17*, 112, doi:10.1186/s12874-017-0381-x.

108. Intramuscular Injections. In *The Royal Marsden manual of clinical nursing procedures*; Dougherty, L., Lister, S.E., Eds.; John Wiley & Sons Inc: Chichester, West Sussex ; Hoboken, NJ, 2015; p. p40 ISBN 978-1-118-74592-2.

109. GE Healthcare Scalable Process for Adenovirus Production.

110. Food and Drug Administration *Guidance for Industry: Clinical Considerations for Therapeutic Cancer Vaccines*; 2011;

111. Bhatt, K.; Verma, S.; Ellner, J.J.; Salgame, P. Quest for Correlates of Protection against Tuberculosis. *Clin Vaccine Immunol* **2015**, *22*, 258–266, doi:10.1128/CVI.00721-14.

112. Saadi, N.; Chi, Y.-L.; Ghosh, S.; Eggo, R.M.; McCarthy, C.V.; Quaife, M.; Dawa, J.; Jit, M.; Vassall, A. Models of COVID-19 Vaccine Prioritisation: A Systematic Literature Search and Narrative Review. *BMC Med* **2021**, *19*, 318, doi:10.1186/s12916-021-02190-3.

113. Salomon, J.A.; Haagsma, J.A.; Davis, A.; Noordhout, C.M. de; Polinder, S.; Havelaar, A.H.; Cassini, A.; Devleesschauwer, B.; Kretzschmar, M.; Speybroeck, N.; et al. Disability Weights for the Global Burden of Disease 2013 Study. *The Lancet Global Health* **2015**, *3*, e712–e723, doi:10.1016/S2214-109X(15)00069-8.

114. Drolet, M.; Laprise, J.-F.; Martin, D.; Jit, M.; Bénard, É.; Gingras, G.; Boily, M.-C.; Alary, M.; Baussano, I.; Hutubessy, R.; et al. Optimal Human Papillomavirus Vaccination Strategies to Prevent Cervical Cancer in Low-Income and Middle-Income Countries in the Context of Limited Resources: A Mathematical Modelling Analysis. *Lancet Infect Dis* **2021**, *21*, 1598–1610, doi:10.1016/S1473-3099(20)30860-4.

115. European Medicines Agency Annex to Vaxzevria Art.5.3 - Visual Risk Contextualisation 2021.

116. Funk, P.R.; Yogurtcu, O.N.; Forshee, R.A.; Anderson, S.A.; Marks, P.W.; Yang, H. Benefit-Risk Assessment of COVID-19 Vaccine, MRNA (Comirnaty) for Age 16–29 Years. *Vaccine* **2022**, *40*, 2781–2789, doi:10.1016/j.vaccine.2022.03.030.

117. McDonald, S.A.; Nijsten, D.; Bollaerts, K.; Bauwens, J.; Praet, N.; van der Sande, M.; Bauchau, V.; de Smedt, T.; Sturkenboom, M.; Hahné, S. Methodology for Computing the Burden of Disease of Adverse Events Following Immunization. *Pharmacoepidemiol Drug Saf* **2018**, *27*, 724–730, doi:10.1002/pds.4419.

118. McDermott, A. Core Concept: Herd Immunity Is an Important—and Often Misunderstood—Public Health Phenomenon. *Proc Natl Acad Sci U S A* **2021**, *118*, e2107692118, doi:10.1073/pnas.2107692118.

119. MacIntyre, C.R.; Costantino, V.; Trent, M. Modelling of COVID-19 Vaccination Strategies and Herd Immunity, in Scenarios of Limited and Full Vaccine Supply in NSW, Australia. *Vaccine* **2022**, *40*, 2506–2513, doi:10.1016/j.vaccine.2021.04.042.

120. Chang, S.L.; Piraveenan, M.; Pattison, P.; Prokopenko, M. Game Theoretic Modelling of Infectious Disease Dynamics and Intervention Methods: A Review. *Journal of Biological Dynamics* **2020**, *14*, 57–89, doi:10.1080/17513758.2020.1720322.

121. Yong, J.C.; Choy, B.K.C. Noncompliance With Safety Guidelines as a Free-Riding Strategy: An Evolutionary Game-Theoretic Approach to Cooperation During the COVID-19 Pandemic. *Front Psychol* **2021**, *12*, 646892, doi:10.3389/fpsyg.2021.646892.

122. Bauch, C.T.; Earn, D.J.D. Vaccination and the Theory of Games. *Proceedings of the National Academy of Sciences* **2004**, *101*, 13391–13394, doi:10.1073/pnas.0403823101.

123. Bhattacharyya, S.; Vutha, A.; Bauch, C.T. The Impact of Rare but Severe Vaccine Adverse Events on Behaviour-Disease Dynamics: A Network Model. *Sci Rep* **2019**, *9*, 7164, doi:10.1038/s41598-019-43596-7.

124. King, J.; Ferraz, O.L.M.; Jones, A. Mandatory COVID-19 Vaccination and Human Rights. *Lancet* **2022**, *399*, 220–222, doi:10.1016/S0140-6736(21)02873-7.

125.   Li, G.; Sun, J.; Rana, M.N.A.; Song, Y.; Liu, C.; Zhu, Z. Optimizing High-Dimensional Functions with an Efficient Particle Swarm Optimization Algorithm. *Mathematical Problems in Engineering* **2020**, *2020*, e5264547, doi:10.1155/2020/5264547.

126.   De Boer, R.J.; Perelson, A.S. Quantifying T Lymphocyte Turnover. *J Theor Biol* **2013**, *327*, 45–87, doi:10.1016/j.jtbi.2012.12.025.

127.   Homann, D.; Teyton, L.; Oldstone, M.B.A. Differential Regulation of Antiviral T-Cell Immunity Results in Stable CD8+ but Declining CD4+ T-Cell Memory. *Nat Med* **2001**, *7*, 913–919, doi:10.1038/90950.

128.   Quiel, J.; Caucheteux, S.; Laurence, A.; Singh, N.J.; Bocharov, G.; Ben-Sasson, S.Z.; Grossman, Z.; Paul, W.E. Antigen-Stimulated CD4 T-Cell Expansion Is Inversely and Log-Linearly Related to Precursor Number. *Proc Natl Acad Sci U S A* **2011**, *108*, 3312–3317, doi:10.1073/pnas.1018525108.

129.   Chung, B.; Stuge, T.B.; Murad, J.P.; Beilhack, G.; Andersen, E.; Armstrong, B.D.; Weber, J.S.; Lee, P.P. Antigen-Specific Inhibition of High-Avidity T Cell Target Lysis by Low-Avidity T Cells via Trogocytosis. *Cell Reports* **2014**, *8*, 871–882, doi:10.1016/j.celrep.2014.06.052.

130.   Kumbhari, A.; Rose, D.; Lee, P.P.; Kim, P.S. A Minimal Model of T Cell Avidity May Identify Subtherapeutic Vaccine Schedules. *Math Biosci* **2021**, *334*, 108556, doi:10.1016/j.mbs.2021.108556.

131.   Demidenko, E.; Miller, T.W. Statistical Determination of Synergy Based on Bliss Definition of Drugs Independence. *PLOS ONE* **2019**, *14*, e0224137, doi:10.1371/journal.pone.0224137.

132.   Sun, L.Z.; Wu, C.; Li, X.; Chen, C.; Schmidt, E.V. Independent Action Models and Prediction of Combination Treatment Effects for Response Rate, Duration of Response and Tumor Size Change in Oncology Drug Development. *Contemporary Clinical Trials* **2021**, *106*, 106434, doi:10.1016/j.cct.2021.106434.

133.   Lederer, S.; Dijkstra, T.M.H.; Heskes, T. Additive Dose Response Models: Defining Synergy. *Frontiers in Pharmacology* **2019**, *10*.

134.   Lee, W.; Cheung, K.C.; Ahn, J.Y. Multivariate Countermonotonicity and the Minimal Copulas. *Journal of Computational and Applied Mathematics* **2017**, *317*, 589–602, doi:10.1016/j.cam.2016.12.032.

135.   Musuamba, F.; Manolis, E.; Holford, N.; Cheung, S.; Friberg, L.; Ogungbenro, K.; Posch, M.; Yates, J.; Berry, S.; Thomas, N.; et al. Advanced Methods for Dose and Regimen Finding During Drug Development: Summary of the EMA/EFPIA Workshop on Dose Finding (London 4–5 December 2014). *CPT Pharmacometrics Syst Pharmacol* **2017**, *6*, 418–429, doi:10.1002/psp4.12196.

136.   Orsini, N. Weighted Mixed-Effects Dose–Response Models for Tables of Correlated Contrasts. *The Stata Journal* **2021**, *21*, 320–347, doi:10.1177/1536867X211025798.

137.   Crippa, A.; Discacciati, A.; Bottai, M.; Spiegelman, D.; Orsini, N. One-Stage Dose–Response Meta-Analysis for Aggregated Data. *Stat Methods Med Res* **2019**, *28*, 1579–1596, doi:10.1177/0962280218773122.

138.   Chen, J.; Liu, T. Statistical Considerations on Implementing the MCP-Mod Method for Binary Endpoints in Clinical Trials. *Contemp Clin Trials Commun* **2020**, *19*, 100641, doi:10.1016/j.conctc.2020.100641.

139.   Cinar, O.; Umbanhowar, J.; Hoeksema, J.D.; Viechtbauer, W. Using Information-Theoretic Approaches for Model Selection in Meta-Analysis. *Res Synth Methods* **2021**, *12*, 537–556, doi:10.1002/jrsm.1489.

140.   Pham, H. A New Criterion for Model Selection. *Mathematics* **2019**, *7*, 1215, doi:10.3390/math7121215.

141.   Rhodes, S.J. The Development of a Mathematical Modelling Framework to Translate TB Vaccine Responses between Species and Predict the Most Immunogenic Dose in Humans Using Animal Data. doctoral, London School of Hygiene & Tropical Medicine, 2018.

142.   Raftery, A.E. Bayesian Model Selection in Social Research. *Sociological Methodology* **1995**, *25*, 111–163, doi:10.2307/271063.

367

143. Schorning, K.; Bornkamp, B.; Bretz, F.; Dette, H. Model Selection versus Model Averaging in Dose Finding Studies. *Statistics in Medicine* **2016**, *35*, 4021–4040, doi:10.1002/sim.6991.

144. Lyauk, Y.K.; Jonker, D.M.; Lund, T.M. Dose Finding in the Clinical Development of 60 US Food and Drug Administration–Approved Drugs Compared With Learning vs. Confirming Recommendations. *Clin Transl Sci* **2019**, *12*, 481–489, doi:10.1111/cts.12641.

145. Friedman, L.M.; Furberg, C.; DeMets, D.L. Adaptive Designs. In *Fundamentals of clinical trials*; Springer: New York, 2010; pp. p114-115 ISBN 978-1-4419-1585-6.

146. Zhu, F.-C.; Li, Y.-H.; Guan, X.-H.; Hou, L.-H.; Wang, W.-J.; Li, J.-X.; Wu, S.-P.; Wang, B.-S.; Wang, Z.; Wang, L.; et al. Safety, Tolerability, and Immunogenicity of a Recombinant Adenovirus Type-5 Vectored COVID-19 Vaccine: A Dose-Escalation, Open-Label, Non-Randomised, First-in-Human Trial. *The Lancet* **2020**, *395*, 1845–1854, doi:10.1016/S0140-6736(20)31208-3.

147. Pollock, K.M.; Cheeseman, H.M.; Szubert, A.J.; Libri, V.; Boffito, M.; Owen, D.; Bern, H.; O'Hara, J.; McFarlane, L.R.; Lemm, N.-M.; et al. Safety and Immunogenicity of a Self-Amplifying RNA Vaccine against COVID-19: COVAC1, a Phase I, Dose-Ranging Trial. *EClinicalMedicine* **2022**, *44*, 101262, doi:10.1016/j.eclinm.2021.101262.

148. Lattimore, T.; Szepesvári, C. Pure Exploration. In *Bandit algorithms*; Cambridge University Press: Cambridge ; New York, NY, 2020; pp. p353-369 ISBN 978-1-108-48682-8.

149. Thompson, W.R. On the Likelihood That One Unknown Probability Exceeds Another in View of the Evidence of Two Samples. *Biometrika* **1933**, *25*, 285–294, doi:10.2307/2332286.

150. Aziz, M.; Kaufmann, E.; Riviere, M.-K. On Multi-Armed Bandit Designs for Dose-Finding Clinical Trials. *J. Mach. Learn. Res.* **2021**, *22*, 14:686-14:723.

151. Liu, M.; Li, Q.; Lin, J.; Lin, Y.; Hoffman, E. Innovative Trial Designs and Analyses for Vaccine Clinical Development. *Contemp Clin Trials* **2021**, *100*, 106225, doi:10.1016/j.cct.2020.106225.

152. Chen, Y.H.J.; Gesser, R.; Luxembourg, A. A Seamless Phase IIB/III Adaptive Outcome Trial: Design Rationale and Implementation Challenges. *Clinical Trials* **2015**, *12*, 84–90, doi:10.1177/1740774514552110.

153. Kurzrock, R.; Lin, C.-C.; Wu, T.-C.; Hobbs, B.P.; Pestana, R.C.; Hong, D.S. Moving Beyond 3+3: The Future of Clinical Trial Design. *American Society of Clinical Oncology Educational Book* **2021**, e133–e144, doi:10.1200/EDBK_319783.

154. Rahma, O.E.; Gammoh, E.; Simon, R.M.; Khleif, S.N. Is the "3+3" Dose-Escalation Phase I Clinical Trial Design Suitable for Therapeutic Cancer Vaccine Development? A Recommendation for Alternative Design. *Clinical Cancer Research* **2014**, *20*, 4758–4767, doi:10.1158/1078-0432.CCR-13-2671.

155. Adaptive Designs Working Group of the MRC Network of Hubs for Trials Methodology Research A Quick Guide Why Not to Use A+B Designs. 2017.

156. van Brummelen, E.M.J.; Huitema, A.D.R.; van Werkhoven, E.; Beijnen, J.H.; Schellens, J.H.M. The Performance of Model-Based versus Rule-Based Phase I Clinical Trials in Oncology. *J Pharmacokinet Pharmacodyn* **2016**, *43*, 235–242, doi:10.1007/s10928-016-9466-0.

157. O'Quigley, J.; Pepe, M.; Fisher, L. Continual Reassessment Method: A Practical Design for Phase 1 Clinical Trials in Cancer. *Biometrics* **1990**, *46*, 33–48.

158. O'Quigley, J.; Conaway, M. Continual Reassessment and Related Dose-Finding Designs. *Stat Sci* **2010**, *25*, 202–216, doi:10.1214/10-STS332.

159. O'quigley, J. Non-Parametric Optimal Design in Dose Finding Studies. *Biostatistics* **2002**, *3*, 51–56, doi:10.1093/biostatistics/3.1.51.

160. Wheeler, G.M.; Mander, A.P.; Bedding, A.; Brock, K.; Cornelius, V.; Grieve, A.P.; Jaki, T.; Love, S.B.; Odondi, L.; Weir, C.J.; et al. How to Design a Dose-Finding Study Using the Continual Reassessment Method. *BMC Medical Research Methodology* **2019**, *19*, 18, doi:10.1186/s12874-018-0638-z.

161. Fedorov, V.V.; Leonov, S.L. *Optimal Design for Nonlinear Response Models*; CRC Press: Boca Raton, 2013; ISBN 978-0-429-10398-8.

162. Silvey, S. *Optimal Design: An Introduction to the Theory for Parameter Estimation*; Springer Netherlands, 1980; ISBN 978-0-412-22910-7.

163.	Ryeznik, Y.; Sverdlov, O.; Hooker, A.C. Adaptive Optimal Designs for Dose-Finding Studies with Time-to-Event Outcomes. *AAPS J* **2017**, *20*, 24, doi:10.1208/s12248-017-0166-5.

164.	Zang, Y.; Lee, J.J.; Yuan, Y. Adaptive Designs for Identifying Optimal Biological Dose for Molecularly Targeted Agents. *Clin Trials* **2014**, *11*, 319–327, doi:10.1177/1740774514529848.

165.	James, G.D.; Symeonides, S.N.; Marshall, J.; Young, J.; Clack, G. Continual Reassessment Method for Dose Escalation Clinical Trials in Oncology: A Comparison of Prior Skeleton Approaches Using AZD3514 Data. *BMC Cancer* **2016**, *16*, 703, doi:10.1186/s12885-016-2702-6.

166.	Pan, H.; Yuan, Y. A Default Method to Specify Skeletons for Bayesian Model Averaging Continual Reassessment Method for Phase I Clinical Trials. *Stat Med* **2017**, *36*, 266–279, doi:10.1002/sim.6941.

167.	Pallmann, P.; Wan, F.; Mander, A.P.; Wheeler, G.M.; Yap, C.; Clive, S.; Hampson, L.V.; Jaki, T. Designing and Evaluating Dose-Escalation Studies Made Easy: The MoDEsT Web App. *Clinical Trials* **2020**, *17*, 147–156, doi:10.1177/1740774519890146.

168.	Lee, S.M.; Cheung, Y.K. Calibration of Prior Variance in the Bayesian Continual Reassessment Method. *Stat Med* **2011**, *30*, 2081–2089, doi:10.1002/sim.4139.

169.	Thall, P.F.; Herrick, R.C.; Nguyen, H.Q.; Venier, J.J.; Norris, J.C. Effective Sample Size for Computing Prior Hyperarameters in Bayesian Phase I-II Dose-Finding. *Clin Trials* **2014**, *11*, 657–666, doi:10.1177/1740774514547397.

170.	Barrett, J.R.; Belij-Rammerstorfer, S.; Dold, C.; Ewer, K.J.; Folegatti, P.M.; Gilbride, C.; Halkerston, R.; Hill, J.; Jenkin, D.; Stockdale, L.; et al. Phase 1/2 Trial of SARS-CoV-2 Vaccine ChAdOx1 NCoV-19 with a Booster Dose Induces Multifunctional Antibody Responses. *Nat Med* **2021**, *27*, 279–288, doi:10.1038/s41591-020-01179-4.

171.	Bosaeed, M.; Balkhy, H.H.; Almaziad, S.; Aljami, H.A.; Alhatmi, H.; Alanazi, H.; Alahmadi, M.; Jawhary, A.; Alenazi, M.W.; Almasoud, A.; et al. Safety and Immunogenicity of ChAdOx1 MERS Vaccine Candidate in Healthy Middle Eastern Adults (MERS002): An Open-Label, Non-Randomised, Dose-Escalation, Phase 1b Trial. *The Lancet Microbe* **2022**, *3*, e11–e20, doi:10.1016/S2666-5247(21)00193-2.

172.	West, G.B.; Brown, J.H. The Origin of Allometric Scaling Laws in Biology from Genomes to Ecosystems: Towards a Quantitative Unifying Theory of Biological Structure and Organization. *Journal of Experimental Biology* **2005**, *208*, 1575–1592, doi:10.1242/jeb.01589.

173.	Nair, A.B.; Jacob, S. A Simple Practice Guide for Dose Conversion between Animals and Human. *J Basic Clin Pharm* **2016**, *7*, 27–31, doi:10.4103/0976-0105.177703.

174.	Savage, V.M.; Deeds, E.J.; Fontana, W. Sizing Up Allometric Scaling Theory. *PLoS Comput Biol* **2008**, *4*, e1000171, doi:10.1371/journal.pcbi.1000171.

175.	Herati, R.S.; Wherry, E.J. What Is the Predictive Value of Animal Models for Vaccine Efficacy in Humans? *Cold Spring Harb Perspect Biol* **2018**, *10*, a031583, doi:10.1101/cshperspect.a031583.

176.	De Cock, R.F.W.; Piana, C.; Krekels, E.H.J.; Danhof, M.; Allegaert, K.; Knibbe, C.A.J. The Role of Population PK–PD Modelling in Paediatric Clinical Research. *Eur J Clin Pharmacol* **2011**, *67*, 5–16, doi:10.1007/s00228-009-0782-9.

177.	Joerger, M. Covariate Pharmacokinetic Model Building in Oncology and Its Potential Clinical Relevance. *AAPS J* **2012**, *14*, 119–132, doi:10.1208/s12248-012-9320-2.

178.	Abbas-Aghababazadeh, F.; Lu, P.; Fridley, B.L. Nonlinear Mixed-Effects Models for Modeling in Vitro Drug Response Data to Determine Problematic Cancer Cell Lines. *Sci Rep* **2019**, *9*, 14421, doi:10.1038/s41598-019-50936-0.

179.	Joerger, M.; Huitema, A.D.R.; van den Bongard, D.H.J.G.; Schellens, J.H.M.; Beijnen, J.H. Quantitative Effect of Gender, Age, Liver Function, and Body Size on the Population Pharmacokinetics of Paclitaxel in Patients with Solid Tumors. *Clinical Cancer Research* **2006**, *12*, 2150–2157, doi:10.1158/1078-0432.CCR-05-2069.

180.	Wicha, S.G.; Märtson, A.-G.; Nielsen, E.I.; Koch, B.C.P.; Friberg, L.E.; Alffenaar, J.-W.; Minichmayr, I.K.; the International Society of Anti-Infective Pharmacology (ISAP), I.D. (EPASG), the PK/PD study group of the European Society of Clinical Microbiology From Therapeutic Drug

Monitoring to Model-Informed Precision Dosing for Antibiotics. *Clinical Pharmacology & Therapeutics* **2021**, *109*, 928–941, doi:10.1002/cpt.2202.

181. Maier, C.; de Wiljes, J.; Hartung, N.; Kloft, C.; Huisinga, W. A Continued Learning Approach for Model-informed Precision Dosing: Updating Models in Clinical Practice. *CPT Pharmacometrics Syst Pharmacol* **2022**, *11*, 185–198, doi:10.1002/psp4.12745.

182. Svensson, R.J.; Niward, K.; Davies Forsman, L.; Bruchfeld, J.; Paues, J.; Eliasson, E.; Schön, T.; Simonsson, U.S.H. Individualised Dosing Algorithm and Personalised Treatment of High-dose Rifampicin for Tuberculosis. *Br J Clin Pharmacol* **2019**, *85*, 2341–2350, doi:10.1111/bcp.14048.

183. James, G.D.; Symeonides, S.; Marshall, J.; Young, J.; Clack, G. Assessment of Various Continual Reassessment Method Models for Dose-Escalation Phase 1 Oncology Clinical Trials: Using Real Clinical Data and Simulation Studies. *BMC Cancer* **2021**, *21*, 7, doi:10.1186/s12885-020-07703-6.

184. Morris, T.P.; White, I.R.; Crowther, M.J. Using Simulation Studies to Evaluate Statistical Methods. *Stat Med* **2019**, *38*, 2074–2102, doi:10.1002/sim.8086.

185. Smith, H.; Sweeting, M.; Morris, T.; Crowther, M.J. A Scoping Methodological Review of Simulation Studies Comparing Statistical and Machine Learning Approaches to Risk Prediction for Time-to-Event Data. *Diagnostic and Prognostic Research* **2022**, *6*, 10, doi:10.1186/s41512-022-00124-y.

186. Lattimore, T.; Szepesvári, C. Asymtotic Lower Bounds for Stochastic Linear Bandits. In *Bandit algorithms*; Cambridge University Press: Cambridge ; New York, NY, 2020; pp. p258-259 ISBN 978-1-108-48682-8.

187. R: The R Project for Statistical Computing Available online: https://www.r-project.org/ (accessed on 23 August 2022).

188. Virtanen, P.; Gommers, R.; Oliphant, T.E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat Methods* **2020**, *17*, 261–272, doi:10.1038/s41592-019-0686-2.

189. Dormann, C.F.; Calabrese, J.M.; Guillera-Arroita, G.; Matechou, E.; Bahn, V.; Bartoń, K.; Beale, C.M.; Ciuti, S.; Elith, J.; Gerstner, K.; et al. Model Averaging in Ecology: A Review of Bayesian, Information-Theoretic, and Tactical Approaches for Predictive Inference. *Ecological Monographs* **2018**, *88*, 485–504, doi:10.1002/ecm.1309.

190. Benest, J.; Rhodes, S.; Afrough, S.; Evans, T.; White, R. Response Type and Host Species May Be Sufficient to Predict Dose-Response Curve Shape for Adenoviral Vector Vaccines. *Vaccines* **2020**, *8*, 155, doi:10.3390/vaccines8020155.

191. Dolgin, E. Could Computer Models Be the Key to Better COVID Vaccines? *Nature* **2022**, *604*, 22–25, doi:10.1038/d41586-022-00924-8.

192. Benest, J.; Rhodes, S.; Quaife, M.; Evans, T.G.; White, R.G. Optimising Vaccine Dose in Inoculation against SARS-CoV-2, a Multi-Factor Optimisation Modelling Study to Maximise Vaccine Safety and Efficacy. *Vaccines* **2021**, *9*, 78, doi:10.3390/vaccines9020078.

193. Thall, P.F. BAYESIAN ADAPTIVE DOSE-FINDING BASED ON EFFICACY AND TOXICITY. *Journal of Statistical Research* **2012**, *46*, 187–202.

194. Benest, J.; Rhodes, S.; Evans, T.G.; White, R.G. Mathematical Modelling for Optimal Vaccine Dose Finding: Maximising Efficacy and Minimising Toxicity. *Vaccines (Basel)* **2022**, *10*, 756, doi:10.3390/vaccines10050756.

195. Goetschalckx, R.; Poupart, P.; Hoey, J. Continuous Correlated Beta Processes. In Proceedings of the Proceedings of the Twenty-Second international joint conference on Artificial Intelligence - Volume Volume Two; AAAI Press: Barcelona, Catalonia, Spain, July 16 2011; pp. 1269–1274.

196. Diniz, M.A.; Tighiouart, M.; Rogatko, A. Comparison between Continuous and Discrete Doses for Model Based Designs in Cancer Dose Finding. *PLOS ONE* **2019**, *14*, e0210139, doi:10.1371/journal.pone.0210139.

197. Zhong, W.; Koopmeiners, J.S.; Carlin, B.P. A Trivariate Continual Reassessment Method for Phase I/II Trials of Toxicity, Efficacy, and Surrogate Efficacy. *Stat Med* **2012**, *31*, 3885–3895, doi:10.1002/sim.5477.

198. Mendez, W.; Shao, K.; Lee, J.S.; Cote, I.; Druwe, I.L.; Davis, A.; Gift, J.S. Model Averaging Methods for the Evaluation of Dose-Response Model Uncertainty When Assessing the Suitability of Studies for Estimating Risk. *Environment International* **2020**, *143*, 105857, doi:10.1016/j.envint.2020.105857.

199. Wheeler, M.W.; Blessinger, T.; Shao, K.; Allen, B.C.; Olszyk, L.; Davis, J.A.; Gift, J.S. Quantitative Risk Assessment: Developing a Bayesian Approach to Dichotomous Dose–Response Uncertainty. *Risk Analysis* **2020**, *40*, 1706–1722, doi:10.1111/risa.13537.

200. Onar, A.; Kocak, M.; Boyett, J.M. Continual Reassessment Method vs. Traditional Empirically-Based Design: Modifications Motivated by Phase I Trials in Pediatric Oncology by the Pediatric Brain Tumor Consortium. *J Biopharm Stat* **2009**, *19*, 437–455, doi:10.1080/10543400902800486.

201. Chiuzan, C.; Shtaynberger, J.; Manji, G.A.; Duong, J.K.; Schwartz, G.K.; Ivanova, A.; Lee, S.M. Dose-Finding Designs for Trials of Molecularly Targeted Agents and Immunotherapies. *J Biopharm Stat* **2017**, *27*, 477–494, doi:10.1080/10543406.2017.1289952.

202. Braun, T.M. The Bivariate Continual Reassessment Method: Extending the CRM to Phase I Trials of Two Competing Outcomes. *Controlled Clinical Trials* **2002**, *23*, 240–256, doi:10.1016/S0197-2456(01)00205-7.

203. Kim, J.H.; Hotez, P.; Batista, C.; Ergonul, O.; Figueroa, J.P.; Gilbert, S.; Gursel, M.; Hassanain, M.; Kang, G.; Lall, B.; et al. Operation Warp Speed: Implications for Global Vaccine Security. *The Lancet Global Health* **2021**, *9*, e1017–e1021, doi:10.1016/S2214-109X(21)00140-6.

204. Reverdy, P.; Leonard, N.E. Parameter Estimation in Softmax Decision-Making Models With Linear Objective Functions. *IEEE Transactions on Automation Science and Engineering* **2016**, *13*, 54–67, doi:10.1109/TASE.2015.2499244.

205. Berger-Tal, O.; Nathan, J.; Meron, E.; Saltz, D. The Exploration-Exploitation Dilemma: A Multidisciplinary Framework. *PLoS One* **2014**, *9*, e95693, doi:10.1371/journal.pone.0095693.

206. Villar, S.S.; Bowden, J.; Wason, J. Multi-Armed Bandit Models for the Optimal Design of Clinical Trials: Benefits and Challenges. *Stat Sci* **2015**, *30*, 199–215, doi:10.1214/14-STS504.

207. Hobbs, B.F.; Hepenstal, A. Is Optimization Optimistically Biased? *Water Resources Research* **1989**, *25*, 152–160, doi:10.1029/WR025i002p00152.

208. Ito, S.; Yabe, A.; Fujimaki, R. Unbiased Objective Estimation in Predictive Optimization. In Proceedings of the Proceedings of the 35th International Conference on Machine Learning; PMLR, July 3 2018; pp. 2176–2185.

209. Hasselt, H. van; Guez, A.; Silver, D. Deep Reinforcement Learning with Double Q-Learning. In Proceedings of the Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence; AAAI Press: Phoenix, Arizona, February 12 2016; pp. 2094–2100.

210. O'Quigley, J.; Iasonos, A.; Bornkamp, B. Dose-Response Functions. In *Handbook of Methods for Designing, Monitoring, and Analyzing Dose-Finding Trials: Handbooks of Modern Statistical Methods*; 2017; p. 199 ISBN 978-1-315-15198-4.

211. Di Veroli, G.Y.; Fornari, C.; Goldlust, I.; Mills, G.; Koh, S.B.; Bramhall, J.L.; Richards, F.M.; Jodrell, D.I. An Automated Fitting Procedure and Software for Dose-Response Curves with Multiphasic Features. *Sci Rep* **2015**, *5*, 14701, doi:10.1038/srep14701.

212. Wu, C.-F. Asymptotic Theory of Nonlinear Least Squares Estimation. *The Annals of Statistics* **1981**, *9*, 501–513.

213. Genser, B.; Cooper, P.J.; Yazdanbakhsh, M.; Barreto, M.L.; Rodrigues, L.C. A Guide to Modern Statistical Analysis of Immunological Data. *BMC Immunol* **2007**, *8*, 27, doi:10.1186/1471-2172-8-27.

214. Crippa, A.; Orsini, N. Dose-Response Meta-Analysis of Differences in Means. *BMC Medical Research Methodology* **2016**, *16*, 91, doi:10.1186/s12874-016-0189-0.

215. Brock, K.; Homer, V.; Soul, G.; Potter, C.; Chiuzan, C.; Lee, S. Is More Better? An Analysis of Toxicity and Response Outcomes from Dose-Finding Clinical Trials in Cancer. *BMC Cancer* **2021**, *21*, 777, doi:10.1186/s12885-021-08440-0.

216. Cunanan, K.; Koopmeiners, J.S. Evaluating the Performance of Copula Models in Phase I-II Clinical Trials under Model Misspecification. *BMC Medical Research Methodology* **2014**, *14*, 51, doi:10.1186/1471-2288-14-51.

217. Zohar, S.; Chevret, S. The Continual Reassessment Method: Comparison of Bayesian Stopping Rules for Dose-Ranging Studies. *Statistics in Medicine* **2001**, *20*, 2827–2843, doi:10.1002/sim.920.

218. Fink, A.L.; Klein, S.L. Sex and Gender Impact Immune Responses to Vaccines Among the Elderly. *Physiology (Bethesda)* **2015**, *30*, 408–416, doi:10.1152/physiol.00035.2015.

219. Liu, G.; Carter, B.; Bricken, T.; Jain, S.; Viard, M.; Carrington, M.; Gifford, D.K. Computationally Optimized SARS-CoV-2 MHC Class I and II Vaccine Formulations Predicted to Target Human Haplotype Distributions. *Cell Syst* **2020**, *11*, 131-144.e6, doi:10.1016/j.cels.2020.06.009.

220. Green, M.S.; Peer, V.; Magid, A.; Hagani, N.; Anis, E.; Nitzan, D. Gender Differences in Adverse Events Following the Pfizer-BioNTech COVID-19 Vaccine. *Vaccines* **2022**, *10*, 233, doi:10.3390/vaccines10020233.

221. Nauta, J. Threshold of Protection. In *Statistics in clinical vaccine trials*; Springer: Heidelberg, 2010; pp. p113-115 ISBN 978-3-642-14690-9.

222. Else, H. How a Torrent of COVID Science Changed Research Publishing — in Seven Charts. *Nature* **2020**, *588*, 553–553, doi:10.1038/d41586-020-03564-y.

223. Cromer, D.; Steain, M.; Reynaldi, A.; Schlub, T.E.; Wheatley, A.K.; Juno, J.A.; Kent, S.J.; Triccas, J.A.; Khoury, D.S.; Davenport, M.P. Neutralising Antibody Titres as Predictors of Protection against SARS-CoV-2 Variants and the Impact of Boosting: A Meta-Analysis. *The Lancet Microbe* **2022**, *3*, e52–e61, doi:10.1016/S2666-5247(21)00267-6.

224. Cai, C.; Rahbar, M.H.; Hossain, M.M.; Yuan, Y.; Gonzales, N.R. A Placebo-Controlled Bayesian Dose Finding Design Based on Continuous Reassessment Method with Application to Stroke Research. *Contemp Clin Trials Commun* **2017**, *7*, 11–17, doi:10.1016/j.conctc.2017.05.002.

225. EMA ICH E10 Choice of Control Group in Clinical Trials Available online: https://www.ema.europa.eu/en/ich-e10-choice-control-group-clinical-trials (accessed on 15 August 2022).

226. Research, C. for D.E. and E10 Choice of Control Group and Related Issues in Clinical Trials Available online: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/e10-choice-control-group-and-related-issues-clinical-trials (accessed on 15 August 2022).

227. Bonate, P.L. Building Credible Systmes through Verification and Validation. In *Pharmacokinetic-pharmacodynamic modeling and simulation*; Springer: New York, 2011; p. p557 ISBN 978-1-4419-9484-4.

228. Mateus, J.; Dan, J.M.; Zhang, Z.; Rydyznski Moderbacher, C.; Lammers, M.; Goodwin, B.; Sette, A.; Crotty, S.; Weiskopf, D. Low-Dose MRNA-1273 COVID-19 Vaccine Generates Durable Memory Enhanced by Cross-Reactive T Cells. *Science* **2021**, *374*, eabj9853, doi:10.1126/science.abj9853.

229. Więcek, W.; Ahuja, A.; Chaudhuri, E.; Kremer, M.; Simoes Gomes, A.; Snyder, C.M.; Tabarrok, A.; Tan, B.J. Testing Fractional Doses of COVID-19 Vaccines. *Proceedings of the National Academy of Sciences* **2022**, *119*, e2116932119, doi:10.1073/pnas.2116932119.

230. Bonate, P.L. Covariate Detection in Population Pharmacokinetics Using Partially Linear Mixed Effects Models. *Pharm Res* **2005**, *22*, 541–549, doi:10.1007/s11095-005-2492-z.

231. Thomas, M.; Bornkamp, B.; Ickstadt, K. Identifying Treatment Effect Heterogeneity in Dose-Finding Trials Using Bayesian Hierarchical Models: Identifying Treatment Effect Heterogeneity in Dose-Finding Trials. *Pharm Stat* **2022**, *21*, 17–37, doi:10.1002/pst.2150.

232.   Wijesinha, M.C.; Piantadosi, S. Dose-Response Models with Covariates. *Biometrics* **1995**, *51*, 977–987, doi:10.2307/2532998.

233.   Rolland, P.; Kavis, A.; Immer, A.; Singla, A.; Cevher, V. Efficient Learning of Smooth Probability Functions from Bernoulli Tests with Guarantees. In Proceedings of the Proceedings of the 36th International Conference on Machine Learning; PMLR, May 24 2019; pp. 5459–5467.

234.   Buckley, P.R.; Alden, K.; Coccia, M.; Chalon, A.; Collignon, C.; Temmerman, S.T.; Didierlaurent, A.M.; van der Most, R.; Timmis, J.; Andersen, C.A.; et al. Application of Modeling Approaches to Explore Vaccine Adjuvant Mode-of-Action. *Frontiers in Immunology* **2019**, *10*.

235.   Pettini, E.; Pastore, G.; Fiorino, F.; Medaglini, D.; Ciabattini, A. Short or Long Interval between Priming and Boosting: Does It Impact on the Vaccine Immunogenicity? *Vaccines (Basel)* **2021**, *9*, 289, doi:10.3390/vaccines9030289.

236.   Shaw, R.H.; Liu, X.; Stuart, A.S.V.; Greenland, M.; Aley, P.K.; Andrews, N.J.; Cameron, J.C.; Charlton, S.; Clutterbuck, E.A.; Collins, A.M.; et al. Effect of Priming Interval on Reactogenicity, Peak Immunological Response, and Waning after Homologous and Heterologous COVID-19 Vaccine Schedules: Exploratory Analyses of Com-COV, a Randomised Control Trial. *The Lancet Respiratory Medicine* **2022**, *0*, doi:10.1016/S2213-2600(22)00163-1.

237.   Castiglione, F.; Mantile, F.; De Berardinis, P.; Prisco, A. How the Interval between Prime and Boost Injection Affects the Immune Response in a Computational Model of the Immune System. *Computational and Mathematical Methods in Medicine* **2012**, *2012*, e842329, doi:10.1155/2012/842329.

238.   O'Quigley, J.; Shen, L.Z. Continual Reassessment Method: A Likelihood Approach. *Biometrics* **1996**, *52*, 673–684.

239.   Onar-Thomas, A.; Xiong, Z. A Simulation Based Comparison of the Traditional Method, Rolling-6 Design and a Frequentist Version of the Continual Reassessment Method with Special Attention to Trial Duration in Pediatric Phase I Oncology Trials. *Contemp Clin Trials* **2010**, *31*, 259–270, doi:10.1016/j.cct.2010.03.006.

240.   Thall, P.F. Bayesian Models and Decision Algorithms for Complex Early Phase Clinical Trials. *Stat Sci* **2010**, *25*, 227–244, doi:10.1214/09-STS315.

241.   Baek, J.; Farias, V.F. TS-UCB: Improving on Thompson Sampling With Little to No Additional Computation 2021.

242.   Wang, J. Response-Adaptive Trial Designs with Accelerated Thompson Sampling. *Pharm Stat* **2021**, *20*, 645–656, doi:10.1002/pst.2098.

243.   Holland-Letz, T.; Kopp-Schneider, A. Optimal Experimental Designs for Dose–Response Studies with Continuous Endpoints. *Arch Toxicol* **2015**, *89*, 2059–2068, doi:10.1007/s00204-014-1335-2.

244.   Parker, S.M.; Gennings, C. Penalized Locally Optimal Experimental Designs for Nonlinear Models. *Journal of Agricultural, Biological, and Environmental Statistics* **2008**, *13*, 334–354.

245.   Bretz, F.; Dette, H.; Pinheiro, J. Practical Considerations for Optimal Designs in Clinical Dose Finding Studies. *Stat Med* **2010**, *29*, 10.1002/sim.3802, doi:10.1002/sim.3802.

246.   Hailemichael, Y.; Dai, Z.; Jaffarzad, N.; Ye, Y.; Medina, M.A.; Huang, X.-F.; Dorta-Estremera, S.M.; Greeley, N.R.; Nitti, G.; Peng, W.; et al. Persistent Antigen at Vaccination Sites Induces Tumor-Specific CD8+ T Cell Sequestration, Dysfunction and Deletion. *Nat Med* **2013**, *19*, 465–472, doi:10.1038/nm.3105.

247.   Lixoft SAS, a Simulations Plus company Monolix 2021R2.

248.   Phoenix™ PK/PD Platform Available online: https://www.certara.com/software/phoenix-pkpd/ (accessed on 15 August 2022).

249.   Beal SL, Sheiner LB, Boeckmann AJ, Bauer RJ NONMEM 7.5 Users Guides.

250.   Brock, K.; University, T. of C. Trialr: Clinical Trial Designs in "Rstan" 2020.

251.   Mo, Q. CRM: Continual Reassessment Method (CRM) for Phase I Clinical Trials 2018.

252.   EMA Clinical Efficacy and Safety: Clinical Pharmacology Pharmacokinetics Available online: https://www.ema.europa.eu/en/human-regulatory/research-development/scientific-

373

guidelines/clinical-efficacy-safety-clinical-pharmacology-pharmacokinetics (accessed on 15 August 2022).

253.   EMA Clinical Evaluation of New Vaccines Available online: https://www.ema.europa.eu/en/clinical-evaluation-new-vaccines (accessed on 15 August 2022).

254.   Chen, H.-W.; Wong, W.K.; Xu, H. Data-Driven Desirability Function to Measure Patients' Disease Progression in a Longitudinal Study. *J Appl Stat* **2016**, *43*, 783–795, doi:10.1080/02664763.2015.1077378.

255.   ISIDLSHTM ISIDesire 2022.

256.   Xie, X.; Kou, S.C.; Brown, L.D. SURE Estimates for a Heteroscedastic Hierarchical Model. *J Am Stat Assoc* **2012**, *107*, 1465–1479, doi:10.1080/01621459.2012.728154.

374

Appendices

Appendix A: Statistical Dose-Response Functions

Given that I suggest that modelling may be beneficial for the purposes of selecting optimal vaccine dose, I include in this appendix a list of some example statistical models for dose-response that could be applied in vaccines. Any of these could be used as a statistical model for dose-efficacy of dose-toxicity relationships, though it may be preferable to use an increasing or saturating model for dose-toxicity. Note that again it may be reasonable to use a weighted average of the models described below, as described in chapter 5. Note also that this does not represent an exhaustive list, as there have been a large number of proposed dose-response functions [1,2].

Throughout this appendix I use $y_{i,j}$ as the outcome observed for response y for individual j receiving dose $dose_i$.

The below include only statistical models. Mechanistic models for dose-immunological response may be beneficial as discussed in the main body of this thesis, but these likely depend on vaccine platform and immune response of interest. The below are also primarily models of single administration dose-response.

I preface these dose-response models with a brief description of what is meant by 'dose' and 'response'.

A.A.1. Dose and Dose Transformation

It may be reasonable to transform the 'raw' value of doses before modelling. This can help with stabilising computation [3]. I list here some common dose transformations. These are of the form

$$x_i = transform(dose_i)$$

Where x_i is the transformed dose value that will be used for modelling and again $dose_i$ is the 'raw' dose value. In particular, using a log10 transform is common in drugs and in vaccines [4,5]

Untransformed:

$$x_i = transform(dose_i) = dose_i$$

Log10:

$$x_i = transform(dose_i) = log_{10}(dose_i)$$

Minmax

$$x_i = transform(dose_i) = \frac{dose_i - dose_{min}}{dose_{max} - dose_{min}}$$

Log10 Minmax

$$x_i = transform(dose_i) = \frac{log_{10}(dose_i) - log_{10}(dose_{min})}{log_{10}(dose_{max}) - log_{10}(dose_{min})}$$

Where $dose_{min}$ and $dose_{max}$ are respectively the smallest and largest possible dose that could be given.

For chapters 2-4 the Log10 transform was used. Chapters 5 and 6 could be considered to be using any transform due to the nature of being simulation studies. However, specifically chapter 5 discussed a Log10 transform, and chapter 6 assumed that either the Minmax or Log10 Minmax transforms were used.

There is also the so called 'codified' transformation that can be used for discrete dosing domains [6]. This transform was not used in this work but is included due to its prevalence in the Continual Reassessment Method literature. For the D potential discretised raw doses the codified transform is given by.

Codified:

$$x_i = log_{10}(dose_i) - \sum_{d=1}^{D} \frac{log_{10}(dose_d)}{D}$$

In Figure A.A.1.1. I show the effect of these transformations using an example discretized dosing domain of 7 discretised potential doses: 100, 1000, 10000, 100000, 1000000, 5000000, 10000000. The units are unimportant but could for example be Viral Particles.
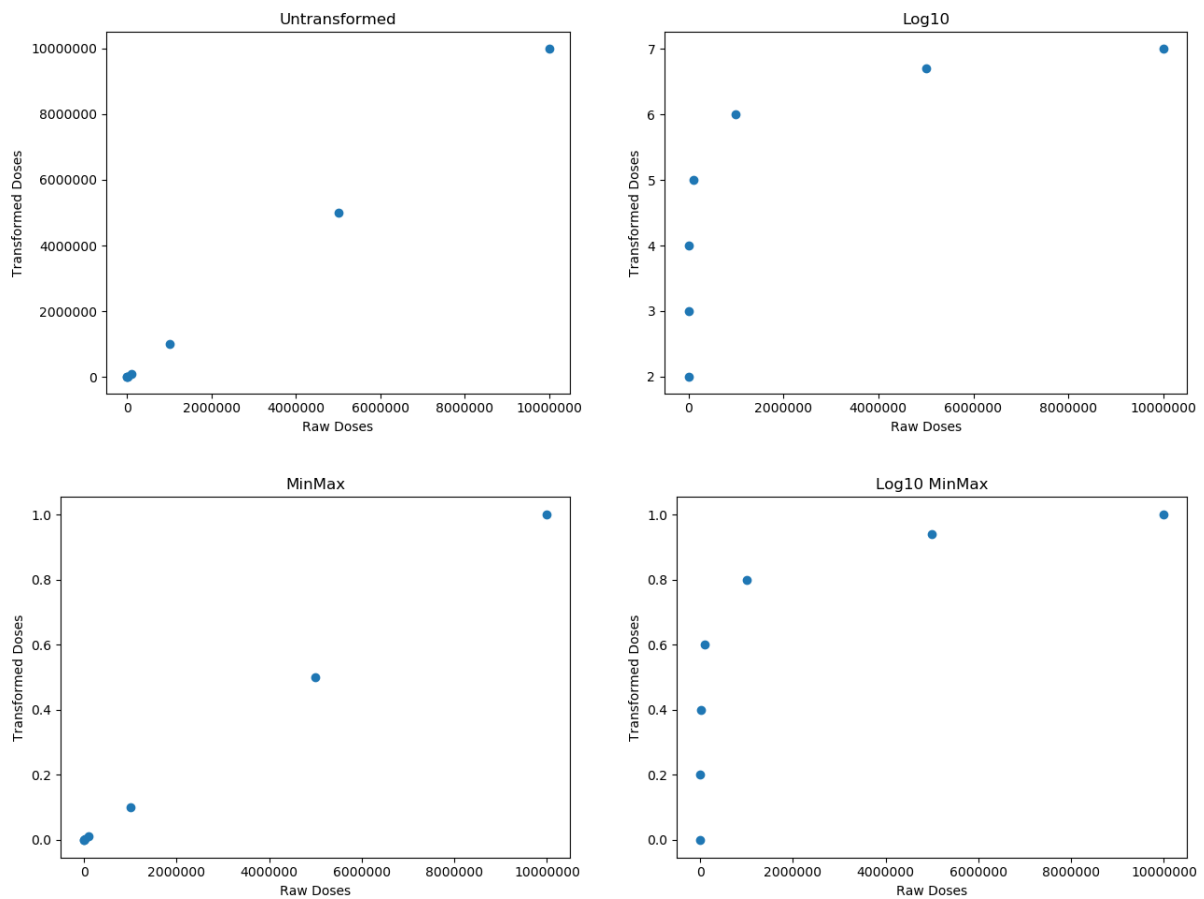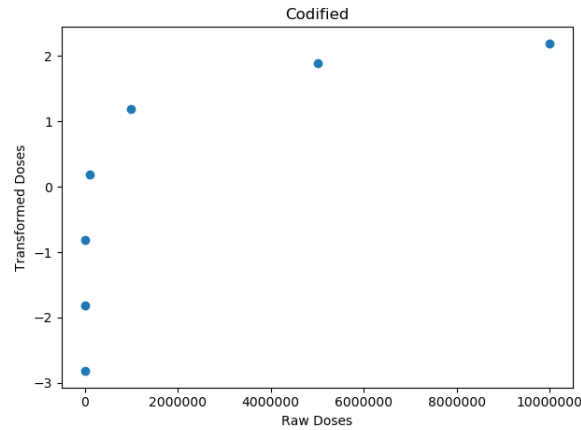
Figure A.A.1.1 Visualisation of dose transformations.

A.A.2. Response

In this work I considered data for which response was either discrete binary/ordinal or could be considered to be approximated for any given dose by a normal distribution. Further modelling techniques are available for other response distributions, for example if data are Poisson distributed [7], but will not be discussed here.

Binary/Ordinal outcomes

Here let there be K potential discrete and exclusive outcomes, and assume that the observations $y_{i,j}$ are generated by

$$P(y_{i,j} = k) = f_k(x_i)$$

With the model $f_k()$ defining the probability of observing the Kth discrete outcome. Note that for all x_i it must be true that

$$0 \leq f_k(x_i) \leq 1$$

$$\sum_{k=1}^{K} f_k(x_i) = 1$$

For calibration of model parameters, typically the parameters are chosen that maximise the likelihood over the J individual data points

378

$$Likelihood = \prod_{J=1}^{J} \sum_{k=1}^{K} P(y_{i,j} = k)\delta(y_{i,j} = k)$$

Where $\delta(y_{i,j} = k)$ is the Kronecker delta function which equals 1 if $y_{i,j} = k$ and otherwise equals 0.

Normally Continuous Outcomes

Here assume that the observations $y_{i,j}$ are generated by l

$$y_{i,j} = f(x_i) + \epsilon_{i,j}$$

$$\epsilon_{i,j} \stackrel{\text{iid}}{\sim} N(0, \sigma^2)$$

With dose-response model $f()$. For calibration of model parameters, σ can often be ignored by choosing to minimise the sum of square error over the J individual data points

$$SSE = \sum_{j=1}^{J} (y_{i,j} - f(x_i))^2$$

This is not appropriate if heteroscedasticity is expected, but homoscedasticity is often assumed [8,9].

A.A.3. Dose-Response Models

Here I list a non-exhaustive subset of dose-response models. For each model I give a brief description of the model and its potential use, the formula $f(x_i)$ that defines it, and a visualisation of the model under three different parameterisations.

Note that many of these models use so-called 'link' functions to map a latent function into another range. These link functions do not have parameters that need to be estimated. These link functions are given here for brevity in the rest of this section.

The logit⁻¹(or sometimes simply 'logit') link function maps from $(-\infty, \infty)$ onto $(0, 1)$ and is given by [10]

$$logit^{-1}(z) = \frac{e^z}{1 + e^z}$$

The SoftPlus link function [11] maps from $(-\infty, \infty)$ onto $(0, \infty)$ and is given by

$$softplus(z) = log(1 + e^z)$$

Figure A.A.3.1 visualises these link functions.
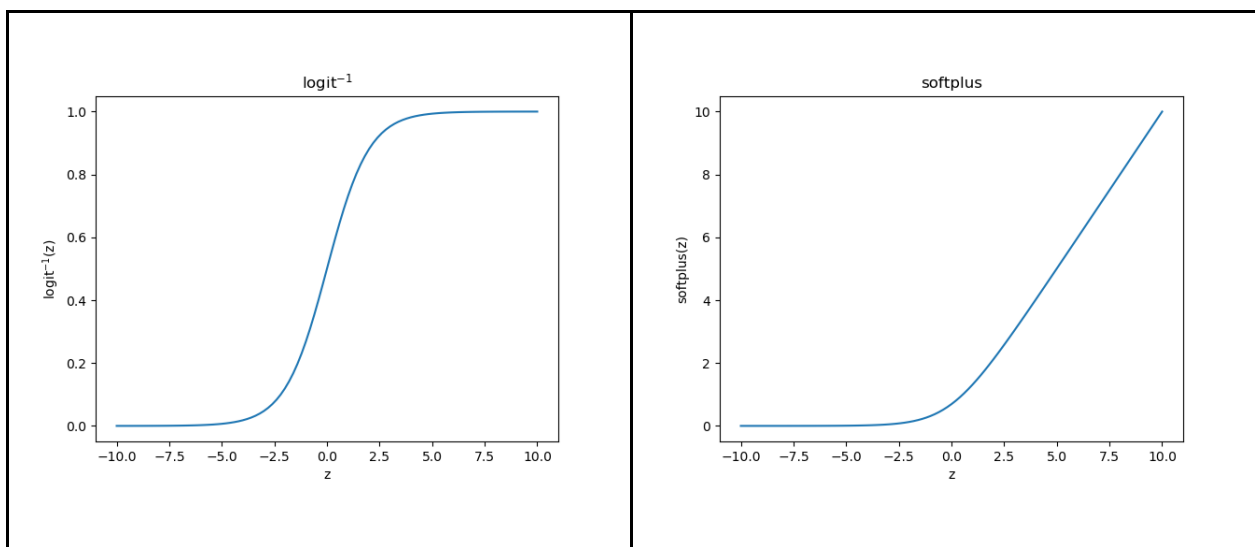
Figure A.A.3.1. Visualisation of two link functions.

For the below models, a, b, c, d are model parameters and e is the mathematical constant. For each model, a visualisation of 3 different parameterisations of that dose-response model is given.

The four-parameter sigmoid saturating model is a flexible model for describing saturating dose-response models[12–14]. This has the form

$$f(x_i) = d + (c - d)logit^{-1}(z)$$

with

$$z = a + bx_i$$

With $c > d \geq 0$.
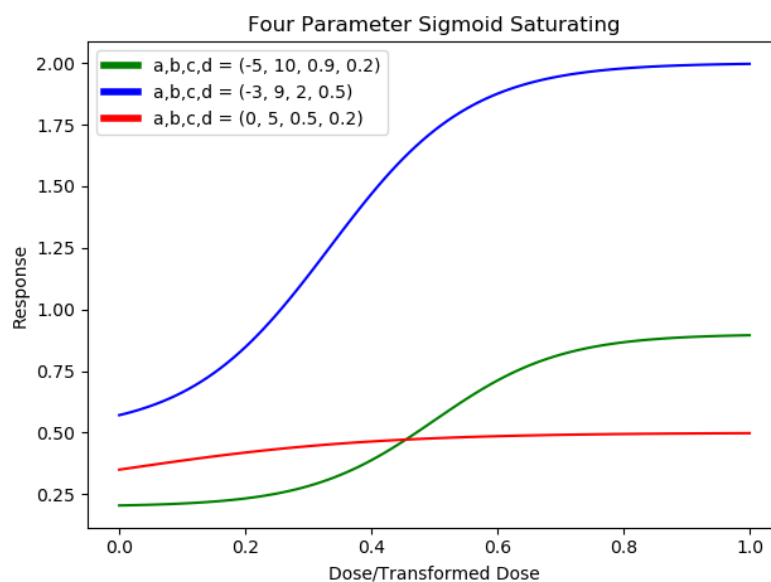
Figure A.A.3.2 Visualisation of the Four Parameter Sigmoid Saturating Model

Three-parameter Sigmoid Saturating

The three-parameter sigmoid saturating model is equivalent to the four-parameter sigmoid saturating model with $d = 0$. This is reasonable if it can be assumed that

$$\lim_{x_i \to -\infty} f(x_i) = 0$$

And is the saturating model assumed throughout this work. This model has the form

$$f(x_i) = c * logit^{-1}(z)$$

with

$$z = a + bx_i$$
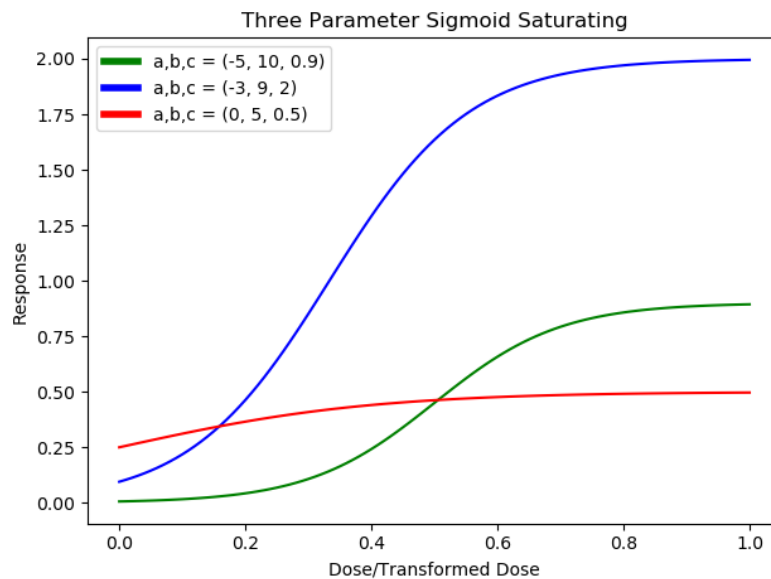
Figure A.A.3.3. Visualisation of the Three-Parameter Sigmoid Saturating Model

Two-Parameter Sigmoid Saturating/Logistic Regression

The two-parameter sigmoid saturating model is equivalent to the three-parameter sigmoid saturating model with $c = 0$. This is reasonable if it can be assumed that

$$\lim_{x_i \to \infty} f(x_i) = 1$$

This assumption may not be reasonable for modelling dose-efficacy for vaccines, as it is possible that some individuals may be non-responders regardless of dose. This is also equivalent to logistic regression if being used to model the probability of a binary outcome. This model has the form

$$f(x_i) = logit^{-1}(z)$$

with

$$z = a + bx_i$$
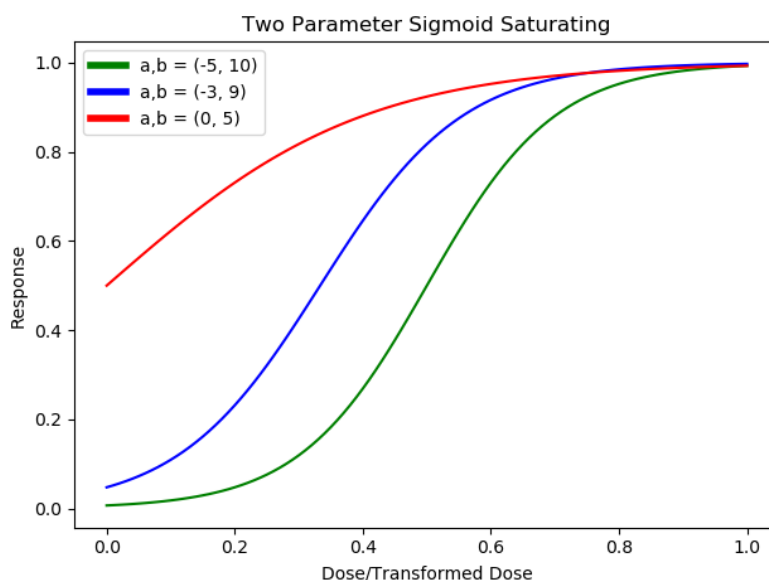
Figure A.A.3.4. Visualisation of the Two Parameter Sigmoid Saturating Model

This may be a reasonable model for describing dose-response where the response can be assumed to increase linearly with dose and the range of the response variable is $[-\infty, \infty]$ [1].

This model has the form

$$f(x_i) = a + b * x_i$$

With $a > 0$ implying that response is increasing with dose.
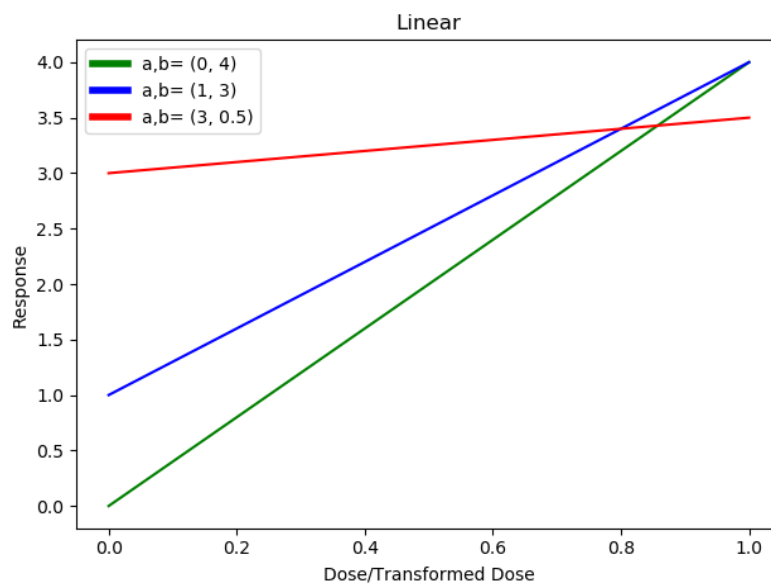
Figure A.A.3.5. Visualisation of the Linear Model

Hill

This model is another example of a potential saturating curve and is sometimes called the 'Hill-Langmuir' model. This has been commonly used in drug dose-response modelling [15,16]. This model should only be used if all transformed doses x_i are greater than 0. This model has the form

$$f(x_i) = \frac{a * x_i^b}{c + x_i^b}$$

with all parameters greater than 0.
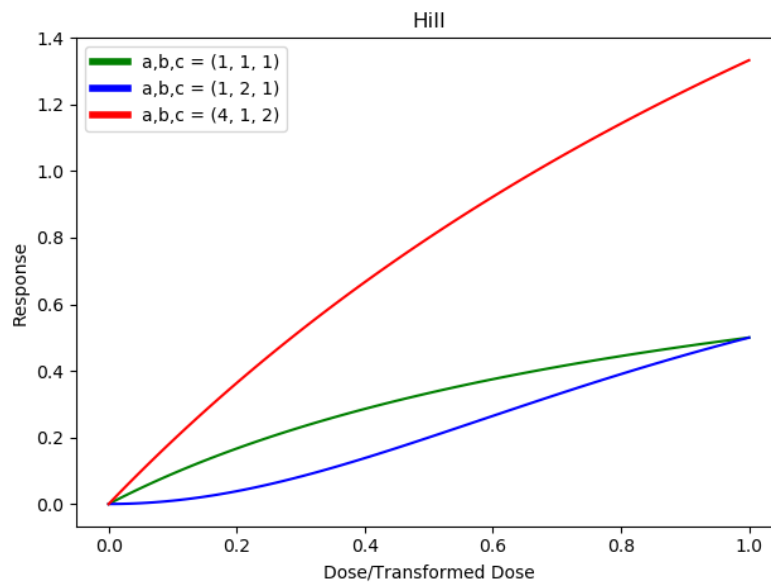
Figure A.A.3.6. Visualisation of the Hill Model

385

Linear Saturating

This model is another example of a potential saturating curve, however this one is not sigmoidal as it has no turning point in its derivative [16]. This is also known as the Michaelis-Menten dose response model. This can also be considered as the Hill-Langmuir model with $b = 1$. This model should only be used if all transformed doses x_i are greater than 0. This model has the form

$$f(x_i) = \frac{a * x_i}{c + x_i}$$

with all parameters greater than 0.
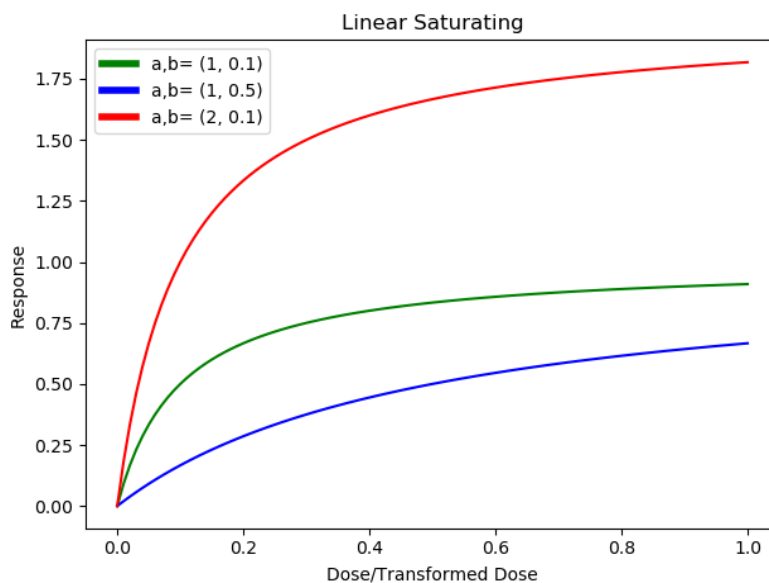
Figure A.A.3.7. Visualisation of the Linear Saturating Model

Quadratic

This may be a reasonable model for describing peaking dose-response where the range of the response variable is $[-\infty, \infty]$ [1]. For modelling binary outcome probabilities or immunological responses where only, positive responses are plausible the latent quadratic or SoftPlus quadratic dose-response models should be used. This model has the form

$$f(x_i) = a + b * x_i + c * x_i^2$$

For the model to be limited to describing a peaking dose response the parameters should be bounded to $b > 0, c < 0$.
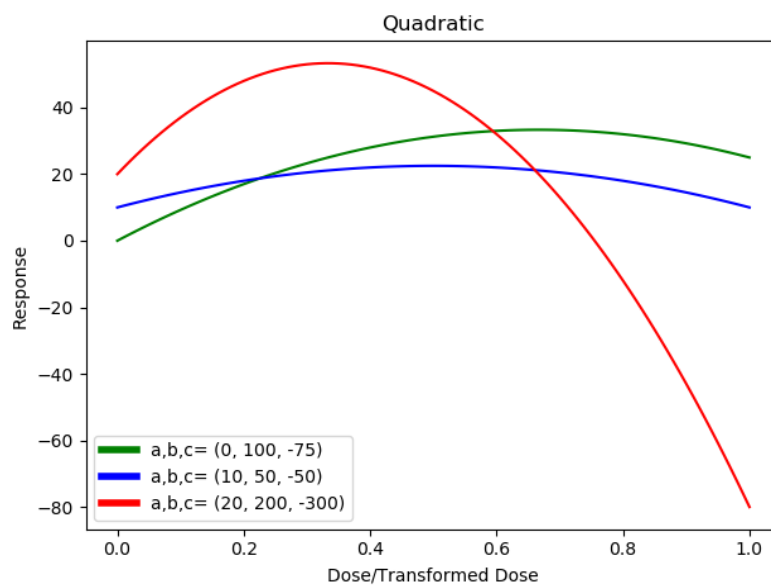
Figure A.A.3.8. Visualisation of the Quadratic Model

This may be a reasonable model for describing dose response where the dose-domain is limited to doses that are too small for any saturating or peaking effects to be observed [1]. This model has the form

$$f(x_i) = d + c * e^{b*x_i - 1}$$

With $b > 0, c > 0, d \geq 0$.

It is possible to set $d = 0$ if it is believed

$$\lim_{x_i \to -\infty} f(x_i) = 0$$
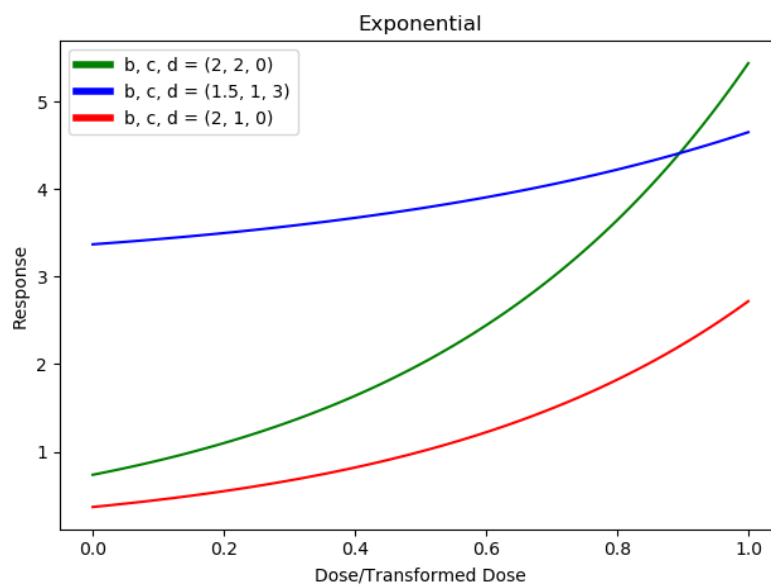
Figure A.A.3.9.Visualisation of the Exponential Model

Power

This may be a reasonable model for describing dose response where the dose-domain is limited to doses that are too small for any saturating or peaking effects to be observed[1]. This model should only be used if all transformed doses x_i are greater than or equal to 0. This model has the form

$$f(x_i) = d + c * x_i^a$$

With $a > 0, c > 0, d \geq 0$.

It is possible to set $d = 0$ if it is believed

$$\lim_{x_i \to -\infty} f(x_i) = 0$$
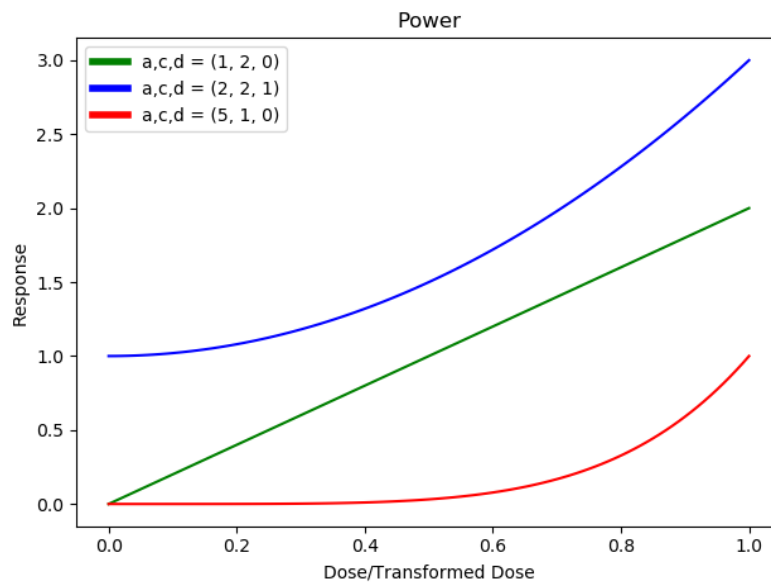
Figure A.A.3.10.Visualisation of the Power Model

Gamma PDF

This is the model that was used to represent a peaking dose-response curve in chapter 2. This has been used historically in [4]. Up to reparameterisation, this model has the form

$$f(x_i) = c * x^{a-1} * e^{-b*x}$$

With all parameters greater than 0.
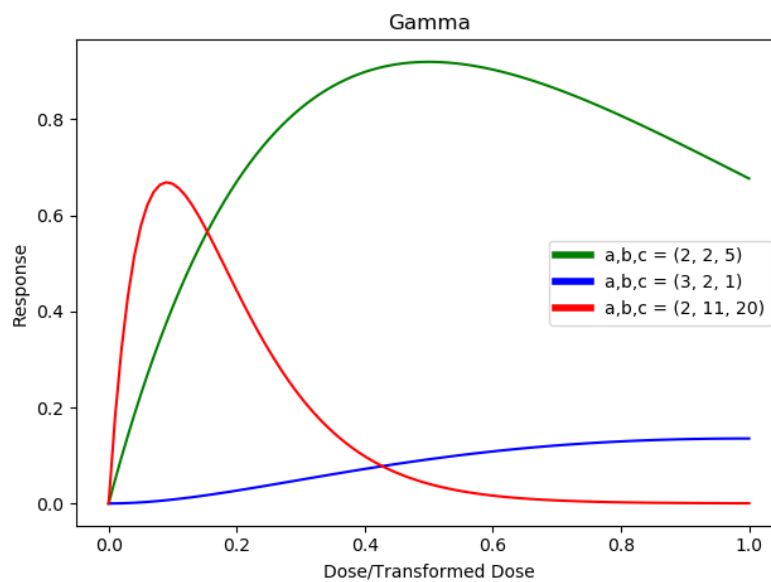
Figure A.A.3.11.Visualisation of the Gamma PDF Model

This may be a reasonable model for describing peaking dose-response where the range of the response variable is $[0, 1]$. For modelling binary outcome probabilities this is reasonable, and thus this model was used in chapters 5 and 6 to model dose-efficacy. This model has the form

$$f(x_i) = logit^{-1}(z)$$

with

$$z = a + b * x_i + c * x_i^2$$

For the model to be limited to describing a peaking dose response the parameters should be bounded to $b > 0, c < 0$.
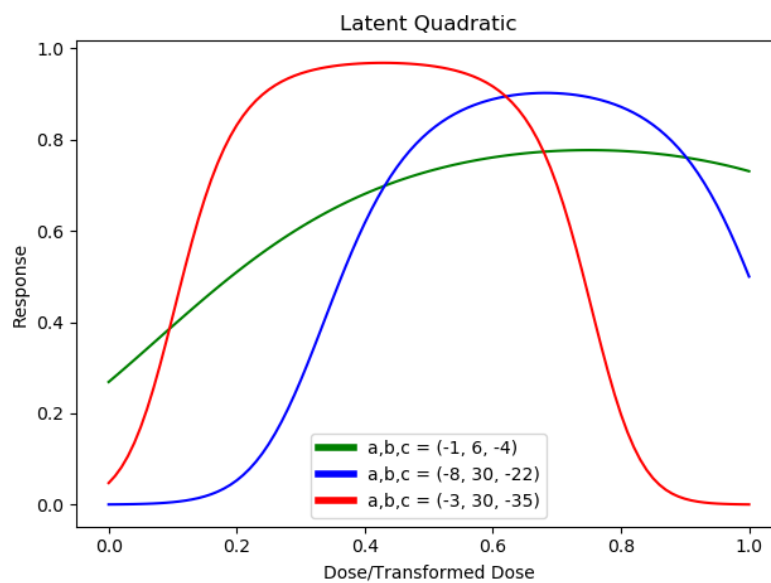
Figure A.A.3.12.Visualisation of the Latent Quadratic Model

This may be a reasonable model for describing dose-response where the response can be assumed to increase linearly with dose but cannot take a non-positive value. For modelling immunological responses where only positive responses are plausible this is reasonable. I am not aware of any dose-response modelling studies where this model has been used but seems intuitively reasonable. This model has the form

$$f(x_i) = softplus(z)$$

with

$$z = a + b * x_i$$

With $a > 0$ implying that response is increasing with dose.
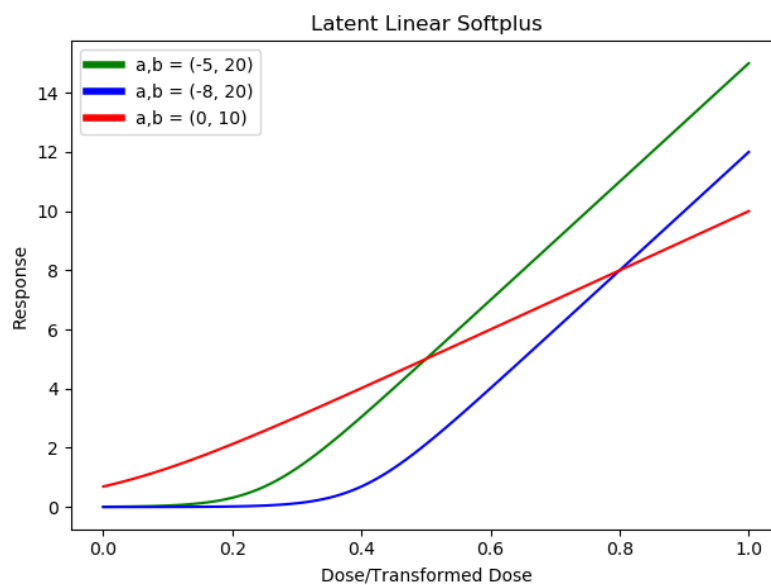
Figure A.A.3.13.Visualisation of the Latent Linear SoftPlus Model

This may be a reasonable model for describing peaking dose-response where the range of the response variable is $(0, \infty]$. For modelling immunological responses where only positive responses are plausible this is reasonable. I am not aware of any dose-response modelling studies where this model has been used but seems intuitively reasonable. This model has the form

$$f(x_i) = softplus(z)$$

with

$$z = a + b * x_i + c * x_i^2$$

or the model to be limited to describing a peaking dose response the parameters should be bounded to $b > 0, c < 0$
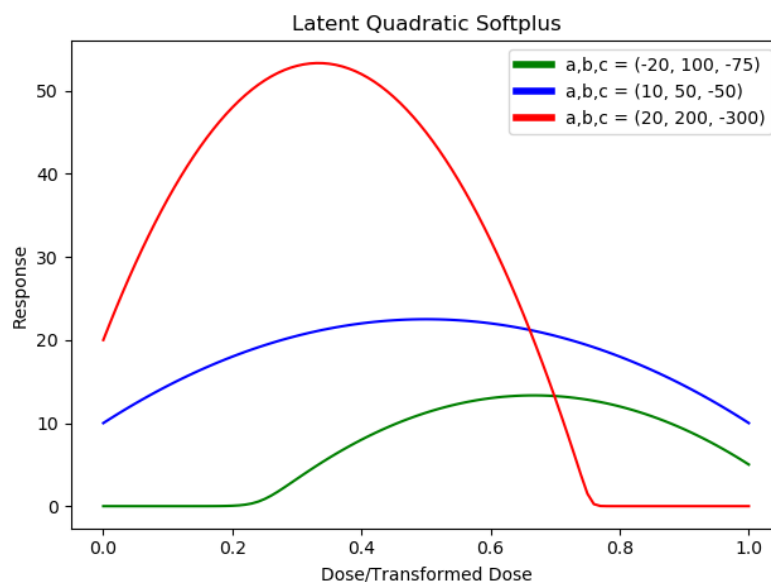
Figure A.A.3.14.Visualisation of the Latent Quadratic SoftPlus Model

Piecewise linear

Piecewise linear models can be used to describe dose-response through linear interpolation between so-called 'knots' [17].

This model has the form:

$$
f(x_i) = \begin{cases} \beta_1 & \text{for } x_i < \alpha_1, \\ \beta_j + (\frac{\beta_{j+1}-\beta_j}{\alpha_{j+1}-\alpha_j})(x_i - \alpha_j) & \text{for } \alpha_j \leq x_i < \alpha_{j+1}, \\ \beta_z & \text{for } \alpha_z \leq x_i \end{cases}
$$

Where α_j, β_j are respectively the jth (transformed) dose and response values for the z knots.
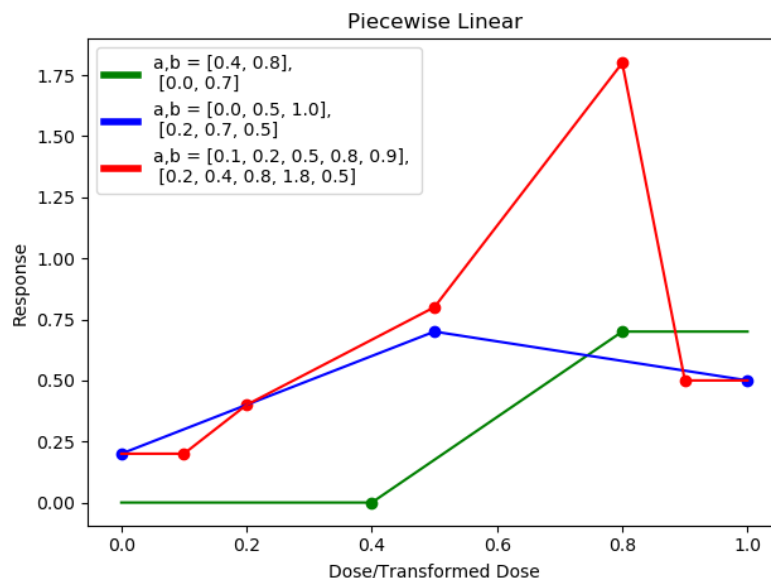
Figure A.A.3.15. Visualisation of the Piecewise Linear Model. The coloured points show the knots for their respective parameter values

Note that often the α parameters are chosen a-priori to reduce the number of parameters and hence reduce the risk of overfitting to limited data.

Artificial Neural Network

Artificial neural networks have become highly relevant mathematical models for prediction for many areas of science over the last decade. The implementation and architecture of such models can vary [18]. They are so called 'universal function

approximators', meaning that with sufficient data and sufficient computation they can approximate any function, which could include dose-response curves. These models typically have a large number of parameters, which given the small number of dosing groups typically investigated in vaccine dose ranging could lead to issues related to overfitting.

Despite this issue, such models have previously been suggested for modelling the dose-response relationship between dose of food-borne pathogens and probability of infection [19]. They found that these could outperform other statistical dose-response models for a collection of four test datasets. Additionally, recent research by Nakkiran et al. into the 'double descent' hypothesis might imply large neural network models do not show the loss of predictive validity that overfitting is typically associated with. These methods have been more recently suggested again with particular interest in incorporating individual covariates into dose-response predictions [20].

CCBP

This is the non-parametric model used in chapter 6. See chapter 6 for details.

Gaussian Process and Product of Beta Prior

These are non-parametric models like the CCBP model discussed in chapter 6. Given the complexity of these models, they will not be detailed here. Please see [21] for details.

Isotonic and Double Isotonic Regression

This is another form of non-parametric modelling, where isotonic regression is potentially useful when it can be assumed that the relationship between dose and response is strictly increasing but also not able to be effectively approximated by established parametric models. Intuitively, for the D discretised and transformed doses $x_1, ..., x_D$, the model predicts the mean response that has previously been observed at that dose. If any model predictions between adjacent doses would violate monotonicity, then the predictions of the model are replaced by a weighted

average of the model predictions that violate monotonicity. If there are no data for a dose, then typically linear interpolation is used.

Double isotonic regression is similar but specifies a specific 'peak' dose x_j for which dose-response is assumed to be increasing below and decreasing beyond. Therefore, this makes the assumption of a peaking dose-response curve, but does not use a specific parametric form. Rather than specifying only one dose as this threshold, D double isotonic regression models would be calibrated to the data, with each of the D potential dose having exactly one model for which it is the peak dose. Whichever of these D models best describes the observed data is used to predict optimal dose.

Figure A.A.3.16 shows an example of a double-isotonic regression model being used to describe simulated dose-response data. A more detailed description and discussion of the application of double isotonic regression in conducting dose-finding trials is given by Zang, Lee, and Yuang [22], and where this form of modelling was shown to be comparable or better than a parametric modelling strategy for many of the scenarios in their simulation study.
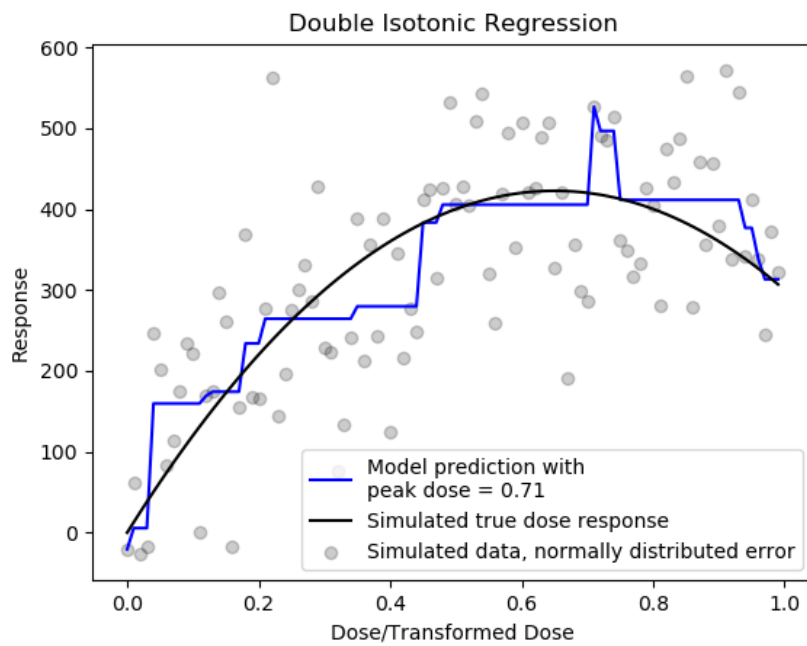
Figure A.A.3.16.Visualisation of the double-isotonic regression model being used to describe a simulated set of dose-response data (simulated true dose-response curve in black).

Probit and Proportional Odds

These models are used for describing ordinal outcome variables. Given the complexity of these models, they will not be detailed here. Please see [23] for details. The probit model was used in chapter 5 to model ordinal toxicity.

The 'power skeleton' dose-response model is distinct from the other models discussed in that it is entirely dependent on expert predictions [24]. This model can only be used for a discrete dosing domain, and can only be used for predicting probabilities of events. Rather than using a dose-transformation as described above, for each discrete dose $dose_j$ an expert prediction p_j of the probability of response for that dose must be elicited. This is called the 'skeleton', and there has been previous discussion of how best to choose this skeleton.

The model then has the form

$$P(response|dose_j) = p_j{}^{exp(a)}$$

With a as the only parameter, that can take any real number value. $a = 0$ implies that the expert predictions were correct for all doses, $a < 0$ implies the expert underestimated the probability of response for all doses and $a > 0$ implies the expert overestimated the probability of response for all doses. This model is commonly used for dose-toxicity studies which use the continual reassessment method.
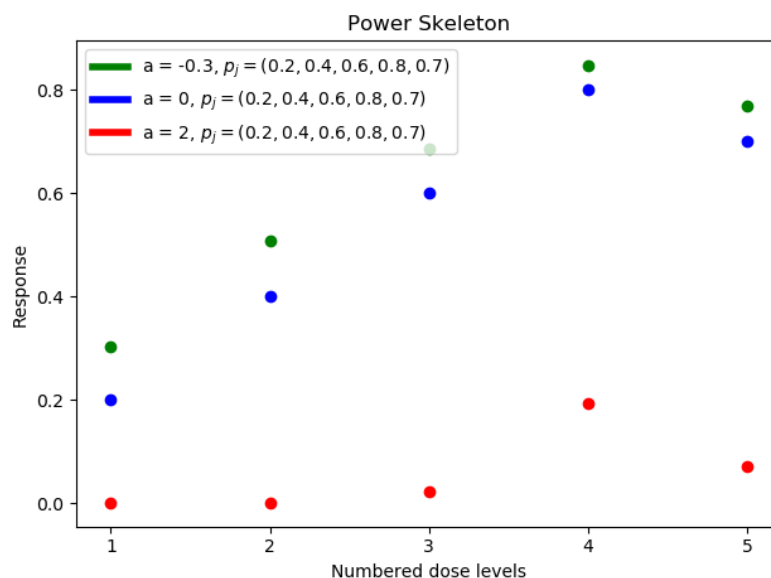
Figure A.A.3.17. Visualisation of the power skeleton model.

Appendix B: Pareto Optimality

In this work I have discussed the concept of multi-objective optimisation problems in vaccine dose-optimisation, and largely propose the use of utility functions to define 'optimal dose'. This was beneficial for chapters 5 and 6, for which I was interested in how well mathematical modelling-based dose-optimisation approaches can locate optimal dose. However, when discussing utility functions and multi-objective optimisation it is important to understand the concept of Pareto optimality as this underpins all theory regarding multi-objective optimisation problems[25–27].

A.B.1. Dominated Solutions and the Pareto Fronts

Optimisation problems (OPs) exist as a class of problem that involves choosing a 'best' option over some space of possibilities. These problems can be written as

$$\text{optimise w.r.t } x \in D$$

$$f(x)$$

Where D is the decision space of possible options, x is a possible option in this space, and f() is the function that I wish to optimise, sometimes called the objective function. 'Optimise' typically means maximise or minimise. There is typically no distinction between maximising and minimising, as any optimistion problem written as

$$\text{maximise w.r.t } x \in D$$

$$f(x)$$

Can equally be considered as

$$\text{minimise w.r.t } x \in D$$

$$g(x) = -f(x)$$

An example vaccine dose-optimisation problem could be

maximise w.r.t dose $\in D$

$$\text{efficacy(dose)}$$

There are many OPs where there is not a single optimisation function that must be considered, but multiple. Suppose that I have two OPs.

maximise w.r.t $x \in D$

$$f_1(x)$$

maximise w.r.t $x \in D$

$$f_2(x)$$

It is possible that I would like to choose some $x \in D$ to optimise both OPs. In this section I discuss such OPs, which are referred to as multi-factoral optimisation problems (MOPs).

We may consider the writing the above MOP as

maximise w.r.t $x \in D$

$$f_1(x), f_2(x)$$

Which is to say that I would like to choose some $x \in D$ that both maximises $f_1(x)$ and $f_2(x)$. This is not a well-defined MOP unless there exists such an $x \in D$ that does so.

To highlight this problem, consider table A.B.1.1. Clearly x_1 is optimal compared to x_2, as $f_1(x_1) < f_1(x_2)$ and $f_2(x_1) < f_2(x_2)$.

x	$f_1(x)$	$f_2(x_2)$
x_1	50	30
x_2	70	40

Table A.B.1.1. First example of Pareto optimality

x	$f_1(x)$	$f_2(x_2)$
x_1	60	50
x_2	80	20

Table A.B.1.2. Second example of Pareto optimality

Now instead consider table A.B.1.2. Here x_1 nor x_2 neither can be considered truly 'optimal' relative to the other, as x_1 better maximises $f_2(x)$.and x_2 better maximises $f_1(x)$. Hence there is no clear 'optimal', and I can only consider 'Pareto optimality' and 'domination'.

We say that ' x_i dominates x_j' to mean that, for all objectives, x_i is at least as good as x_j, and that there is at least one objective for which x_i is preferable to x_j. x_i is 'Pareto optimal' for a MOP if it is not dominated by any other $x \in D$. So, in table A.B.1.1, only x_2 is Pareto optimal. In table A.B.1.2, both x_1 and x_2 are Pareto optimal. In table A.B.1.3, x_1 is dominated by x_2 and x_3, x_2 is not dominated by any other $\in D$, x_3 is not dominated by any other $x \in D$, and x_4 is dominated by all other $\in D$. Hence here only x_2 and x_3 are Pareto optimal.

x	$f_{1(x)}$	$f_{2(x)}$
x_1	60	50
x_2	60	60
x_3	70	55
x_4	20	10

Table A.B.1.3. Third example of Pareto optimality

The set of points that are Pareto optimal is called the Pareto front. The main purpose of considering the Pareto front is that only solutions within this set can be called optimal, and that any solution within this set could be considered optimal depending on the utility function [28]. This will be relevant for the question of maximising efficacy and minimising toxicity.

A.B.2. Pareto optimality in multi-objective optimisation for single administration vaccines

Pareto optimality is important when considering single dose optimisation and informs why a utility function is needed. Considering figure A.B.2.1, I see that when dose-efficacy and dose-toxicity are saturating, all doses are Pareto optimal, and so defining optimal dose as the dose which maximises efficacy and minimises toxicity is not sufficient to choose a dose as 'optimal' using these dose-response curves. When dose-efficacy is peaking, all doses greater than the dose which maximises efficacy are dominated by the dose that maximises efficacy, and hence all below this point are Pareto optimal [figure A.B.2.2].
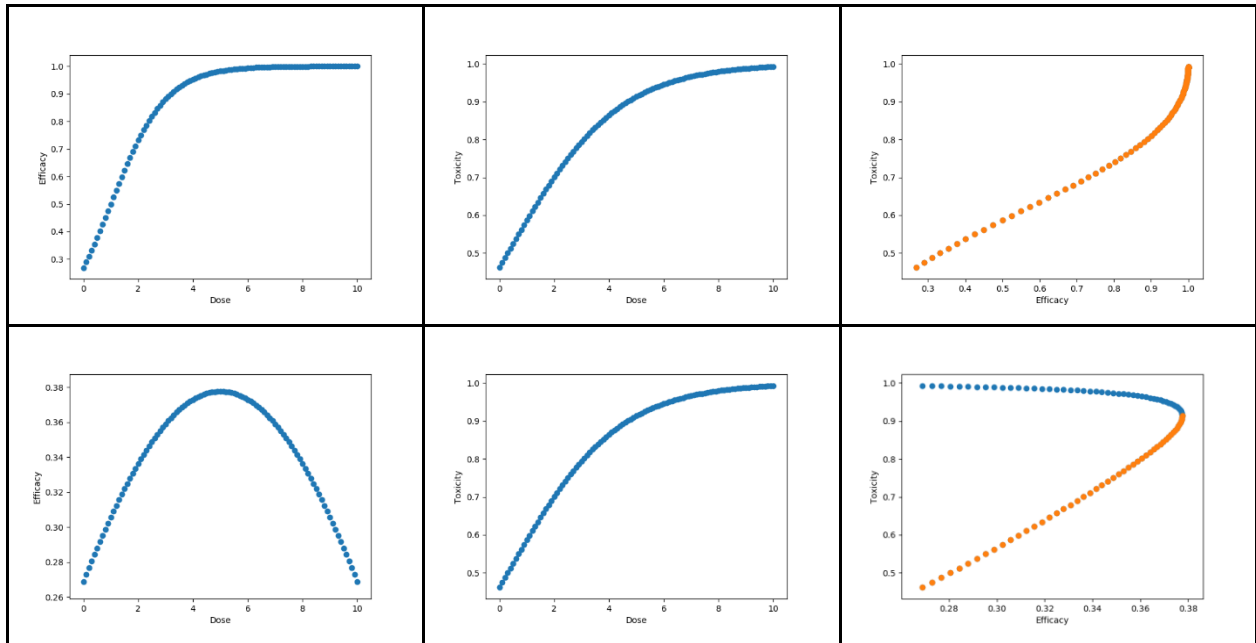
Figure A.B.2.2. Example peaking dose-efficacy (left), dose-toxicity(middle), and resultant efficacy-toxicity (right) plots. For the efficacy-toxicity plot, Pareto optimal doses are in orange and non-Pareto optimal doses are in blue (here all doses less than that which maximises efficacy are Pareto optimal).

A.B.3. Pareto optimality in multi-objective optimisation for prime-boost vaccines

I here provide an example Pareto front for a prime-boost administration vaccine. I assume that dose-efficacy has been modelled using the 2 dimensional variant of the latent quadratic model as described in chapter 6, and after calibration the model parameters have been estimated as a, b_1, b_2, c_1, c_2 = -1, 5, 4, 7, 4. Similarly, assume that dose-toxicity has been modelled using the 2 dimensional variant of the latent linear model as described in chapter 6, and after calibration the model parameters have been estimated as a, b_1, b_2, = -2, 1, 3. Thus the results predicted dose-efficacy and dose-toxicity contours are as shown in Figure A.B.3.1.
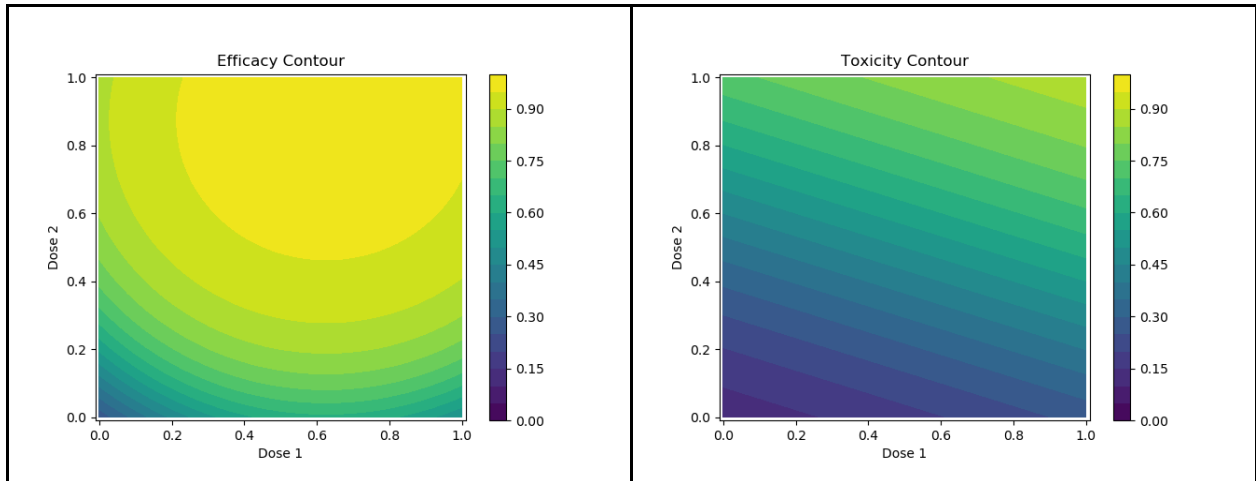
FigureA.B.3.1. Example peaking dose-efficacy (left) and dose-toxicity(right) contours. Yellow regions represent increased probability.

The Pareto front can be calculated and is shown in figure A.B.3.2. For any dose not on the Pareto front (red line) a dose can be chosen that is either more efficacious with equal/lower toxicity or less toxic with equal/greater efficacy. For example, a non-Pareto optimal dose of (dose 1, dose 2) = (0.0, 0.6) would have 85% efficacy and 45% toxicity. (dose 1, dose 2) = (0.45, 0.45) would have the same toxicity (=45%), but 95% efficacy. Thus, assuming that there are no other objectives other than efficacy and toxicity, no decision maker could rationally select a dose of (0.0, 0.6) as 'optimal' regardless of how they weighted efficacy and toxicity.

Note that, as expected, the dose that is 'optimal' as defined by the utility contour utility function defined in paper 5 indeed does lie on the Pareto front [figure A.B.3.2.c].
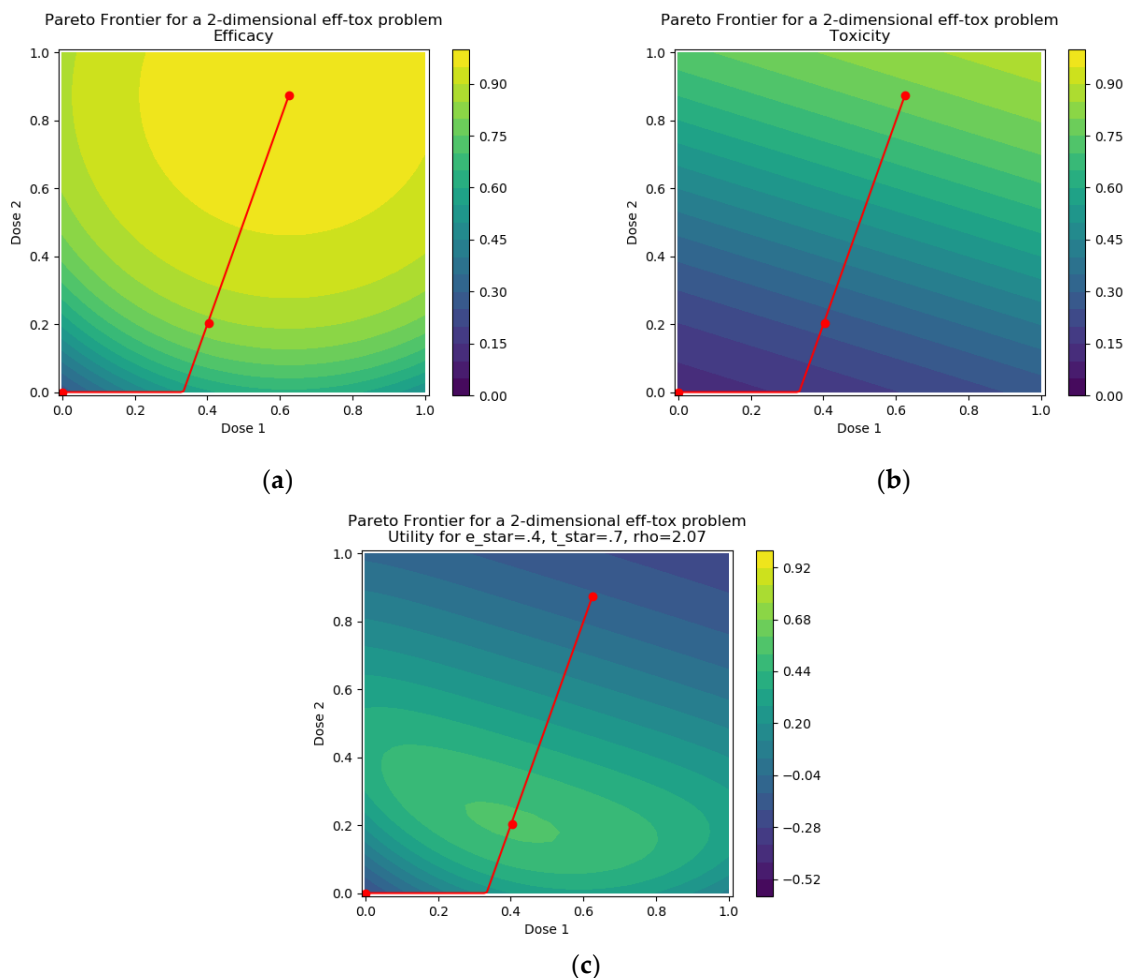
Figure A.B.3.2. Example peaking dose-efficacy (a) and dose-toxicity(b) contours. The dose-utility contour (c) is shown for the case where utility is defined by the utility contour utility function used in chapter 6. Yellow regions represent increased probability/utility. The red line shows the Pareto front, and the red dots show: the dose that minimises toxicity irrespective of efficacy, the dose that maximises efficacy irrespective of toxicity, and the dose that maximises the utility contour utility function used in chapter 6.

The doses and resulting Pareto front can also be plotted on an efficacy/toxicity curve [Figure A.B.3.3]. This may better show what is meant by the Pareto optimal front. Note that the majority of doses are not on the Pareto optimal front, in contrast to the observations in the single-administration case. A utility function is still needed to meaningfully define 'optimal' dose.
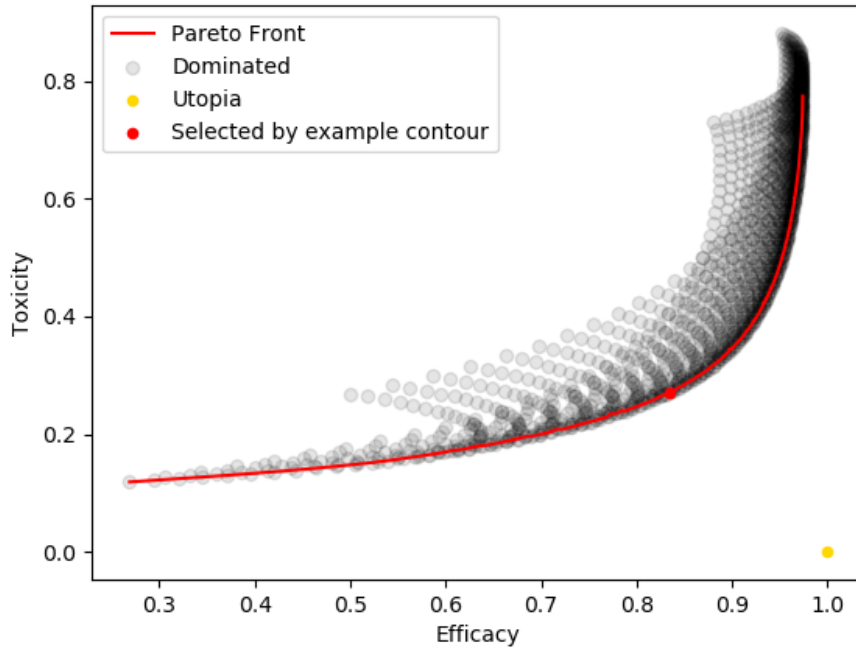
Figure A.B.3.3. Example efficacy-toxicity plots and Pareto front for the example in AB.3. The gold point is the theoretical utopia point, with 100% efficacy and 0% toxicity.

A.B.4. Utility functions in relation to the Pareto front.

Throughout this work I argue for the use of utility functions to define 'optimal dose' in MOPs, and here discuss the relationship of these to the Pareto front. For the efficacy-toxicity and efficacy-toxicity-cost utility functions in paper 3, the scalar aggregation utility function used in paper 4 and the utility contour utility functions used in paper 5, any the solution ('dose') that maximises the utility function will lie on the Pareto front according to model predictions.

First note that these utility functions satisfy:

$$\text{for all i } \in [1, n], f_i \prec f_i^* \iff U(f_1, \ldots, f_i, \ldots, f_n) \prec U(f_1, \ldots, f_i^*, \ldots, f_n)$$

Where $\prec$ means 'is less preferable', so $f_i \prec f_i^*$ means that f_i is less preferable that f_i^* for objective i. I here prove that for utility functions satisfying this statement, a solution that uniquely maximises the utility function of model predictions will be a solution on the Pareto front according to the model predictions.

For a proof by contradiction assume that x is the solution that uniquely amongst the set of doses D maximises U satisfying the above statement but does not lie on the Pareto front of elements of D according to model predictions. This implies that there exists a solution $y \in D$ s.t.

$$f_i(x) \preccurlyeq f_i(y) \, for \, all \, i \in [1,n]$$

And

$$f_i(x) \prec f_i(y) \, for \, all \, i \in [1,n]$$

Then, as

$$for \, all \, i \in [1,n], f_i \prec f_i^* \iff U(f_i, \dots, f_i, \dots, f_n) \prec U(f_i, \dots, f_i^*, \dots, f_n)$$

We have that

$$U(f_1(x),\dots,f_i(x),\dots,f_n(x)) \prec U(f_i(y),\dots,f_i(y),\dots,f_n(y))$$

And so x does not uniquely maximise U amongst set D, a contradiction.

Thus all 'optimal doses' that were calculated within the chapters of this work must be Pareto optimal according to model predictions and data. Note that this does not guarantee that such doses were Pareto optimal for the 'true' efficacy and toxicity probabilities. Where predicted efficacy/toxicity are not accurate, the predicted optimal dose might not lie on the true Pareto front.

Appendix C: Utility Functions

Throughout this work I have suggested multiple 'utility functions' that could be used to choose optimal vaccine dose. These allow for a quantitative definition of a dose being 'optimal', particularly in the multi-objective optimisation setting. Given that one of the key recommendations based on chapter 4 is that a utility function should be chosen a-priori of a clinical trial, it seems relevant to provide some examples of potential utility functions.

A.C.1. Threshold Functions

A threshold utility function involves setting threshold bounds on each response or value of interest. These bounds should define whether a certain response or value of interest is acceptable. If all of the responses or values of interest are within their specific bounds for some dose, then that dose would be considered 'acceptable'. Formally this is given by

$$Utility(dose, \bar{\theta}_l, \bar{\theta}_u) = \prod_{r=1}^{R} \delta[\theta_u^r < \pi^r(dose) < \theta_u^r]$$

Where for each response/value of interest r of the R responses/values of interest, $\pi^r(dose)$ is the model predicted probability of that response occurring for some dose or model predicted value at some dose, $\bar{\theta}_l = (\theta_l^1, ...\theta_l^R)$ is the 1xR vector of lower bounds for the R responses, and $\bar{\theta}_u = (\theta_u^1, ...\theta_u^R)$ is the 1xR vector of upper bounds for the R responses. $\delta(a < b < c)$ is the Kronecker delta function which equals 1 if $a < b < c$ and otherwise equals 0.

For each response, statistical modelling or analysis is used to either predict a maximum likelihood estimate of or posterior distribution description of each response for each dose.

If a maximum likelihood estimate is used, then any dose for which the maximum likelihood estimate would be considered 'acceptable' is also considered acceptable., see figure A.C.1.1. An optimal dose is then chosen from this set of acceptable doses, usually by choosing either the smallest or largest such dose for the acceptable set, or by choosing the dose within the acceptable set that maximises

some secondary utility function (for example maximising some probability of efficacy). See [23] for an example of this.
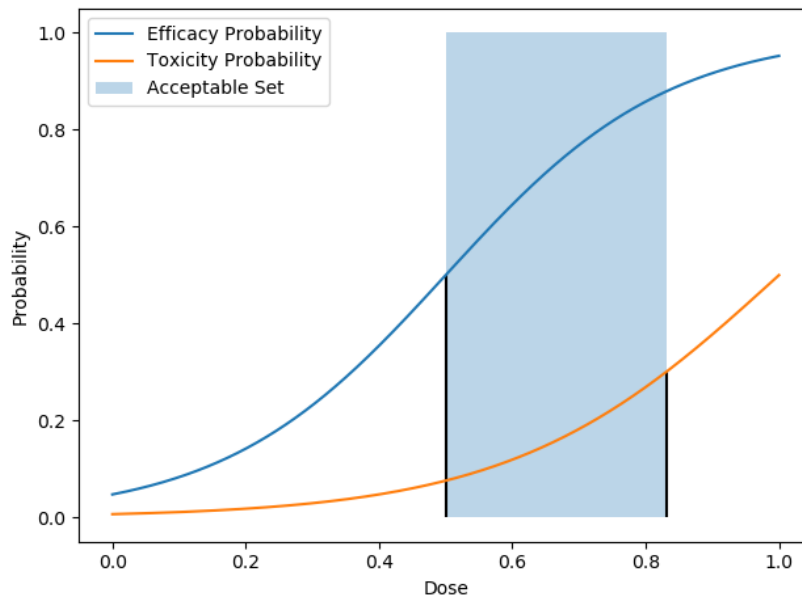
Figure A.C.1.1. Acceptable set for P(e)>0.5 and P(t)<0.3.

If a posterior distribution of model parameters is used rather than calculating the maximum likelihood or maximum a-posteriori estimates or parameters; then instead a Bayesian decision protocol [29,30] can be used. The probability that a dose is 'acceptable' can be found by sampling from the posterior distributions. The optimal dose is then defined either as

- the dose which is predicted to have the maximum posterior probability of being acceptable.
- The dose which maximises some other function of the response/values of interest whilst having the posterior probability of being acceptable that is greater than some threshold.

See [31] for a further description.

Note that one theoretical pitfall with using threshold utility functions is that there may be discontinuity in the dose-utility function [figure A.C.1.2]. This is where a relatively small change in dose leads to a large change in utility. This also means that a tiny

change in model parameters could lead to the same dose being predicted as optimal or unacceptable. It is possible that discontinuity in the dose-utility function or model predicted utility for certain doses being sensitive to small parameter value changes may not be practically relevant for dose optimisation, but I note this here for consideration. Further, sampling from a posterior distribution of model parameters may be reasonable to reduce this parameter sensitivity.
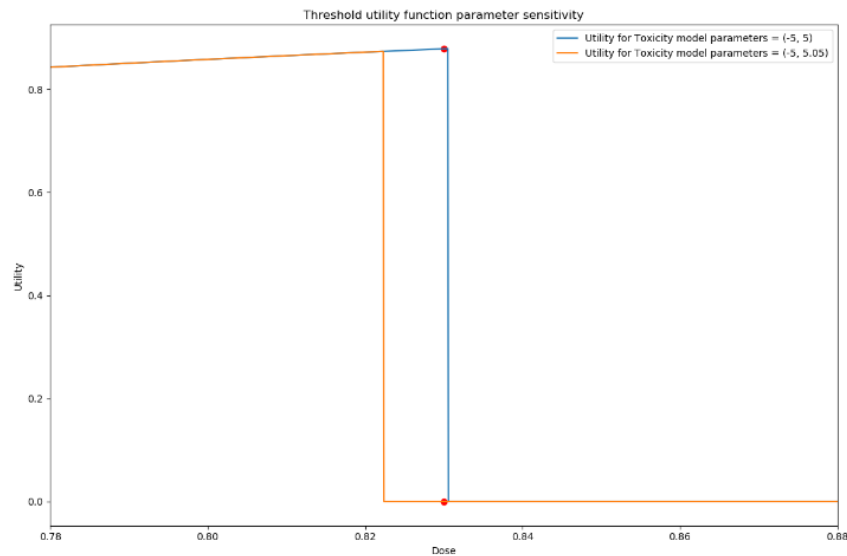
Figure A.C.1.2. A small change in parameter values can lead to a large change in utility for a dose when using a threshold utility function. Red dots show utility for the same dose for two models of toxicity with small parameter differences. One is nearly optimal, one has the worst possible utility.

A.C.2. Scaled Threshold Functions

Takahashi and Suzuki recommend a utility function for determining optimal dose for biologic agents [21]. This function uses thresholds similar to the above functions but penalises doses for being outside of these bounds rather than deeming them as entirely unacceptable. The utility function is given by

$$Utility(dose, \theta^T, \theta^E) = w_t \delta[\pi^T(dose) > \theta^T](\pi^T(dose) - \theta^T)^2 + w_E \delta[\pi^E(dose) < \theta^E](\pi^E(dose) - \theta^E)^2$$

Where w_t, w_e are the relative importance of ensuring toxicity and efficacy are within the bounds, $\pi^T(dose)$, $\pi^E(dose)$ are the model predicted probabilities of efficacy and toxicity for some dose, θ^T is the threshold for maximum acceptable probability of toxicity and θ^E is the threshold for minimum acceptable probability of efficacy.

Note that for Takahashi and Suzuki's utility function a lower value of $Utility(dose)$ means that a dose is more optimal. The best possible value for $Utility(dose)$ is 0. The authors of this function estimate the posterior probabilities $\pi^T(dose)$ and $\pi^E(dose)$ and choose as the optimal dose is the dose that minimises the expected value of $Utility(dose, \theta^T + 0.1, \theta^E - 0.1)$.

This utility function does not have these issues with discontinuity in the dose-utility functions that are observed with normal threshold functions but does mean that it does not improve utility for a dose to have increased efficacy above the threshold value or decreased toxicity below the threshold value.

A.C.3. Scalar aggregation utility functions

In chapter 5 I discussed a simple, interpretable utility function that is calculated as the weighted sum of the probabilities of binary response outcomes being observed. This was an example of a 'scalar aggregation' utility function [32]. The formula for this utility function is

$$Utility(dose) = \sum_{i=1}^{k} w_i * \pi^i(dose)$$

Where $\pi^i(dose)$ is the probability of observing response i for a given dose and w_i is the relative importance of the response i. $w_i > 0$ implies that response i is considered to be beneficial, and $w_i < 0$ implies that response i is considered to be detrimental. Therefore, for this utility function an increased value of $Utility(dose)$ means that a dose is more optimal.

There are two main benefits to this type of utility function. Firstly, the dose that maximises this utility function is guaranteed to be on the Pareto optimal front [32]. Secondly, the values of w_i are interpretable. If $w_{efficacy} = -2 * w_{toxicity}$, then this is easily interpreted as saying that one efficacy response would be required to counterbalance two individuals experiencing vaccine related toxicity. Likewise, if f $w_{cellular\_efficacy} = 3 * w_{humoural\_efficacy}$, then this is easily interpreted as saying that I should consider observing some binary cellular immune response outcome to be

three times more important that observing some binary humoral immune response for a given vaccine. This interpretability could be beneficial, as it may mean that vaccine developers can trust that maximising this utility is important.

A.C.4. Desirability functions

Desirability functions have been used in various bodies of literature as utility functions [33–35]. These allow for any number of potential response outcomes to be weighed against each other in a way that is interpretable. Another useful aspect of these desirability functions is that they are usable regardless of whether response outcomes are probabilities of binary responses or continuous outcome measures. These are detailed here [35] but are explained in the vaccine dosing context below.

A desirability function $d_i(r_i)$ transforms all possible values for response r_i into the range [0,1] where $d_i(r_i)=0$ denotes that value of r_i is not at all desirable and $d_i(r_i)=1$ denotes that r_i is maximally desirable. Each r_i is for the purpose of this work is a function of dose. Three common functions are used for desirability.

The 'bigger-is-better' desirability function, which aims to maximise response r_i:

$$d_i(r_i) = \begin{cases} 0, & \text{for } r_i \leq r_{i,min} \\ (\frac{r_i - r_{i,min}}{r_{i,max} - r_{i,min}})^{m_i}, & \text{for } r_{i,min} \leq r_i \leq r_{i,max}, 0 \leq m_i \\ 1, & \text{for } r_{i,max} \leq r_i \end{cases}$$

The 'smaller-is-better' desirability function, which aims to minimise response r_i:

$$d_i(r_i) = \begin{cases} 1, & \text{for } r_i \leq r_{i,min} \\ (\frac{r_i - r_{i,max}}{r_{i,min} - r_{i,max}})^{m_i}, & \text{for } r_{i,min} \leq r_i \leq c_i, 0 \leq m_i \\ 0, & \text{for } r_{i,max} \leq r_i \end{cases}$$

And the 'target' desirability function, which aims to have r_i close to some target value c_i:

$$d_i(r_i) = \begin{cases} 0, & \text{for } r_i \leq r_{i,min}, r_{i,max} \leq r_i \\ \left(\frac{r_i - r_{i,min}}{c_i - r_{i,min}}\right)^{m_i}, & \text{for } r_{i,min} \leq r_i \leq c_i, 0 \leq m_i \\ \left(\frac{r_i - r_{i,max}}{c_i - r_{i,max}}\right)^{t_i}, & \text{for } c_i \leq r_i \leq r_{i,max}, 0 \leq t_i \end{cases}$$

In all these functions $r_{i,min}$ and $r_{i,max}$ are 'specification limits' which represent the values for which desirability is not altered by r_i decreasing below or increasing above respectively. m_i and t_i determine the steepness of the desirability curve, see figure A.C.4.1.1

(a)

(b)

(c)
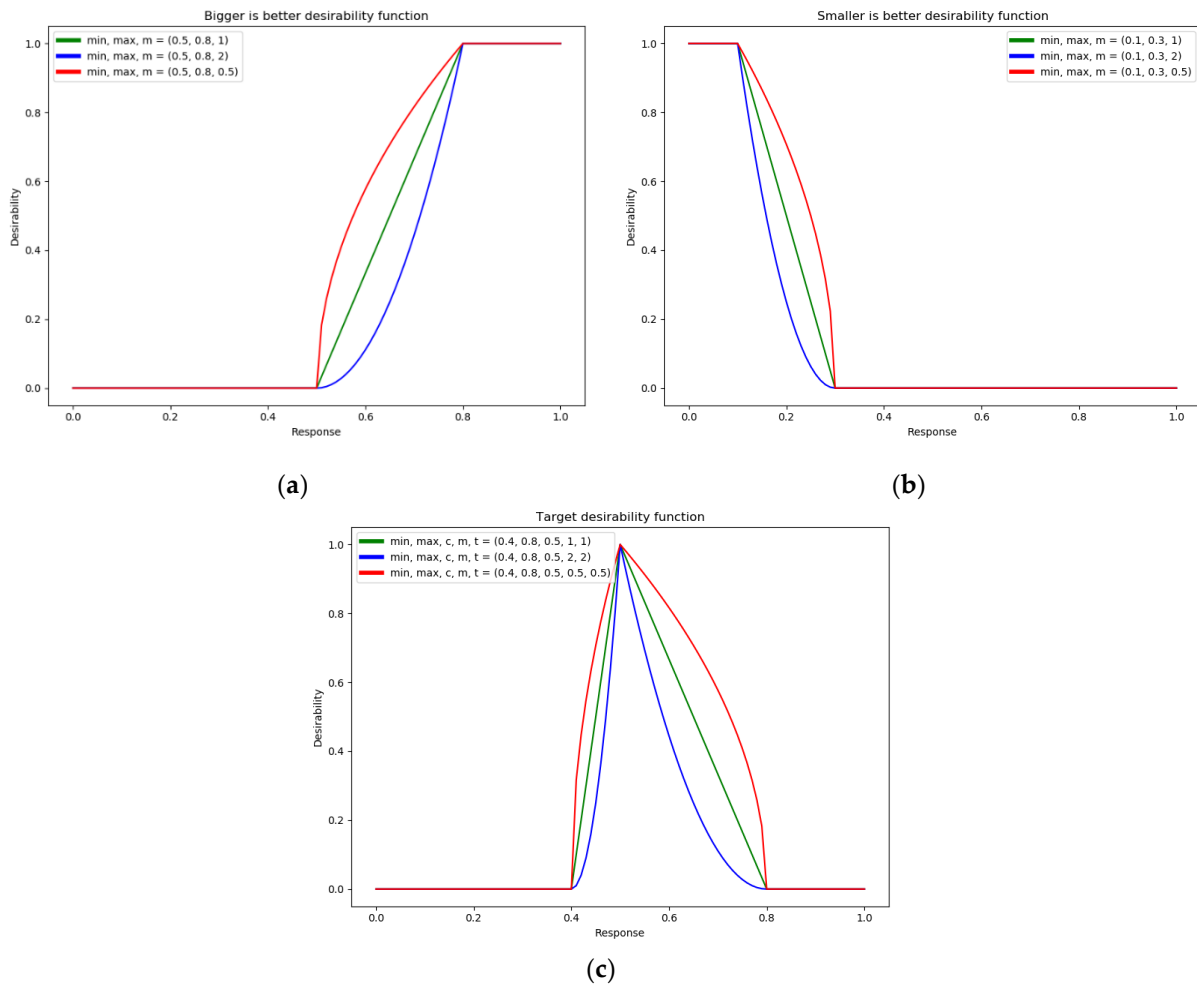
Figure A.C.4.1.1 Visualisation of common desirability functions for three different parameterisations.

An example use case for the 'bigger-is-better' function may be maximising efficacy or some immunogenicity response. An example use case for the 'smaller-is-better' function would be minimising toxicity. An example for using the 'target' desirability function could relate the practicalities of administering a specific dose, where accurate measurements become impossible below a specific dose, and above a specific dose the volume of vaccine cannot be given all at once.

For the 'bigger-is-better' and 'smaller-is-better' desirability functions values for m_i can be respectively calculated as

$$m_i = log(\frac{r_{50} - r_{i,min}}{r_{i,max} - r_{i,min}}) \div log(0.5)$$

And

$$m_i = log(\frac{r_{50} - r_{i,max}}{r_{i,min} - r_{i,max}}) \div log(0.5)$$

Where r_{50} is between $r_{i,min}$ and $r_{i,max}$ and is the response that would be deemed 50% desirable.

These response specific desirability functions can then be combined to form "overall desirability functions". These can be unweighted or weighted.

A.C.4.1. Unweighted Overall Desirability

Harrington suggested the first overall desirability function in 1965. This is given by the geometric means of the k response specific desirability functions, as below

$$Utility(Dose) = D(dose) = (d_1 * d_2 * \ldots * d_k)^{\frac{1}{k}}$$

Where each d_i is the specific desirability of the ith response elicited by that dose, $d_i(r_i(dose))$. Note that this means that if for any of the i responses the desirability $d_i(r_i(dose)) = 0$, then the overall utility/desirability of that dose $Utility(Dose) = D(dose) = 0$.

414

A.C.4.2. Weighted Overall Desirability

The above overall desirability function is reasonable if all K responses are deemed equally of importance. However, in some cases this may not be the case. For example, it is possible in vaccine dose-finding trials that both maximising cellular and humoral response is deemed important. However, it may be that humoral response is deemed a better correlate of protection than cellular response. In this case, it would be preferable for the desirability function to reflect this.

For this purpose, the 'weighted overall desirability' function can be used.

$$Utility(Dose) = D(dose) = (d_1^{w_1} * d_2^{w_2} * \ldots * d_k^{w_k})^{\frac{1}{\sum w_k}}$$

Each of the K response specific desirability's are raised to the power of some $w_i > 0$, with an increase in the value of w_i reflecting an increase in the importance of the ith response relative to the other responses.

A further discussion on the application of weighted overall desirability as a utility function in medicine is given by [36]. This also discusses how values for w_i can be chosen and gives an alternative method of combining the desirability for different responses which is more similar to the utility function used in chapter 5.

I note that the 'utility contour' utility function used in chapter 6 could be considered an example of a 'weighed overall desirability' utility function.

A.C.4.3. An example of using utility and overall desirability functions

Here I give a narrative example of how overall desirability functions as a utility function for optimal vaccine dose selection. Everything presented in this section is meant as demonstrative and does not use real data nor reflect any discussions had with clinicians.

Imagine that a vaccine developer would like to find the optimal dose for a novel single-administration vaccine. The developer determines that there are three binary response outcomes that are of interest. Two of these, $efficacy_1$ and $efficacy_2$ are binary immunological responses that the developer would like to maximise the

probability of. The third, $toxicity$, is a binary toxicological outcome, say occurrence of a grade 3 adverse event, that the developer would like to minimise the probability of.

Through discussion with regulators and key stake holders, the developer is able to determine that for approval of their vaccine the following must be true:

- The probability of observing $efficacy_1$ must be greater than 50%
- The probability of observing $efficacy_2$ must be greater than 40%
- The probability of observing $toxicity$ must be less than 30%

Therefore

$$efficacy_{1,min} = 0.5(50\%), efficacy_{2,min} = 0.4(40\%), toxicity_{max} = 0.3(50\%)$$

Further discussion shows that the key stakeholders would prefer the probability of observing $toxicity$ to be less than 10%, but that reducing this probability below 10% is no more preferable than reducing it to 10%. With regards to the probabilities of $efficacy_1$ and $efficacy_2$, any increase in probability of efficacy is desirable. Therefore $efficacy_{1,max} = efficacy_{2,max} = 1.0(100\%), toxicity_{min} = 0.1(10\%)$.

In order to calculate $m_{efficacy_1}, m_{efficacy_2}, m_{toxicity}$, key stakeholders are asked to propose a value for each response that would be 50% desirable. The key stakeholders decide that $efficacy_{1,50} = 0.75, efficacy_{2,50} = 0.6, toxicity_{50} = 0.15$ would be reasonable. Therefore, these values can be calculated as above to be $m_{efficacy_1} = 1.00, m_{efficacy_2} = 0.63, m_{toxicity} = 2.41$.

Based on this discussion, the following response specific desirability functions are decided upon.

- $d_{efficacy_1}$ should use the 'bigger-is-better' desirability function with $efficacy_{1,min} = 0.5(50\%), efficacy_{1,max} = 1.0(100\%), m = 1.00$
- $d_{efficacy_2}$ should use the 'bigger-is-better' desirability function with $efficacy_{2,min} = 0.4(40\%), efficacy_{2,max} = 1.0(100\%), m = 0.63$

- $d_{toxicity}$ should use the 'smaller-is-better' desirability function with $toxicity_{min} = 0.1(10\%), toxicity_{max} = 0.3(30\%), m = 2.41$

The probability-desirability curves are shown in Figure A.C.4.3.
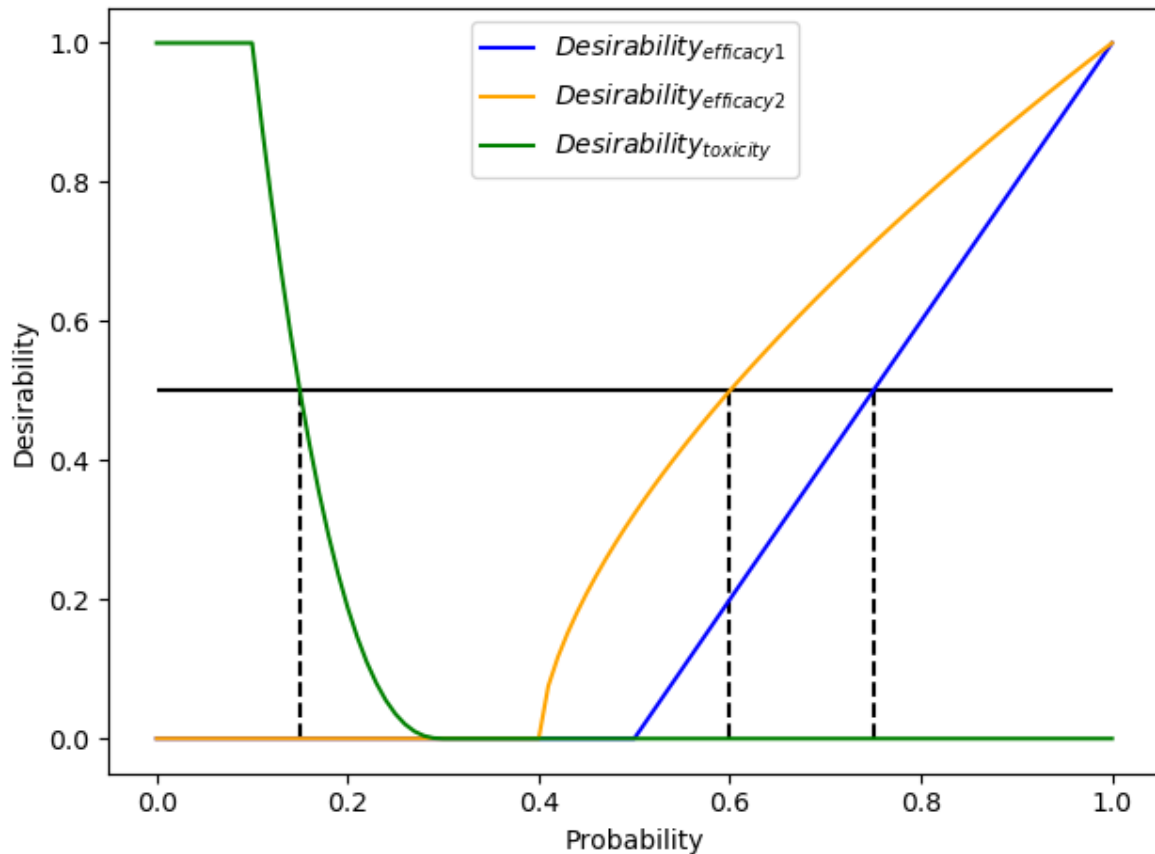
Figure A.C.4.3.1 Probability-Desirability curves for the three responses of interest in the narrative example. The black dotted lines are used to show the 50% desirable values for each response at $efficacy_{1,50} = 0.75, efficacy_{2,50} = 0.6, toxicity_{50} = 0.15$ with the thick black line at desirability = 50%

Through further discussion, the key stakeholders decide to use the unweighted overall desirability function to combine these desirability functions and to define the utility function used to determine the overall 'optimal' dose. One stakeholder believes that minimising $toxicity$ and maximising $efficacy_1$ is more important than maximising $efficacy_2$, and would also like to know which dose would have been selected for a weighted overall desirability function with $w_{efficacy_1} = 4, w_{efficacy_2} = 1, w_{toxicity} = 2$.

417

The dosing domain is decided upon (here I use the range [0,1] as discussed in chapter 6), and a clinical trial is then conducted. Data are gathered and models for dose-$efficacy_1$, dose-$efficacy_2$ and dose-$toxicity$ are found using maximum-likelihood estimates of the parameters. The dose-response curves and resulting response specific dose-desirability curves are shown in figure A.C.4.3.2.
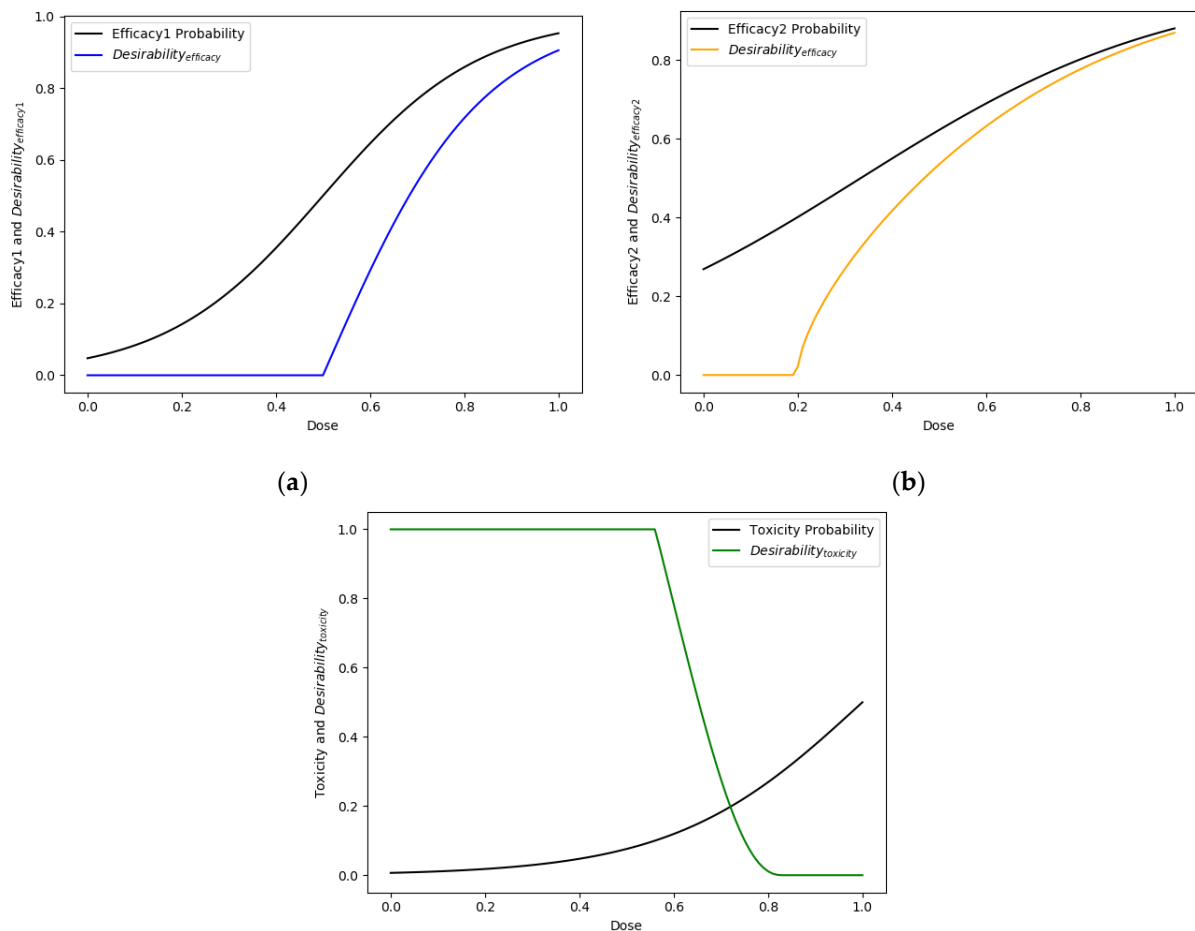
Figure A.C.4.3.2 Predicted dose-response and resulting dose-desirability curves for the three responses of interest in the narrative example. Black lines represent model-predicted dose response. Coloured lines present model-predicted dose-desirability. Saturating functions of dose response for all three responses were used here for simplicity, but these techniques generalise to any dose-response function.

These response specific dose-desirability curves are then combined to find the overall dose-desirability. The dose-desirability curves for both the unweighted overall desirability and weighted desirability are shown in figure A.C.4.3.3
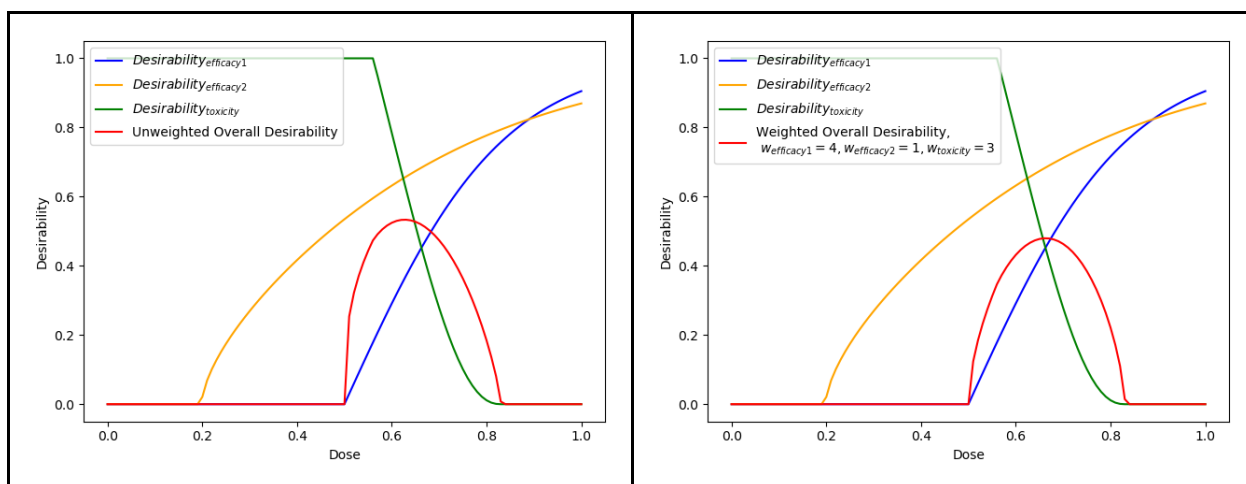
Figure A.C.4.3.3. Overall dose-desirability for the narrative example using an unweighted (left) and weighted (right) overall dose desirability function.

The dose that should then be predicted as optimal using the unweighted overall desirability as the utility function would be 0.63. This would have model-predicted response probabilities 68.6%, 70.9%, and 13.6% with response specific desirabilities of 0.66, 0.62, and 0.53 for $eficacy_1, efficacy_2, toxicity$ respectively.

If the weighted overall desirability function with $w_{efficacy_1} = 4, w_{efficacy_2} = 1, w_{toxicity} = 2$ had been used as the utility function for selecting dose, the dose that would be predicted to be optimal would be 0.66. This would have model-predicted response probabilities 72.3%, 72.7%, and 15.4% with response specific desirabilities of 0.72, 0.73 and 0.48 for $efficacy_1, efficacy_2, toxicity$ respectively. This change in the utility function would have suggested choosing a slightly larger dose relative to the unweighted overall desirability function in this case.

We could also consider this example from the perspective of Pareto Optimality, as discussed in Appendix B. Rather than considering specific utility functions, generating Pareto front plots using the model predictions can be useful to show the trade-off between maximising the probability of the two efficacy outcomes and minimising the probability of the toxicity outcome. These are shown in figure A.C.4.3.4. Note that due to the saturating dose-efficacy and dose-toxicity curves all doses lie on the Pareto optimal front, as noted in Appendix B.

419
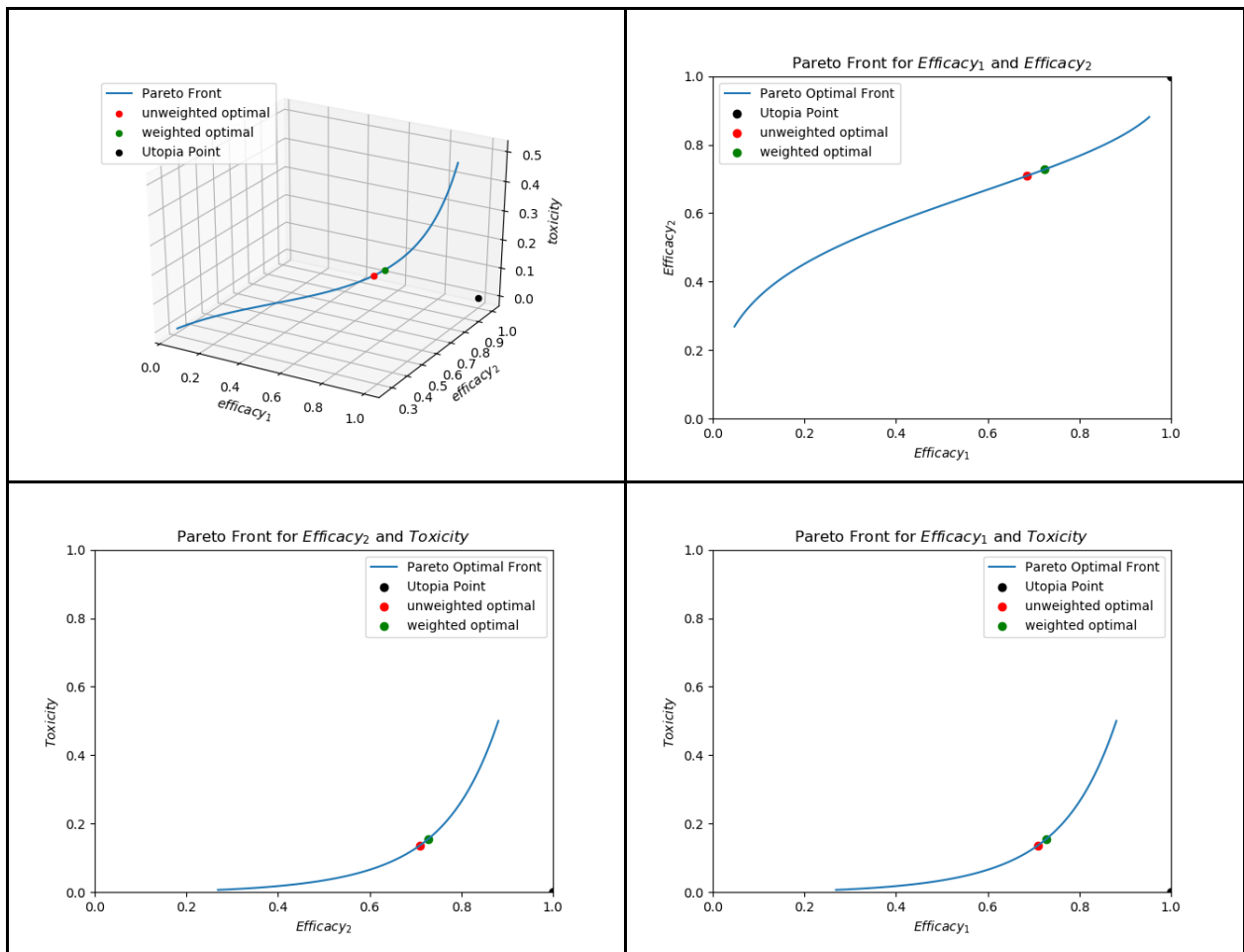
Figure A.C.4.3.4. Pareto fronts for the example. The full Pareto front is shown (top-left) across all three outcome dimensions, with the other plots showing the three 2 dimensional projections of this front. The Pareto front is in blue. The black dot shows the 'utopia point' which is the most optimal point in the outcome space (100% probability of both efficacy outcomes, 0% probability of toxicity outcome), which may not be and in this example indeed is not possible to achieve with any of the possible doses. The red and green dots represent the doses that would be predicted optimal using the unweighted and weighted overall desirability functions discussed above.

Appendix D: Additional supplementary documents

A.D.1: Additional Supplementary documents for paper 2

Supplementary Figures

Response Type: Antibody

Vector Species: B

Host Species: Mouse

Route of Administration: IM

Paper 1269:

Day 56
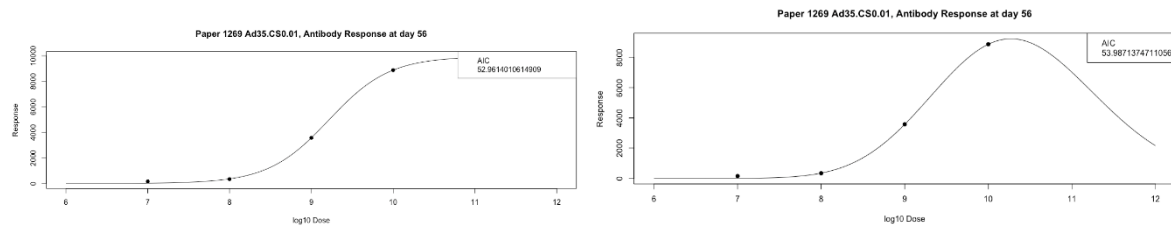
Paper 1492:

Day 14
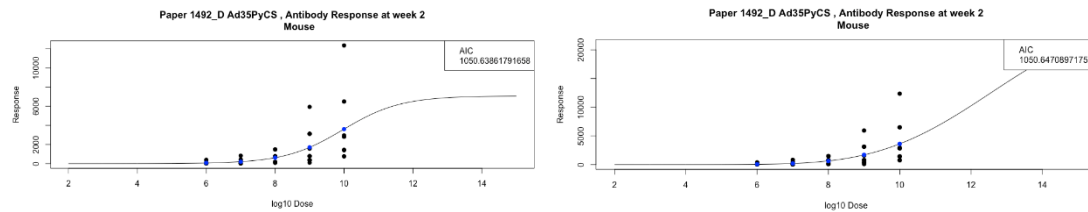
Host Species: Human

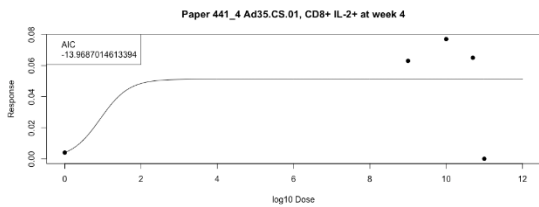Route of Administration: IM

Paper 441:

Day 28

Paper 467:

Day 60

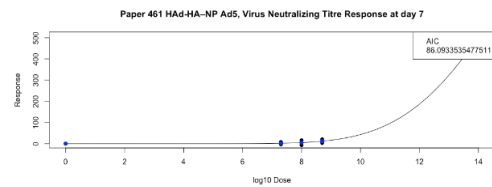Paper 467 Ad-LiGP Ad5, CD8+ Response at day 60

Paper 467 Ad-LiGP Ad5, CD8+ Response at day 60

Paper 633

Day 28

Paper 633 AD35-GRIN/ENV, Antibody Response at day 28

Paper 633 AD35-GRIN/ENV, Antibody Response at day 28
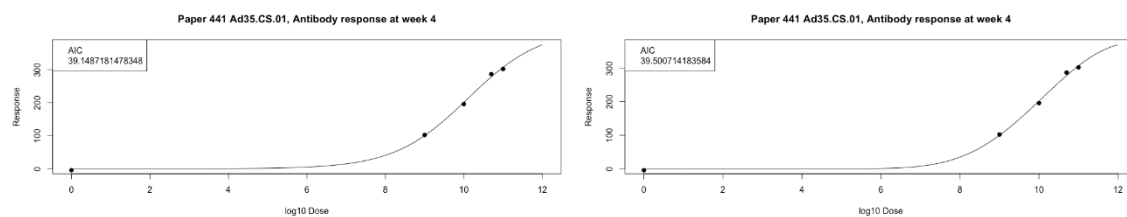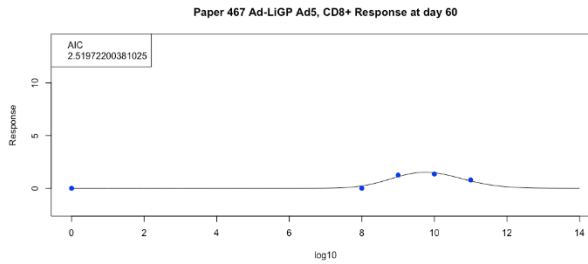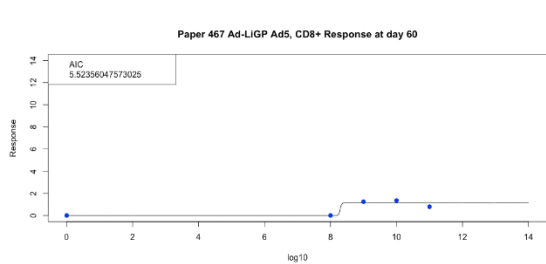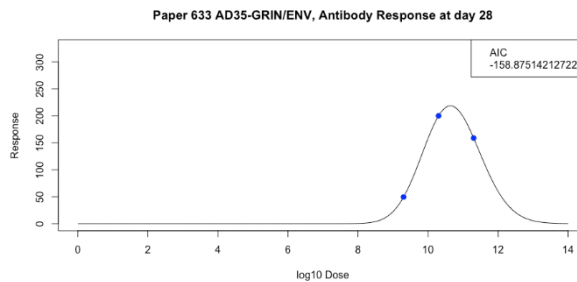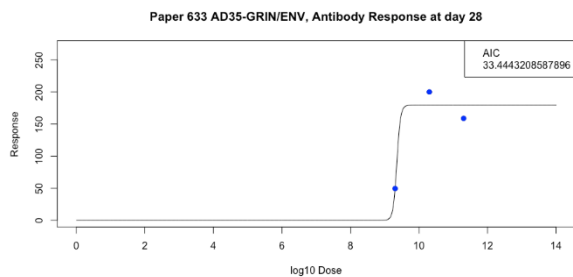
Vector Species: C

Host Species: Rabbit

Route of Administration: IM

Paper 744:

Day 28

Paper 744 rAdV-SFV-E2 (Ad5), Antibody Response at week 4

Paper 744 rAdV-SFV-E2 (Ad5), Antibody Response at week 4
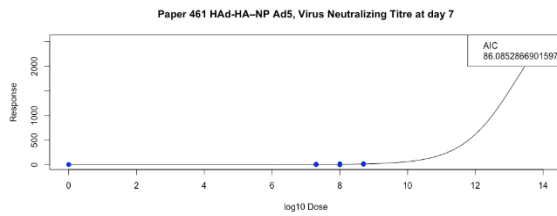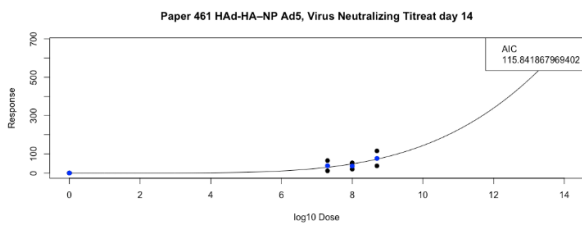
Host Species: Mouse

Route of Administration: IM

Paper 461:

Day 7

Paper 461 HAd-HA–NP Ad5, Antibody Response at day 7

Paper 461 HAd-HA–NP Ad5, Antibody Response at day 7

Day 14

422
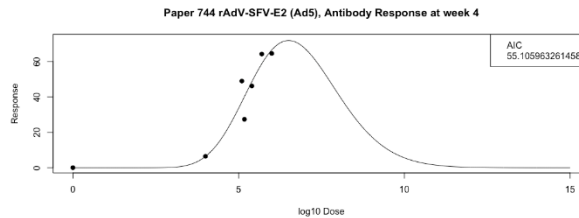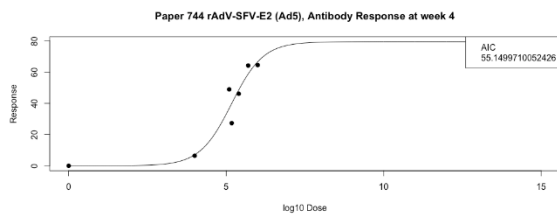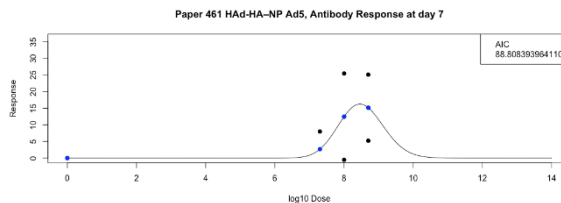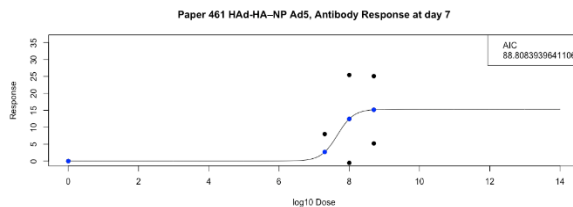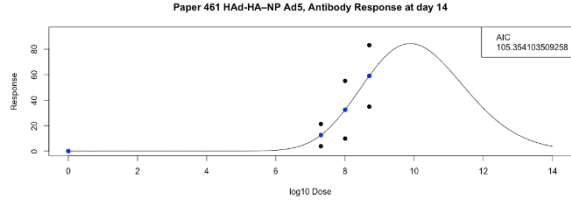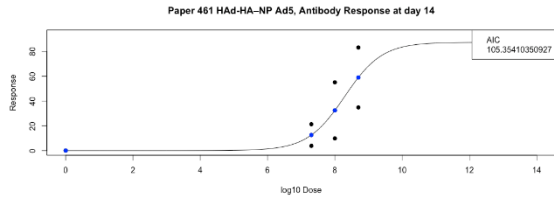
Paper 574:
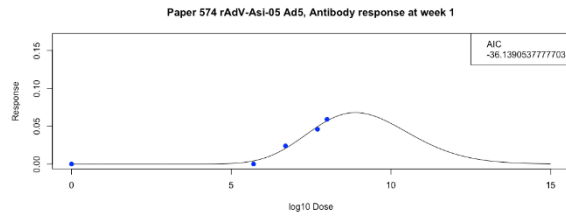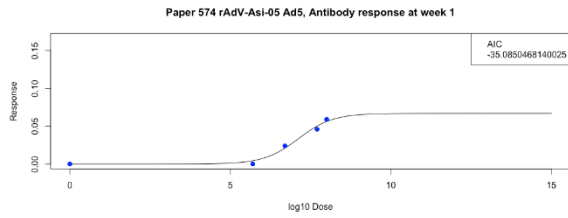
Day 7

Day 21
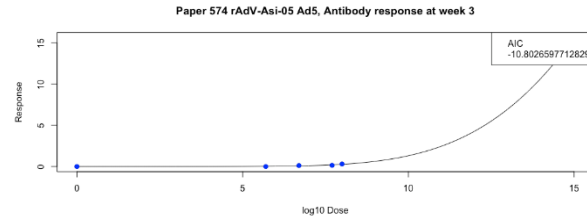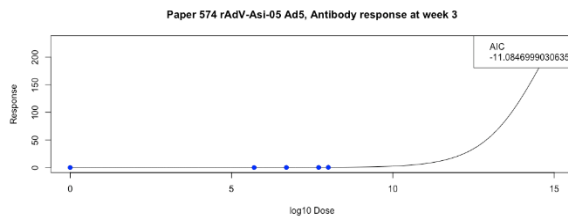
Day 35

Day 49

Day 63

Day 77

423
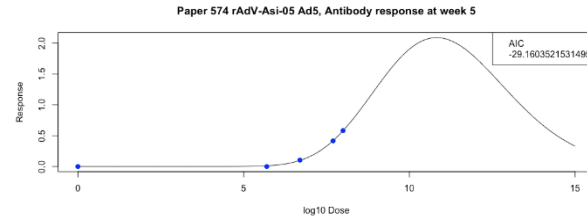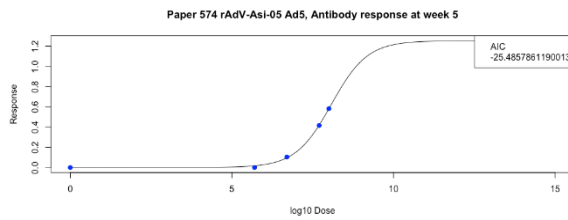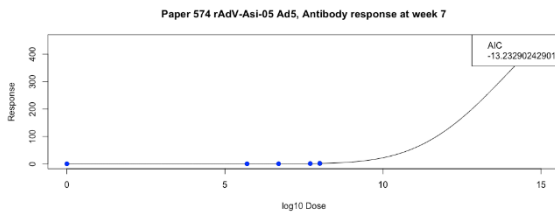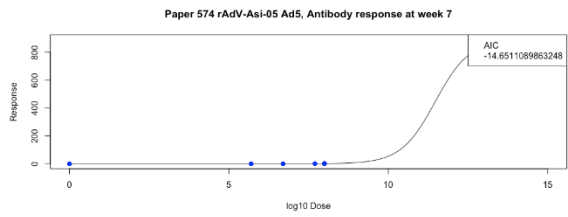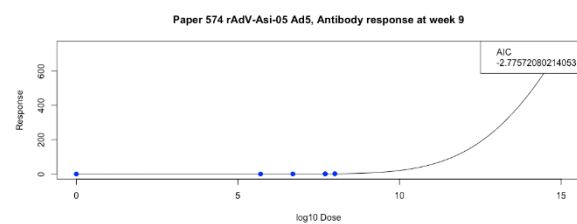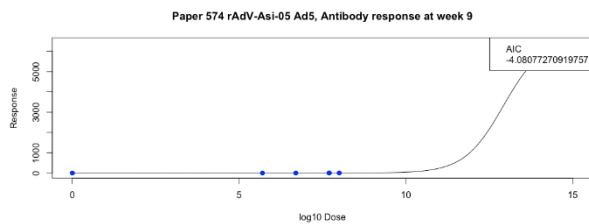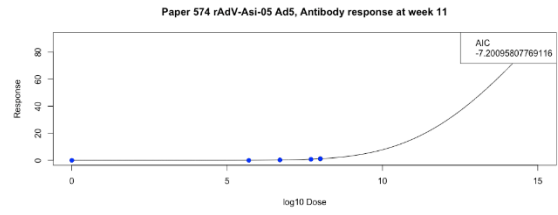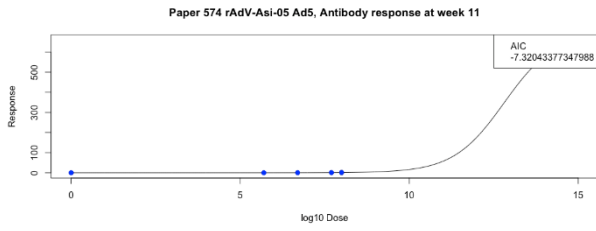
Day 91
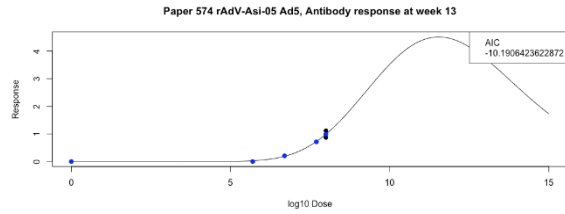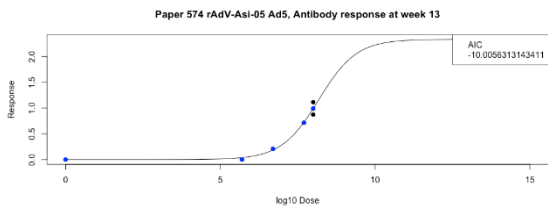
Day 119
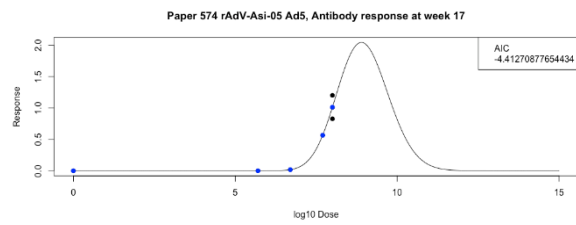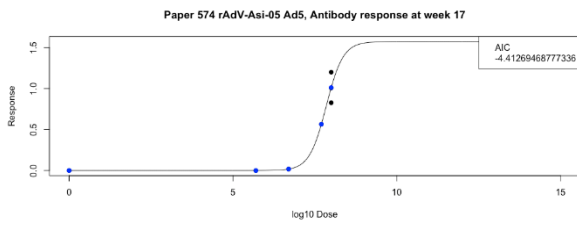
Day 147
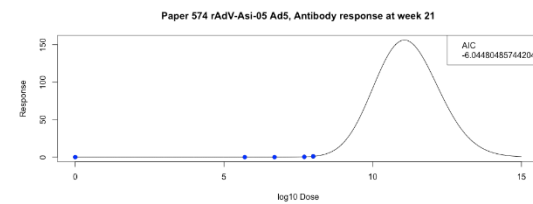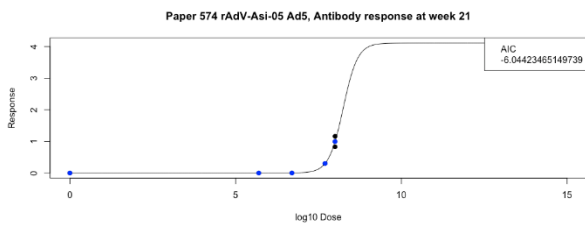
Day 161
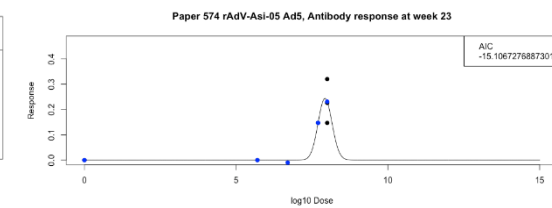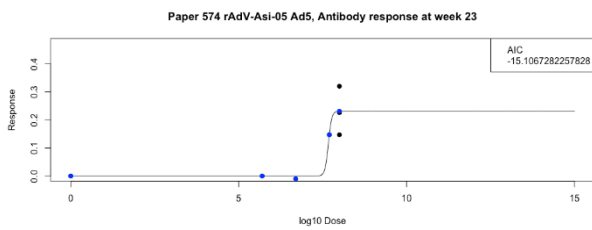
Paper 1492:

Day 14
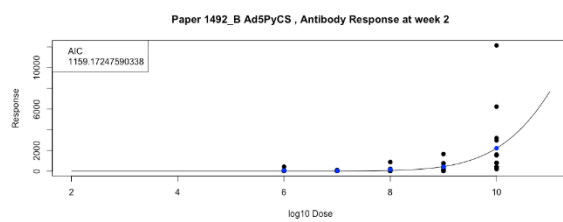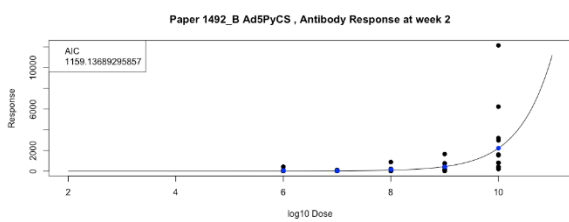
Paper 2531:

Day 10 Ad-gD
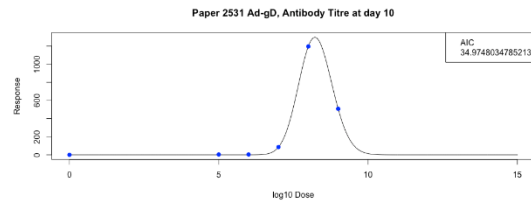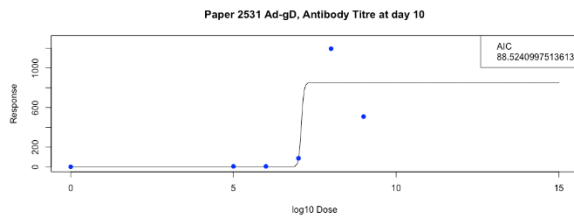
Day 10 Ad-gD-E1A
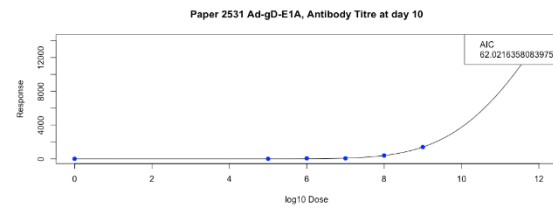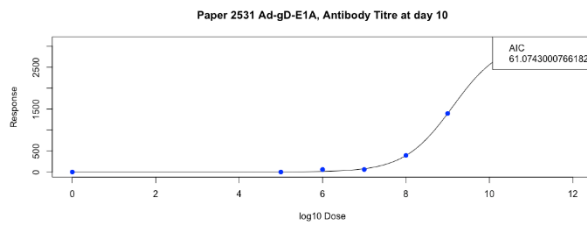
Route of Administration: SQ

Paper 936:

Day 35 - Wildtype
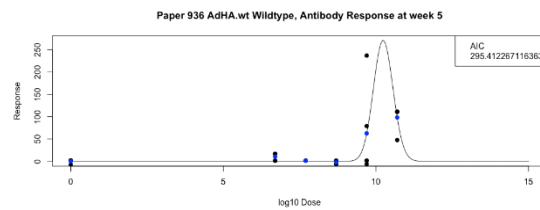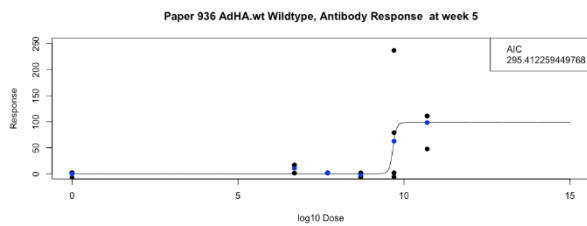
Day 35 - Codon Optimized
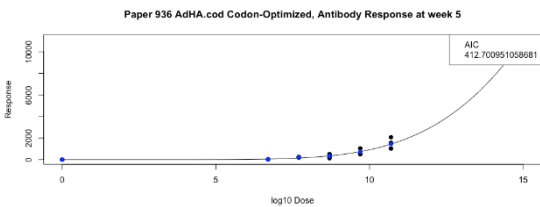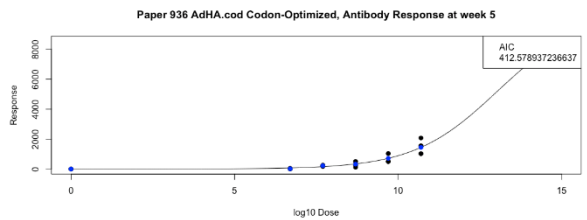
Host Species: Human

Route of Administration: IM

Paper 140

Day 28

Paper 249:

Day 14

425

Paper 249 ChAd3-ZEBOV, Antibody Response at day 14

Day 28

Paper 249 ChAd3-ZEBOV, Antibody Response at day 28

Day 180
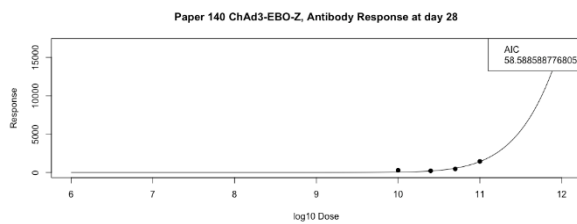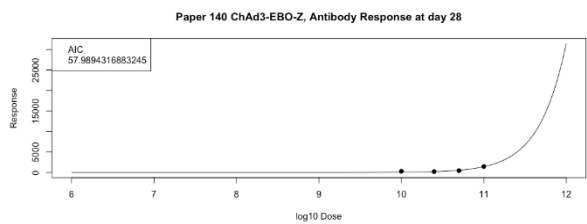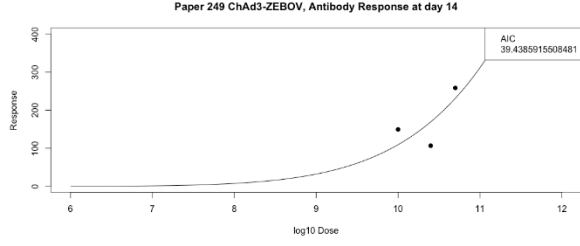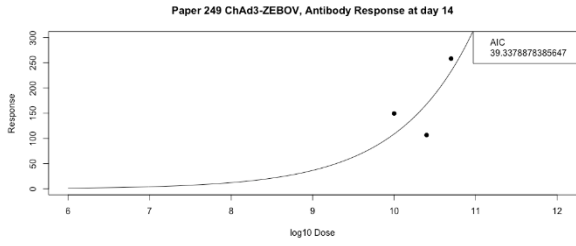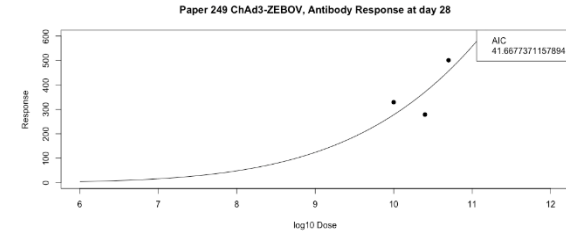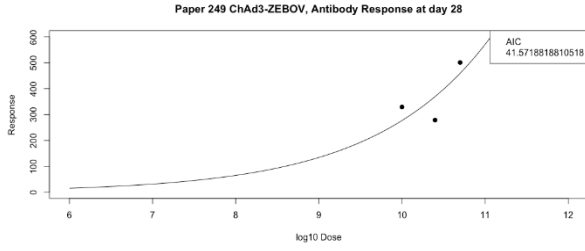
Paper 249 ChAd3-ZEBOV, Antibody Response at day 180
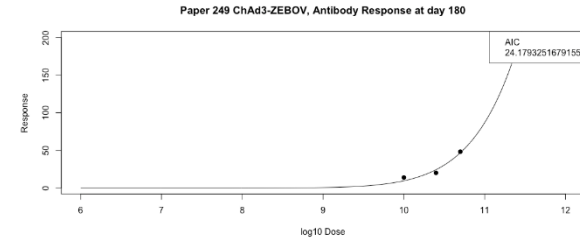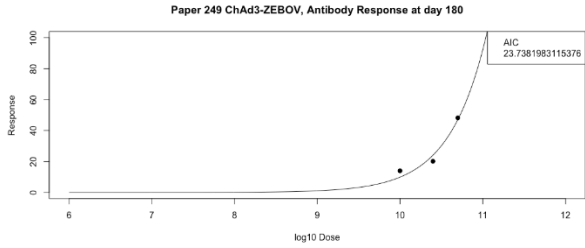
Host Species: Monkey

Route of Administration: IM

Paper 1877:

Day 28

Paper 1877_3B Ad5 Ad5-gag, Antibody Response at week 4

Day 56

Paper 1877_3B Ad5 Ad5-gag, Antibody Response at week 8

Day 84

426
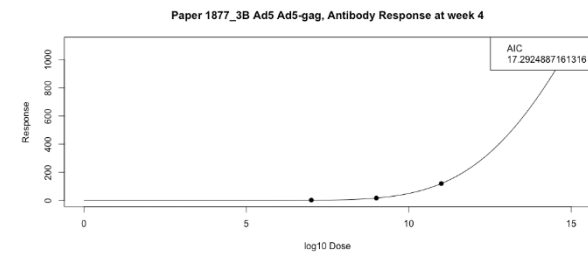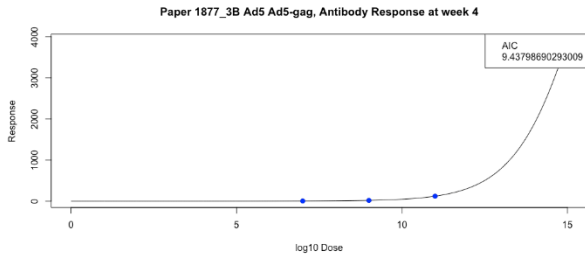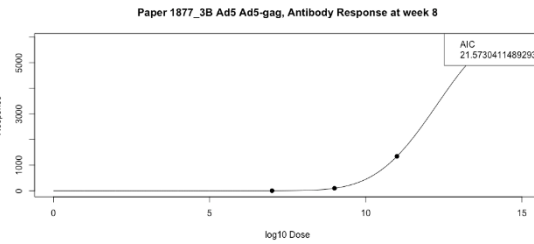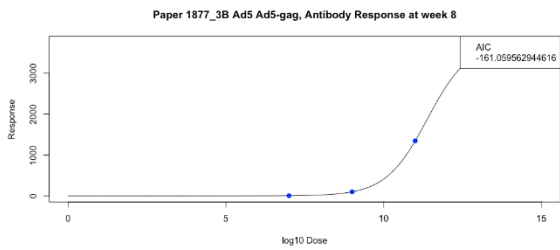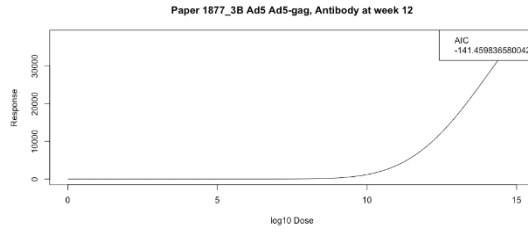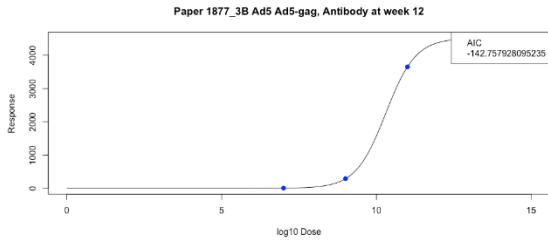
Day 112
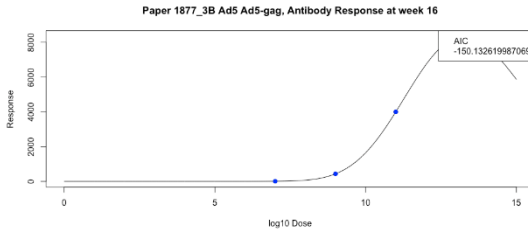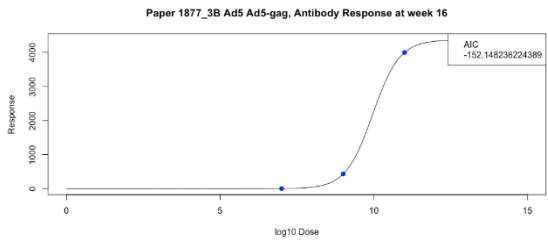
Day 140
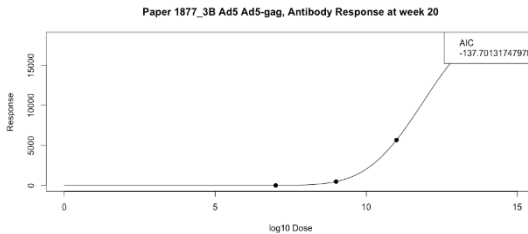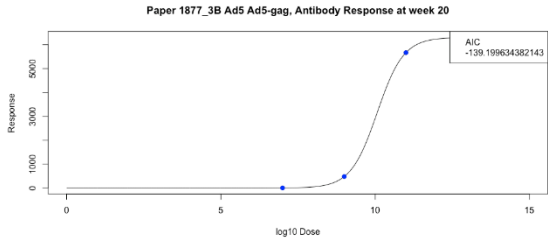
Host Species: Rat

Route of Administration: IM

Paper 2531:

Day 10 Ad-gD
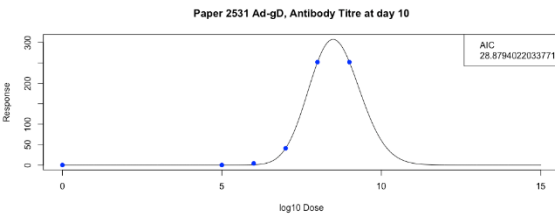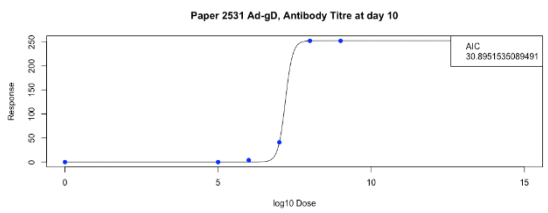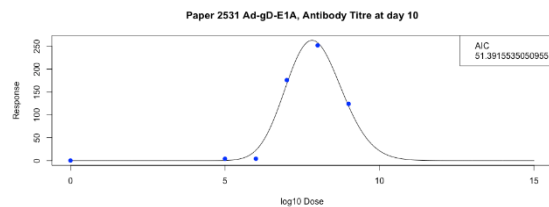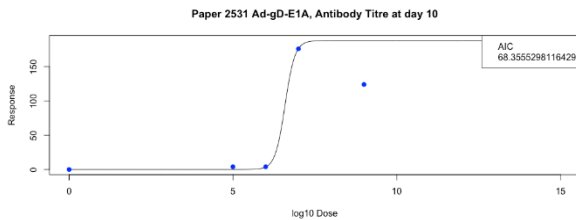
Day 10 Ad-gD-E1A

Vector Species: D

Host Species: Mouse

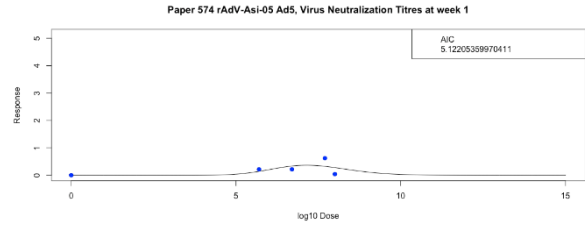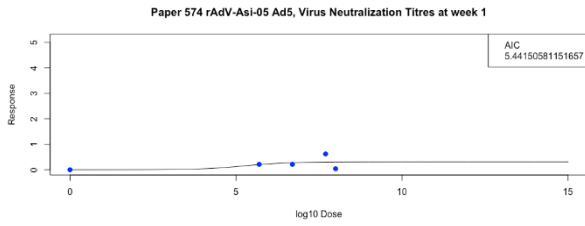Route of Administration: IM

Paper 578:

Day 28 - Ad26 Zaire
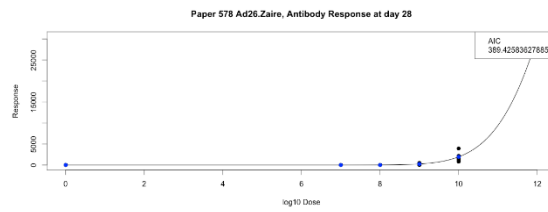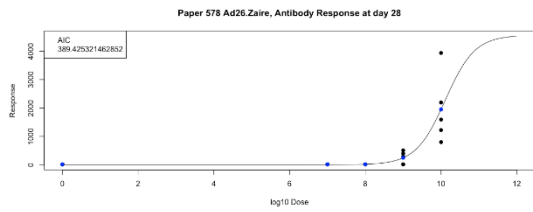
Day 28 - Ad26 Angola
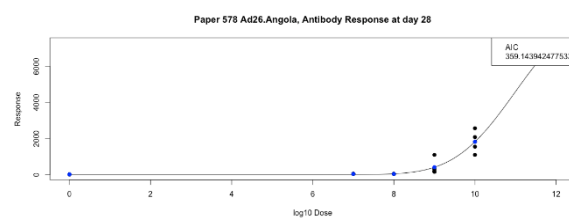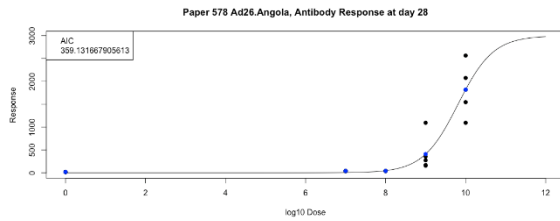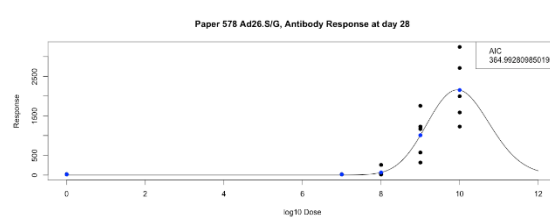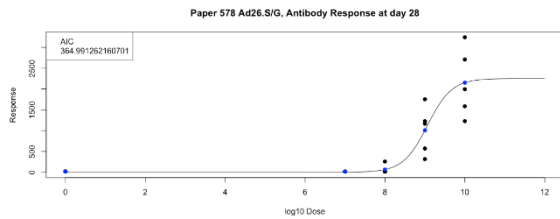
Day 28 - Ad26 S/G
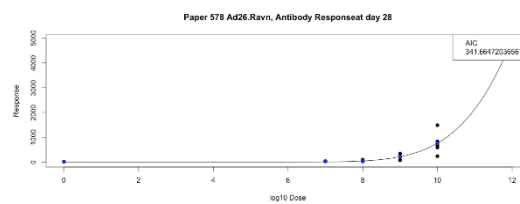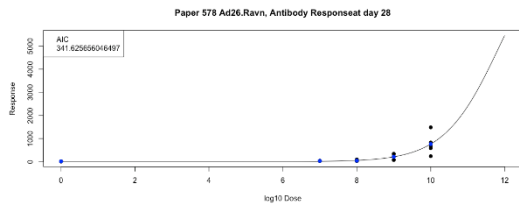
Day 28 - Ad26 Ravn

Day 28 - Ad26 I.C

Host Species: Human

Route of Administration: IM

Paper 594:

Day 14

Paper 594_2 Ad26, Log EnvA ELISA titer at day 14

Paper 594_2 Ad26, Log EnvA ELISA titer at day 14
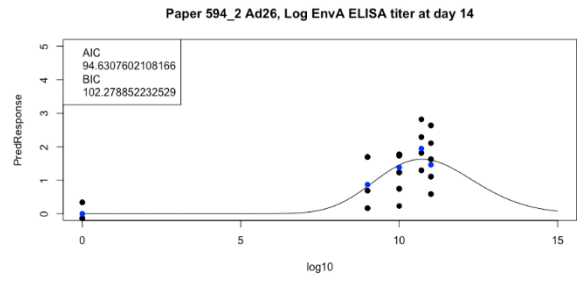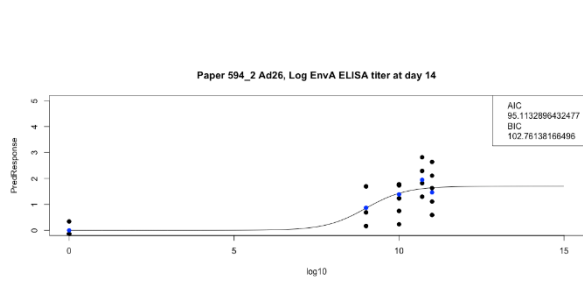
Vector Species: E

Host Species: Mouse

Route of Administration: IM

Paper 1539:

Day 14 - Anti PA

Paper 1539_2A AdC7PA, Antibody Response at day 14

Paper 1539_2A AdC7PA, Antibody Response at day 14

Day 28 - Anti PA

Paper 1539_2A AdC7PA, Antibody Response at day 28

Paper 1539_2A AdC7PA, Antibody Response at day 28

Day 42 - Anti PA

Paper 1539_2A AdC7PA, Antibodie Response at day 42

Paper 1539_2A AdC7PA, Antibody Response at day 42

Day 14 - Anti LT

Paper 1539_2B AdC7PA, Antibodie Response at day 14

Paper 1539_2B AdC7PA, Antibody Response at day 14

Day 28 - Anti LT

429
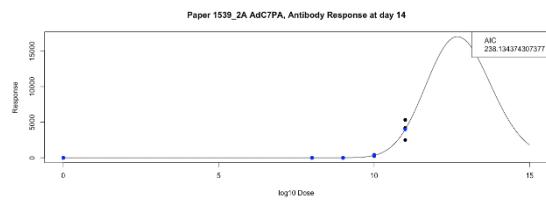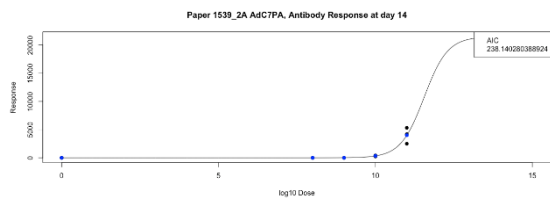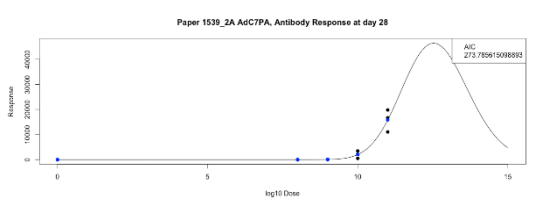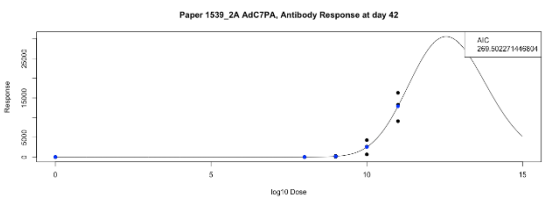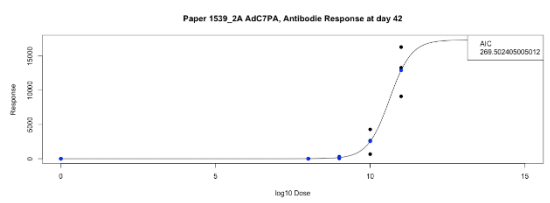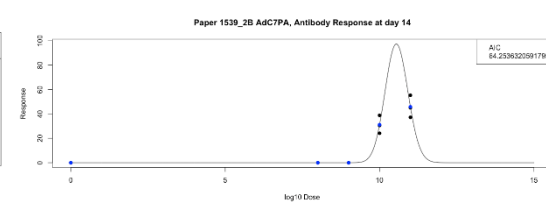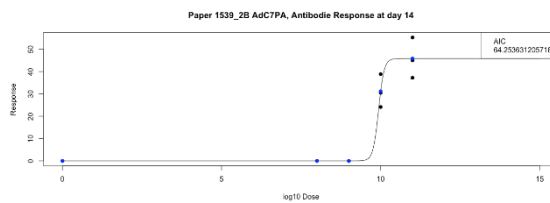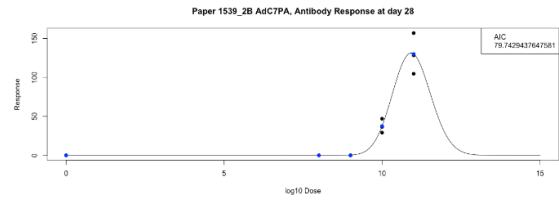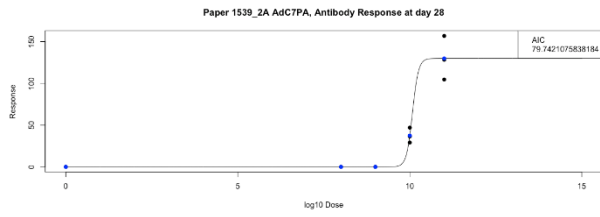
Day 42 - Anti LT
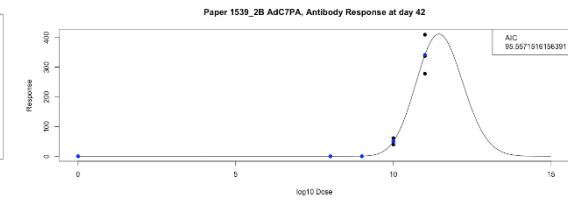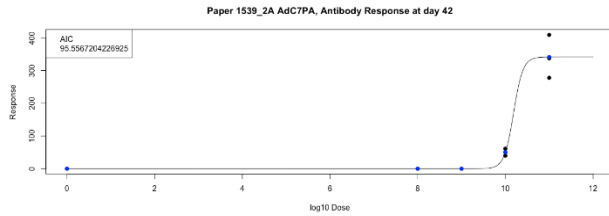
Paper 2919:

Day 20
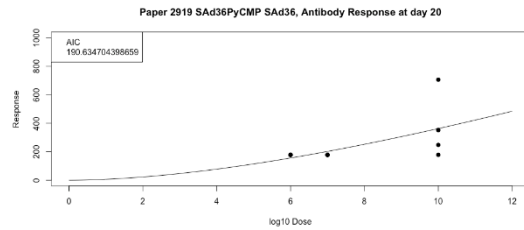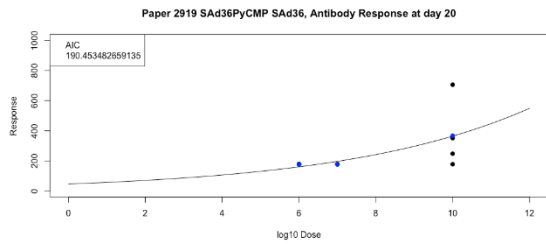
Host Species: Human

Route of Administration: IM

Paper 417:

Day 14
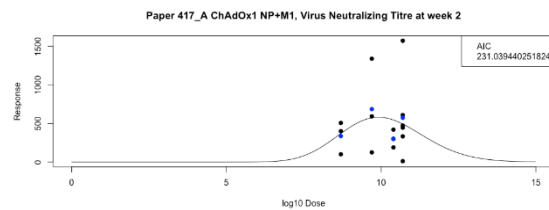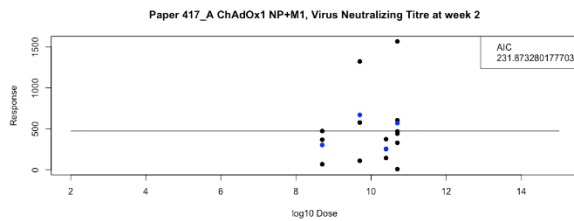
Day 21
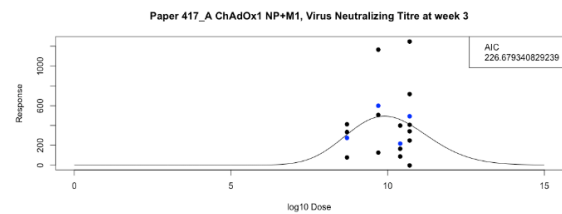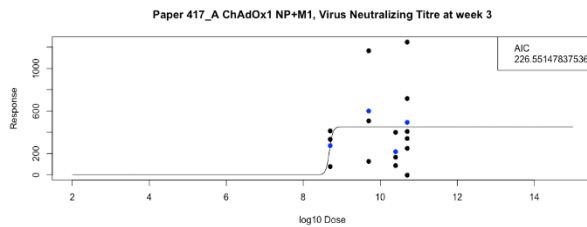
Response Type: T cell

Vector Species: B

Host Species: Mouse

Route of Administration: IM

Paper 578:

Day 28 - Ad35 Zaire
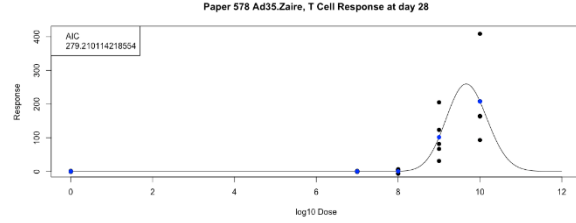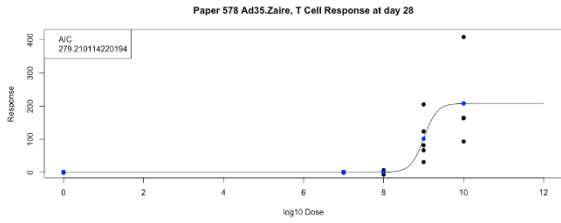
Day 28 - Ad35 Angola
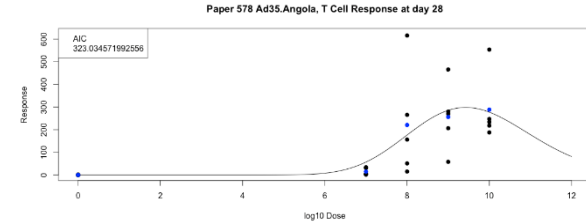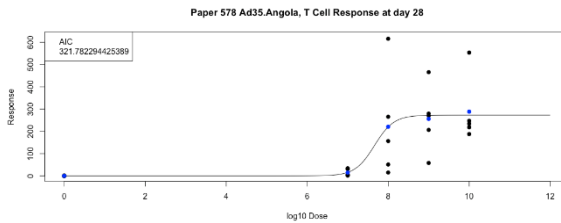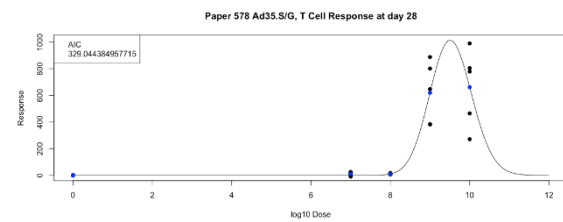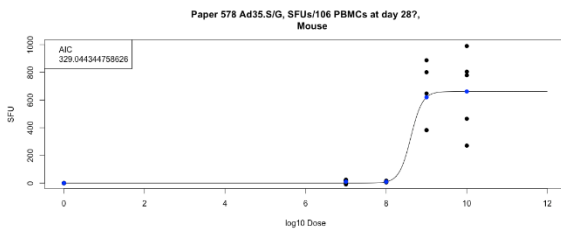
Day 28 - Ad35 S/G
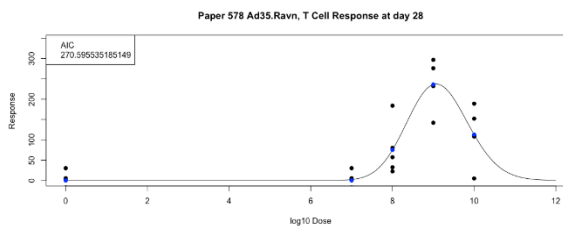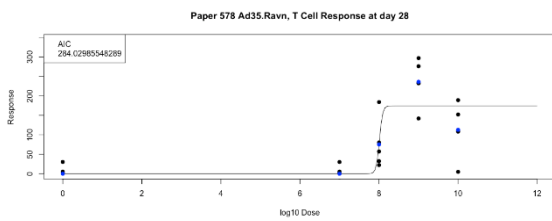
Day 28 - Ad35 Ravn

Day 28 - Ad35 I.C

Paper 1269:

Day 56

431
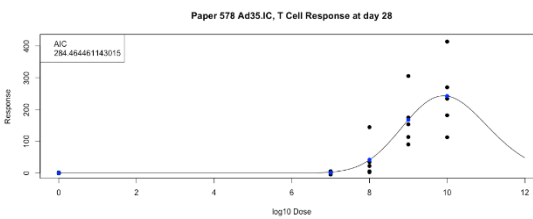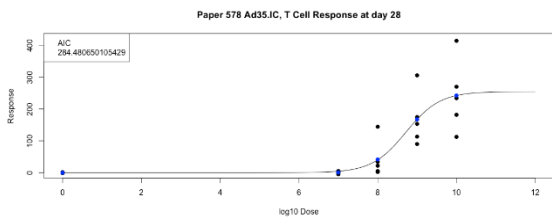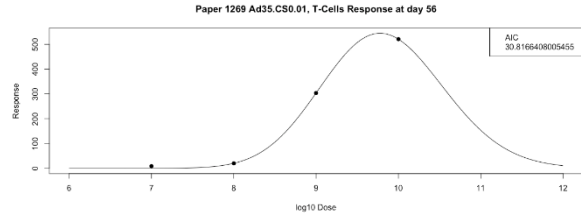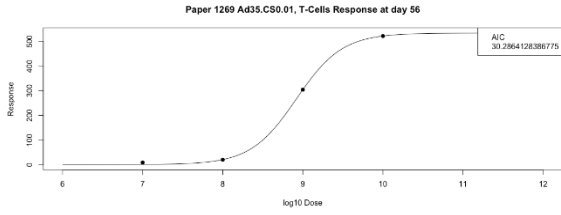
Host Species: Human

Route of Administration: IM

Paper 441:
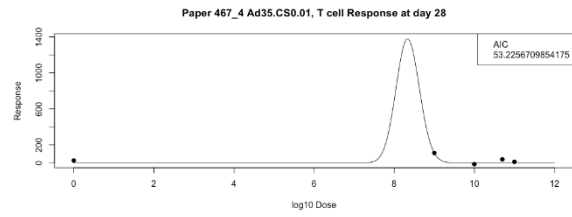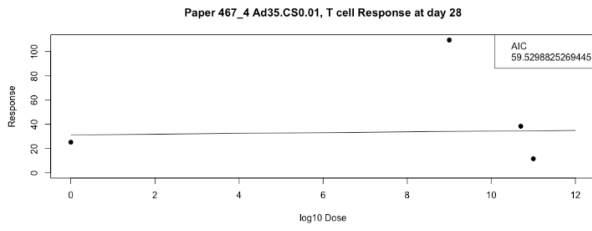
Day 28

Paper 633:

Day 14
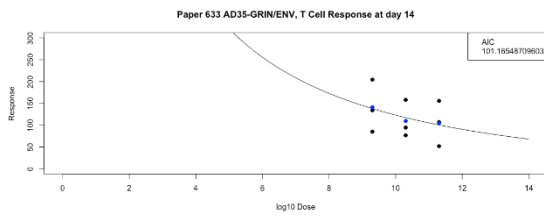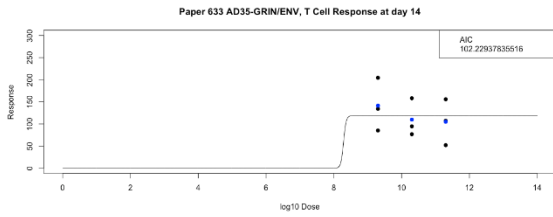
Day 28
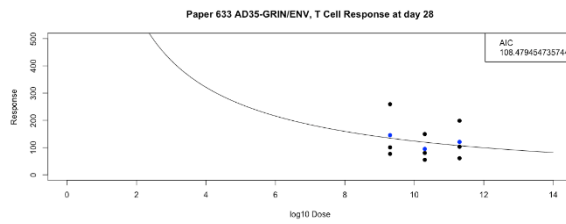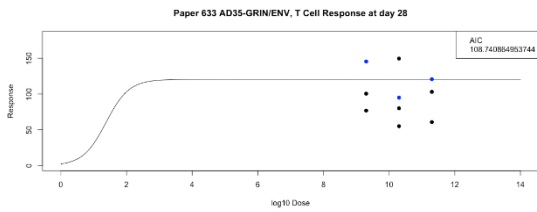
Vector Species: C

Host Species: Mouse

Route of Administration: IM

Paper 2916:

Day 25 - Wildtype pp65

432
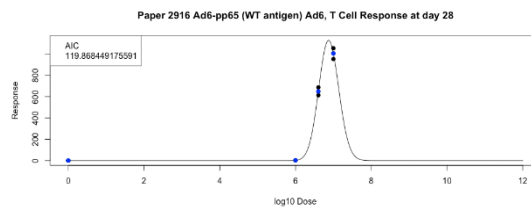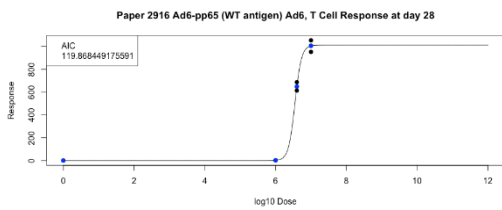
Day 25 - Modified pp65
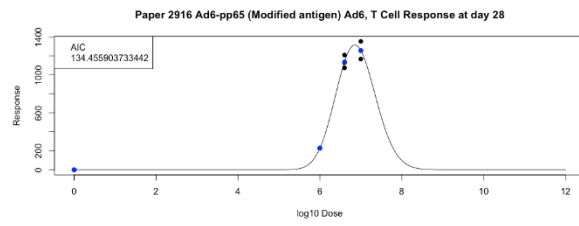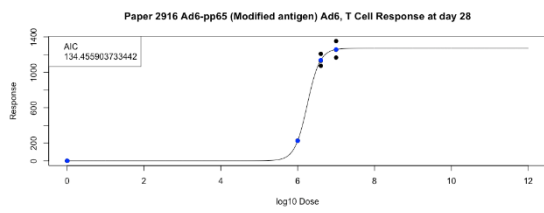
Day 25 - Wildtype Ad-IE1
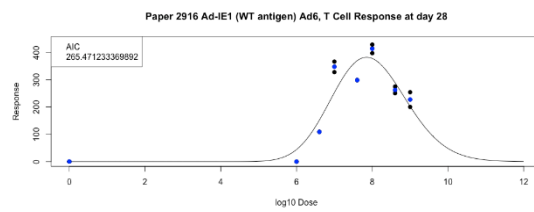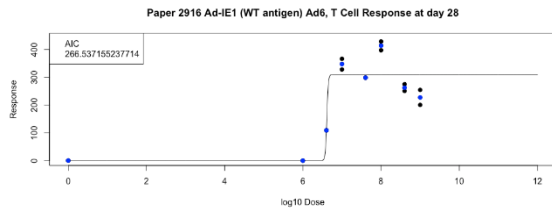
Day 25 - Modified Ad-IE1
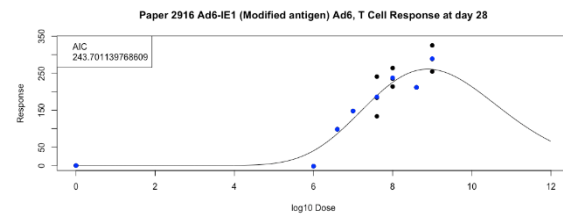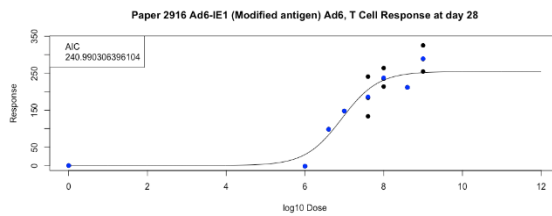
Day 25 - Wildtype Ad-IE2
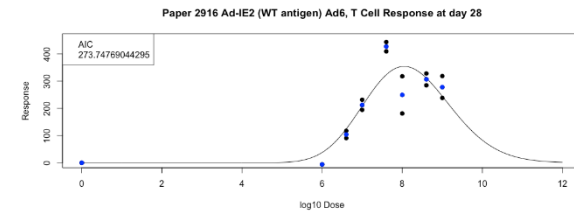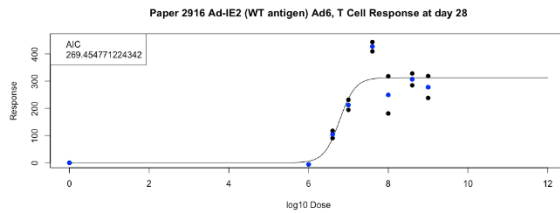
Day 25 - Modified Ad-IE1
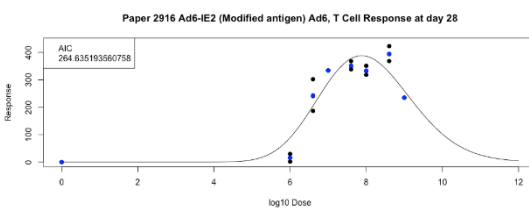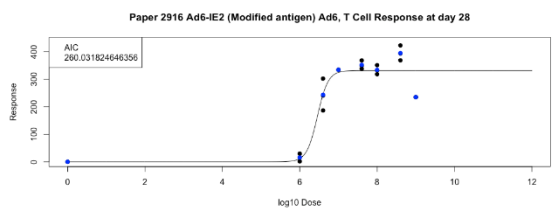
Host Species: Monkey
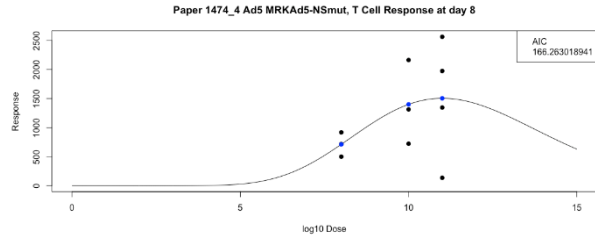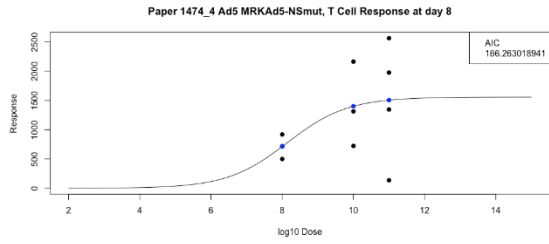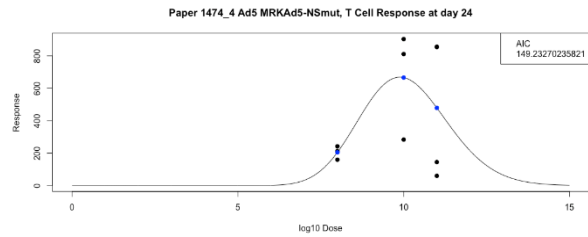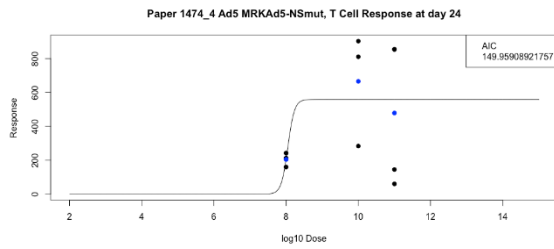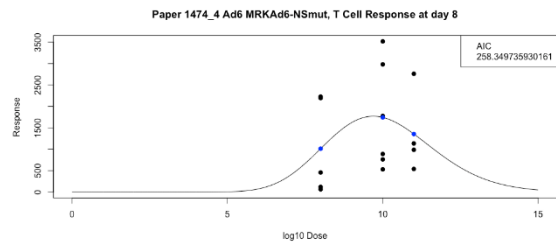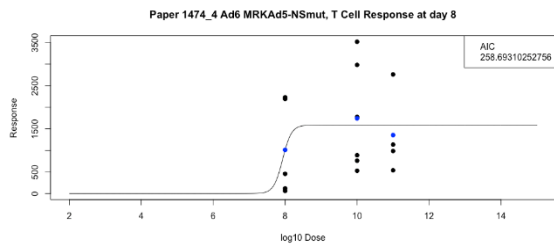
Route of Administration: IM

Paper 1474:

Day 8 - Ad5

433

Paper 1474_4 Ad5 MRKAd5-NSmut, T Cell Response at day 8

Day 24 - Ad5

Paper 1474_4 Ad5 MRKAd5-NSmut, T Cell Response at day 24

Day 8 - Ad6

Paper 1474_4 Ad6 MRKAd5-NSmut, T Cell Response at day 8

Day 24 - Ad6

Paper 1474_4 Ad6 MRKAd6-NSmut, T Cell Response at day 24

Paper 1877:

Day 28

Paper 1877_1 Ad5 Ad5-gag, T Cell IFN+ Response at week 4

Day 168

434
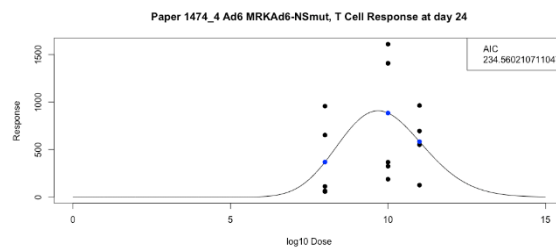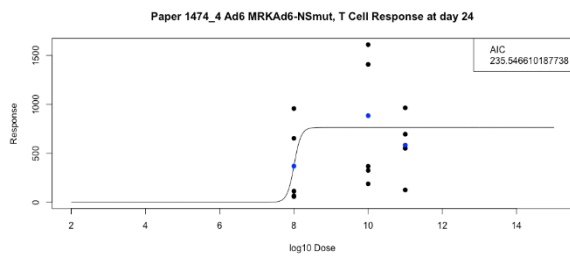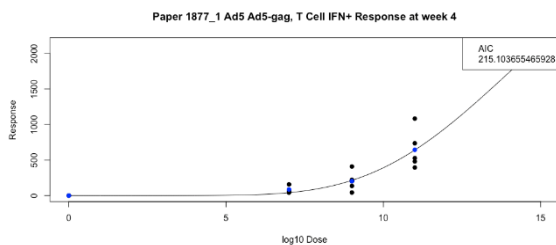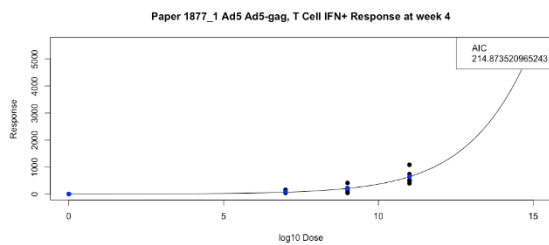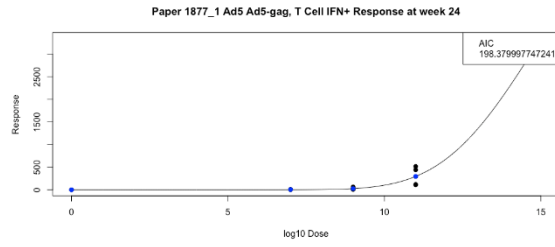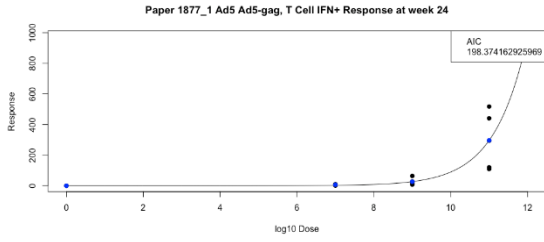
Vector Species: D

Host Species: Mouse

Route of Administration: IM

Paper 578:

Day 28 - Ad26 Zaire
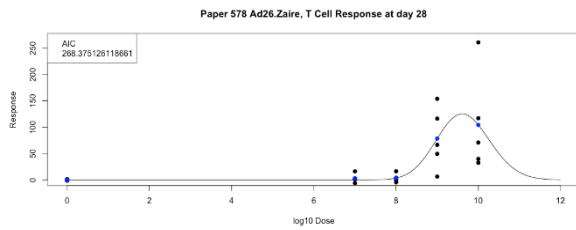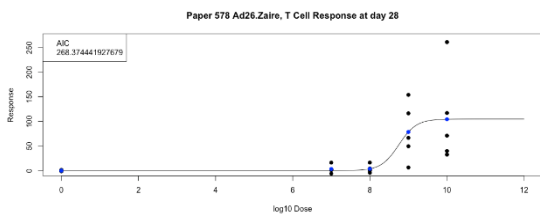
Day 28 - Ad26 Angola
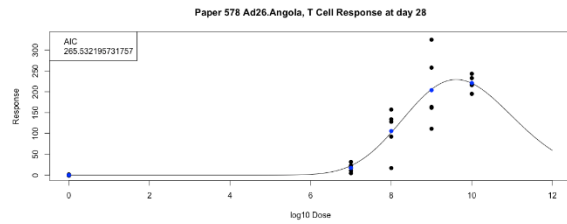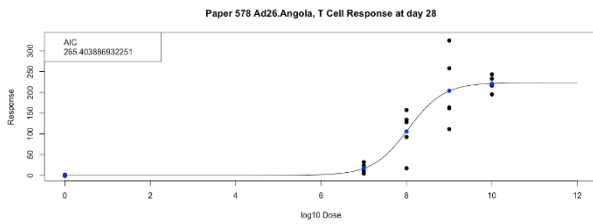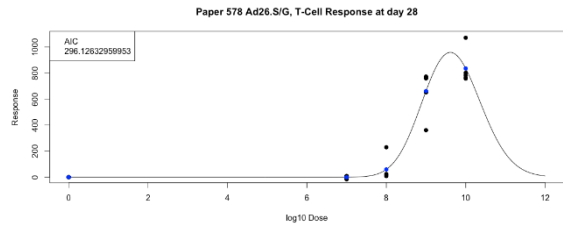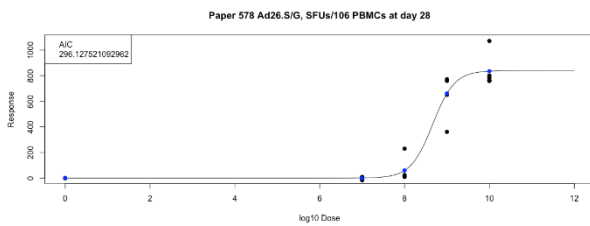
Day 28 - Ad26 S/G
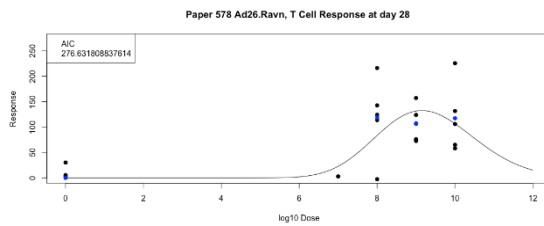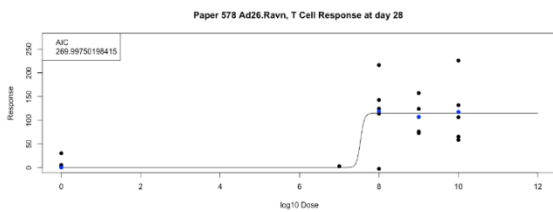
Day 28 - Ad26 Ravn

Day 28 - Ad26 I.C
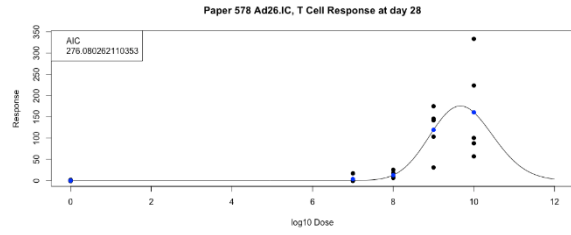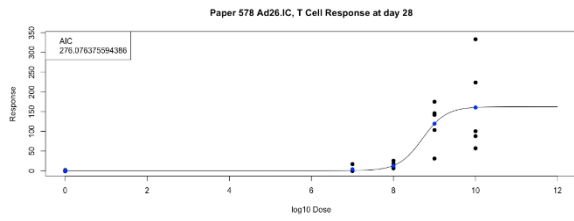
Paper 924

Day 14
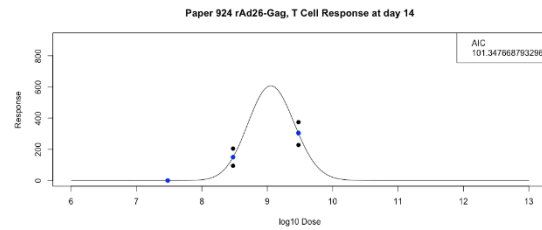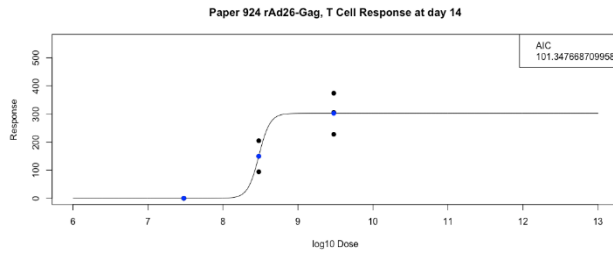
Host Species: Human

Route of Administration: IM

Paper 594:
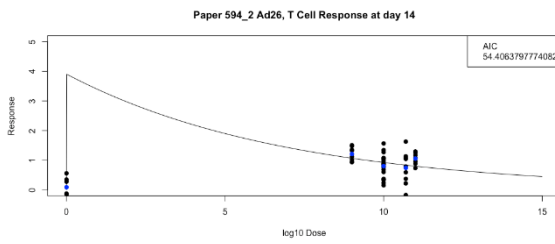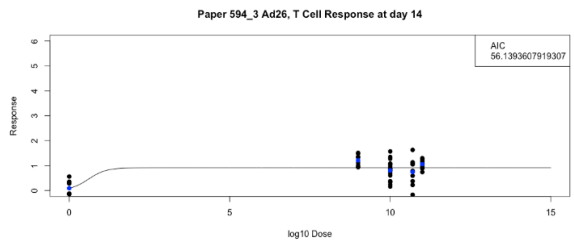
Day 14

Vector Species: E

Host Species: Mouse

Route of Administration: IM

Paper 305:

Day 21
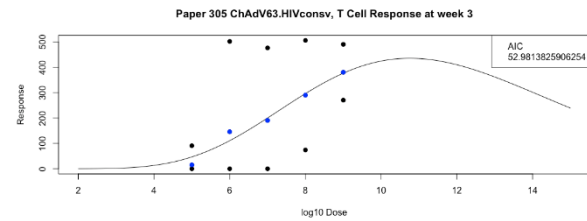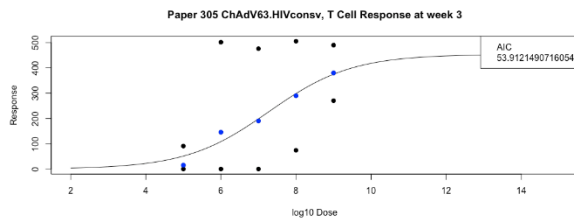
Host Species: Human

Route of Administration: IM

Paper 417:
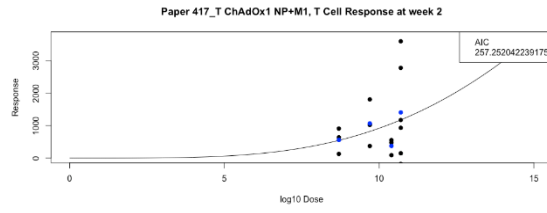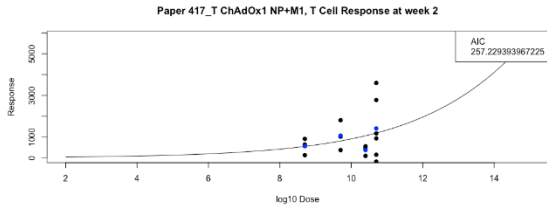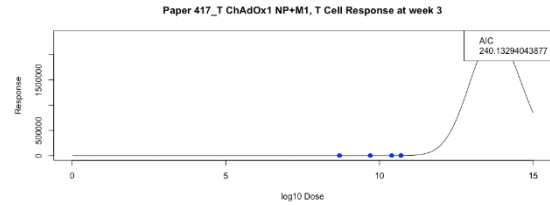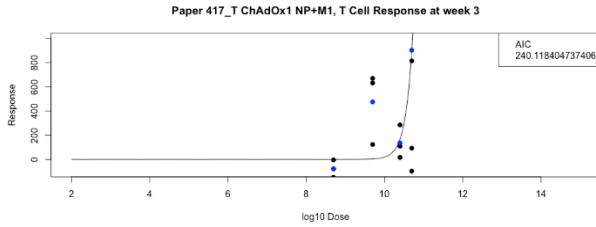
Day 14

Day 21

Paper 686:

Day 14
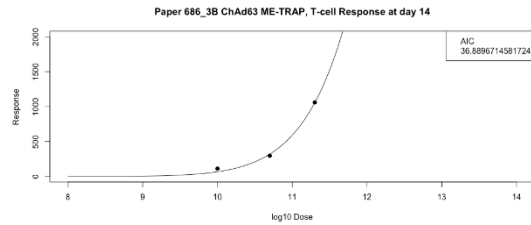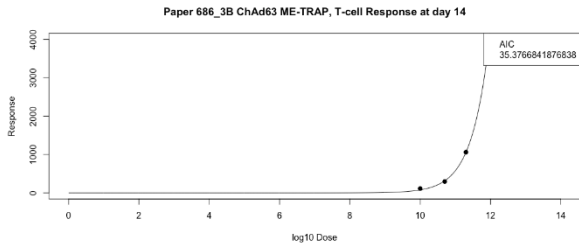
Day 21
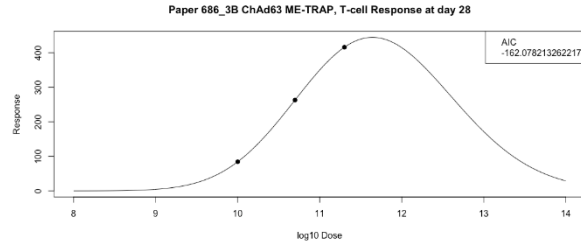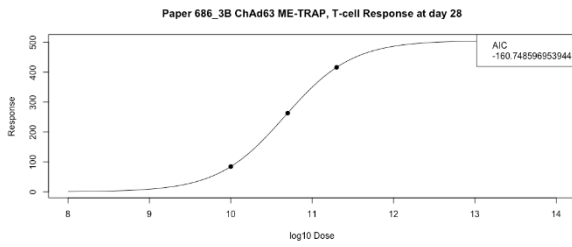
Day 90
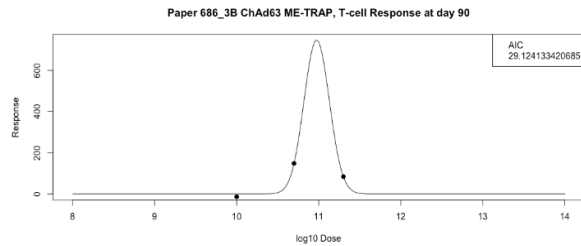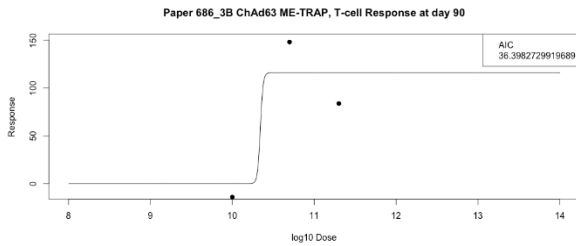
Response Type: CD4

Vector Species: B

Host Species: Mouse

Route of Administration: IM

Paper 1201:

Day 14
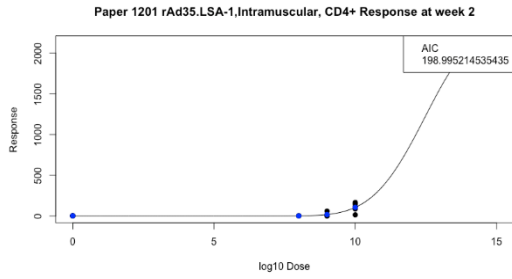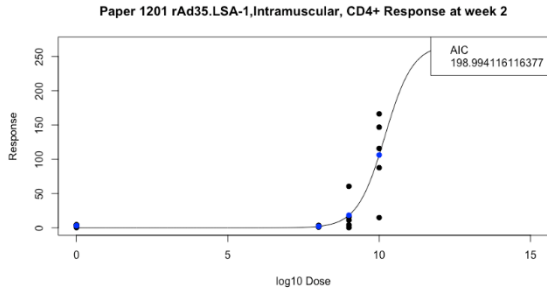
Host Species: Human
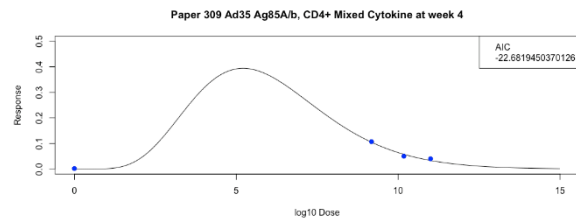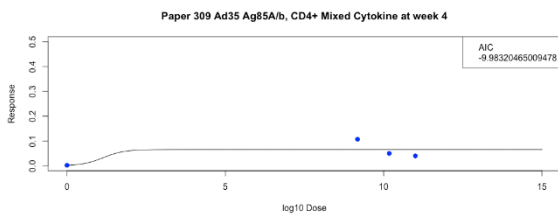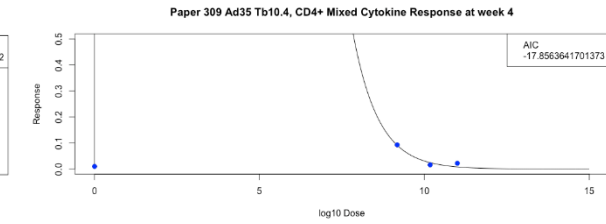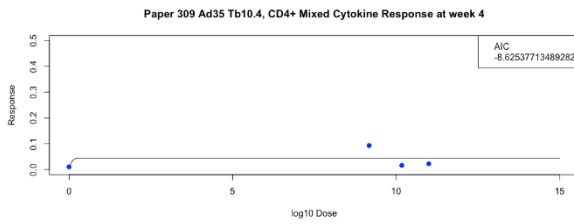
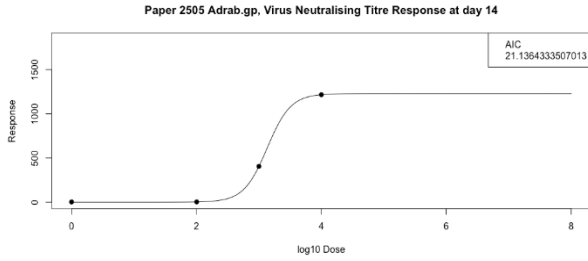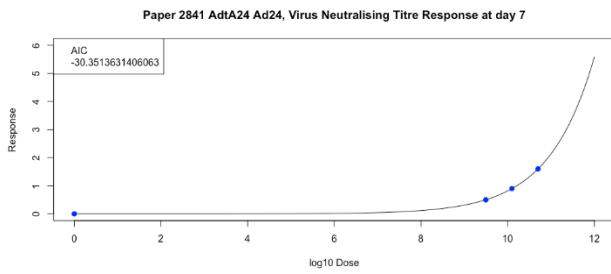Route of Administration: IM

Paper 309:

Day 28 - Ag85A/b

Day 28 - Tb10.4

Vector Species: C

Host Species: Monkey
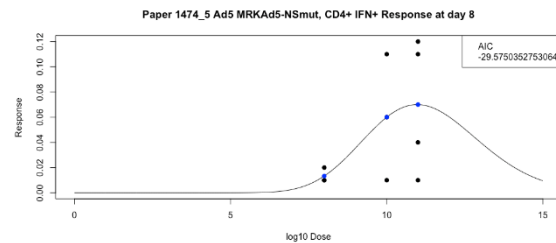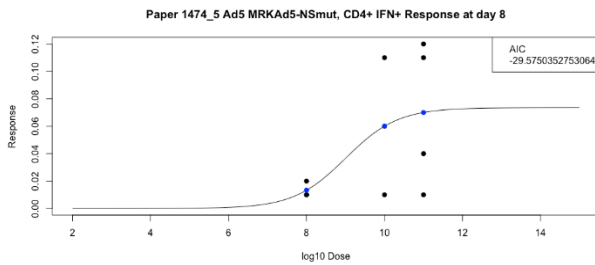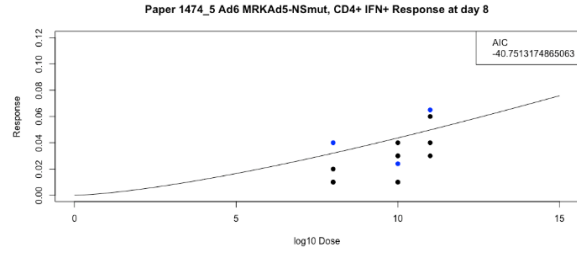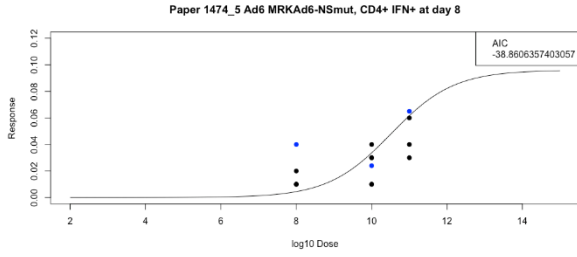
Route of Administration: IM

Paper 1474:

Day 8 - Ad5

Day 8 - Ad6

Paper 1474_5 Ad6 MRKAd6-NSmut, CD4+ IFN+ at day 8

Paper 1474_5 Ad6 MRKAd5-NSmut, CD4+ IFN+ Response at day 8

Vector Species: D

Host Species: Mouse

Route of Administration: IM

Paper 924

Day 14

Paper 924 rAd26-Gag, CD4+ Response at day 14

Paper 924 rAd26-Gag, CD4+ Response at day 14
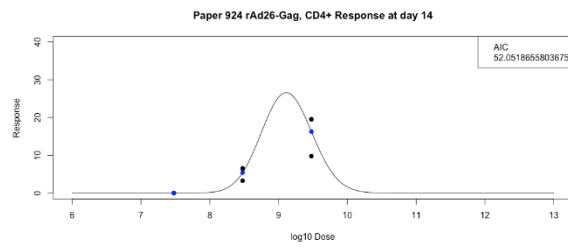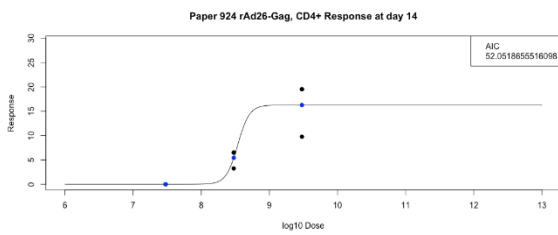
Response Type: CD8

Vector Species: B

Host Species: Mouse

Route of Administration: IM

Paper 1201: Stage Antigen 1

Day 14

Paper 1201 rAd35.LSA-1,Intramuscular, CD8+ Response at week 2

Paper 1201 rAd35.LSA-1,Intramuscular, CD8+ Response at week 2

Paper 1269:

Day 56

Paper 1269 Ad35.CS0.01, CD8+ T-Cell Response at day 56

Paper 1269 Ad35.CS0.01, CD8+ T-Cell Response at day 56

Paper 1492:

439
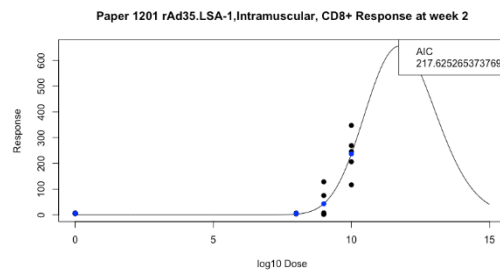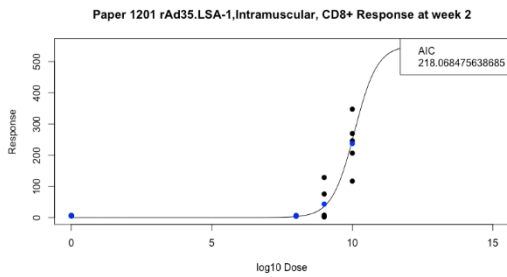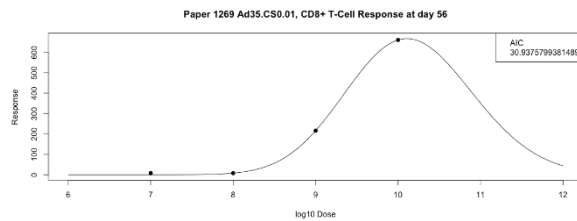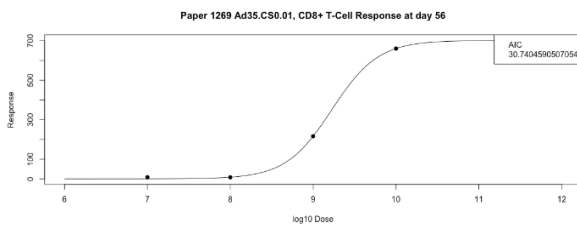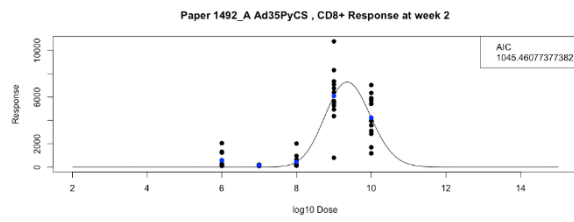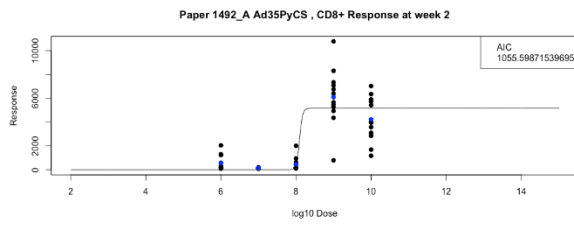
Day 14

Route of Administration: SQ

Paper 555:

rAd35

Day 14 - Tetramer Staining
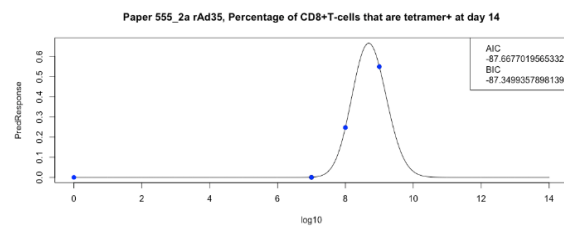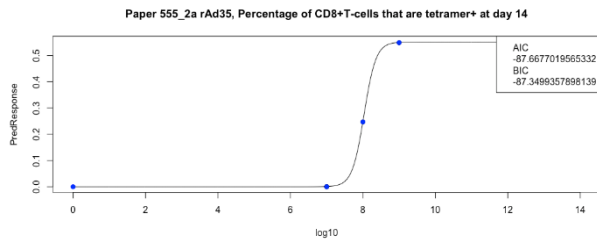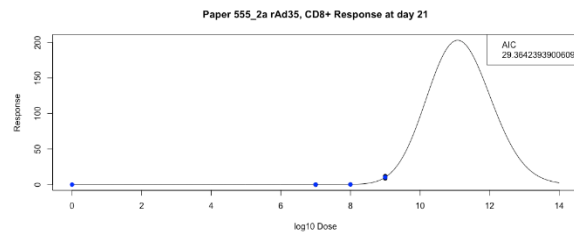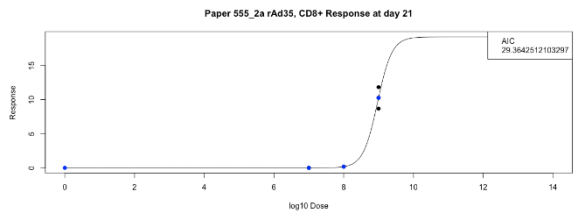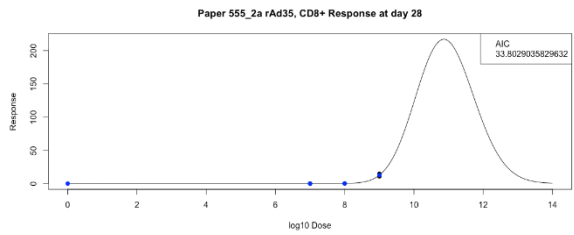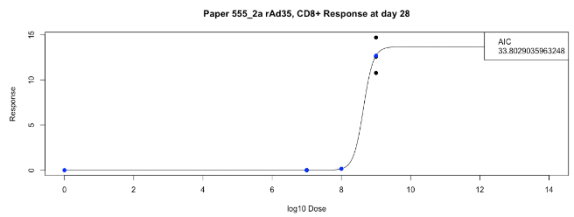
Day 21 - Tetramer Staining

Day 28 - Tetramer Staining

Day 35 - Tetramer Staining
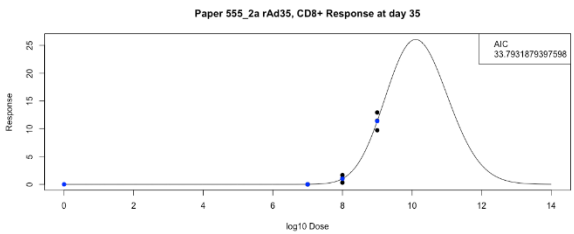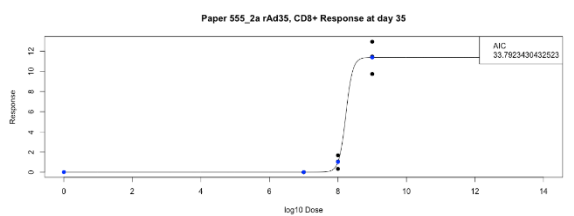
Day 70 - Tetramer Staining
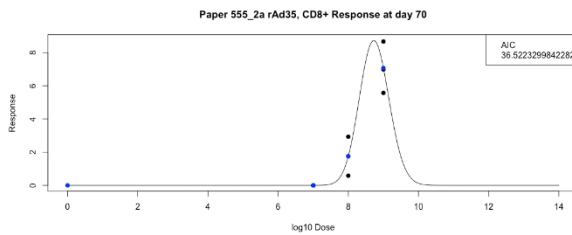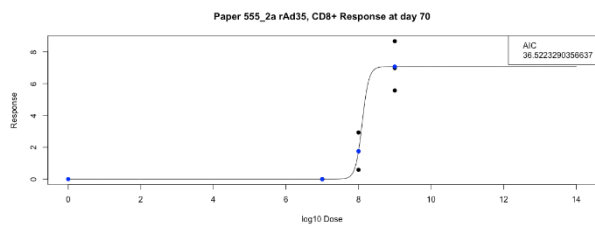
Day 23 - Cytokine Staining
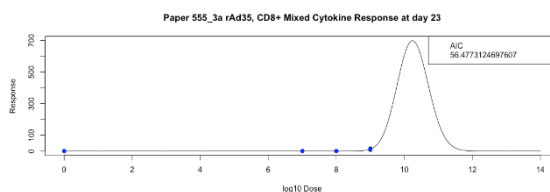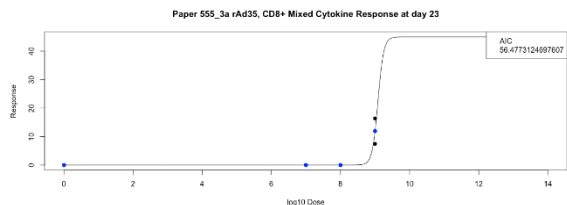
Day 70 -Cytokine Staining
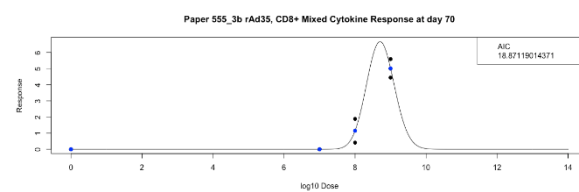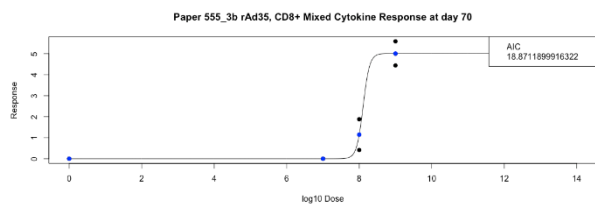
Host Species: Human

Route of Administration: IM

Paper 309:

Day 28 - Ag85A/b
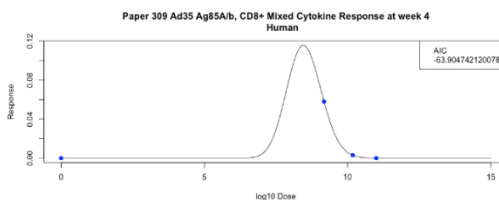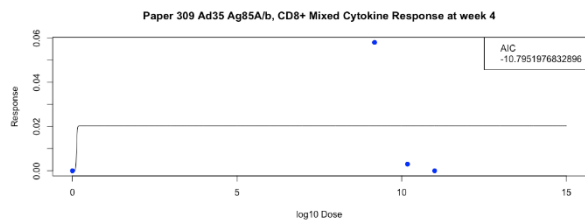
Day 28 - Tb.104
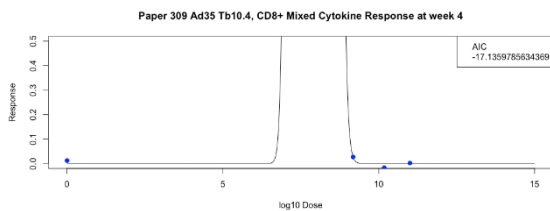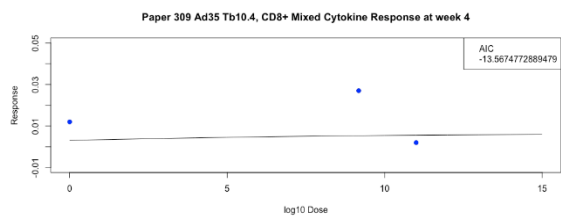
Vector Species: C

Host Species: Mouse

Route of Administration: IM

Paper 461:

Day 7 HA-518

441
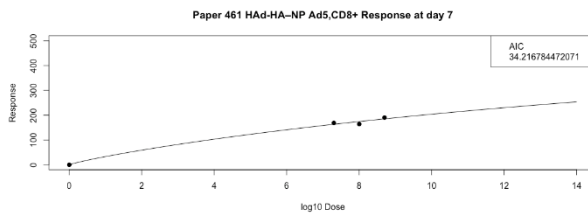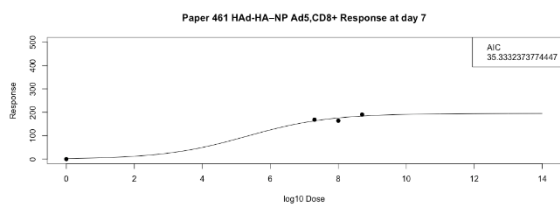
Day 14 HA-518
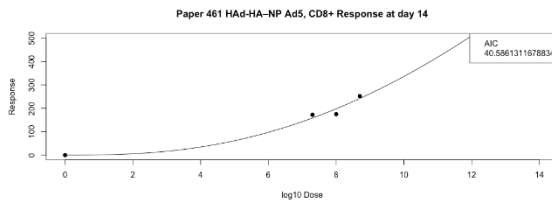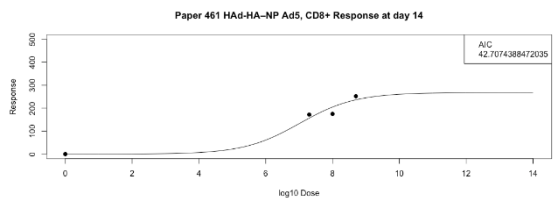
Day 7 NP-147
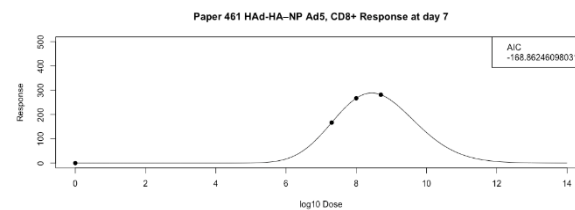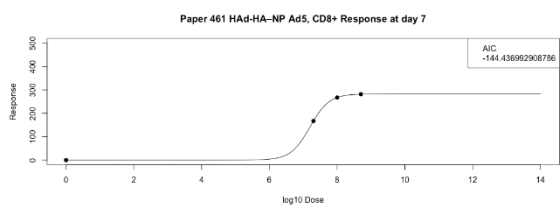
Day 14 NP-147
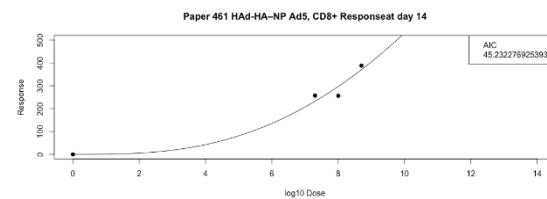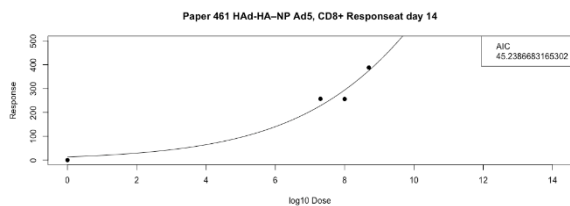
Paper 1039

Day 8 - Ad-CAGoptZGP
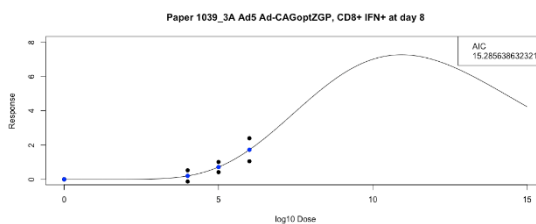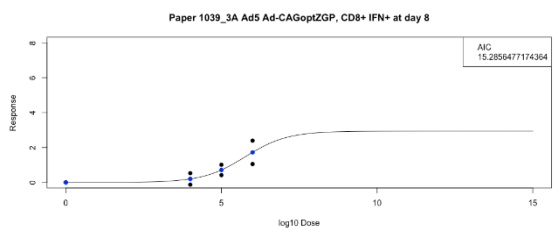
Day 8 - Ad-AdCMVZGP

Paper 1492:

Day 14

442

Paper 1492_C rAd5PyCS , CD8+ Response at week 2

Paper 1492_C rAd5PyCS , CD8+ Response at week 2

Paper 3018

Day 7

Paper 3018 Ad5-SIVmac239Gag, CD8+ Response at day 7

Paper 3018 Ad5-SIVmac239Gag, CD8+ Response at day 7

Day 15

Paper 3018 Ad5-SIVmac239Gag, CD8+ Response at day 15

Paper 3018 Ad5-SIVmac239Gag, CD8+ Response at day 15

Day 30

Paper 3018 Ad5-SIVmac239Gag, CD8+ Response at day 30
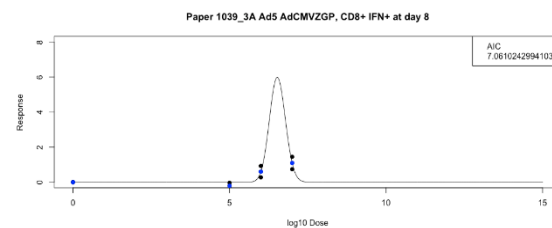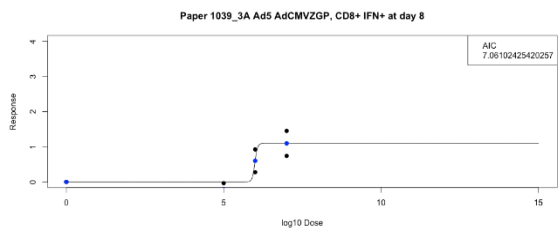
Paper 3018 Ad5-SIVmac239Gag, CD8+ Response at day 30
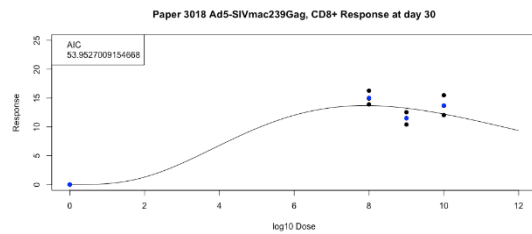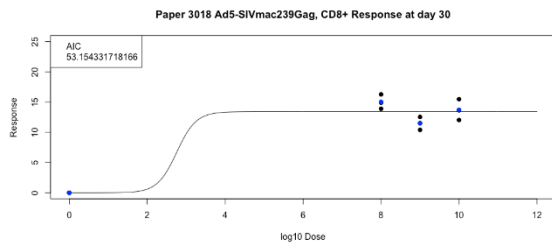
Route of Administration: SQ

Paper 555: C

rAd5

Day 14 - Tetramer Staining
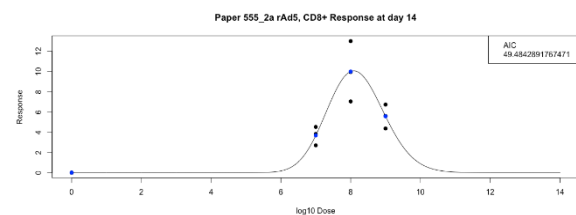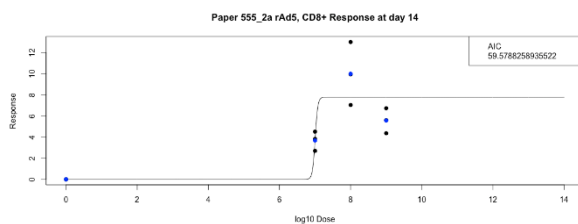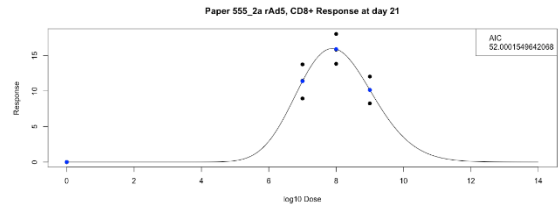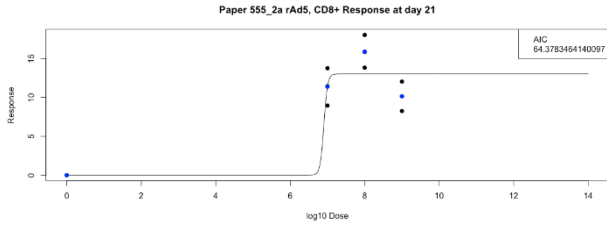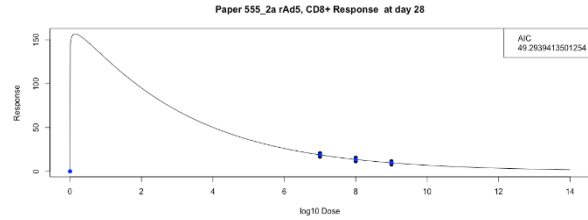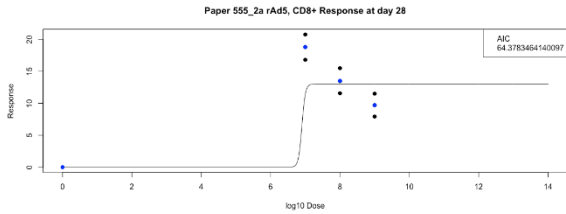
Paper 555_2a rAd5, CD8+ Response at day 14

Paper 555_2a rAd5, CD8+ Response at day 14

Day 21 - Tetramer Staining

443

Day 28 - Tetramer Staining

Day 35 - Tetramer Staining
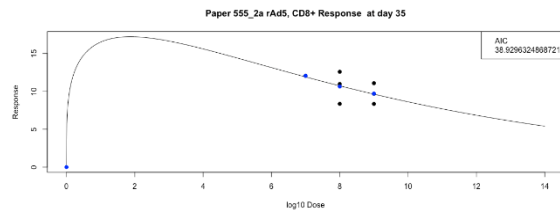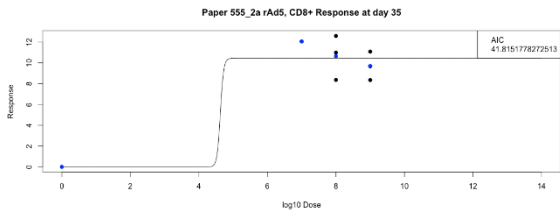
Day 70 - Tetramer Staining
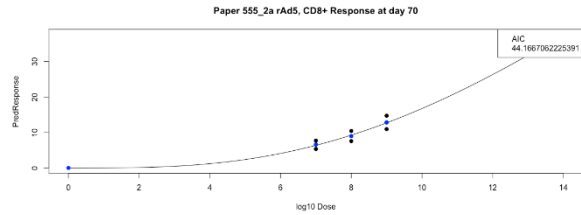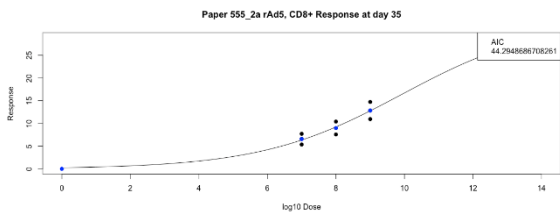
Day 23 - Cytokine Staining
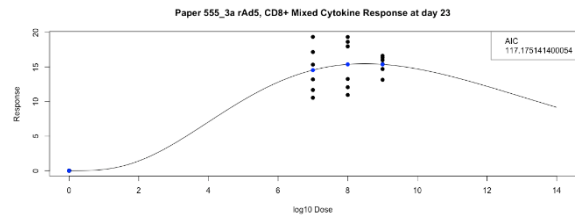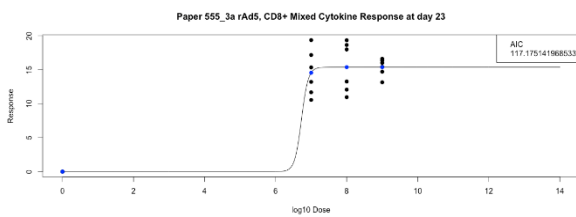
Day 70 -Cytokine Staining
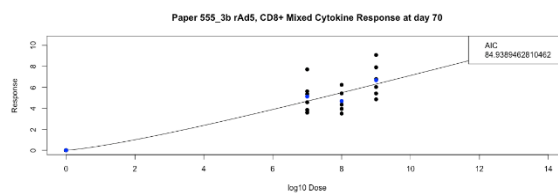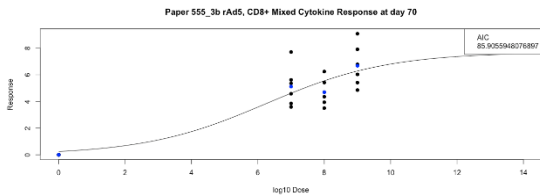
ChAd3

Day 14 - Tetramer Staining
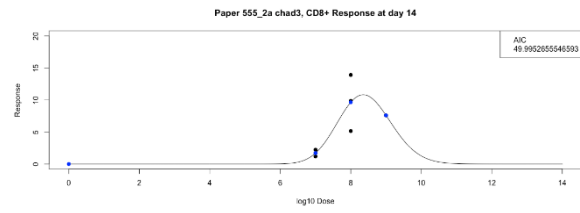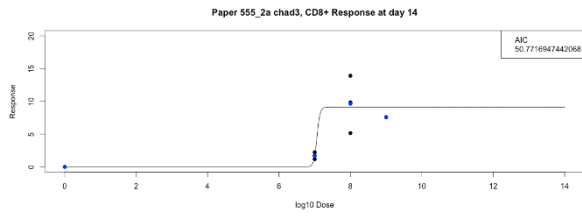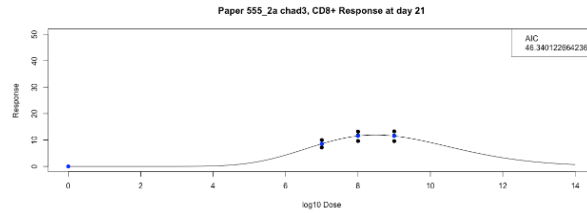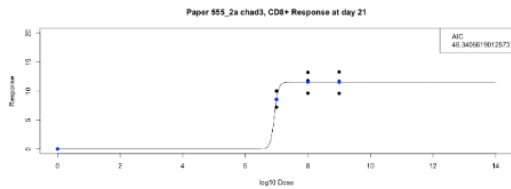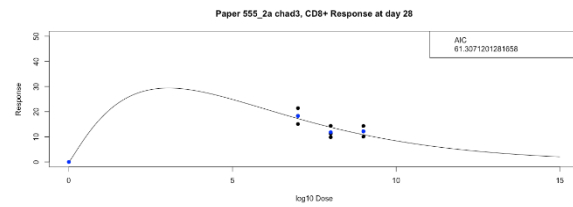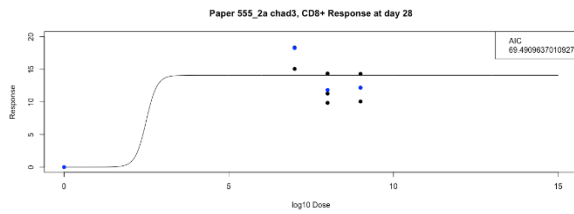
444

Day 21 - Tetramer Staining

Day 28 - Tetramer Staining

Day 35 - Tetramer Staining

Day 70 - Tetramer Staining

Day 23 - Cytokine Staining

Day 70 -Cytokine Staining

445
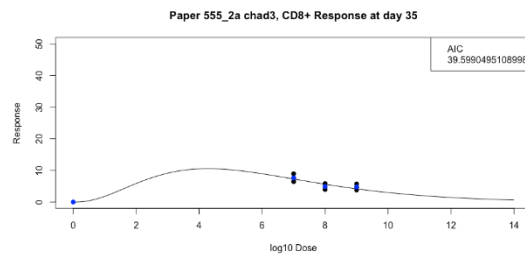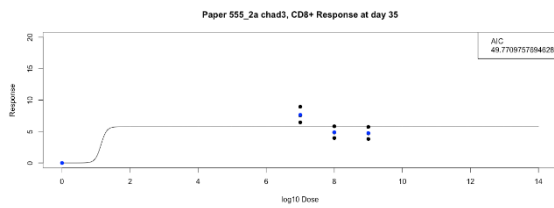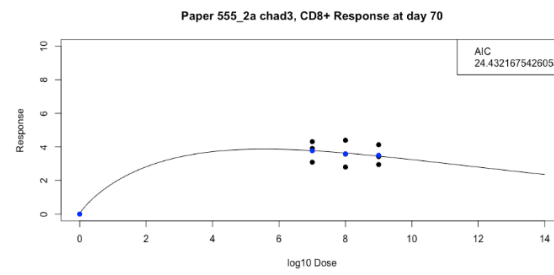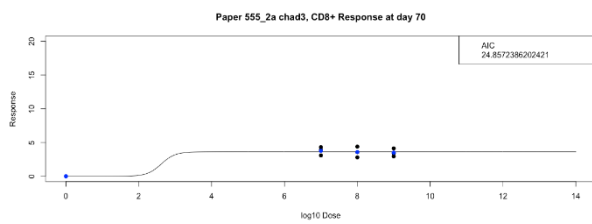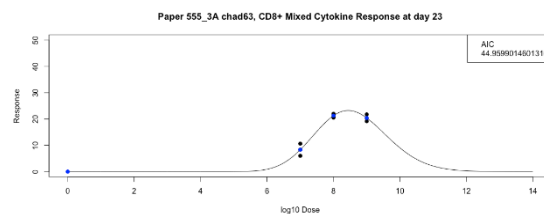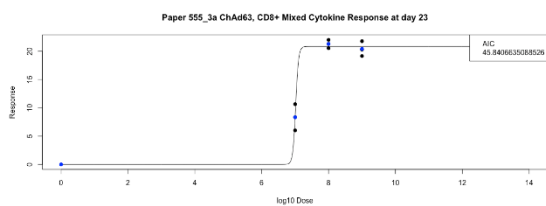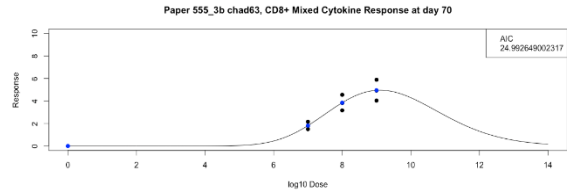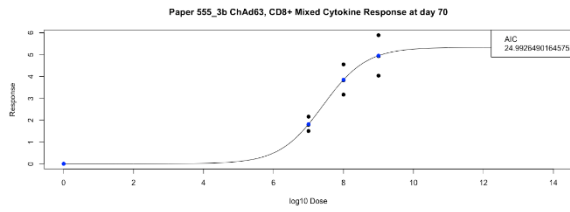
Paper 2980:

Day 60
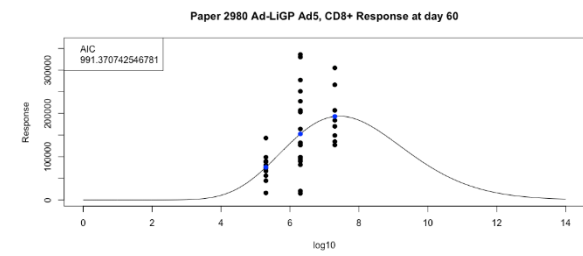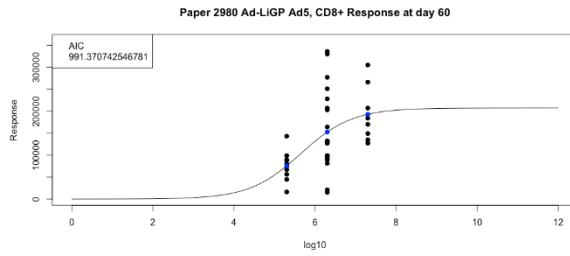
Host Species: Monkey
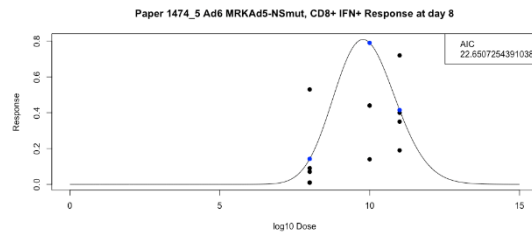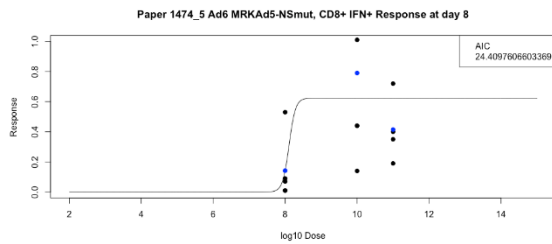
Route of Administration: IM

Paper 1474:

Day 8 - Ad5

Day 8 - Ad6

Vector Species: D

Host Species: Mouse

Route of Administration: IM

Paper 924

Day 14

446
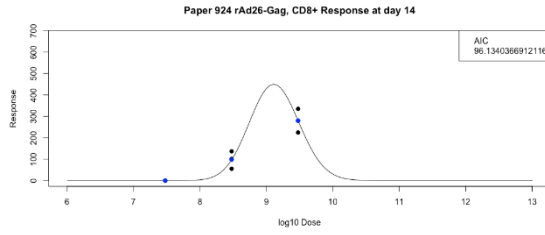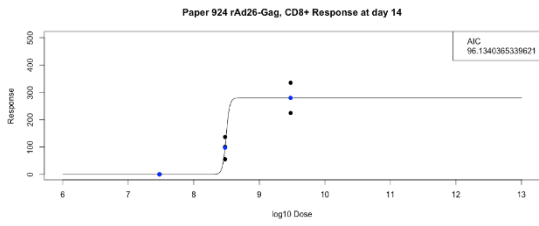
Route of Administration: SQ

Paper 555

rAd28

Day 14 - Tetramer Staining
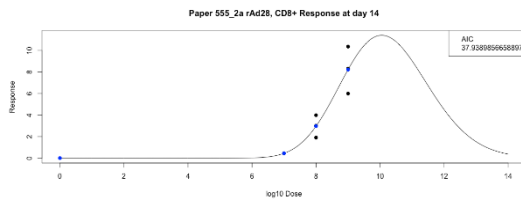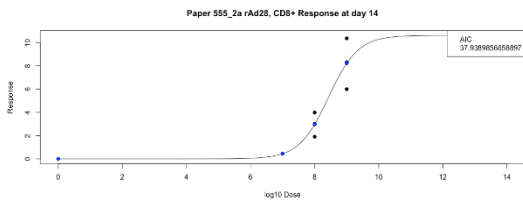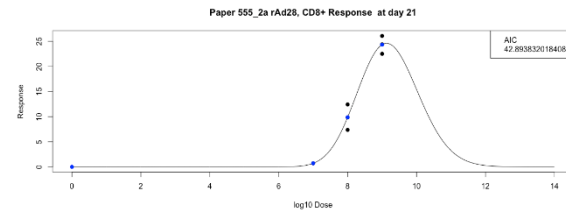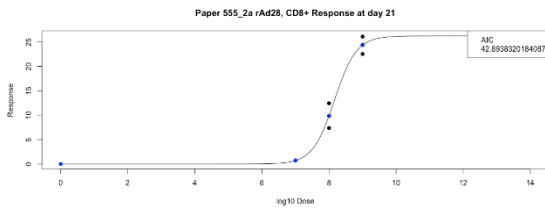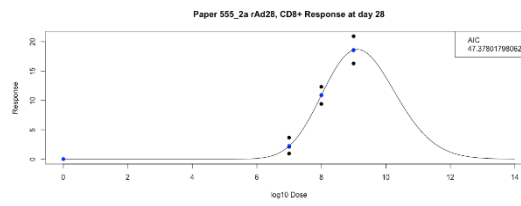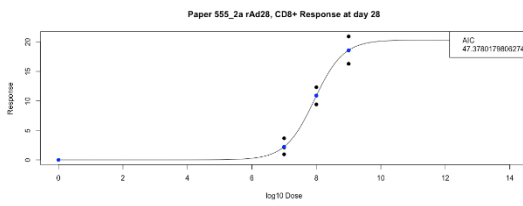
Day 21 - Tetramer Staining

Day 28 - Tetramer Staining

Day 35 - Tetramer Staining

Day 70 - Tetramer Staining

Paper 555_2a rAd28, CD8+ Response at day 70

Day 23 - Cytokine Staining

Paper 555_3a rAd28, CD8+ Mixed Cytokine Response at day 23

Day 70 -Cytokine Staining

Paper 555_3b rAd28, CD8+ Mixed Cytokine Response at day 70
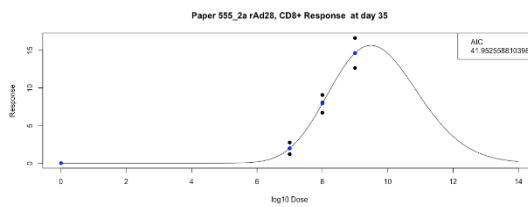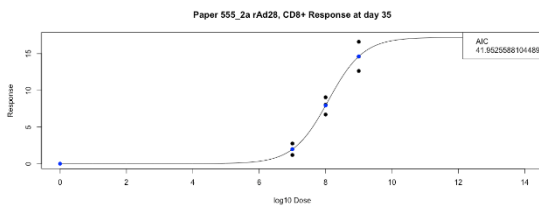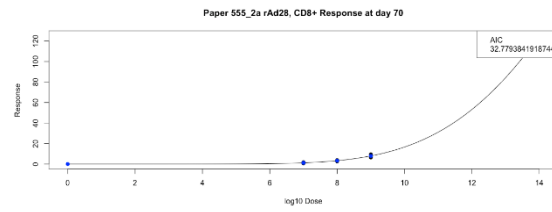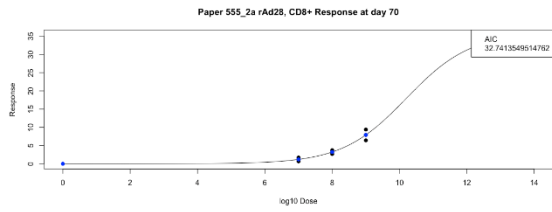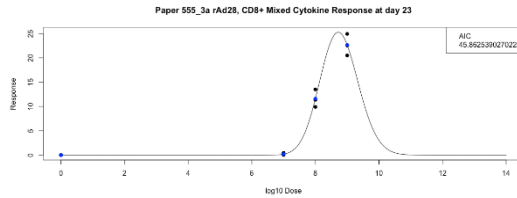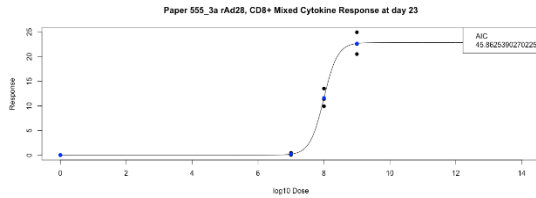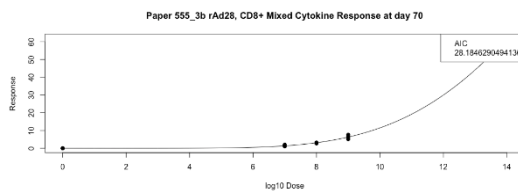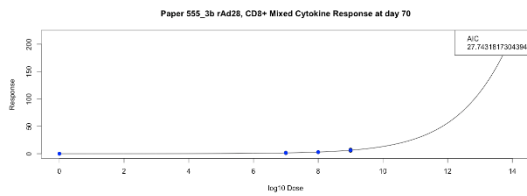
Vector Species: E

Host Species: Mouse

Route of Administration: IM

Paper 2919:

Day 10

Paper 2919 SAd36PyCMP SAd36, CD8+ Response at day 10

Day 20
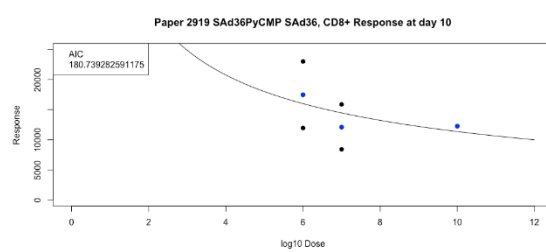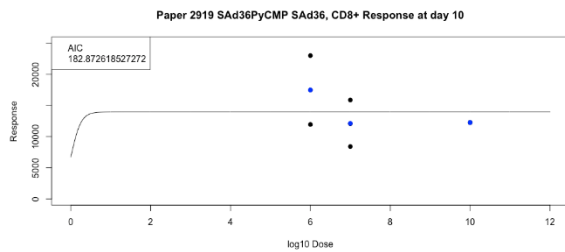
Paper 2919 SAd36PyCMP SAd36, CD8+ Response at day 20
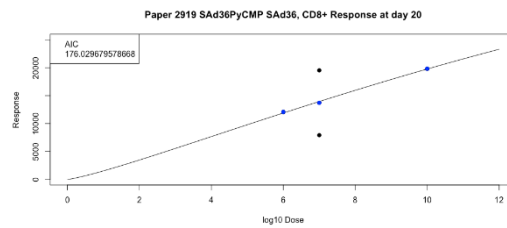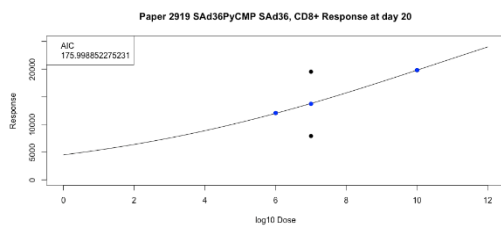
Route of Administration: SQ

Paper 555:

ChAd63

Day 14 - Tetramer Staining
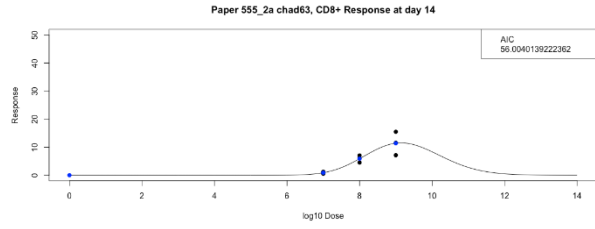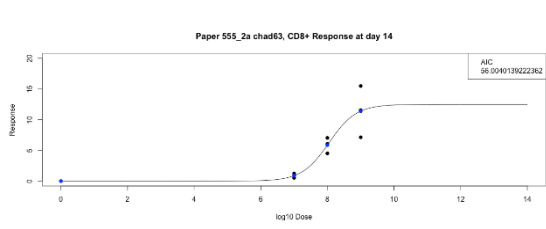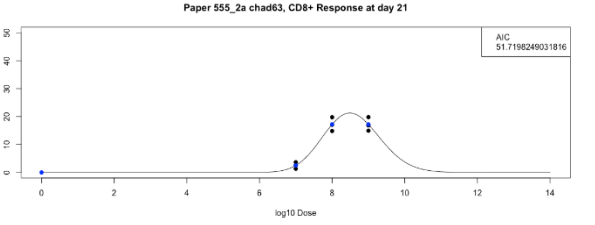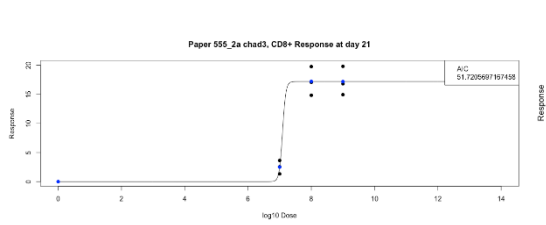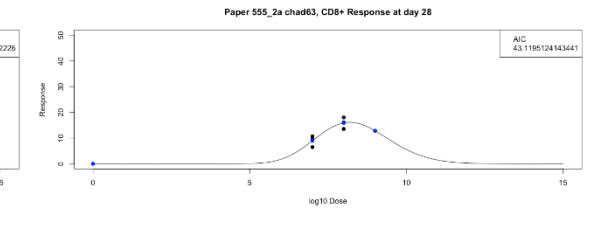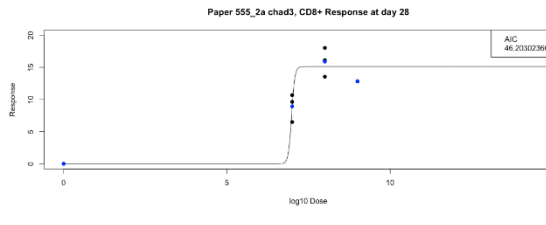
448

Day 21 - Tetramer Staining

Day 28 - Tetramer Staining

Day 35 - Tetramer Staining

Day 70 - Tetramer Staining

Day 23 - Cytokine Staining

Day 70 -Cytokine Staining

449
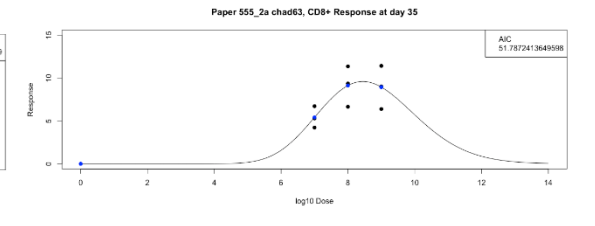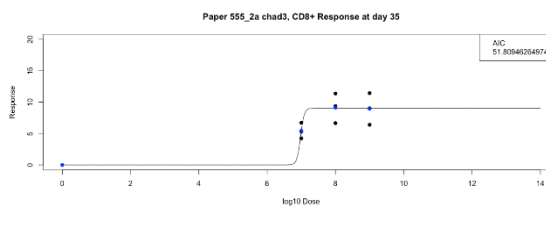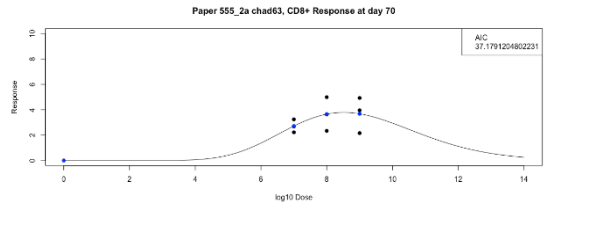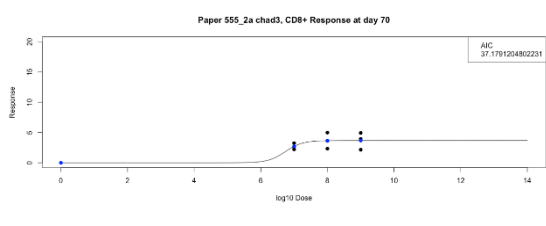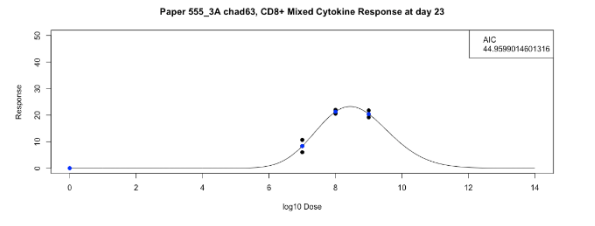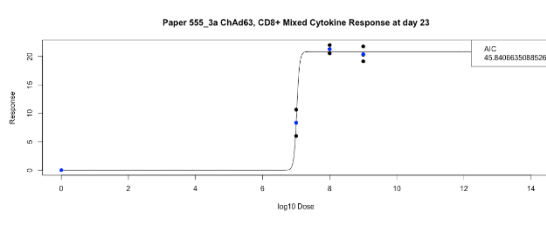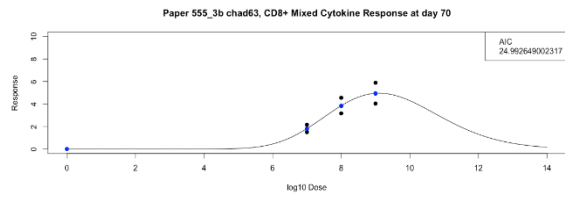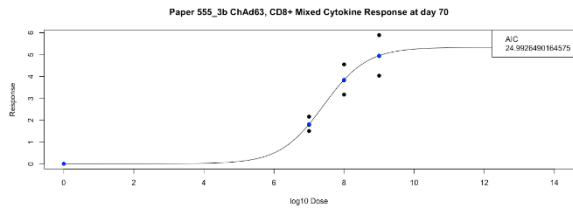
Vector Species: G

Host Species: Mouse

Route of Administration: SQ

Paper 555:

sAd11

Day 14 - Tetramer Staining
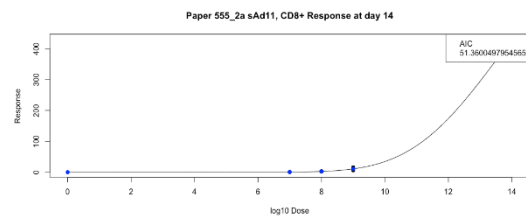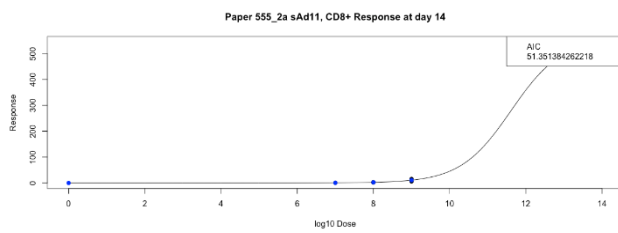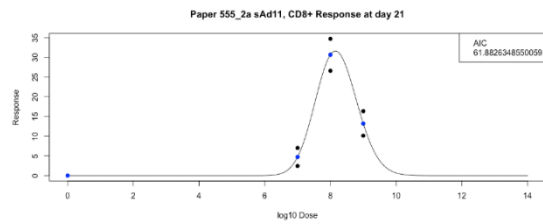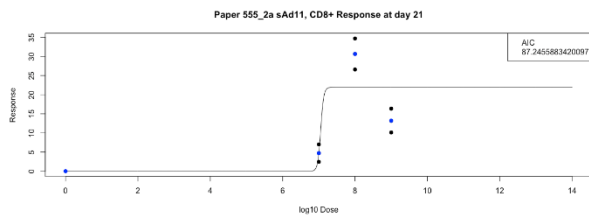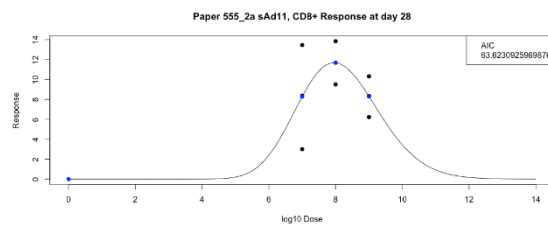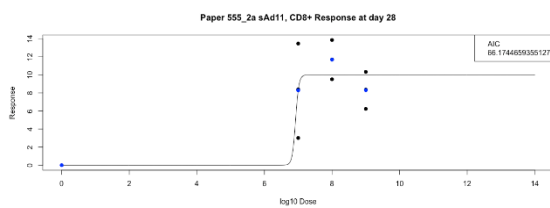
Day 21 - Tetramer Staining

Day 28 - Tetramer Staining

Day 35 - Tetramer Staining

Day 70 - Tetramer Staining

450

Paper 555_2a sAd11, CD8+ Response at day 70

Day 23 - Cytokine Staining

Paper 555_3a sAd11, CD8+ Mixed Cytokine Response at day 23

Day 70 -Cytokine Staining

Paper 555_3b sAd11, CD8+ Mixed Cytokine Response at day 70
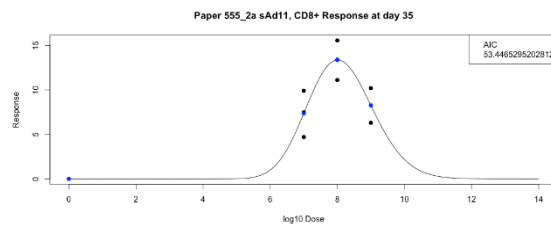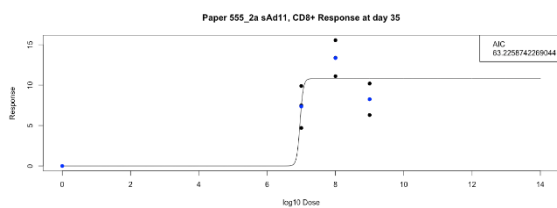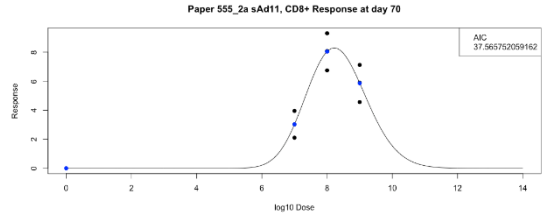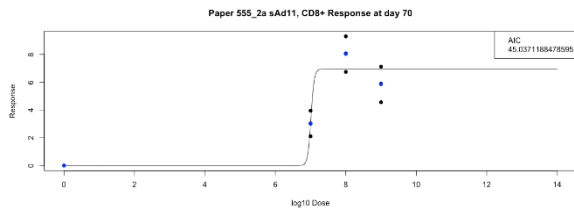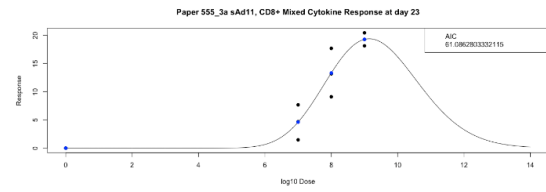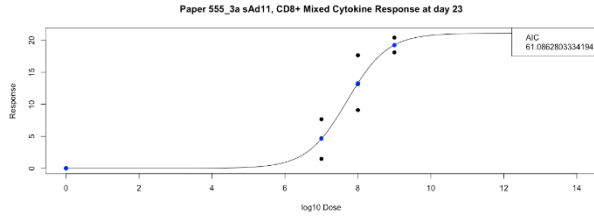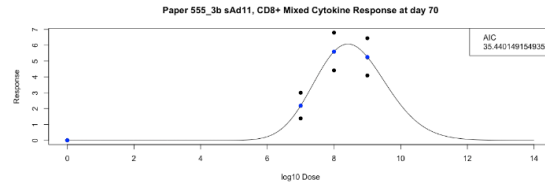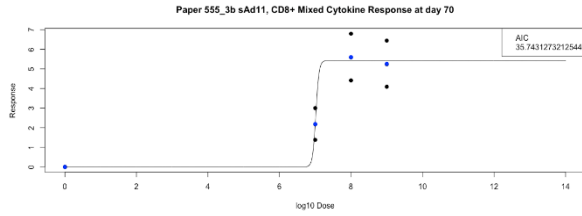
Vector Species: None

Host Species: Mouse

Route of Administration: SQ

Paper 555:

sAd16

Day 14 - Tetramer Staining
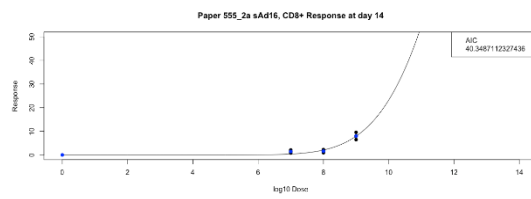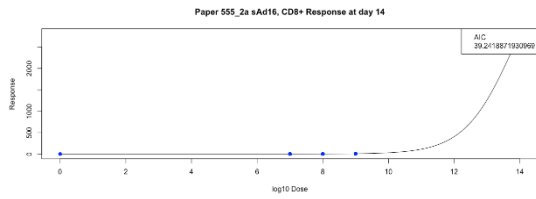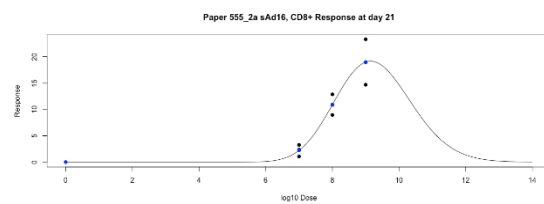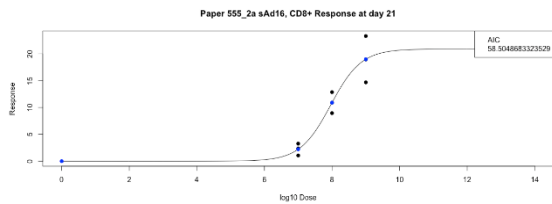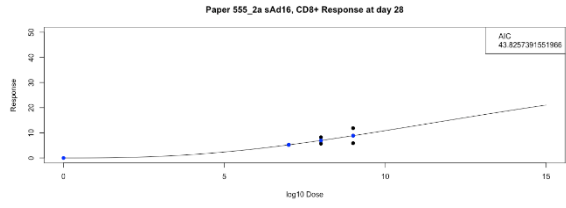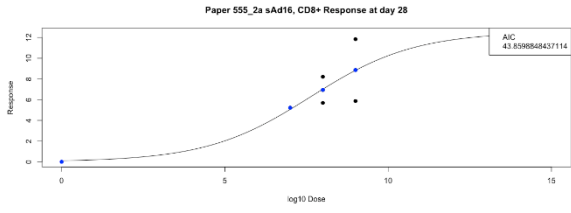
Paper 555_2a sAd16, CD8+ Response at day 14

Day 21 - Tetramer Staining

Paper 555_2a sAd16, CD8+ Response at day 21

Day 28 - Tetramer Staining

Paper 555_2a sAd16, CD8+ Response at day 28

Day 35 - Tetramer Staining

Paper 555_2a sAd16, CD8+ Response at day 35

Day 70 - Tetramer Staining

Paper 555_2a sAd16, CD8+ Response at day 70

Day 23 - Cytokine Staining

Paper 555_3a sAd16, CD8+ Mixed Cytokine Response at day 23

Day 70 - Cytokine Staining

Paper 555_3b sAd16, CD8+ Mixed Cytokine Response at day 70
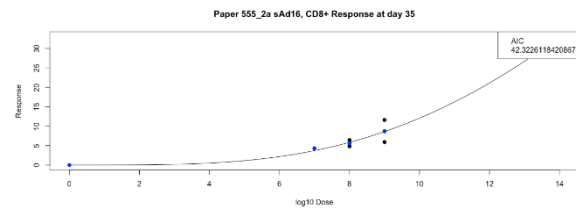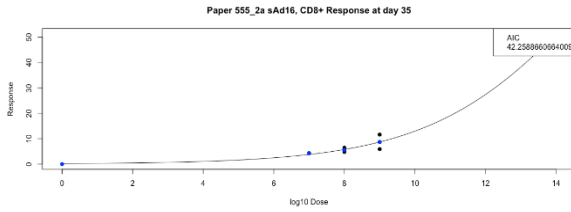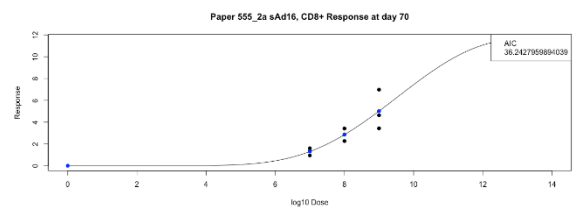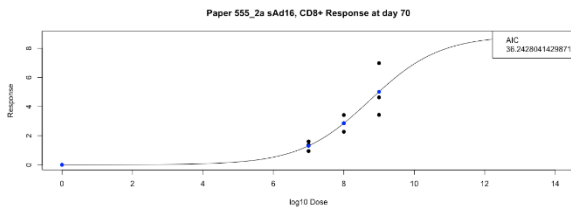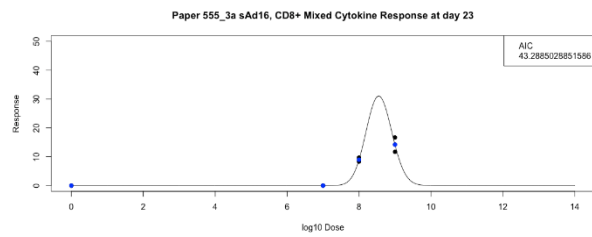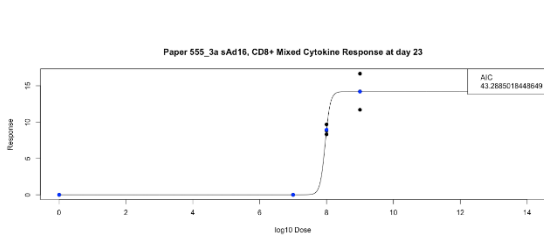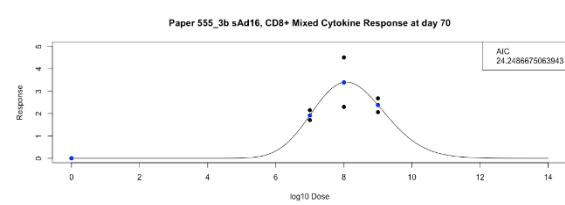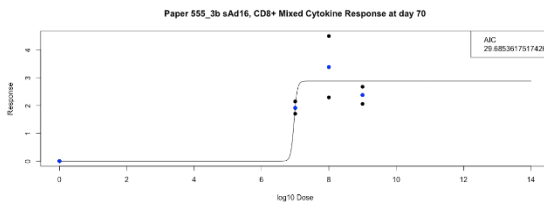
Response Type: CD4+ IFN+%

Vector Species: B

Host Species: Human

Route of Administration: IM

Paper 309:

Day 28

Paper 309 AErAS-402 Intramuscular, CD4+ IFN-γ+ Response at week 4

Paper 441:

Day 28

Paper 467_5 Ad35.CS0.01, CD4+ IFN-γ+ at day 28
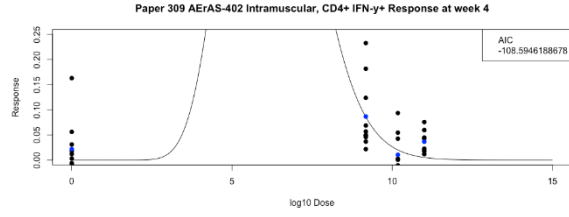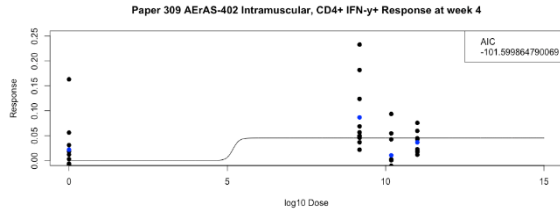
Response Type: CD8+ IFN+%

Vector Species: B

Host Species: Human

Route of Administration: IM

Paper 309:

Day 28

Paper 309 AErAS-402 Intramuscular, CD8+ IFN-γ+ Response at week 4

Paper 441:

Day 28

Paper 441_4 Ad35.CS.01, CD8+ IFN-γ% at week 4
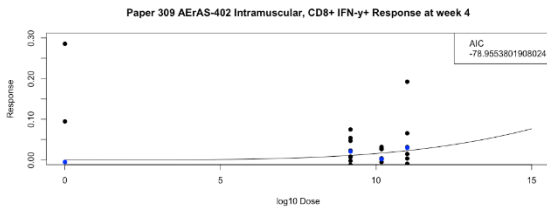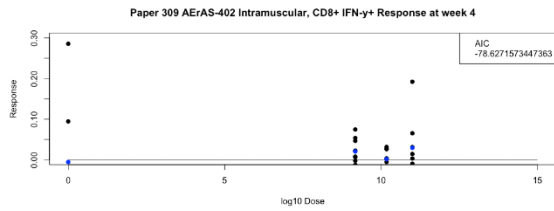
Vector Species: E

Host Species: Mouse

Route of Administration: IM

Paper 1801:

Day 10

Paper 1801 AdC6gag37, Frequency of IFN-y producing splenic CD8+ T Cells at day 10

Paper 1801 AdC6gag37, Frequency of IFN-y producing splenic CD8+ T Cells at day 10
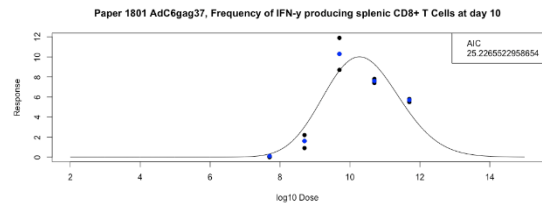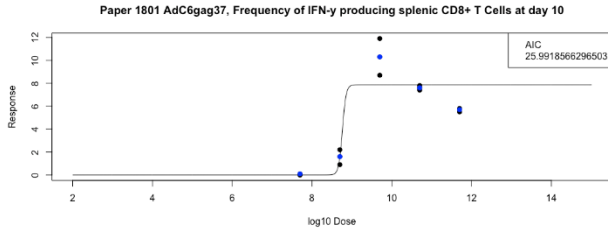
Response Type: CD4+ TNF+%

Vector Species: B

Host Species: Human

Route of Administration: IM

Paper 309:

Day 28

Paper 309 AErAS-402 Intramuscular, CD4+ TFN-a+ Response at week 4

Paper 309 AErAS-402 Intramuscular, CD4+ TFN-a+ Response at week 4

Paper 441:

Day 28

Paper 467_5 Ad35.CS0.01, CD4+ TNF-a+ at day 28

Paper 467_5 Ad35.CS0.01, CD4+ TNF-a+ at day 28
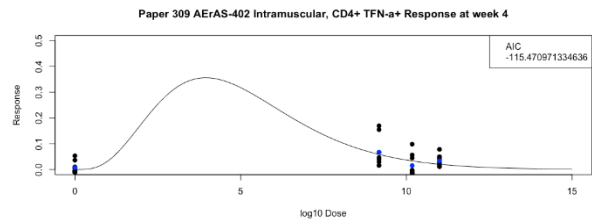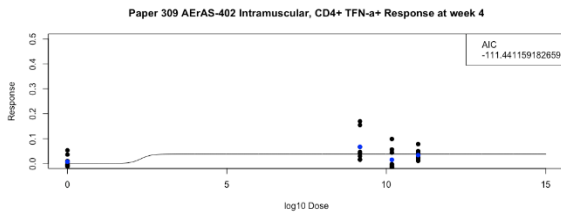
Response Type: CD8+ TNF+%

Vector Species: B
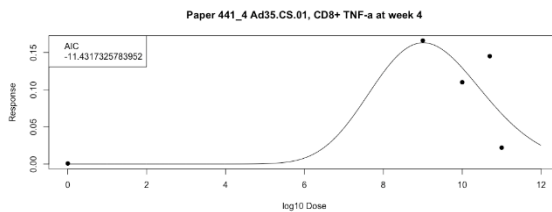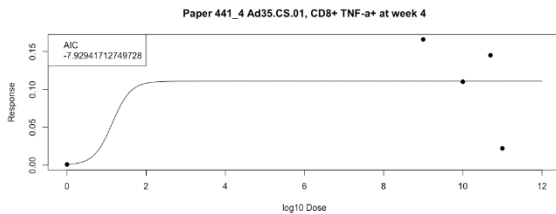
Host Species: Human

Route of Administration: IM

Paper 441:

Day 28

Paper 441_4 Ad35.CS.01, CD8+ TNF-a+ at week 4

Paper 441_4 Ad35.CS.01, CD8+ TNF-a at week 4

Response Type: CD4+ Il2+%

Vector Species: B

Host Species: Human

Route of Administration: IM

Paper 309:

Day 28

Paper 441:

Day 28
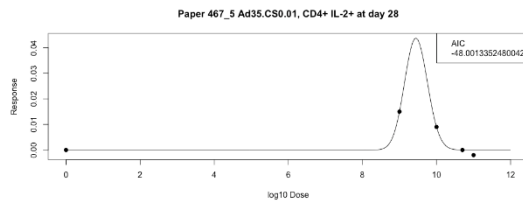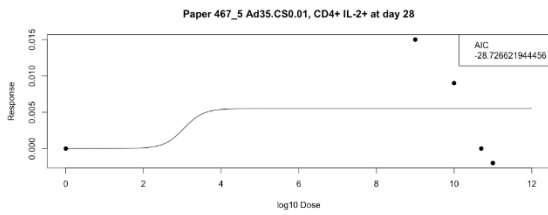
Response Type: CD8+ Il2+%

Vector Species: B

Host Species: Human

Route of Administration: IM

Paper 441:
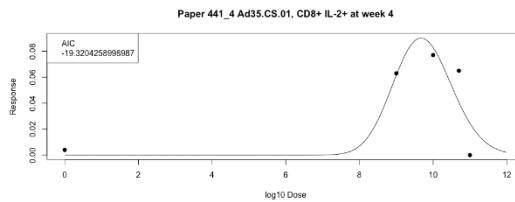
Day 28

Response Type: CD4+ Il17+%

Vector Species: B

Host Species: Human

Route of Administration: IM

Paper 309:

Day 28
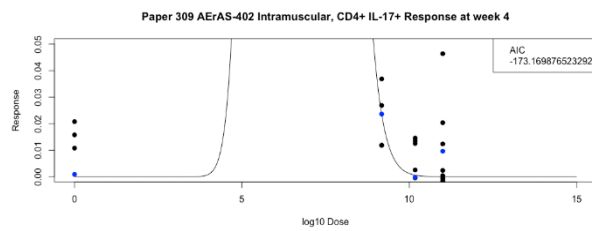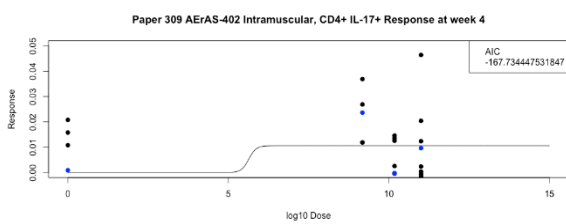
Response Type: Virus Neutralisation Titre

Vector Species: B

455

Host Species: Mouse

Route of Administration: IM

Paper 1269:

Day 56
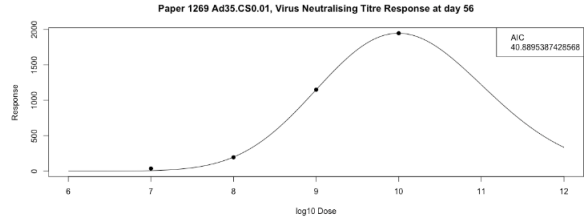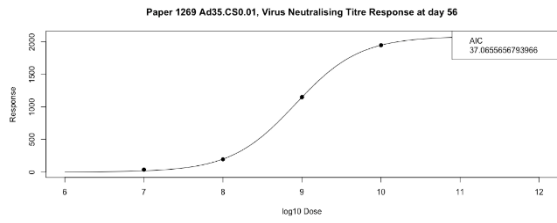
Route of Administration: SQ

Paper 2030:
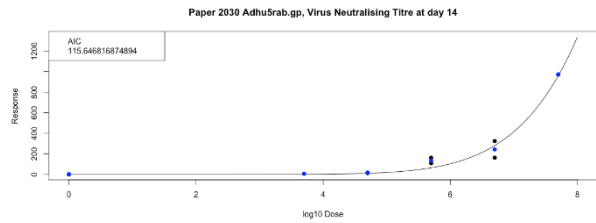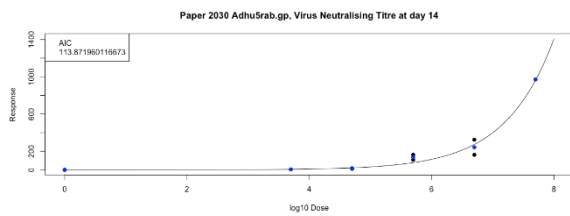
Day 14

Vector Species: C

Host Species: Mouse

Route of Administration: IM

Paper 461:

Day 7

Day 14

Paper 574:

Day 7

Paper 574 rAdV-Asi-05 Ad5, Virus Neutralization Titres at week 1

Day 21

Paper 574 rAdV-Asi-05 Ad5, Virus Neutralization Titres at week 3

Day 35

Paper 574 rAdV-Asi-05 Ad5, Virus Neutralization Titres at week 5

Day 49

Paper 574 rAdV-Asi-05 Ad5, Virus Neutralization Titres at week 7

Day 63

Paper 574 rAdV-Asi-05 Ad5, Virus Neutralization Titres at week 9 Mouse

Day 77

Paper 574 rAdV-Asi-05 Ad5, Virus Neutralization Titres at week 11

Day 119

457
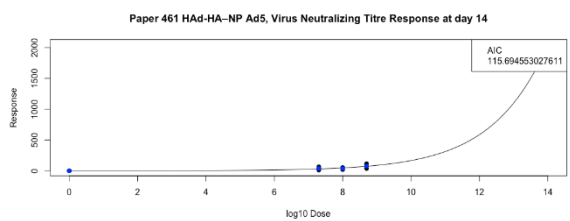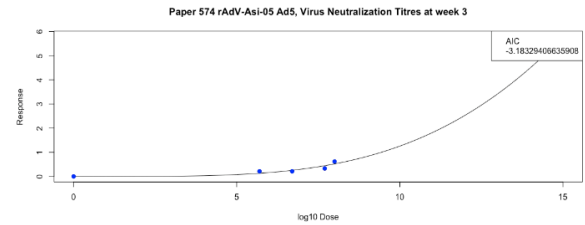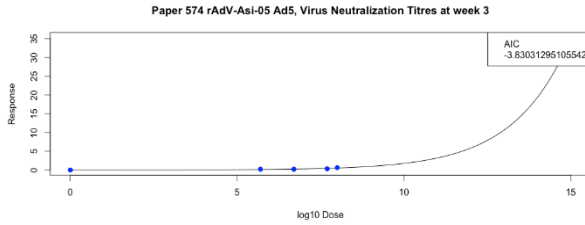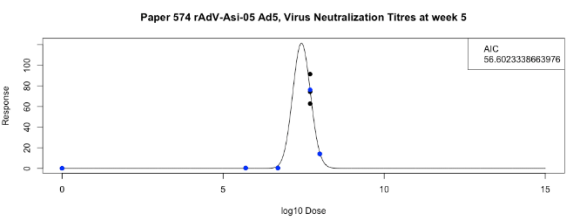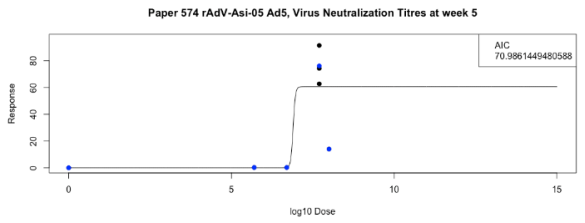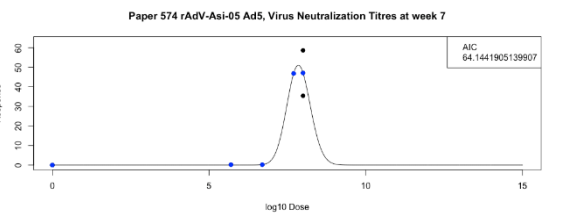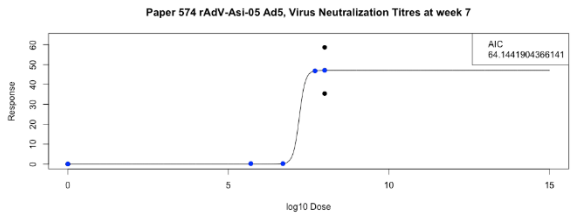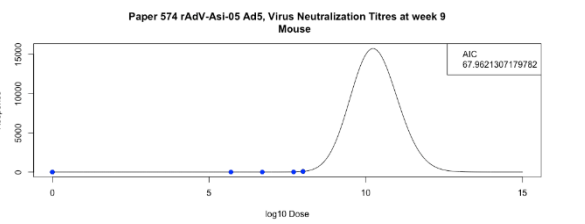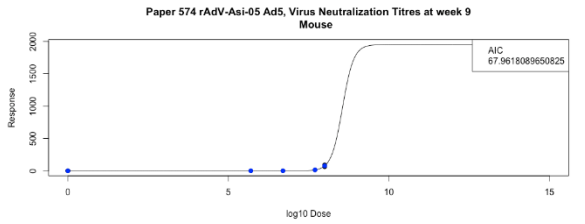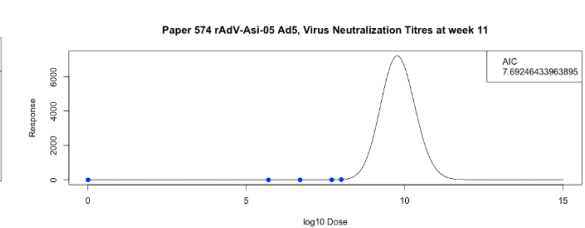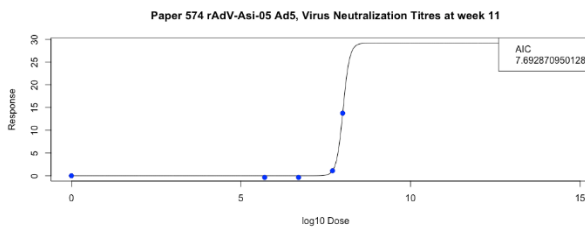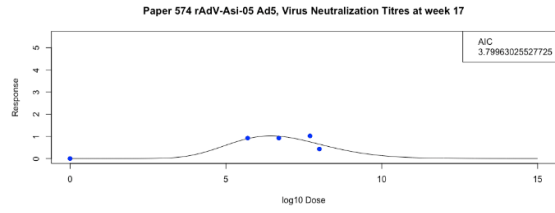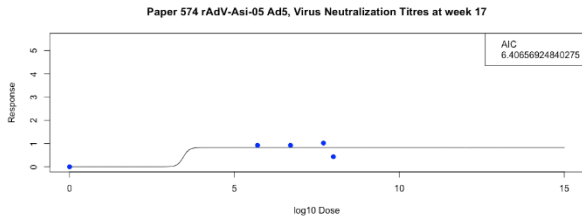
Day 161
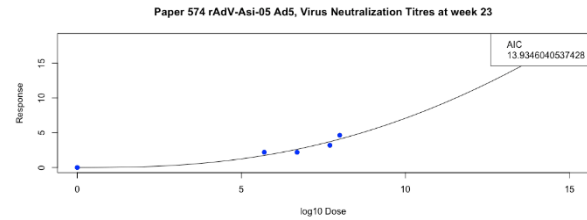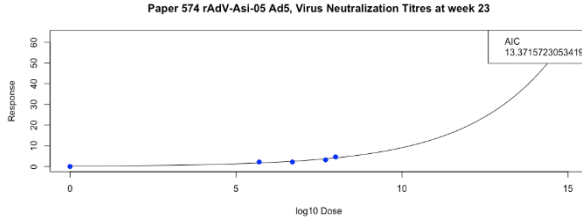
Paper 669:

Day 28
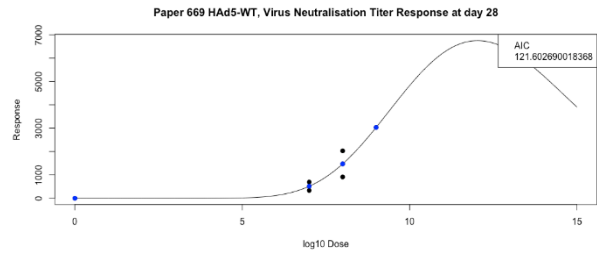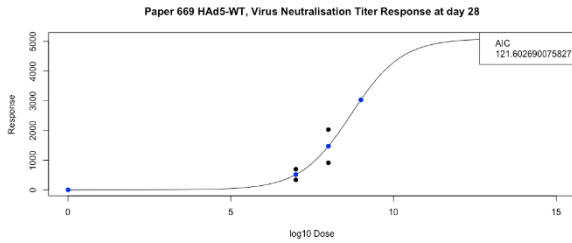
Paper 1039

Day 8 - Ad-CAGoptZGP
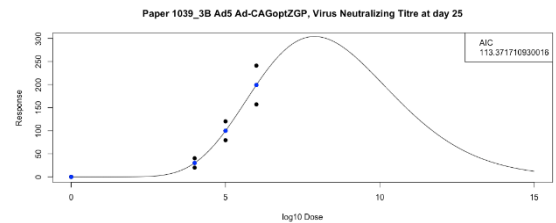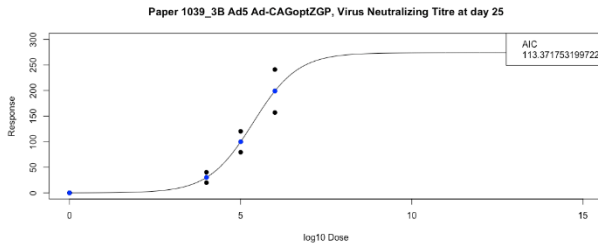
Route of Administration: SQ

Paper 2505:

Day 10
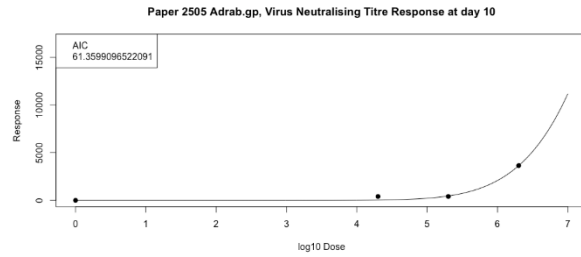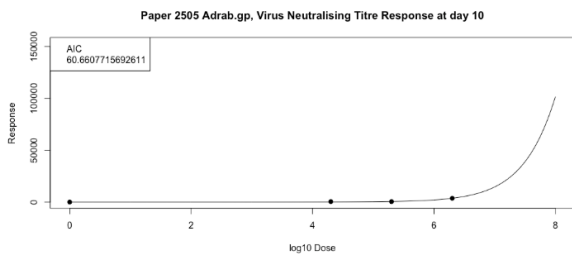
Day 14
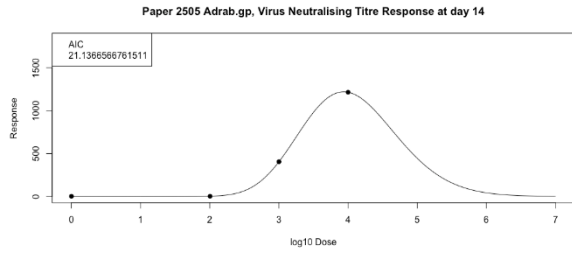
Host Species: Cattle

Route of Administration: IM

Paper 2841:

Day 7
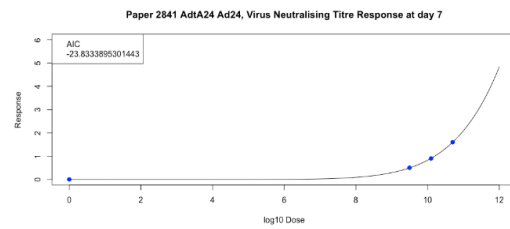
Vector Species: D

Host Species: Human

Route of Administration: IM

Paper 594:
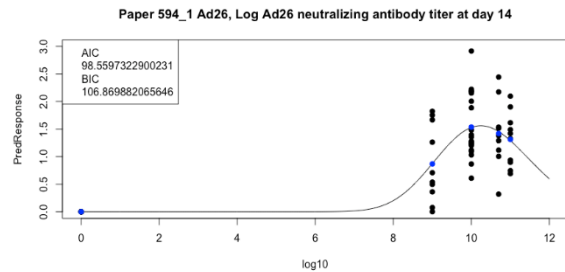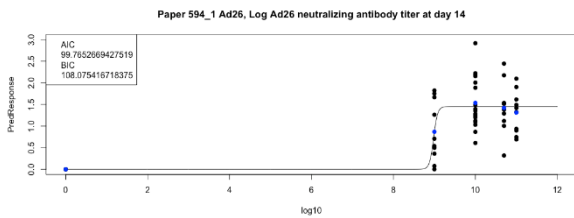
Day 14

Host Species: Rat

Route of Administration: IM

Paper 1343:

Day 28
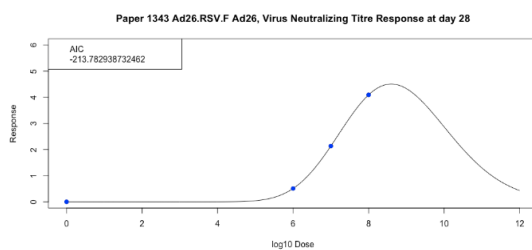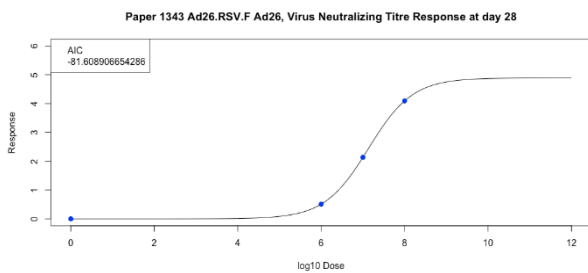
Vector Species: E

Host Species: Mouse

Route of Administration: SQ

Paper 2030:
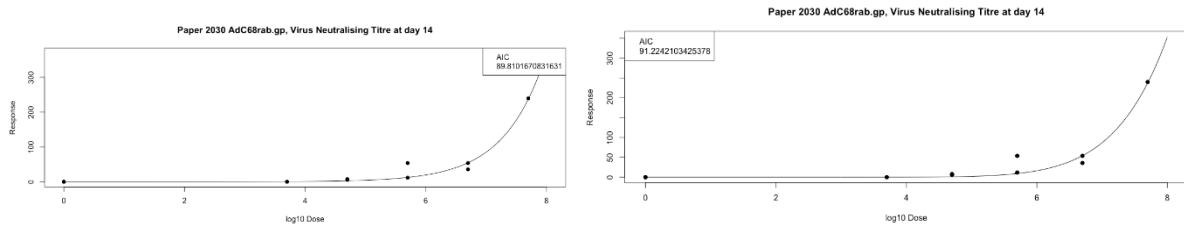
Day 14

Host Species: Human

Route of Administration: IM

Paper 686:

Day 14
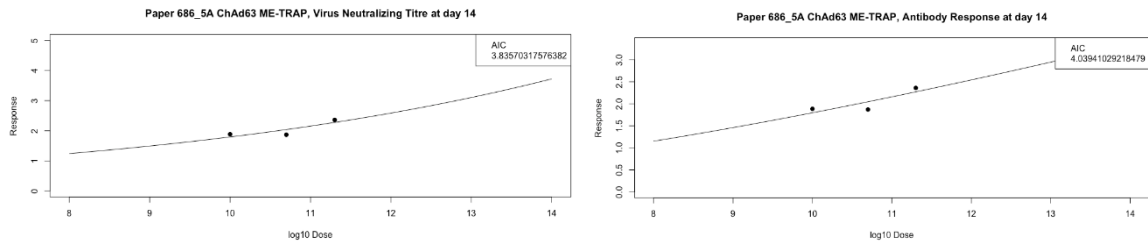
Day 21
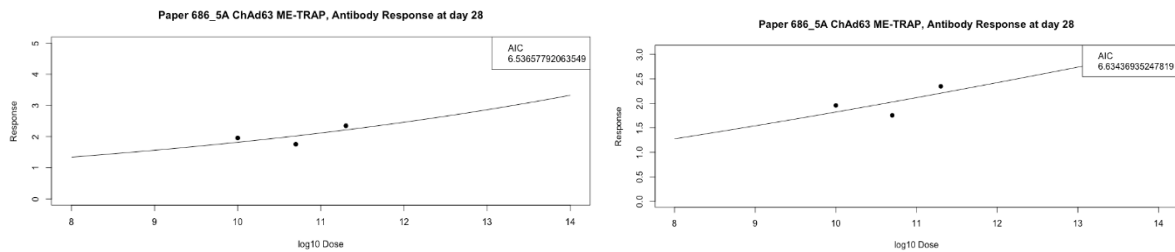
Figure S.Figures.B. The remaining plotted calibrated models. Left shows the saturating curve and right shows the peaking curve.

Supplementary Tables

T Cells									
Group (Adenoviral Species/Host Species/Administration Route)	Paper	Days post inoculation	Absolute Peaking Evidence (<10)	Strong Peaking Evidence (10-6)	Slight Peaking Evidence (2-6)	No Evidence	Slight Saturating Evidence (2-6)	Strong Saturating Evidence (10-6)	Strong Saturating Evidence (<10)
B/Mouse/IM	578	28				x			
		28				x			

460

		28				x			
		28	x						
		28				x			
	1269	56				x			
B/Human/IM	441	28		x					
	633	14				x			
		28				x			
C/Mouse/IM	2916	25				x			
		25				x			
		25				x			
		25					x		
		25					x		
		25					x		
C/Monkey/IM	1474	8				x			
		24				x			
		8				x			
		24				x			
	1877	28				x			
		168				x			
D/Mouse/IM	578	28				x			
		28				x			
		28				x			
		28						x	
		28				x			
	924	14				x			
D/Human/IM	594	14				x			
E/Mouse/IM	305	21				x			
E/Human/IM	417	14				x			
		21				x			
	686	14				x			
		21				x			
		90		x					
Count			1	2	0	27	3	1	0

Table S2. Evidence for T cell response

CD4									
Group (Adenoviral Species/Host Species/Administration Route)	Paper	Days post inoculation	Absolute Peaking Evidence (<10)	Strong Peaking Evidence (10-6)	Slight Peaking Evidence (2-6)	No Evidence	Slight Saturating Evidence (2-6)	Strong Saturating Evidence (10-6)	Strong Saturating Evidence (<10)
B/Mouse/IM	1201	14				x			
B/Human/IM	309	28	x						
		28		x					
C/Monkey/IM	1474	8				x			
		8				x			
D/Mouse/IM	924	14				x			
Count			1	1	0	4	0	0	0

Table S3. Evidence for CD4 response

Group (Adenoviral Species/Host Species/Administration Route)	Paper	Days post inoculation	Absolute Peaking Evidence (<10)	Strong Peaking Evidence (10-6)	Slight Peaking Evidence (2-6)	No Evidence	Slight Saturating Evidence (2-6)	Strong Saturating Evidence (10-6)	Strong Saturating Evidence (<10)
B/Mouse/IM	1201	14				x			
	1269	56				x			
	1492	14	x						
B/Mouse/SQ	555	14				x			
		21				x			
		28				x			
		35				x			
		70				x			
		23				x			
		70				x			
B/Human/IM	309	28	x						
		28			x				
C/Mouse/IM	461	7				x			
		14			x				

462

Group	No.	Day							
463		7	x						
		14				x			
	1039	8				x			
		8				x			
	1492	14	x						
	3018	7	x						
		15	x						
		30				x			
C/Mouse/SQ	555	14	x						
		21	x						
		28	x						
		35			x				
		70				x			
		23				x			
		70				x			
		14				x			
		21				x			
		28		x					
		35	x						
		70				x			
		23				x			
		70				x			
	2980	60				x			
C/Monkey/IM	1474	8				x			
		8				x			
D/Mouse/IM	924	14				x			
D/Mouse/SQ	555	14				x			
		21				x			
		28				x			
		35				x			
		70				x			
		23				x			
		70				x			

Group	Paper	Days	Absolute Peaking (<10)	Strong Peaking (10-6)	Slight Peaking (2-6)	No Evidence	Slight Saturating (2-6)	Strong Saturating (10-6)	Strong Saturating (<10)
E/Mouse/IM	2919	10			x				
		20				x			
E/Mouse/SQ	555	14							
		21							
		28			x				
		35							
		70							
		23							
		70							
G/Mouse/SQ	555	14							
		21	x						
		28			x				
		35		x					
		70		x					
		23				x			
		70				x			
Unknown(sAd16)/ Mouse/SQ	555	14				x			
		21				x			
		28				x			
		35				x			
		70				x			
		23				x			
		70			x				
Count			11	3	7	42	0	0	0

Table S4. Evidence for CD8 response

			CD4 IFN						
Group (Adenoviral Species/Host Species/Administration Route)	Paper	Days post inoculation	Absolute Peaking Evidence (<10)	Strong Peaking Evidence (10-6)	Slight Peaking Evidence (2-6)	No Evidence	Slight Saturating Evidence (2-6)	Strong Saturating Evidence (10-6)	Strong Saturating Evidence (<10)
B/Human/IM	309	28		x					
	441	28	x						
Count			1	1	0	0	0	0	0

Table S5. Evidence for CD5 IFN response

CD8 IFN									
Group (Adenoviral Species/Host Species/Administration Route)	Paper	Days post inoculation	Absolute Peaking Evidence (<10)	Strong Peaking Evidence (10-6)	Slight Peaking Evidence (2-6)	No Evidence	Slight Saturating Evidence (2-6)	Strong Saturating Evidence (10-6)	Strong Saturating Evidence (<10)
B/Human/IM	309	28				x			
B/Human/IM	441	28	x						
E/Mouse/IM	1801	10				x			
Count			1	0	0	2	0	0	0

Table S6. Evidence for CD8 IFN response

CD4 TNF									
Group (Adenoviral Species/Host Species/Administration Route)	Paper	Days post inoculation	Absolute Peaking Evidence (<10)	Strong Peaking Evidence (10-6)	Slight Peaking Evidence (2-6)	No Evidence	Slight Saturating Evidence (2-6)	Strong Saturating Evidence (10-6)	Strong Saturating Evidence (<10)
B/Human/IM	309	28			x				
B/Human/IM	441	28				x			
Count			0	0	1	1	0	0	0

Table S7. Evidence for CD4 TNF response

CD8 TNF									
Group (Adenoviral Species/Host Species/Administration Route)	Paper	Days post inoculation	Absolute Peaking Evidence (<10)	Strong Peaking Evidence (10-6)	Slight Peaking Evidence (2-6)	No Evidence	Slight Saturating Evidence (2-6)	Strong Saturating Evidence (10-6)	Strong Saturating Evidence (<10)
B/Human/IM	441	28			x				
Count			0	0	1	0	0	0	0

Table S8. Evidence for CD8 TNF response

CD4 IL2									
Group (Adenoviral Species/Host Species/Administration Route)	Paper	Days post inoculation	Absolute Peaking Evidence (<10)	Strong Peaking Evidence (10-6)	Slight Peaking Evidence (2-6)	No Evidence	Slight Saturating Evidence (2-6)	Strong Saturating Evidence (10-6)	Strong Saturating Evidence (<10)
B/Human/IM	309	28				x			
	441	28	x						
Count			1	0	0	1	0	0	0

Table S9. Evidence for CD4 IL2 response

CD8 IL2									
Group (Adenoviral Species/Host Species/Administration Route)	Paper	Days post inoculation	Absolute Peaking Evidence (<10)	Strong Peaking Evidence (10-6)	Slight Peaking Evidence (2-6)	No Evidence	Slight Saturating Evidence (2-6)	Strong Saturating Evidence (10-6)	Strong Saturating Evidence (<10)
B/Human/IM	441	28			x				
Count			0	0	1	0	0	0	0

Table S10. Evidence for CD8 IL2 response

CD4 IL17									
Group (Adenoviral Species/Host Species/Administration Route)	Paper	Days post inoculation	Absolute Peaking Evidence (<10)	Strong Peaking Evidence (10-6)	Slight Peaking Evidence (2-6)	No Evidence	Slight Saturating Evidence (2-6)	Strong Saturating Evidence (10-6)	Strong Saturating Evidence (<10)
B/Human/IM	309	28			x				
Count			0	0	1	0	0	0	0

Table S11. Evidence for CD8 IL17 response

Virus Neutralisation Titre									
Group (Adenoviral Species/Host Species/Administration Route)	Paper	Days post inoculation	Absolute Peaking Evidence (<10)	Strong Peaking Evidence (10-6)	Slight Peaking Evidence (2-6)	No Evidence	Slight Saturating Evidence (2-6)	Strong Saturating Evidence (10-6)	Strong Saturating Evidence (<10)
B/Mouse/SQ	2030	14				x			
	1269	56					x		
C/Mouse/IM	461	7				x			
		14				x			
	574	7				x			
		21				x			
		35	x						
		49				x			
		63				x			
		77				x			
		119			x				
		161				x			
	669	28				x			
	1039	8				x			
C/Mouse/SQ	2505	10				x			
		14				x			
C/Cattle/IM	2841	7						x	
D/Human/IM	594	14				x			
D/Rat/IM	594	28	x						
E/Mouse/SQ	2030	14				x			
E/Human/IM	686	14				x			
		21				x			
Count			2	0	1	17	1	1	0

Table S12. Evidence for Virus Neutralisation Titre response

S6. Weighted Utility Functions

We suggested than alternate approach to the utility function, where I weight the expected discomfort of a SARS-CoV-2 infection relative to the expected discomfort of receiving a vaccination. This approach requires defining such a weighting, which would require making additional assumptions and introducing complexity that I did not believe added to the main body of this work. Whilst establishing reasonable weightings are beyond the scope of this work, I suggest potential utility functions with pseudo-arbitrary values for the weighting. Hence, whilst these utility functions would not be useful presently for decision-making, if weights could accurately be determined they may be informative. Hence potential weighted utility functions are proposed. I note that the likely method of determining weighting is through a questionnaire of experts and decision makers or through group discussion, as is typical for determining weightings in multi-criteria decision analysis [9].

S6.1. 2:1 Ratio

The utility functions recommended in 3.4 and 3.5 assume that the only desirable outcome of vaccination is seroconversion without experiencing grade 3+ adverse events. This, implicitly, assumes that both seroconverting and avoiding grade 3+ adverse events are equally as desirable. Alternatively, I could consider all outcomes of vaccination with relative weightings of utility. I (pseudo-arbitrarily) choose a 2:1 weighting, where seroconversion is twice as desirable as avoiding a grade 3+ adverse event. The possible outcomes are namely;

- Not seroconverting or experiencing grade 3+ adverse events.
- Not seroconverting, experiences grade 3+ adverse events.
- Seroconversion, does not experience grade 3+ adverse events.
- Seroconversion, experiences grade 3+ adverse events.

The below table indicates the relative 'scores' of each of these outcomes.

Name	Experiences Seroconversion (+2)	Experience Grade 3+ Adverse Events (-1)	Score
$S'A'$	NO	NO	0 (= 0 + 0)
$S'A$	NO	YES	-1 (= 0 - 1)
SA'	YES	NO	2 (= 2 + 0)
SA	YES	YES	1 (= 2 - 1)

Table S3. Scores for the 2:1 utility function.

So defining PS as the probability of seroconversion, PS'=1- PS as the probability of no seroconversion, PA as the probability of experiencing grade 3+ adverse events, PA'=1- PA as the probability of not experiencing grade 3+ adverse events, I have the following utility function.

$U_{2:1}$(Dose) = Score(SA) ×PSPA +Score(SA') ×PSPA'+ Score(S'A) ×PS'PA+ Score(S'A') ×PS'PA'

Below is the dose-utility function for this utility function. I see that under this function and weighting the utility increases with dose, before decreasing. For sufficiently large doses the utility tends to 1 = Score(SA), as the model predicts 100% of individuals

experience seroconversion and grade 3+ adverse events.
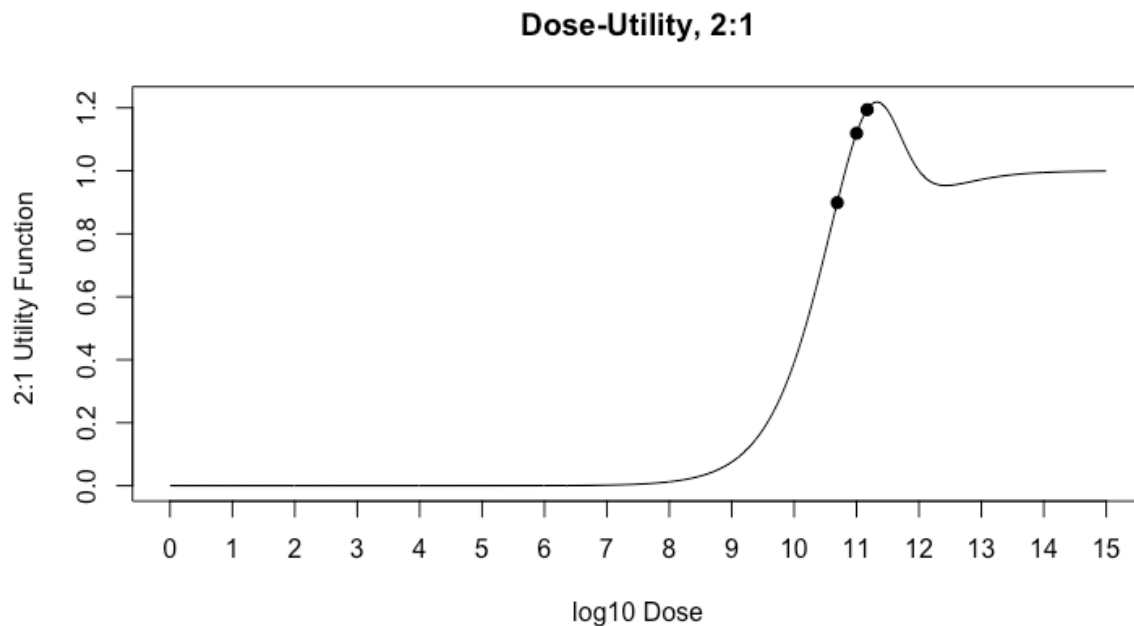
Dose-Utility, 2:1

Figure S13. Dose-Utility for the 2:1 utility function. Black dots represent the empirically tested doses.

S6.2. Expected discomfort

Alternatively, I can look to expressly minimise expected discomfort. I can consider that an individual has two sources of potential discomfort, namely discomfort arising from the vaccination and discomfort arising from the disease that the vaccine aims to prevent or minimise symptoms of.

We consider that for these two sources of potential discomfort, the discomfort could be rated as Mild (Grade 1,2), Severe (Grade 3), Critical without being fatal (Grade 4 if non-fatal), Critical and Fatal (Grade 4 if fatal). The probability for each of these outcomes if an individual's contracts SARS-CoV-2 are derived from literature [43], and the probability of mild or severe adverse events are estimated from the 'dose-any grade adverse event' and 'dose-grade 3+ adverse event' models discussed in the main body of this work. As I have no data to estimate the relationship between dose and the other two outcomes, I assume that the vaccine cannot cause either of these outcomes.

470

These outcomes are each assigned weights for discomfort, which are not based on literature but represent the idea that critical sickness or death are significantly worse outcomes than mild sickness.

Adverse Event	Vaccine probability	SARS-CoV-2 probability	Discomfort weight (pseudo-arbitrary)
Mild	Estimated in the paper as a function of dose, any grade adverse events, PA1	81%	1 = DWM
Severe	Estimated in the paper as a function of dose, grade 3+ adverse events, PA3+	14%	5= DWS
Critical (Non-fatal)	Assumed 0%	2.7%	50 = DWC
Critical (Fatal)	Assumed 0%	2.3%	100 =DWF

Table S4. Scores for the expected discomfort utility function.

We can define the expected discomfort of contracting SARS-CoV-2 as

$$ExpectedDiscomfort_{SARS}$$
$$= DWM \times 0.81 + DWS \times 0.14 \ + \ DWC \times .027 \ + \ DWF \times .023$$

$$ExpectedDiscomfort_{SARS} = 5.16$$

We can also estimate that an individual has a 65.5% (=0.655) (the herd immunity threshold) probability of contracting SARS-CoV-2 if they are not protected. However, this may be reduced depending on the percentage of individuals in the population that have previously contracted or received a vaccine for SARS-CoV-2 (which could be investigated by considering epidemiological models).

Hence a vaccinated individual is predicted to experience expected discomfort as a function of dose:

Expected Discomfort (Dose) = PA1(Dose) × DWM + PA3+(Dose) × DWS + 0.655 × PS'(Dose) × ExpectedDiscomfort$_{Sars}$

Where PS'(Dose) is the probability of not seroconverting and hence not being protected as a function of dose. S14 shows this relationship.
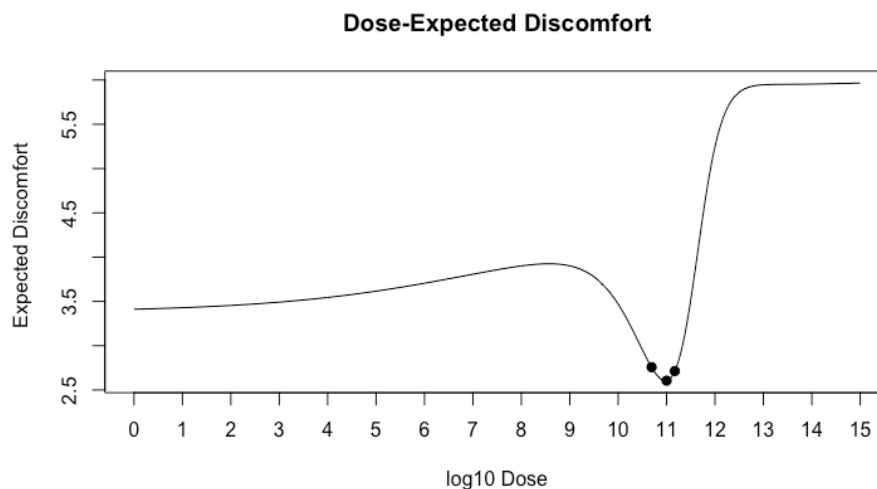
Figure S14. Dose-Utility for the expected discomfort utility function. Black dots represent the empirically tested doses.

For these weights, the following behaviour is observed. As the dose increases from 0, the increasing discomfort of vaccination, whilst small, is not justified by the possible reduction in SARS-CoV-2 risk, as there is no meaningful level of seroconversion. For doses at approximately 10^{11}, I see a reduction in expected discomfort. At higher doses, whilst seroconversion is probable, the probability of grade 3+ adverse events is large enough that vaccination at this dose may be considered to be less comfortable than the average SARS-CoV-2 infection.

We can also consider the expected reduction in discomfort from baseline by subtracting the dose-dependent expected discomfort from the zero-dose expected discomfort.

472

Expected Discomfort Reduction (Dose) = 0.655 ×PS'(0) ×ExpectedDiscomfort$_{Sars}$ - PA1(Dose) × DWM - PA3+(Dose) × DWS - 0.655 ×PS'(Dose) ×ExpectedDiscomfort$_{Sars}$

Dose-Discomfort Reduction
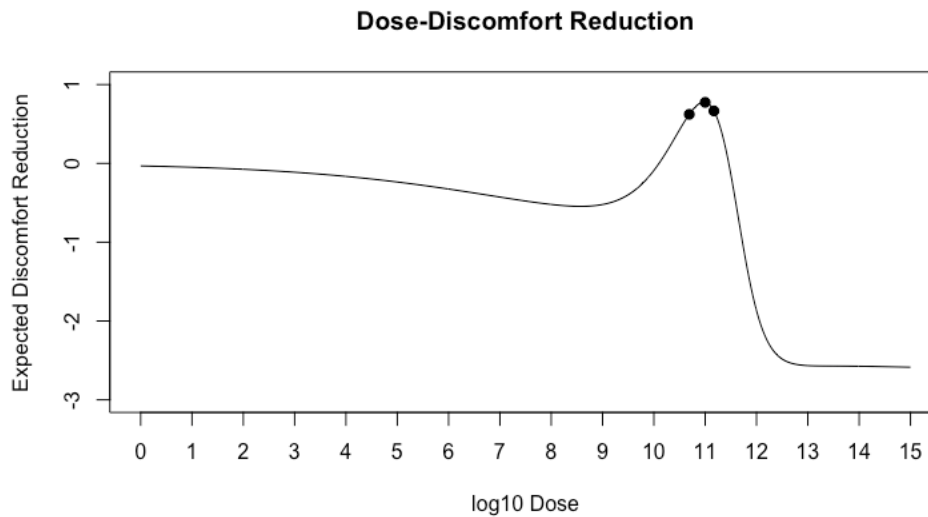
Figure S15. Baseline subtracted Dose-Utility for the expected discomfort utility function. Black dots represent the empirically tested doses.

Further, I can consider only the doses where the discomfort reduction is predicted to be greater than 0. Hence, by dividing by the 'dose-cost' model found in the main body of this work I can also estimate the expected reduction in discomfort per GBP spent on the vaccine at each dosing level.

473

Dose-Discomfort Reduction, Greater than 0, per Cost
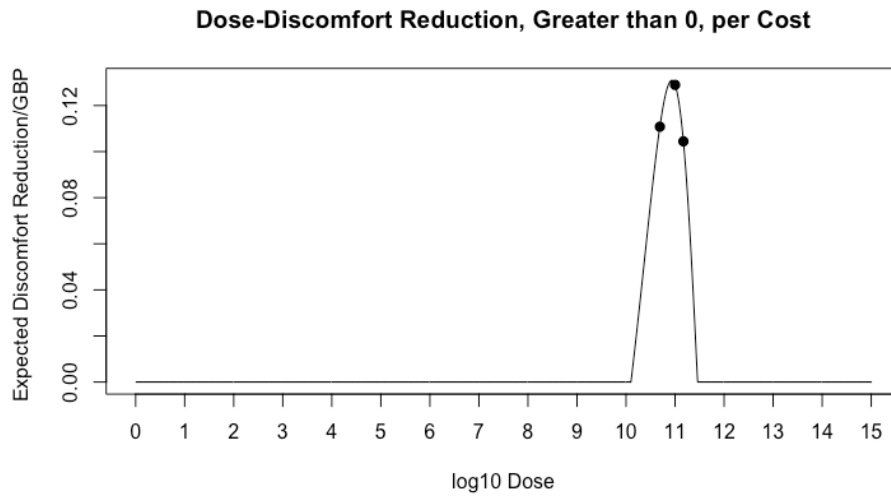
Figure S16. Baseline subtracted Dose-Utility for the expected discomfort utility function, divided by cost and censored if discomfort reduction is less than 0. Black dots represent the empirically tested doses.

S4. Scenarios

Scenario Saturating 2

Qualitatively this scenario had a saturating efficacy curve and a middling optimal dose, with a relatively broad utility curve.

|	Parameter	Value
Efficacy	gradient	1.800
	midpoint	2.500
	maximum	0.900
Toxicity	gradient	0.500
	threshold1	1.000
	threshold2	4.000
	threshold3	5.000
Utility Weights	Weight$_{Efficacy}$	0.133

Table.S.4.2 .Saturating 2. Parameters for the scenario Saturating 2
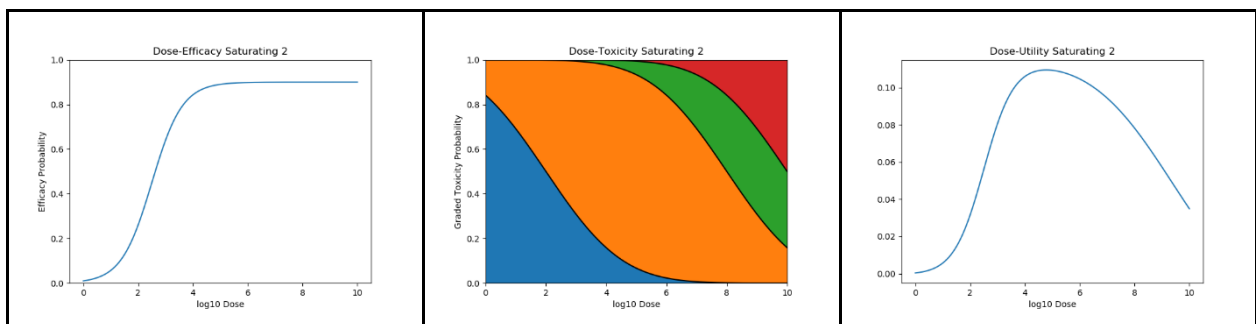
Figure.S.4.2. Saturating 2. Dose-efficacy, dose-toxicity, and dose-utility plots for the scenario Saturating 2

Scenario Saturating 3

Qualitatively this scenario had a saturating efficacy curve and a low optimal dose, with a relatively steep utility curve.

	Parameter	Value
Efficacy	gradient	2.500
	midpoint	2.000
	maximum	0.900
Toxicity	gradient	0.500
	threshold1	0.100
	threshold2	2.500
	threshold3	3.000
Utility Weights	Weight$_{Efficacy}$	0.133

**Table.S.4.3.Saturating 3. Parameters for the scenario Saturating 3**
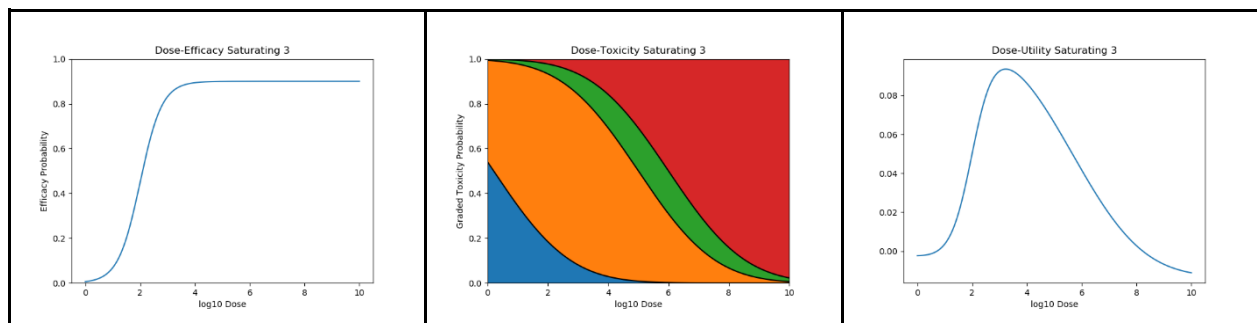
Figure.S.4.3.Saturating 3. Dose-efficacy, dose-toxicity, and dose-utility plots for the scenario Saturating 3

476

Qualitatively this scenario had a saturating efficacy curve and a high optimal dose at 10, representing the case where the 'true optimal' is not within the dosing space.

|	Parameter	Value
Efficacy	gradient	0.700
	midpoint	7.500
	maximum	0.900
Toxicity	gradient	0.500
	threshold1	1.000
	threshold2	2.000
	threshold3	5.000
Utility Weights	Weight$_{Efficacy}$	.266

Table.S.4.4.Saturating 4. Parameters for the scenario Saturating 4
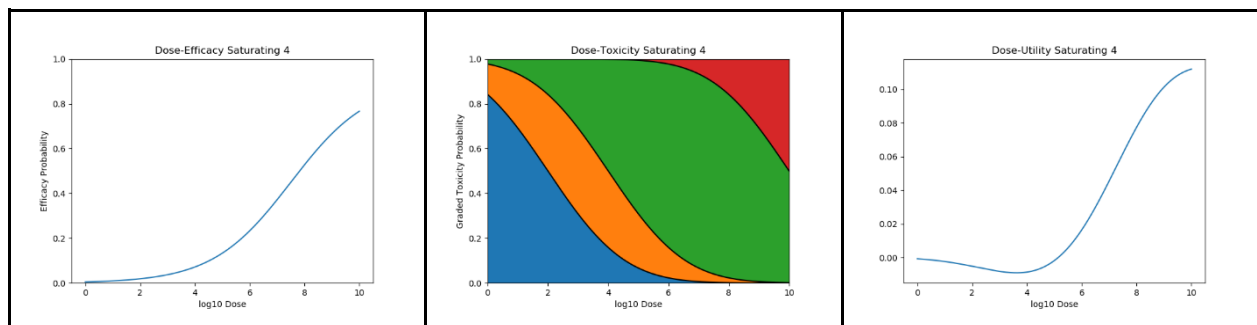
Figure.S.4.4.Saturating 4. Dose-efficacy, dose-toxicity, and dose-utility plots for the scenario Saturating 4

Scenario Saturating 5

Qualitatively this scenario had a saturating efficacy curve, which changed only gradually over the dosing space.

	Parameter	Value
Efficacy	gradient	0.100
	midpoint	8.000
	maximum	1.000
Toxicity	gradient	0.500
	threshold1	1.000
	threshold2	4.000
	threshold3	5.000
Utility Weights	Weight$_{Efficacy}$	.266

Table.S.4.5.Saturating 5. Parameters for the scenario Saturating 5
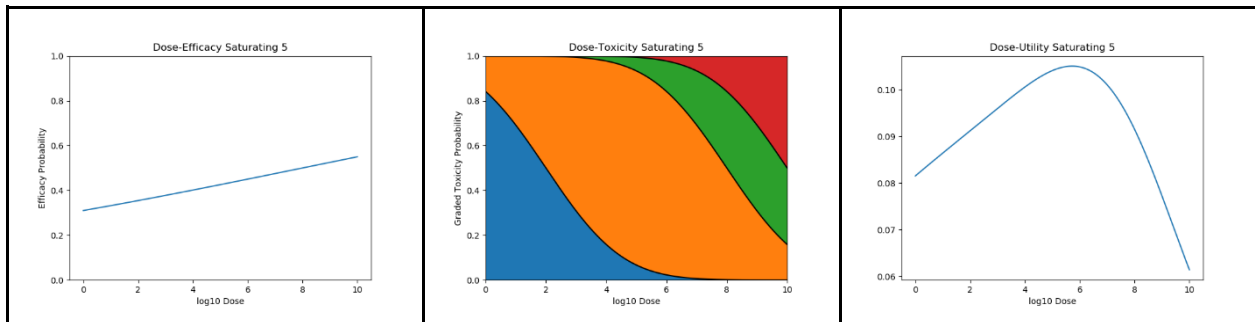
Figure.S.4.5.Saturating 5. Dose-efficacy, dose-toxicity, and dose-utility plots for the scenario Saturating 5

478

Scenario Peaking 1

Qualitatively this scenario had a peaking efficacy curve and a high optimal dose, with a relatively steep utility curve.

	Parameter	Value
Efficacy	base	-9.000
	gradient1	3.000
	gradient2	-0.214
Toxicity	gradient	1.000
	threshold1	3.000
	threshold2	9.000
	threshold3	10.500
Utility Weights	Weight$_{Efficacy}$	0.133

Table.S.4.6.Peaking 1. Parameters for the scenario Peaking 1
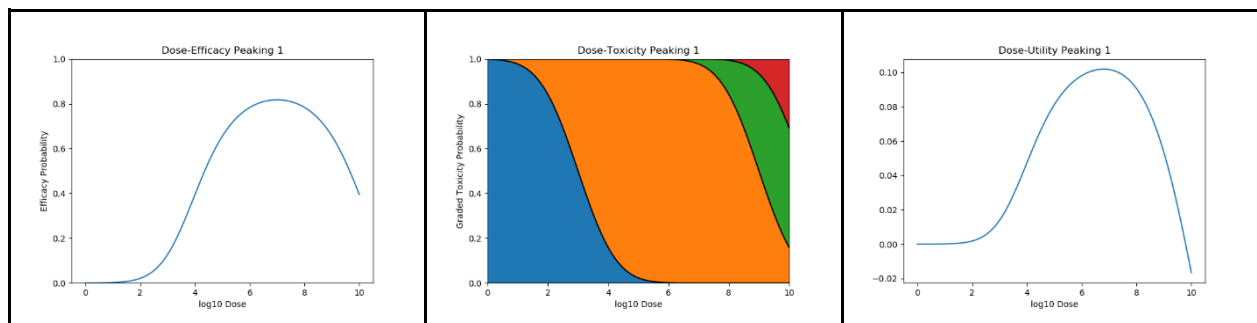
Figure.S.4.6.Peaking 1. Dose-efficacy, dose-toxicity, and dose-utility plots for the scenario Peaking 1

Scenario Peaking 2

Qualitatively this scenario had a peaking efficacy curve and a middling optimal dose, with a relatively broad utility curve.

	Parameter	Value
Efficacy	base	-4.000
	gradient1	2.000
	gradient2	-0.166
Toxicity	gradient	0.100
	threshold1	0.100
	threshold2	0.400
	threshold3	1.500
Utility Weights	Weight$_{Efficacy}$	0.133

Table.S.4.7.Peaking 2. Parameters for the scenario Peaking 2
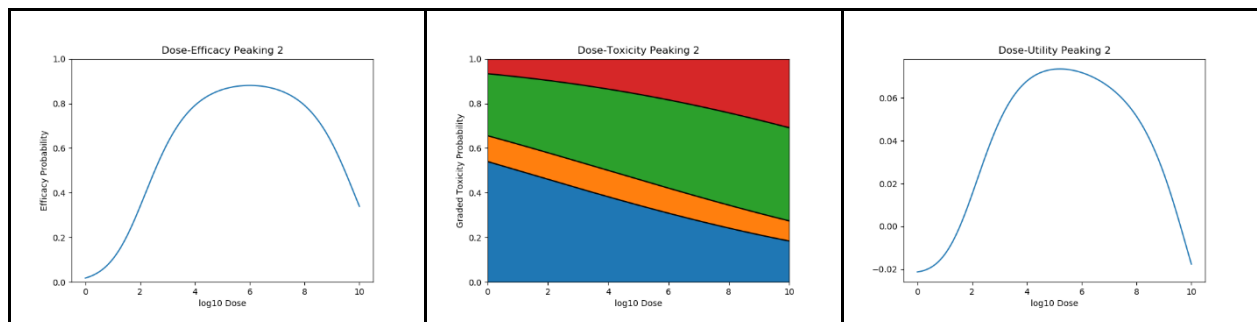
Figure.S.4.7.Peaking 2. Dose-efficacy, dose-toxicity, and dose-utility plots for the scenario Peaking 2

Qualitatively this scenario had a peaking efficacy curve and a high optimal dose at 10, representing the case where the 'true optimal' is not within the dosing space.

Peaking 4 uses a peaking model, but efficacy is still increasing at the maximum dose. Thus this is effectively a saturating scenario, but represents the potential case where dose-efficacy is peaking but this is unimportant within the feasible dosing space.

	Parameter	Value
Efficacy	base	-12.000
	gradient1	2.500
	gradient2	-0.114
Toxicity	gradient	.300
	threshold1	1.000
	threshold2	1.500
	threshold3	2.000
Utility Weights	Weight$_{Efficacy}$	0.266

Table.S.4.9.Peaking 4. Parameters for the scenario Peaking 4
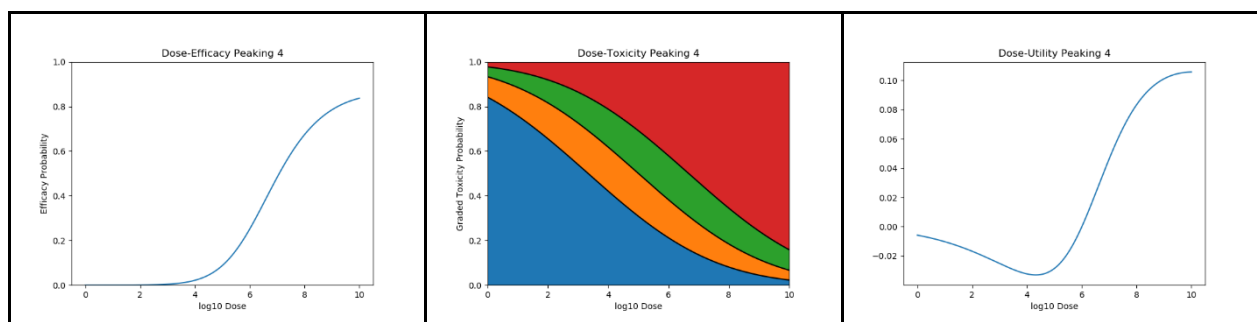
Figure.S.4.9.Peaking 4. Dose-efficacy, dose-toxicity, and dose-utility plots for the scenario Peaking 4

481

Scenario Peaking 5

Qualitatively this scenario had a peaking efficacy curve, which changed only gradually over the dosing space.

	Parameter	Value
Efficacy	base	0.000
	gradient1	0.800
	gradient2	0.067
Toxicity	gradient	0.100
	threshold1	0.100
	threshold2	0.400
	threshold3	1.500
Utility Weights	Weight$_{Efficacy}$	0.133

Table.S.4.10.Peaking 5. Parameters for the scenario Peaking 5
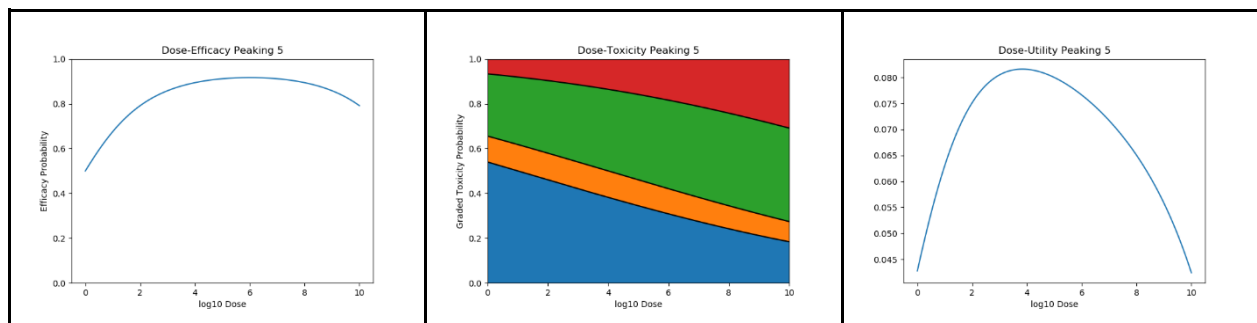
Figure.S.4.10.Peaking 5. Dose-efficacy, dose-toxicity, and dose-utility plots for the scenario Peaking 5

Scenario Other 1

Other 1 represents a vaccine for which nearly zero efficacy is observed for all doses. The efficacy model is given as

Flat(dose) = base

	Parameter	Value
Efficacy	base	0.020
Toxicity	gradient	0.500
	threshold1	0.000
	threshold2	3.000
	threshold3	5.000
Utility Weights	Weight$_{Efficacy}$	0.133

Table.S.4.11.Other 1. Parameters for the scenario Other 1
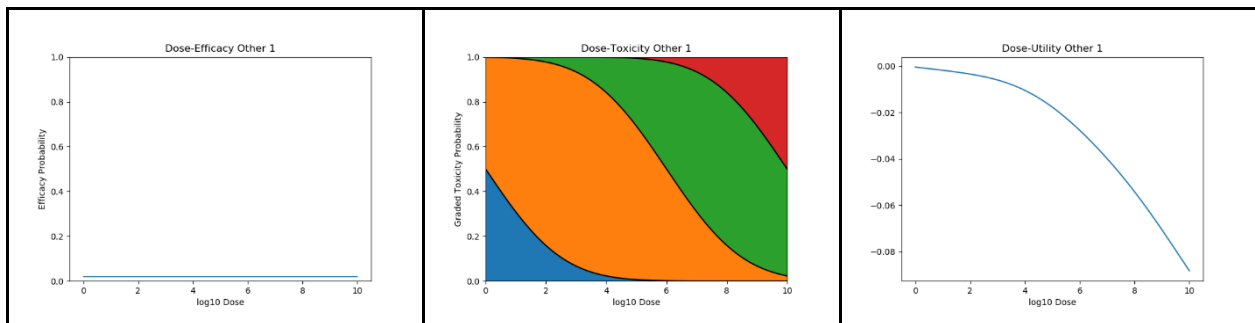
Figure.S.4.11.Other 1. Dose-efficacy, dose-toxicity, and dose-utility plots for the scenario Other 1

483

Other 3 represents a vaccine for which dose-efficacy is fundamentally peaking, but follows a different and more complicated biphasic parametric form to the latent quadratic saturating model assumed elsewhere in this paper.

	Parameter	Value
Efficacy	gradient1	1.000
	gradient2	2.000
	midpoint1	4.000
	midpoint2	7.000
	maximum	0.500
	fraction	2.000
Toxicity	gradient	0.500
	threshold1	1.000
	threshold2	3.000
	threshold3	5.000
Utility Weights	Weight$_{Efficacy}$	0.133

Table.S.4.13.Other 3. Parameters for the scenario Other 3
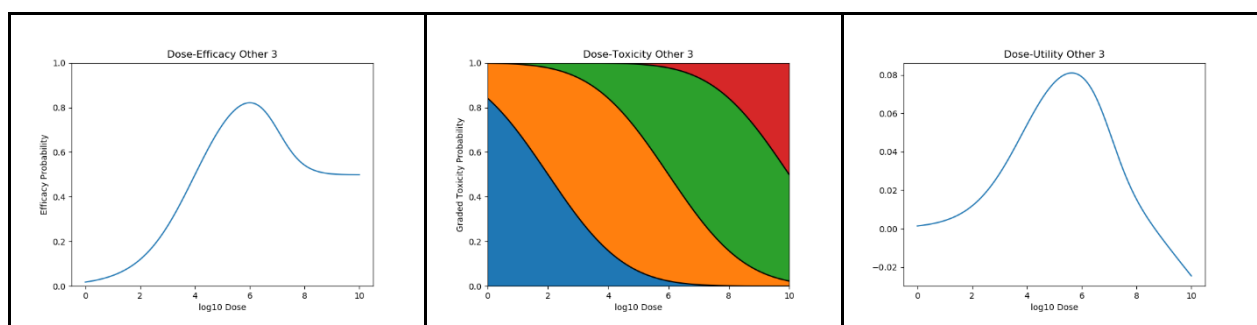
Figure.S.4.13.Other 3. Dose-efficacy, dose-toxicity, and dose-utility plots for the scenario Other 3

Scenario Other 4

Other 4 represents a vaccine for which dose-efficacy is fundamentally saturating, but follows yet another different parametric saturating model, the linear model. The efficacy model is given as

$$linear(dose) = \frac{maximum \times dose}{gradient + dose}$$

	Parameter	Value
Efficacy	gradient	3.000
	maximum	1.200
Toxicity	gradient	0.200
	threshold1	1.000
	threshold2	1.200
	threshold3	2.000
Utility Weights	Weight$_{Efficacy}$	0.133

Table.S.4.14.Other 4. Parameters for the scenario Other 4
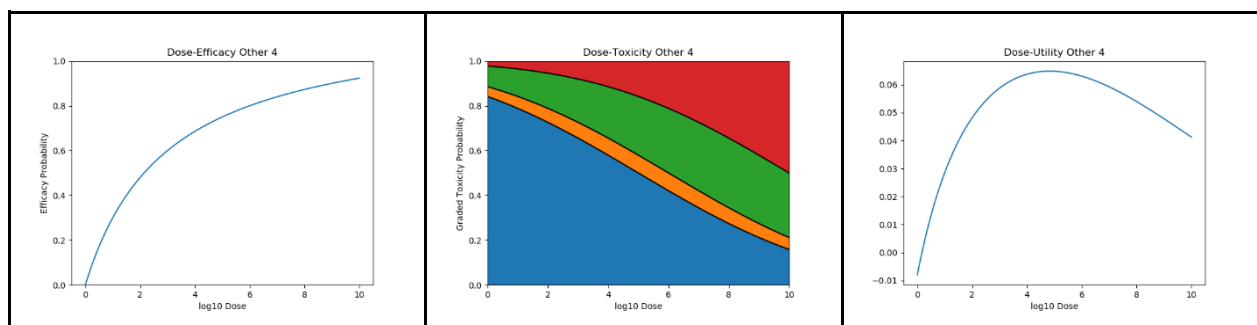
Figure.S.4.14.Other 4. Dose-efficacy, dose-toxicity, and dose-utility plots for the scenario Other 4.

S9.3. Optimistic Bias: Simplified modelling

We show that it is intuitive that overestimation may occur in an example simplified model case. Consider again an example of trying to optimise some continuous function of dose-utility. I use a modified bell curve as the parameters are more interpretable and assume that both the true function and assumed function follow this. The parameters of this function are midpoint, maximum, and scale. Midpoint is the point where the peak of the dose-utility curve occurs, maximum is the value of utility for this dose, and scale widens or shrinks the bell curve around this point. Say the true curve is defined by:

$$\text{Midpoint}_{true} = 5$$

$$\text{Maximum}_{true} = 3$$

$$\text{Scale}_{true} = 1$$

And thus looks like:
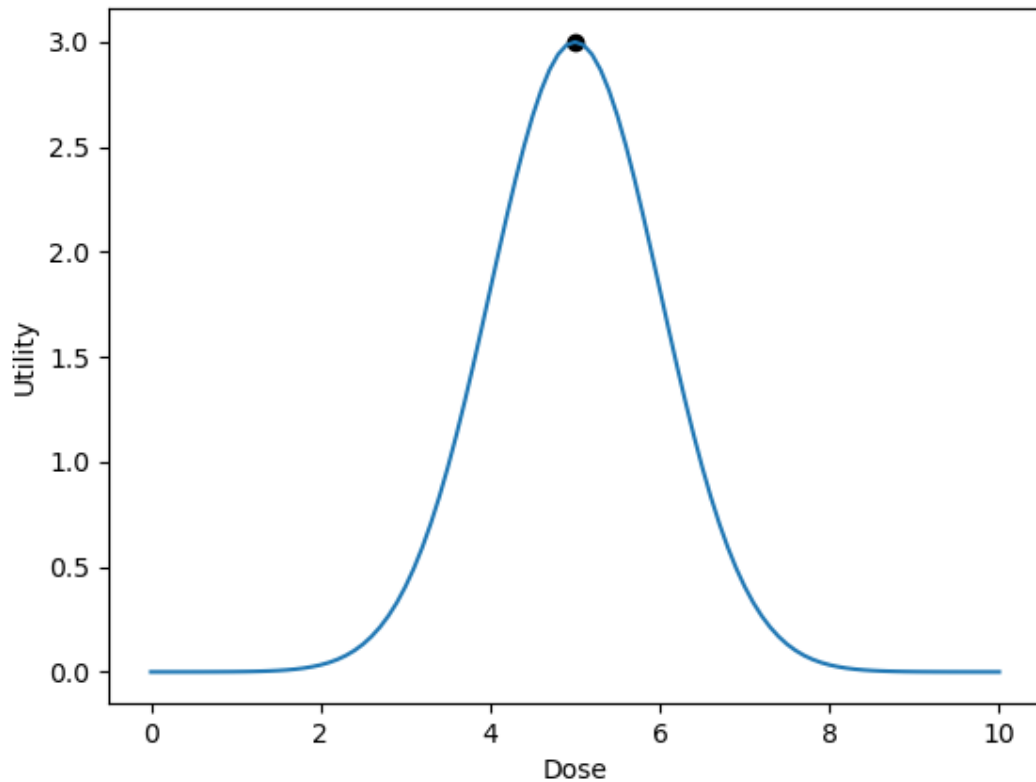
Figure S9.1. True Dose Utility Curve in toy example. Black point represents optimal dose and response.

An experiment is conducted, and estimates are calculated for these model parameters

$$error_i \sim \text{N}(0, errorscale)$$

$$midpoint_{estimate} = midpoint_{true} + error_{midpoint}$$

$$maximum_{estimate} = maximum_{true} + error_{maximum}$$

$$scale_{estimate} = scale_{true} + error_{scale}$$

Results of underestimating/overestimating each of these values individually are shown in the figure below. An overestimation would occur in 3 of the 6 cases, and an underestimation in only 1 of the 6. See
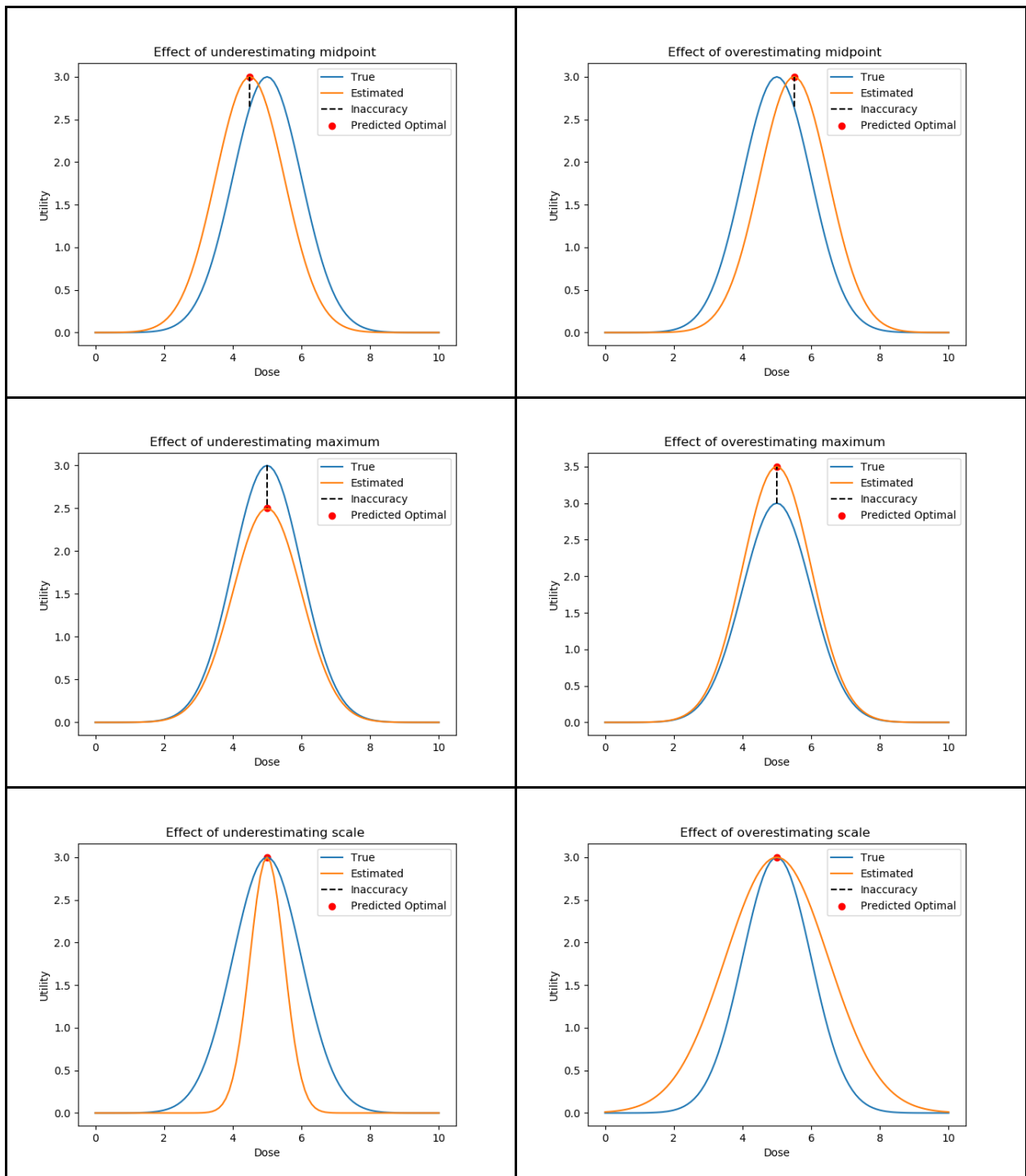
Figure S9.2. Demonstrations for how for some model overestimation can arise from parameter uncertainty.

Running simulations of 10000 clinicals trials with $errorscale$ =0.5 for all parameters, I found that in 68.63% of simulation there was overestimation. The figure below should suggest why the quasi-convex nature of the dose-utility curve might lead to overestimation even if the parameter errors are normally distributed around the true value (unbiased) and the distribution of predicted optimal dose/response is a multivariate normal around the true

488

value.
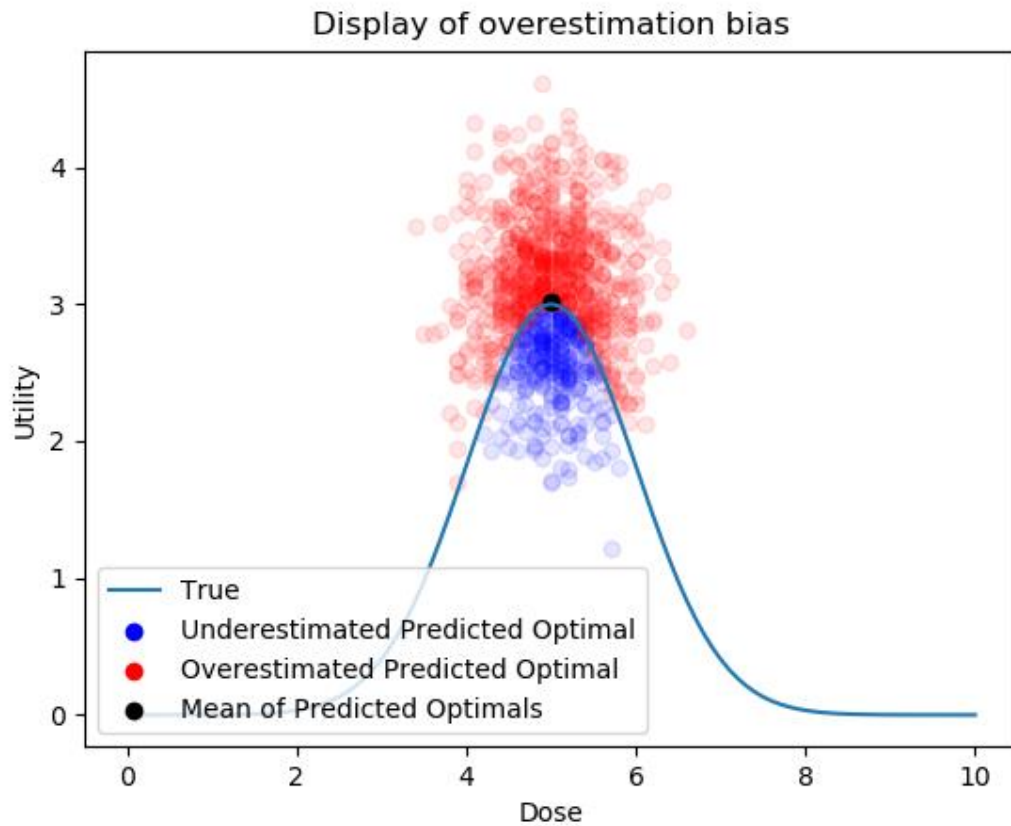
Figure S9.3. Display of overestimation bias. Each point represents a 'clinical trial' where instead of considering uncertainty in parameters I assume that the predicted optimal dose and predicted optimal response are normally distributed around the true optimal. In this case the quasi-convex shape of the utility curve ensures that there is an optimistic bias.

We repeated the simulation of clinical trials with parameter uncertainty for errorscale between 0 and 0.5, and I found that increasing the errorscale increases the overestimation bias, but that at least some bias is observed for all $\text{errorscale} > 0$

489
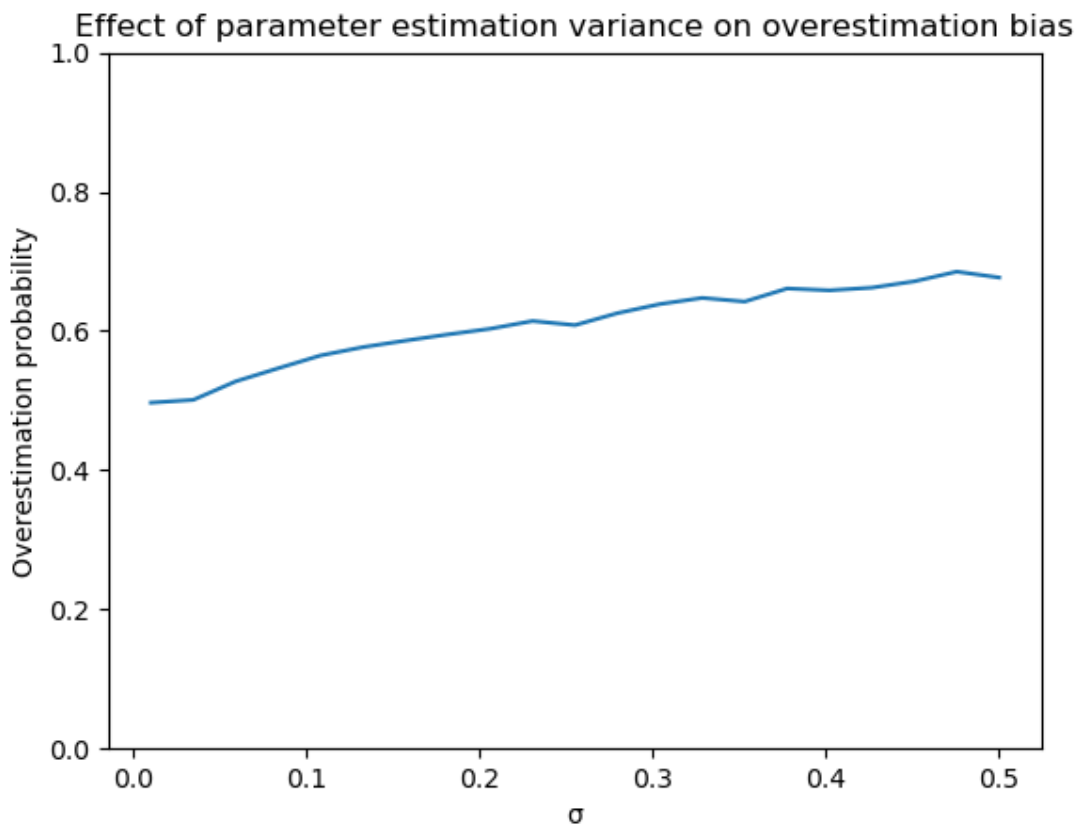
Figure S9.4. Effect on errorscale on overestimation bias in the simplified modelling setting

S9.4. Optimistic Bias: Physical modelling

To take a break from questions of dose to touch on a different example of a non-trivial optimisation problem. Consider the toy problem of attempting to choose optimal battery size for an unmanned aerial vehicle ('drone'), where 'optimal' means the battery such that the drone can fly furthest. A heavily simplified physical model is built, see figure S9.5.
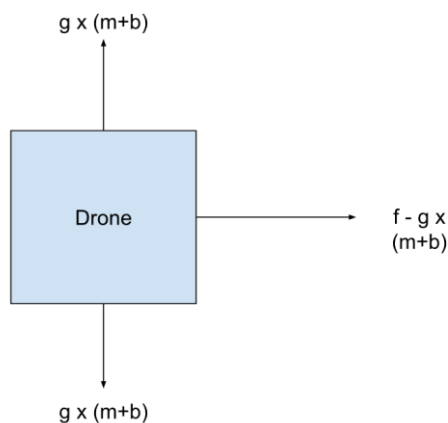
Figure S9.5. A visual depiction of the model forces of drone flight used in this section.

The distance the drone flies is determined by the following 5 parameters.

g: the acceleration due to gravity

m: The base mass of the drone

a: drain rate, the rate at which the battery drains

f: The force that the drone is capable of generating to fly (will be divided between maintaining height and moving forward)

b: the size of the battery, nondimensionalised to reflect both mass and capacity of the battery.

Increasing b increases the time taken to drain the battery, but also increases the weight of the drone, reducing the amount of force that can be used to move forward.

It can be shown that the distance travelled (s) given g, m, a, f, and b is

$$s = \frac{(f - g(m + b))b}{(m + b)a}$$

and hence that optimal b is found for $\frac{\partial s}{\partial b} = 0$ and hence the optimal b is equal to

$$\max\left(-m - \sqrt{\frac{fm}{g}}, -m + \sqrt{\frac{fm}{g}}\right)$$

A researcher estimates g, m, a, and f from available data, the results of which are normally distributed around the true values with some error scale. That is to say, with

$$\{error\}_i \sim N(0, errorscale)$$

So

491

$$g_{\{estimate\}} = g_{\{true\}} + \{error\}_g$$

$$f_{\{estimate\}} = f_{\{true\}} + \{error\}_f$$

$$m_{\{estimate\}} = m_{\{true\}} + \{error\}_m$$

$$a_{\{estimate\}} = a_{\{true\}} + \{error\}_a$$

These estimations are then used to predict optimal battery size and hence the predicted maximal distance the drone should travel with that battery size. Again an overestimation bias observed that again depends on the variance of the parameters estimates from the true values [Figure S9.6].
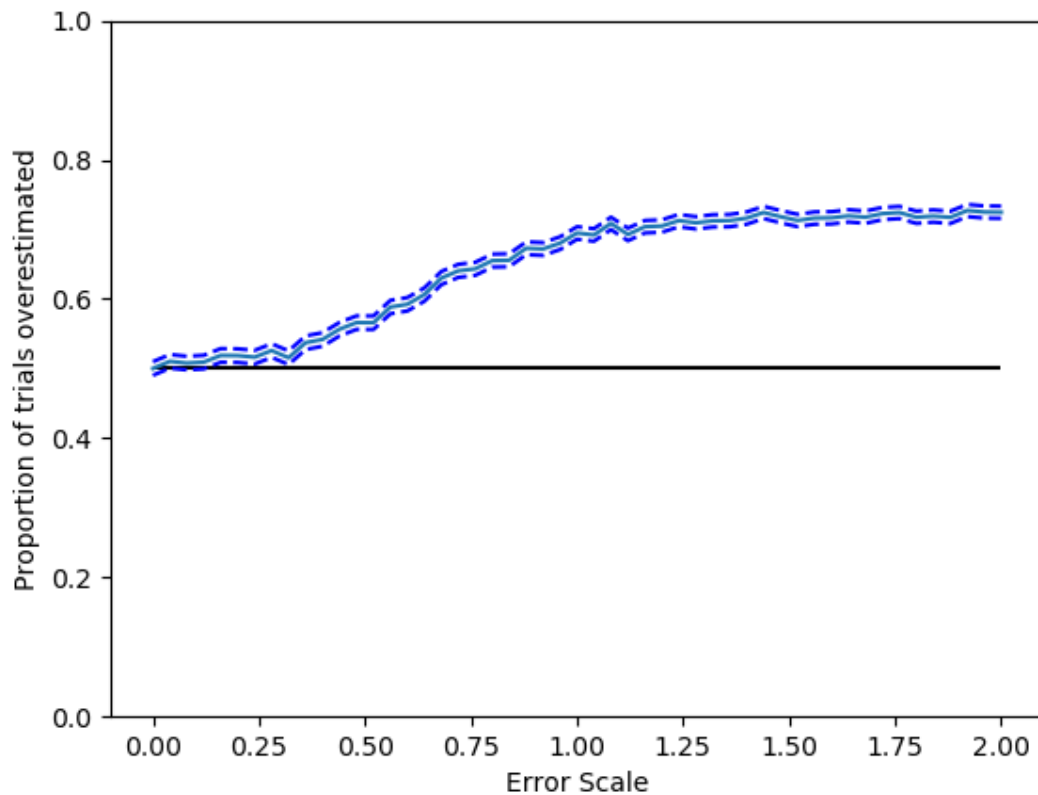
Figure S9.6. The effect of increasing the variance of normally distributed errors in parameters for the drone optimisation problem. 10,000 simulations of parameter estimation, optimisation, and comparison to true flight distance for the chosen battery size were done for each error scale value. 95% confidence bounds on the overestimation proportion are in dashed blue.

Scenario Saturating 2
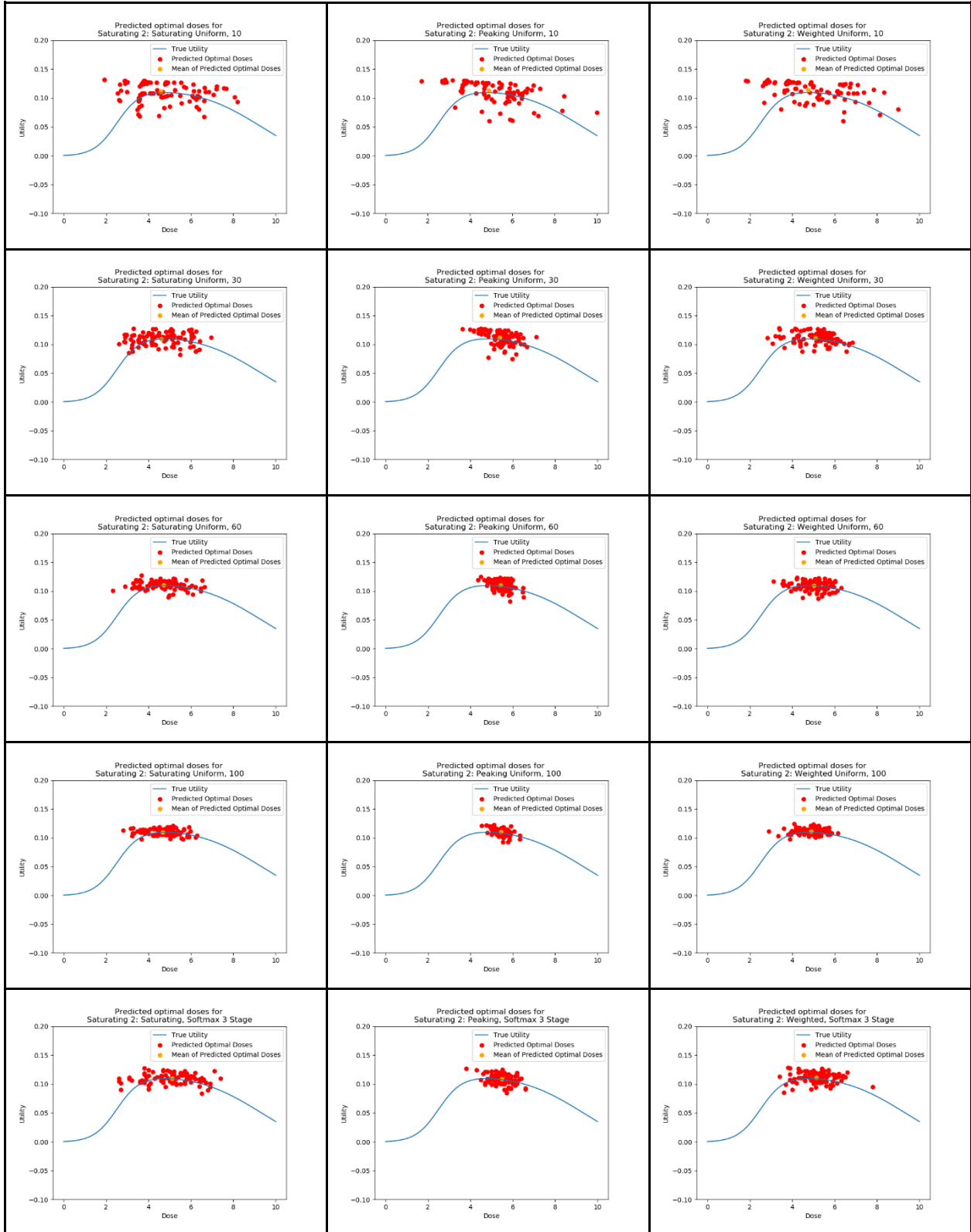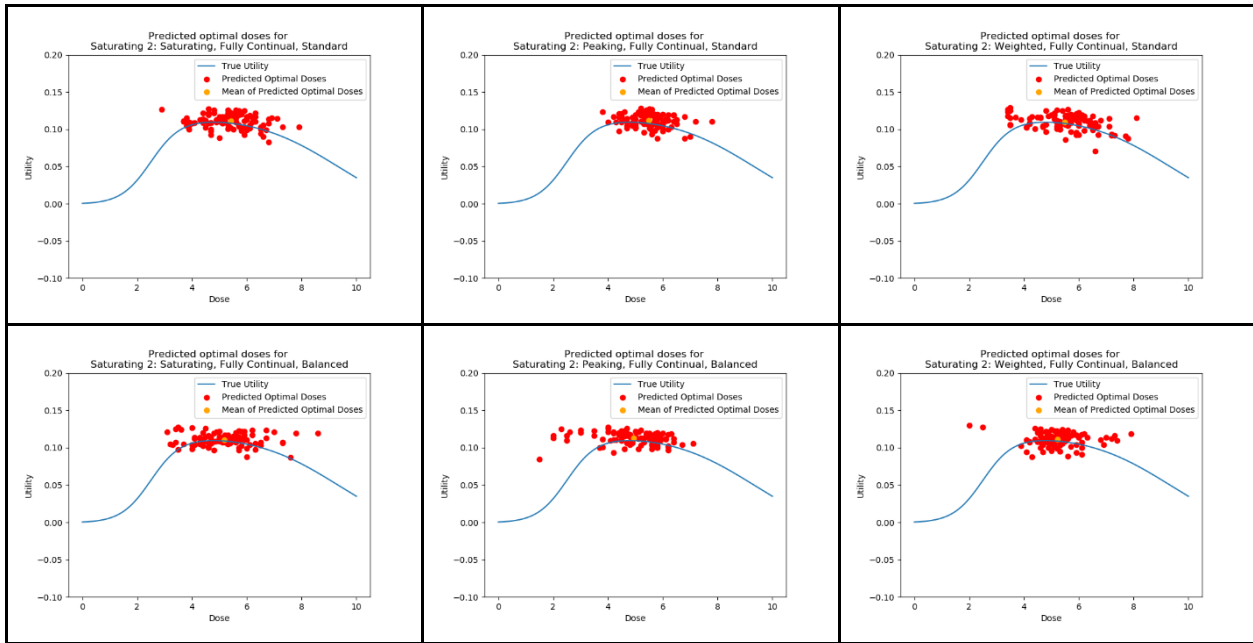
Figure S.10.2. Clinical trials by dose optimisation approach for scenario Saturating 2

Scenario Saturating 3

494
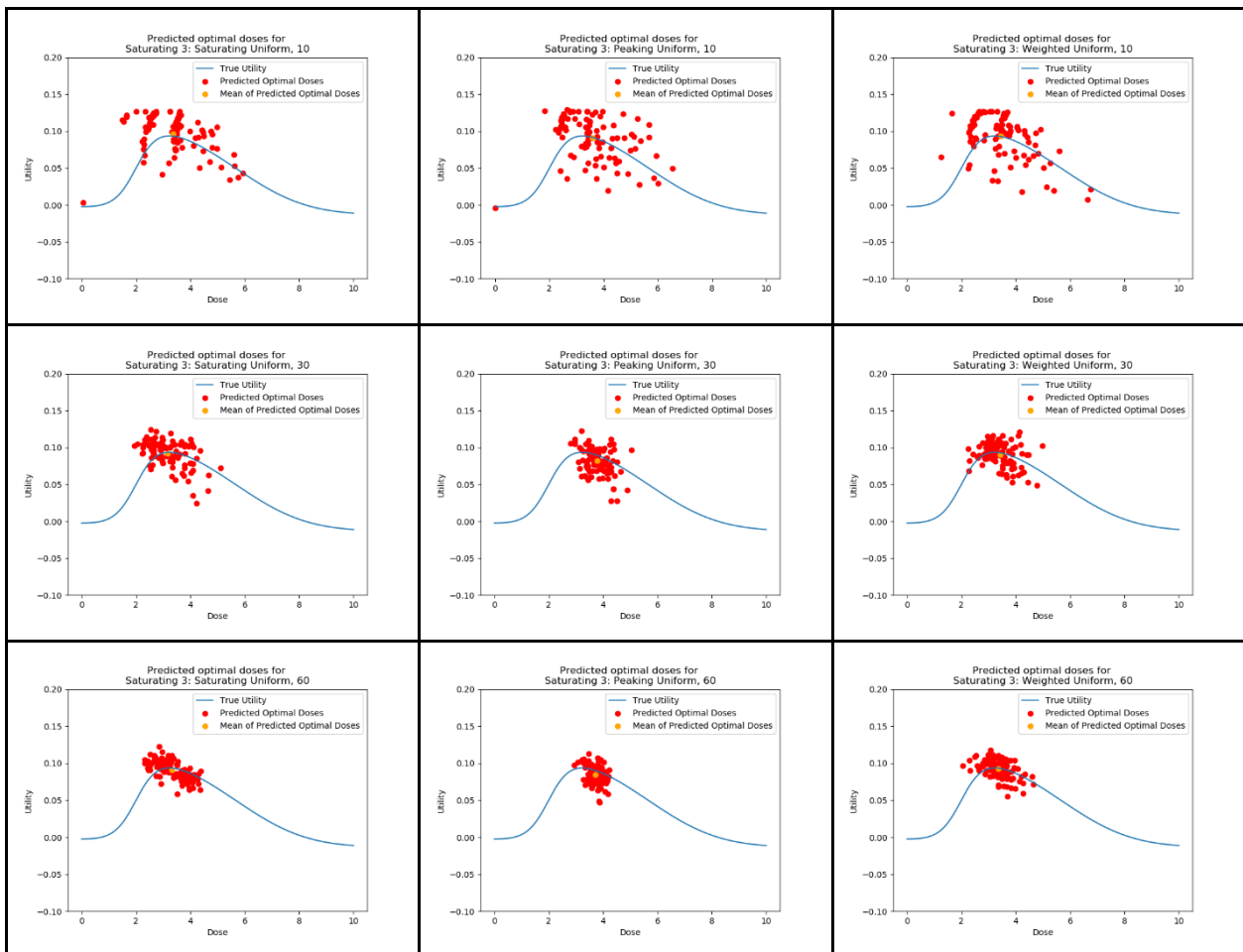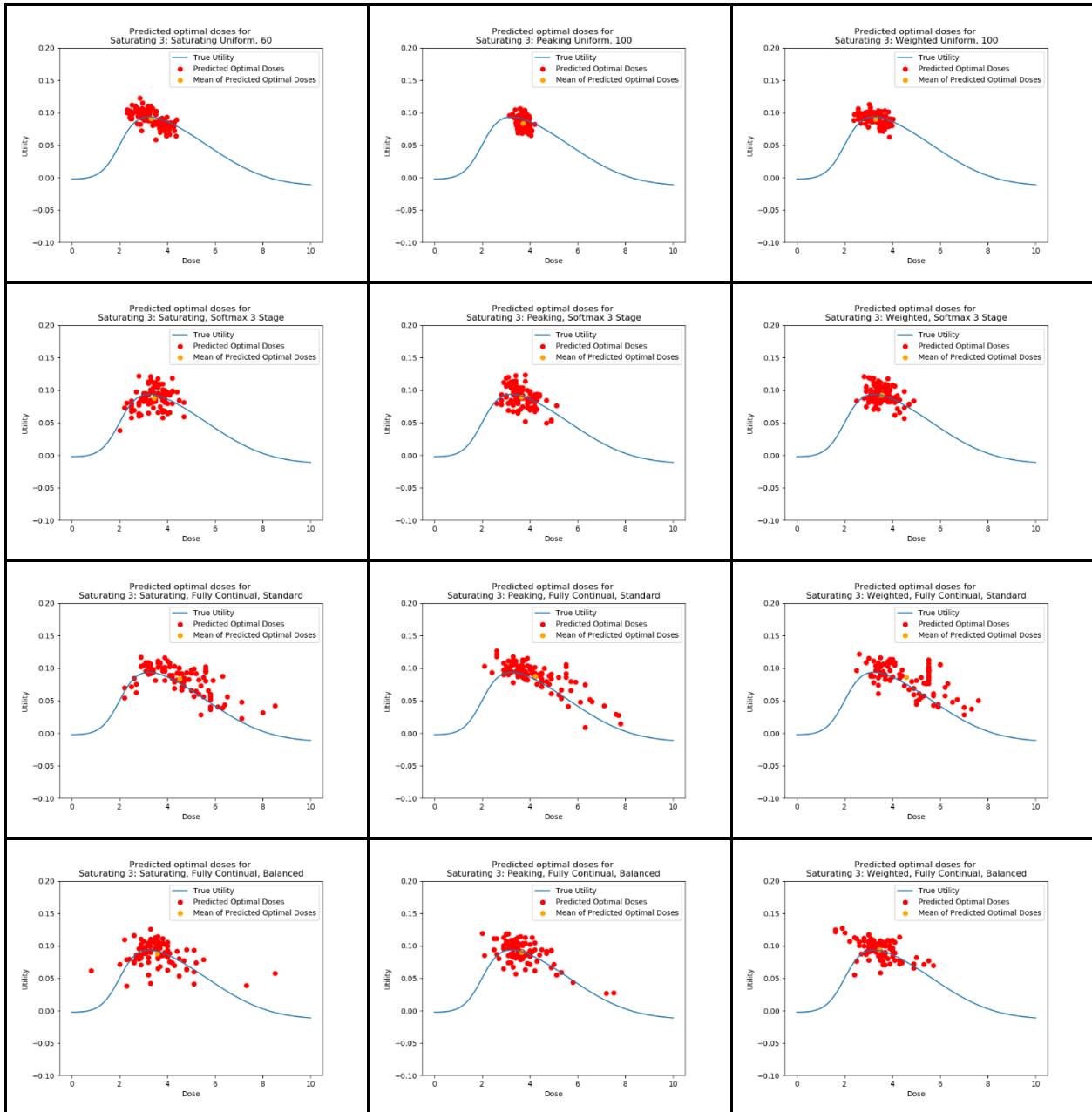
Figure S.10.3. Clinical trials by dose optimisation approach for scenario Saturating 3
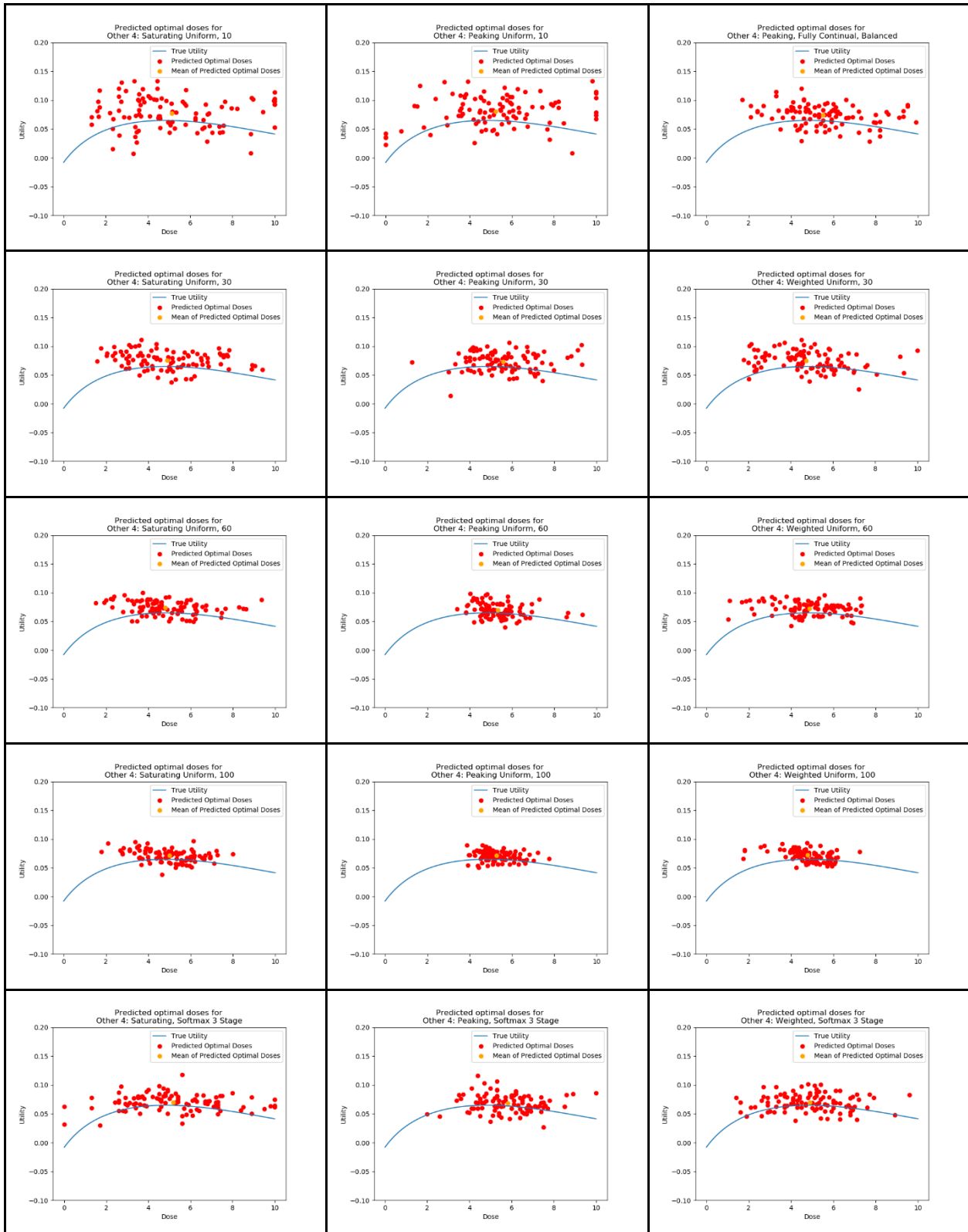
Scenario Saturating 4

496
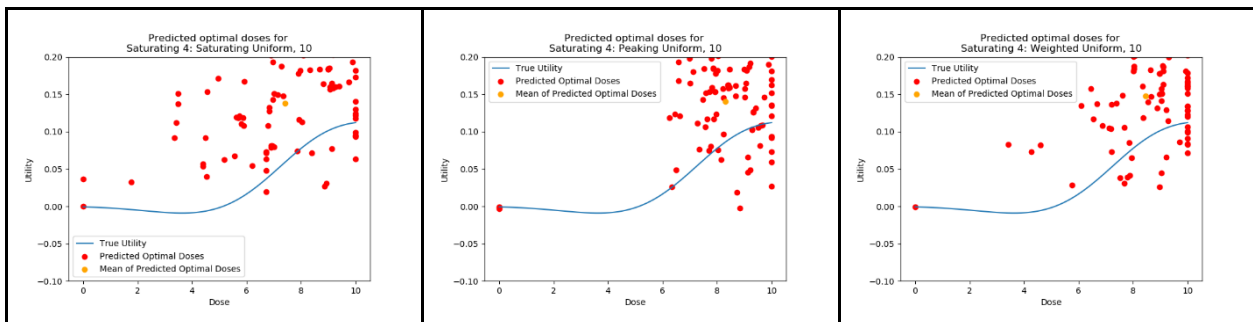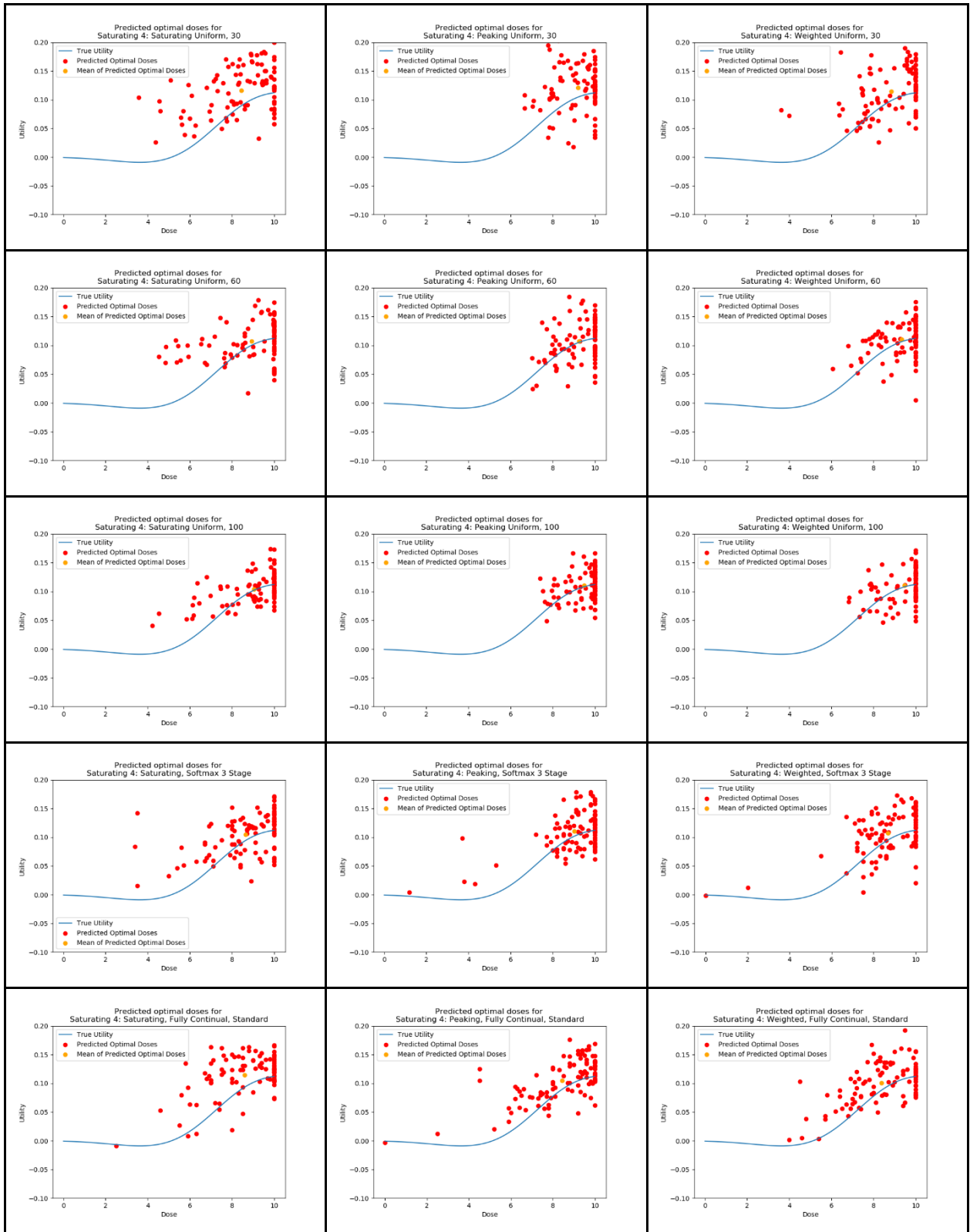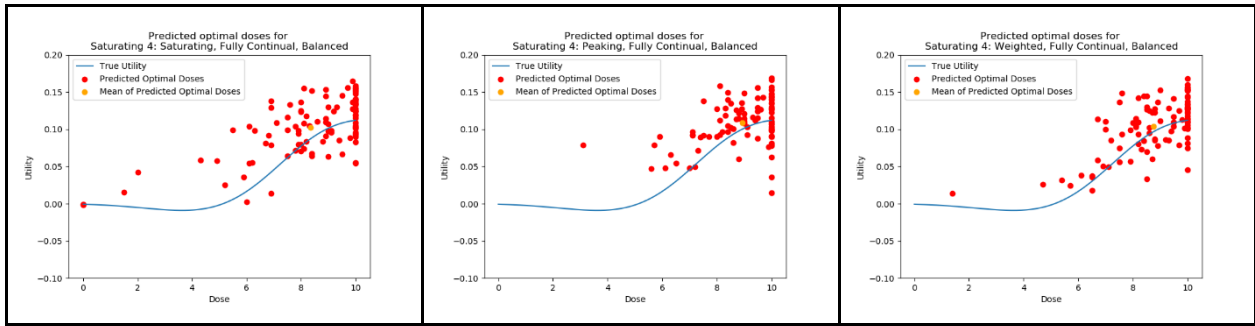
Figure S.10.4. Clinical trials by dose optimisation approach for scenario Saturating 4

Scenario Saturating 5

497
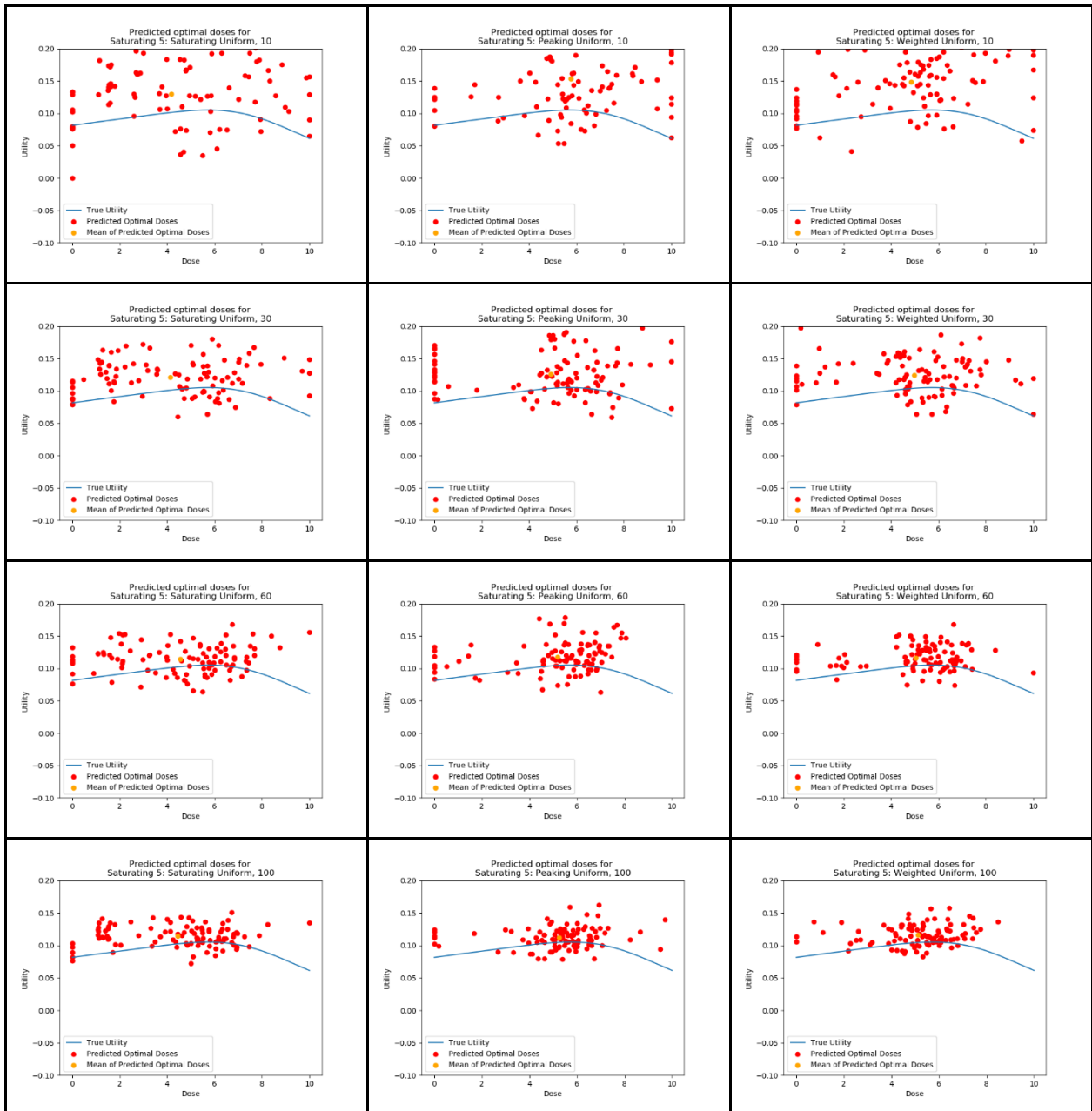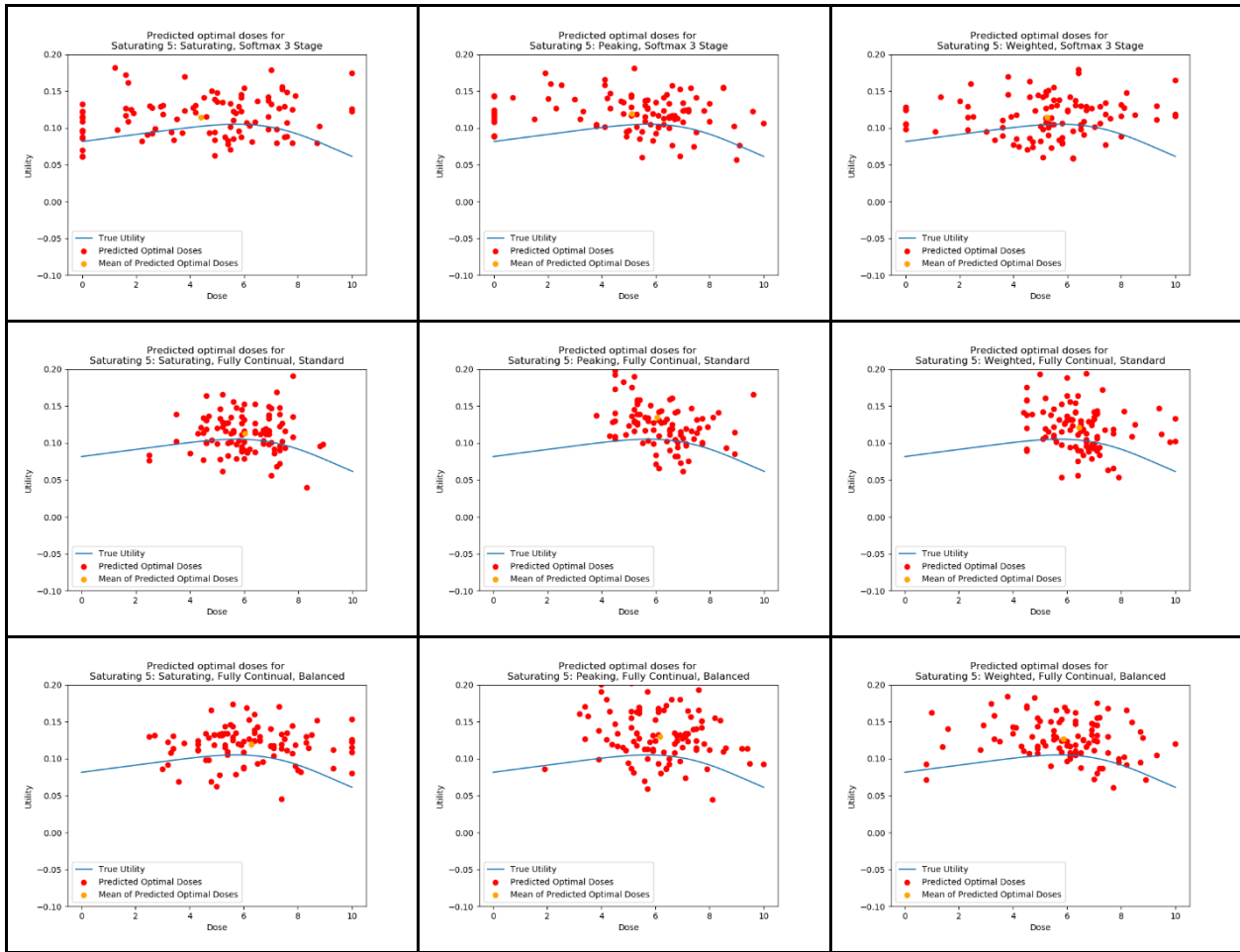
Figure S.10.5. Clinical trials by dose optimisation approach for scenario Saturating 5
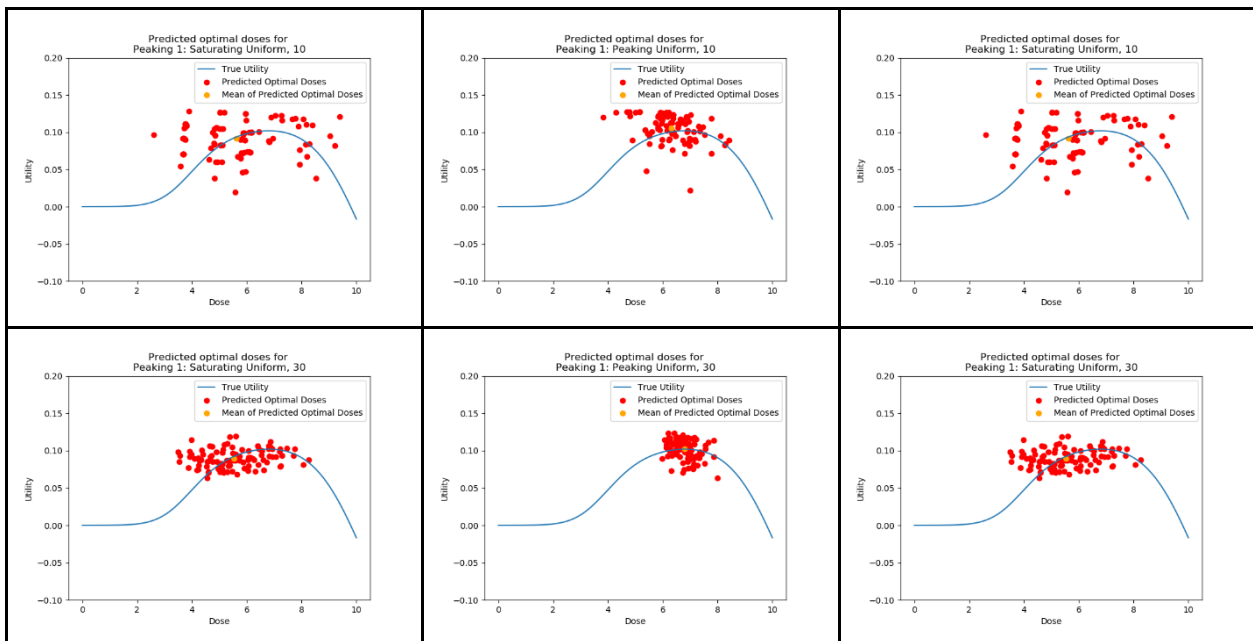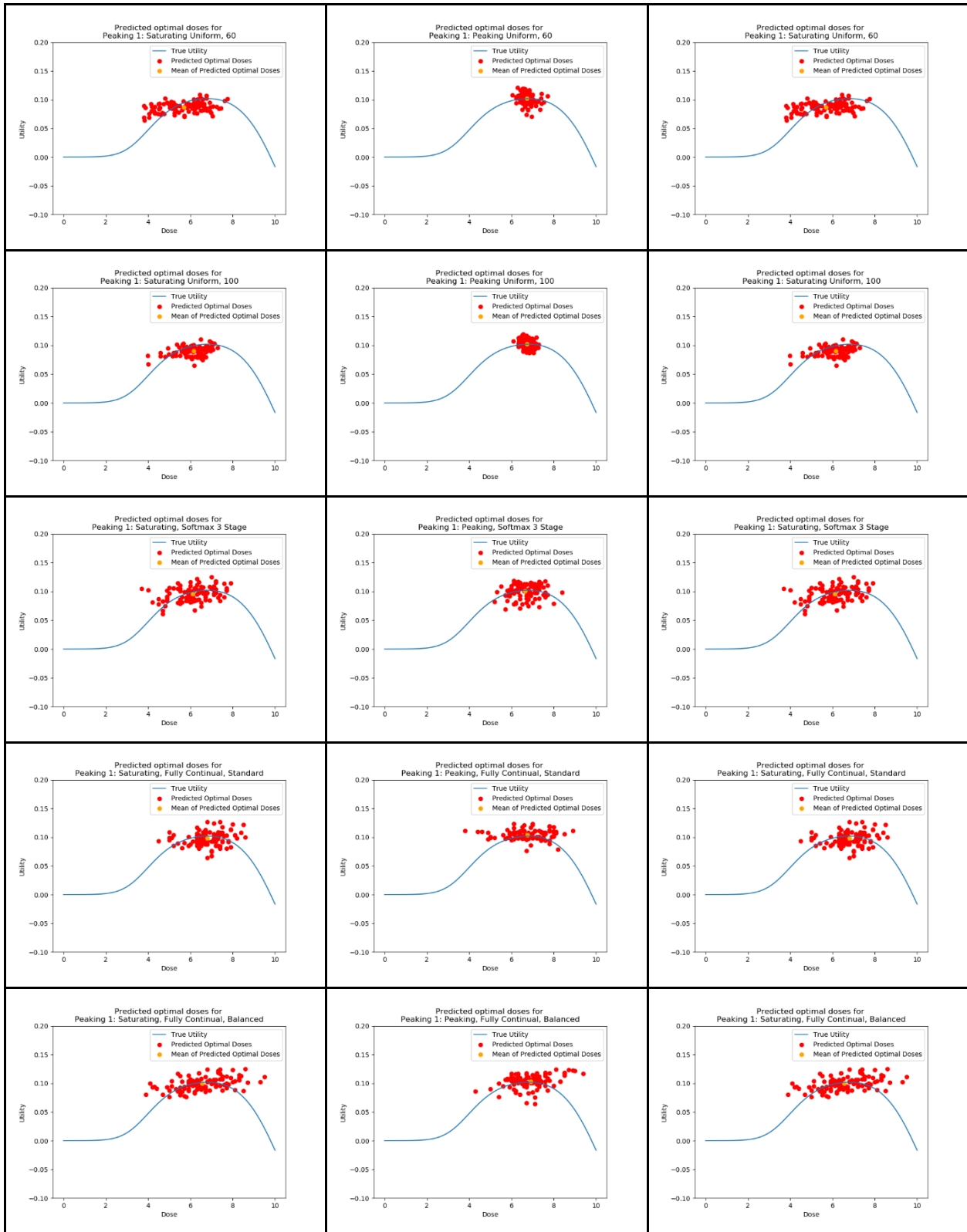
Scenario Peaking 1

Figure S.10.6. Clinical trials by dose optimisation approach for scenario Peaking 1
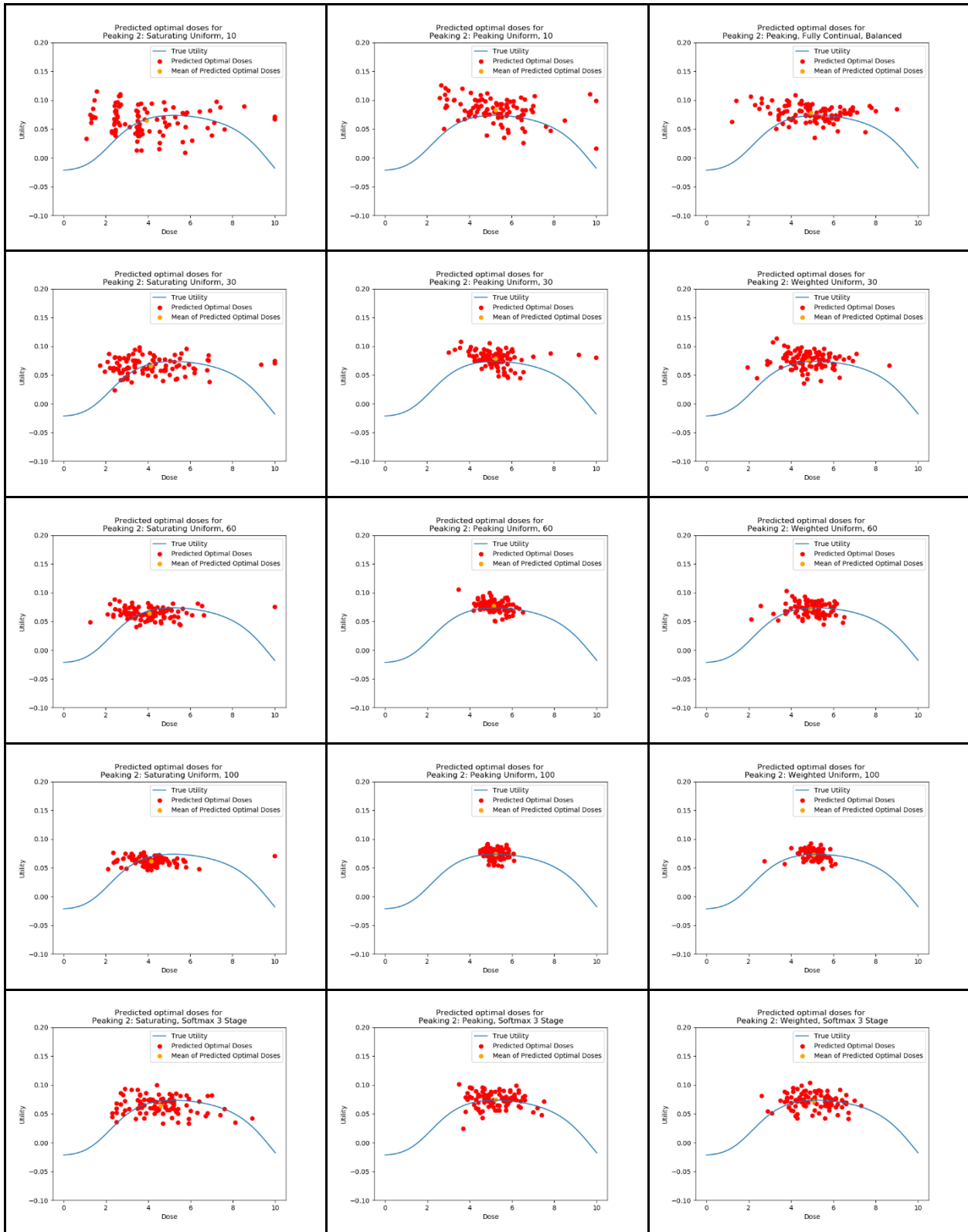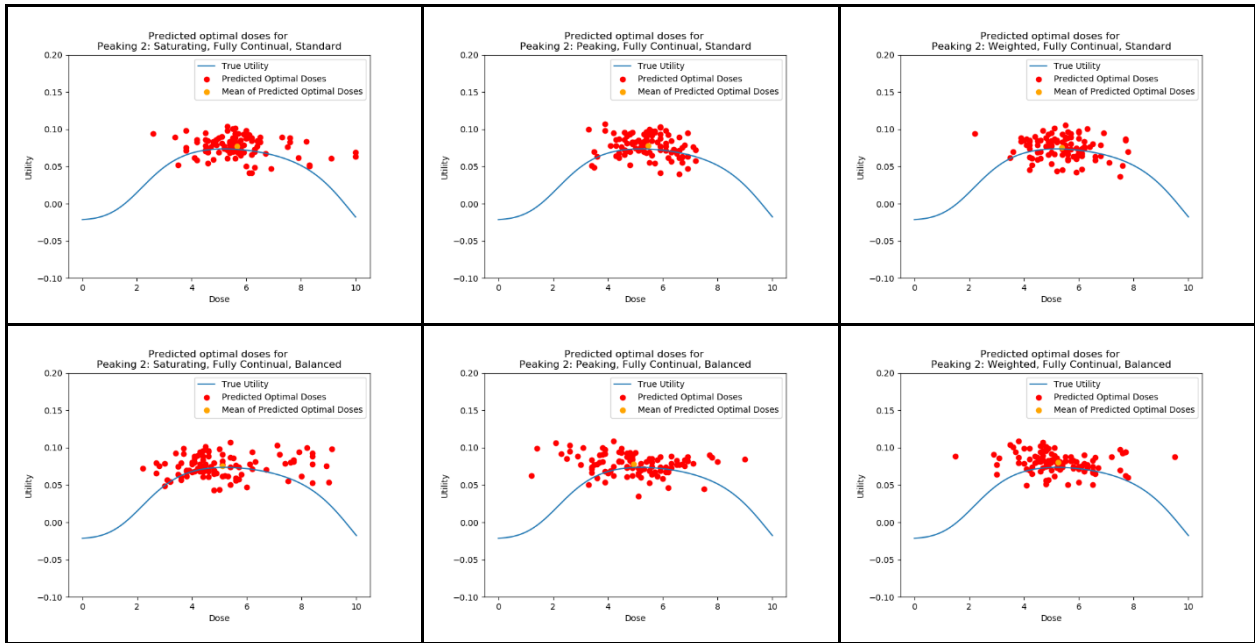
Scenario Peaking 2

Figure S.10.7. Clinical trials by dose optimisation approach for scenario Peaking 2

Scenario Peaking 4

501
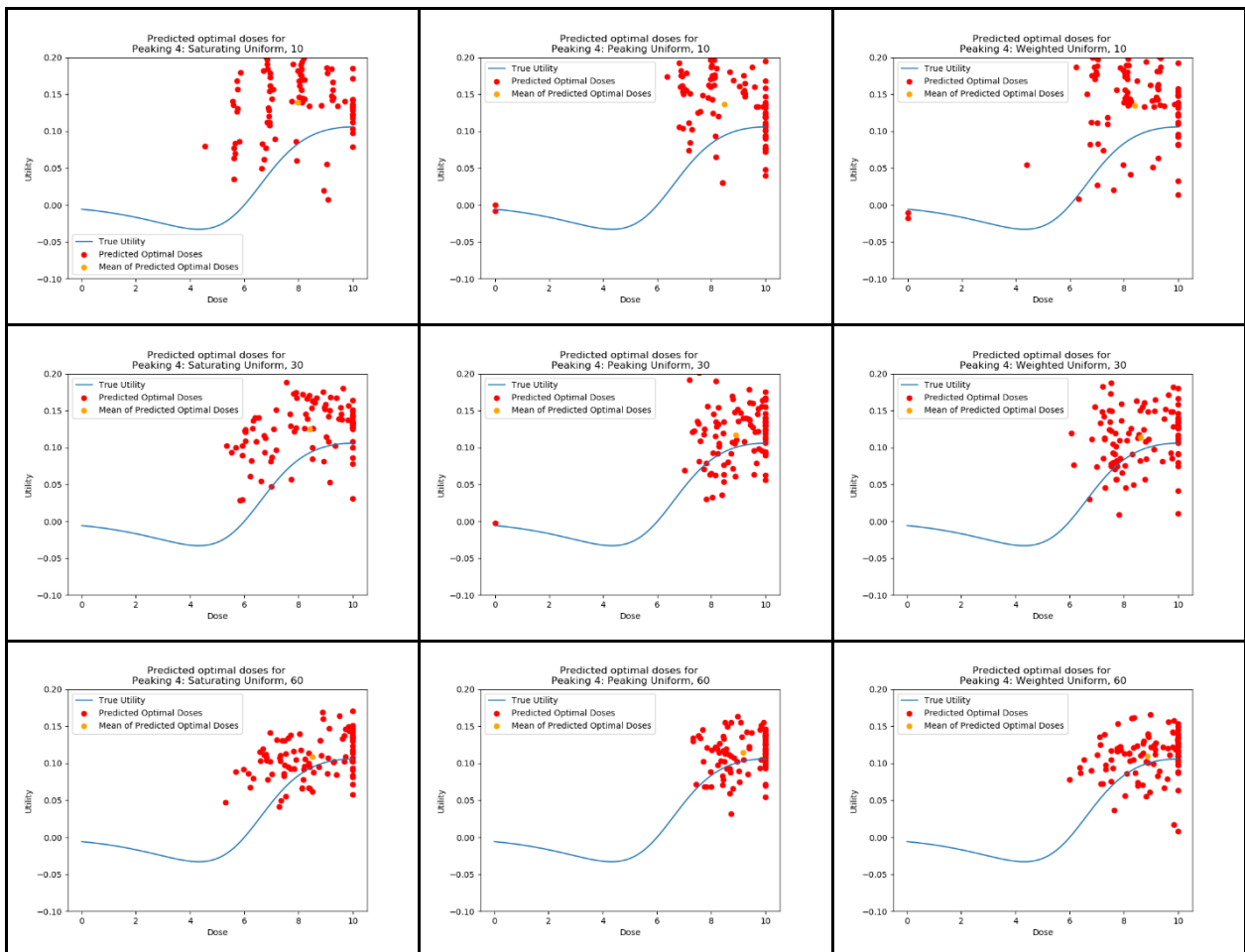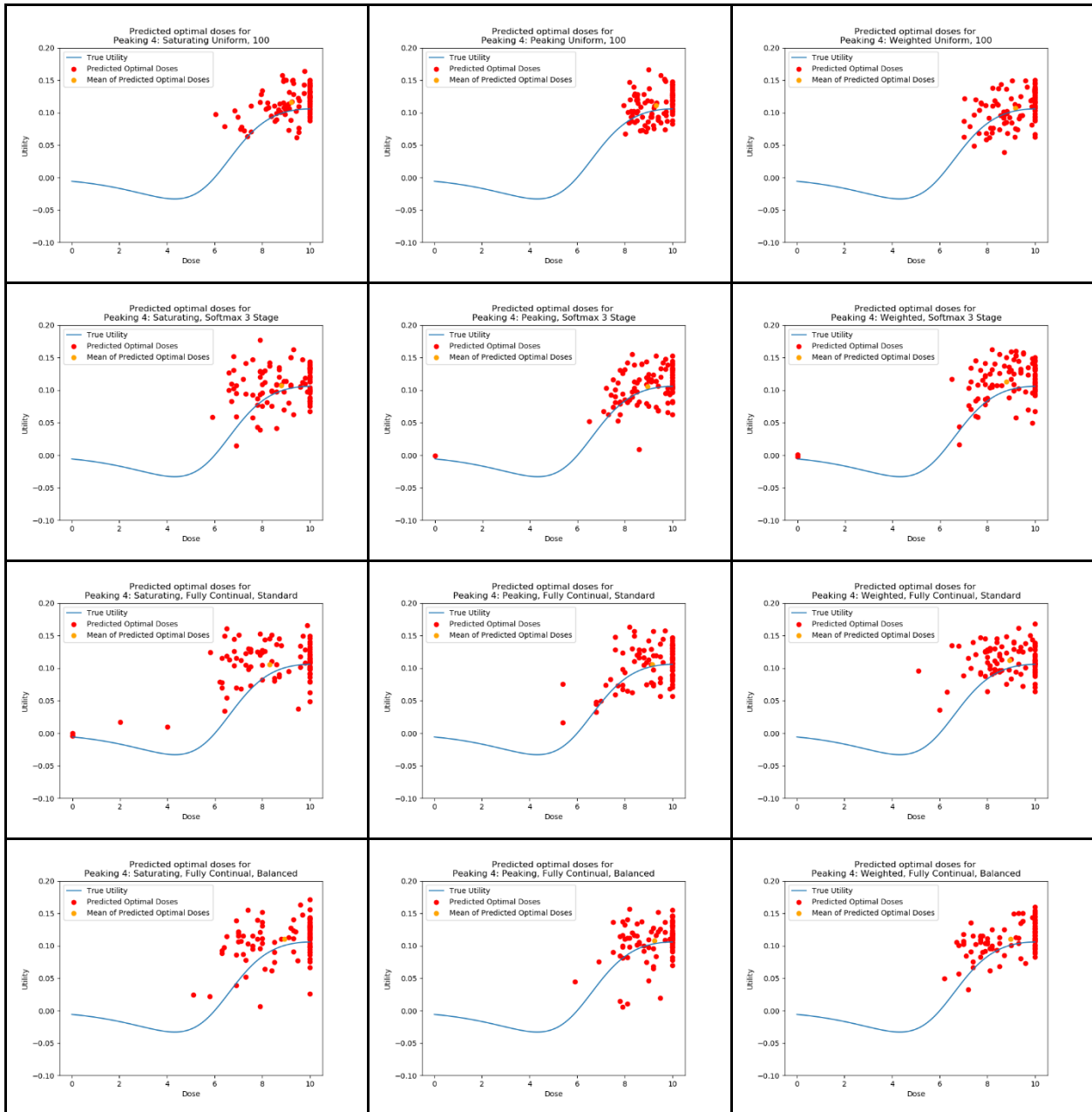
Figure S.10.9. Clinical trials by dose optimisation approach for scenario Peaking 4
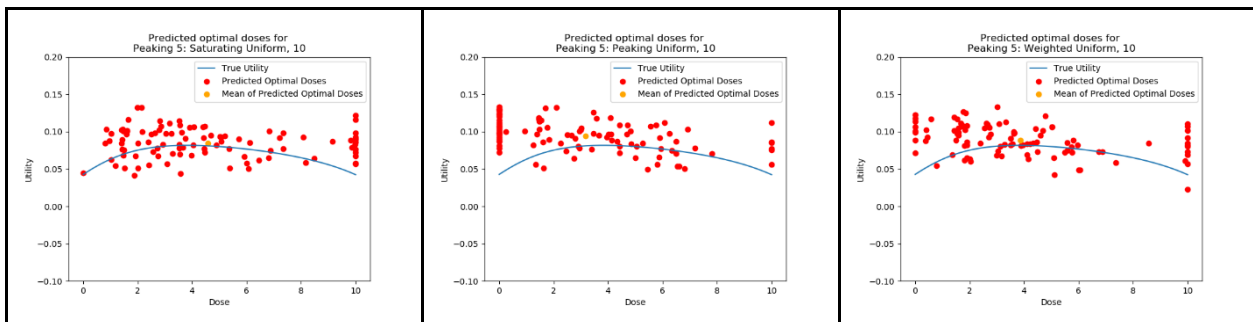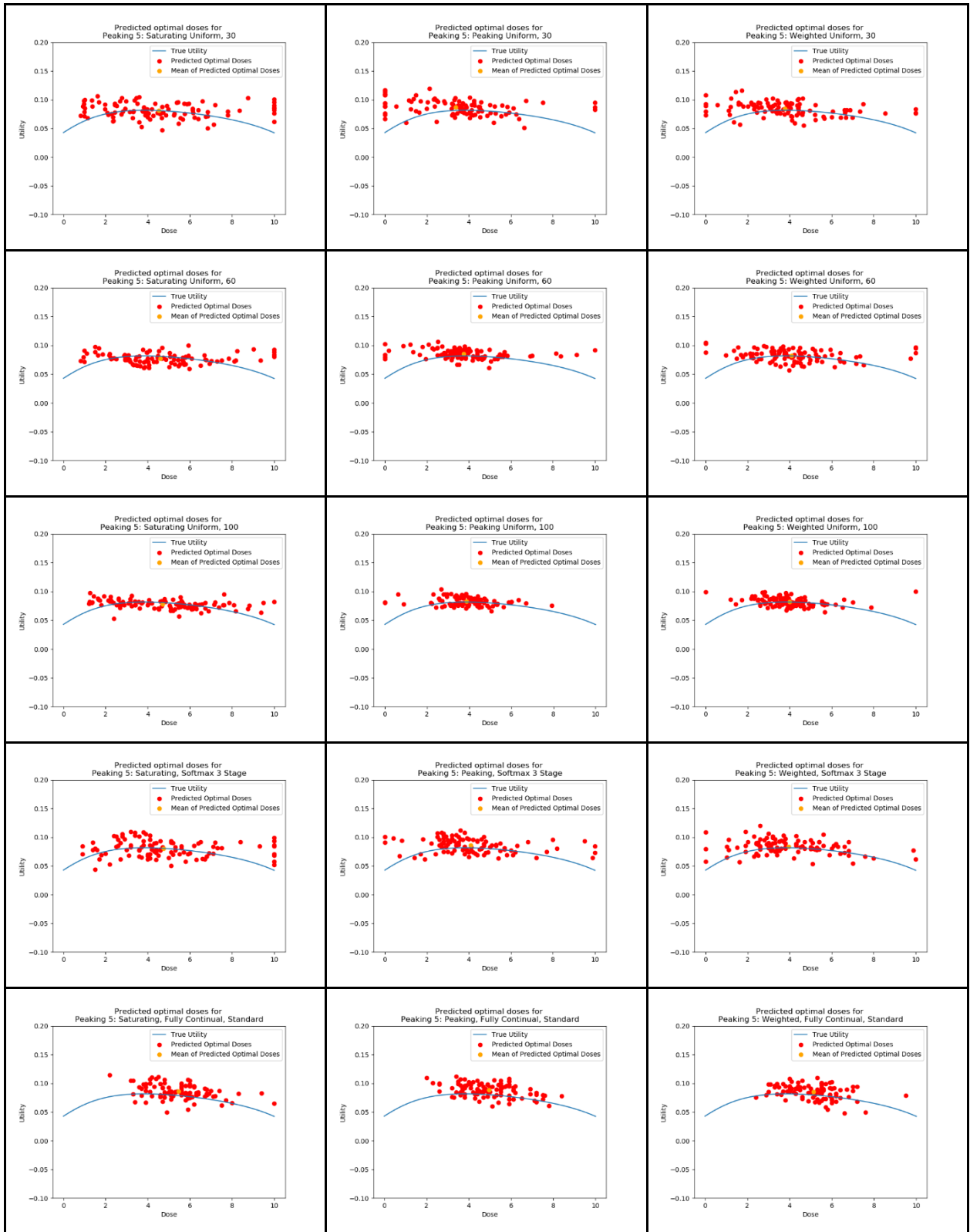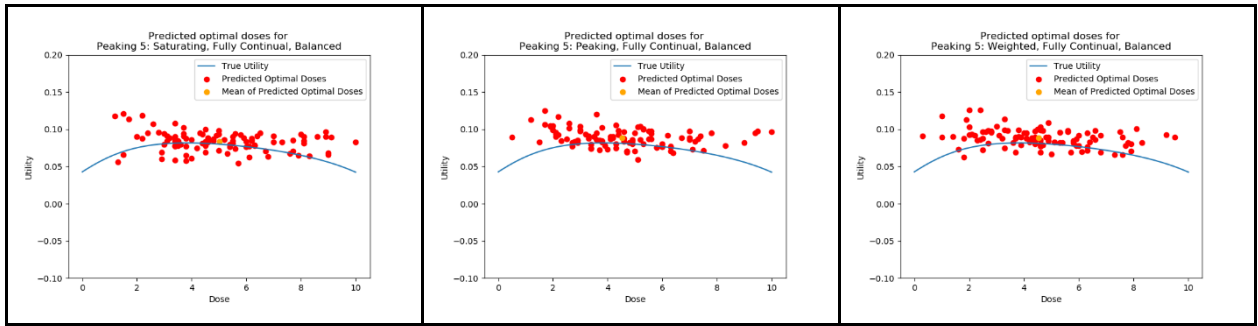
Scenario Peaking 5

Figure S.10.10. Clinical trials by dose optimisation approach for scenario Peaking 5
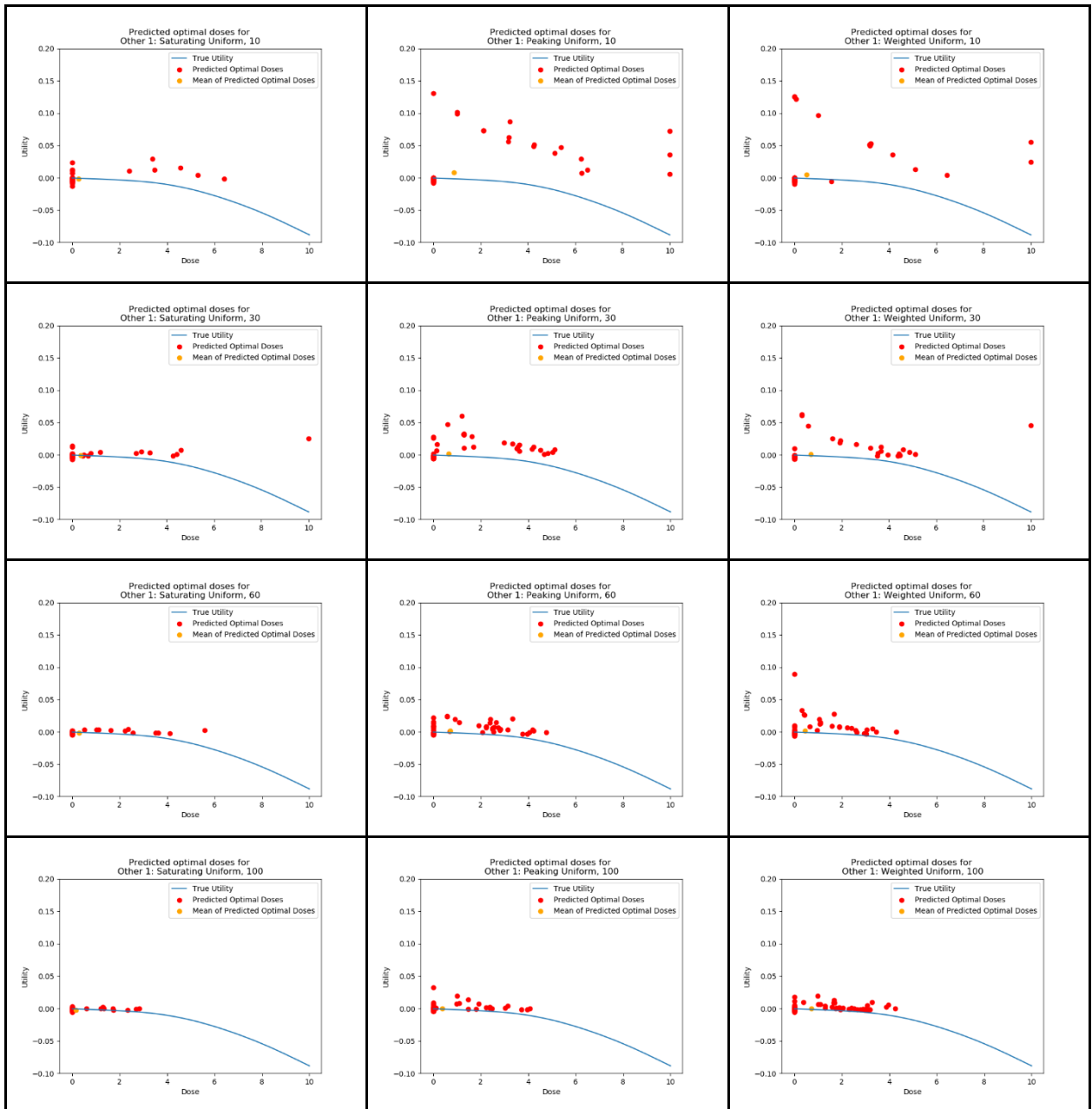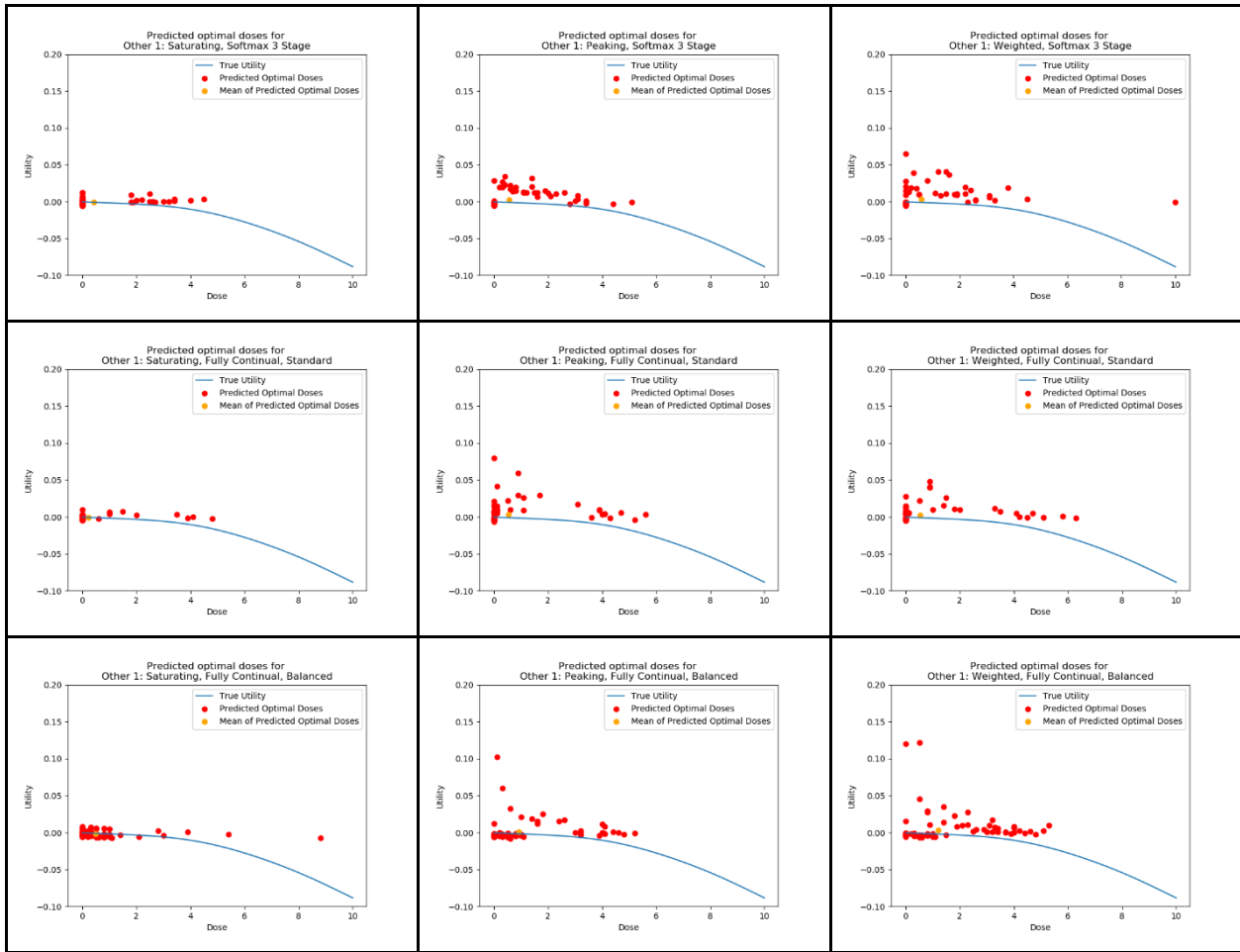
Scenario Other 1

Figure S.10.11. Clinical trials by dose optimisation approach for scenario Other 1

Scenario Other 3
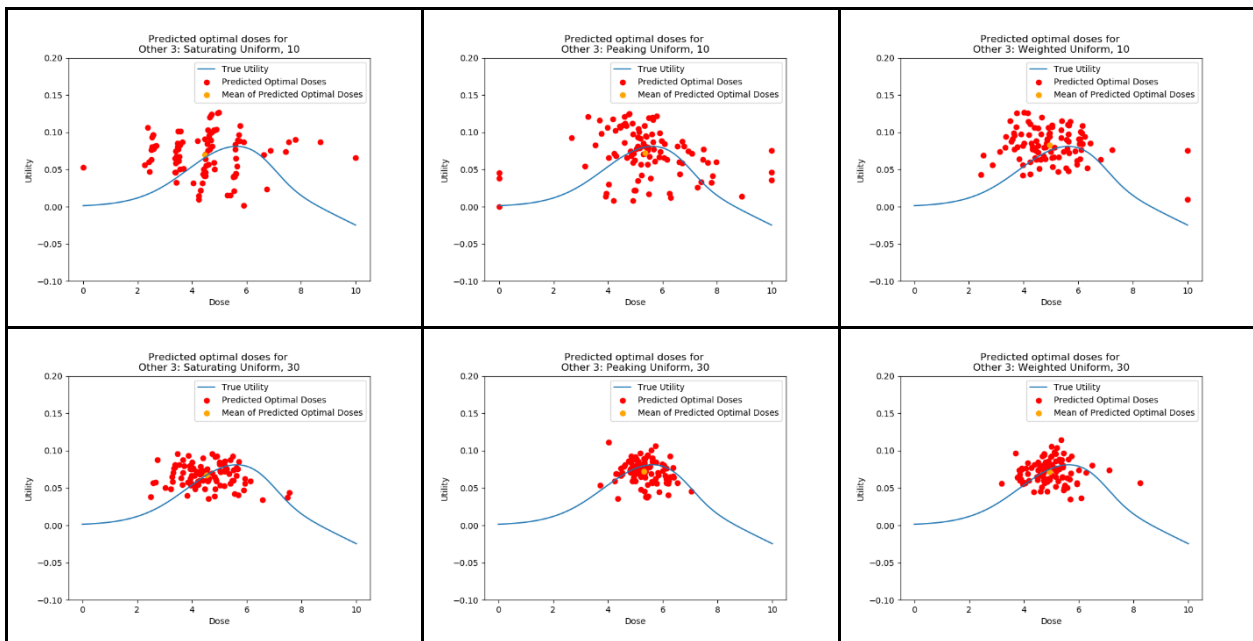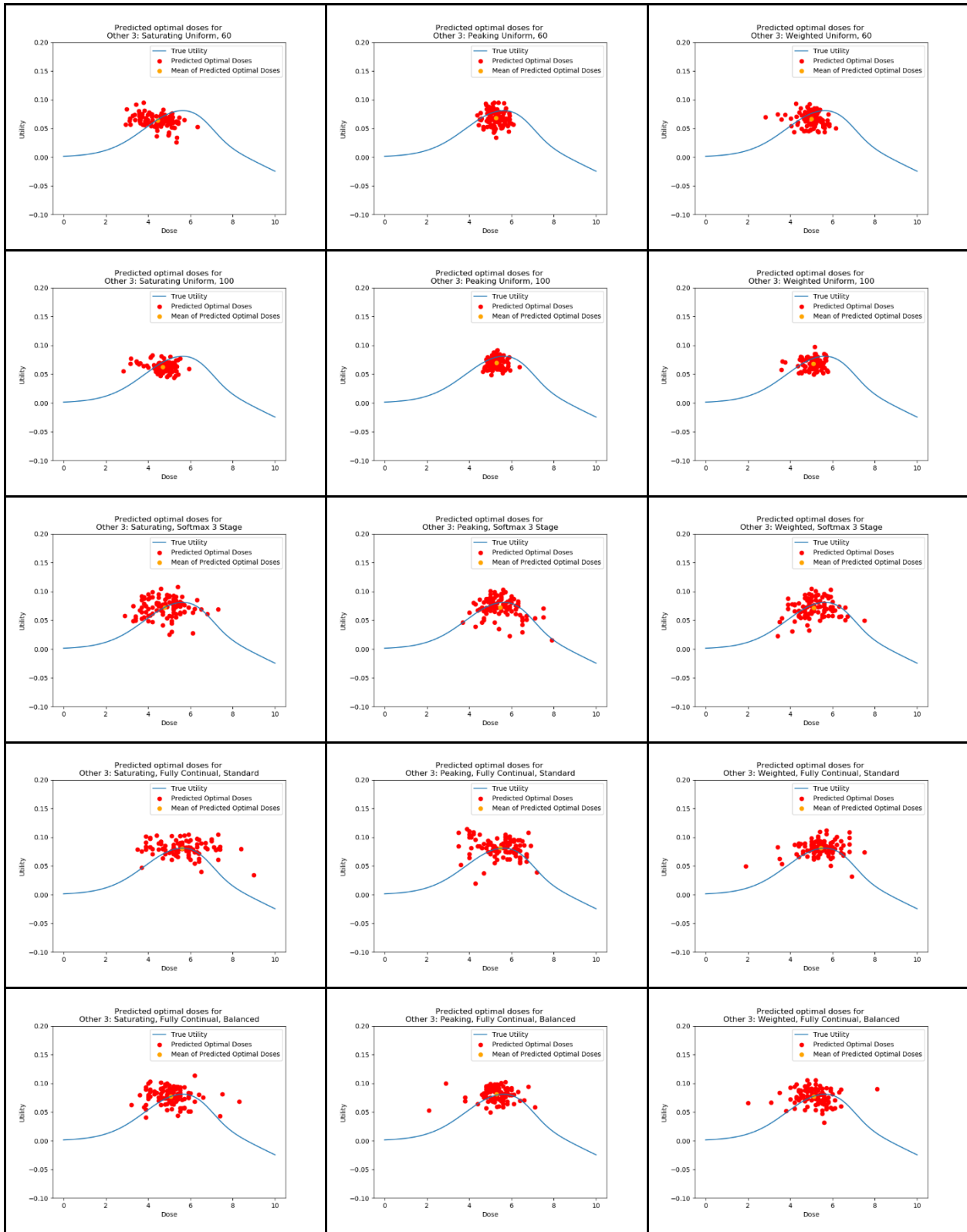
Figure S.10.13. Clinical trials by dose optimisation approach for scenario Other 2

Scenario Other 4
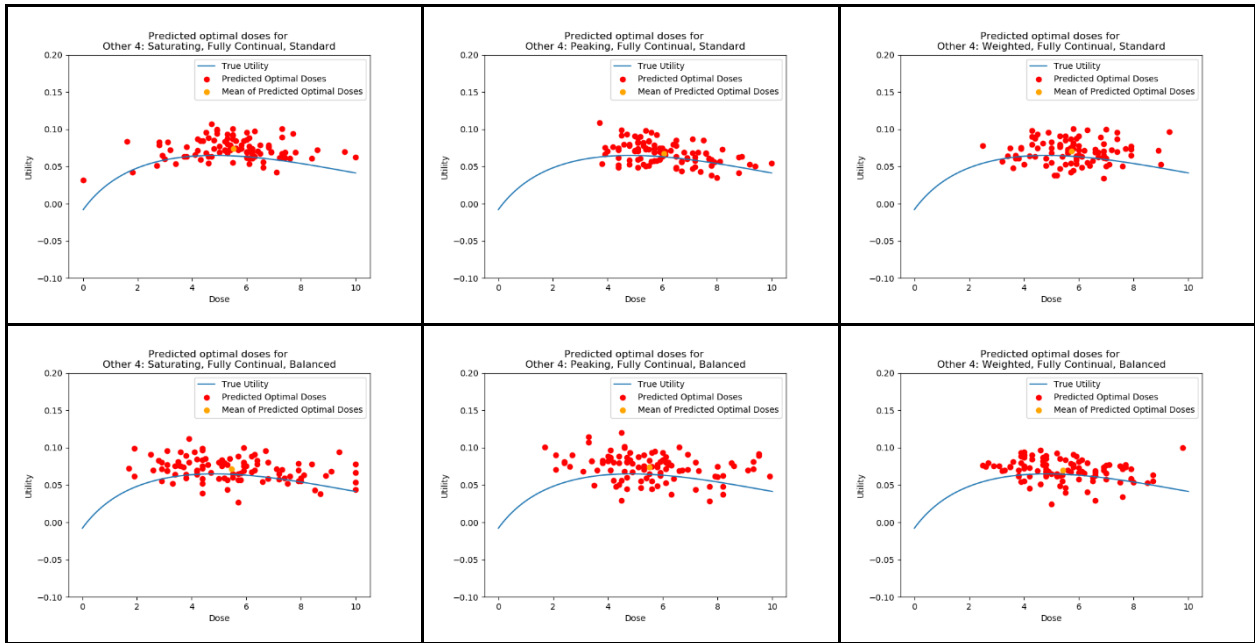
Figure S.10.14. Clinical trials by dose optimisation approach for scenario Other 3

Scenario Saturating 2
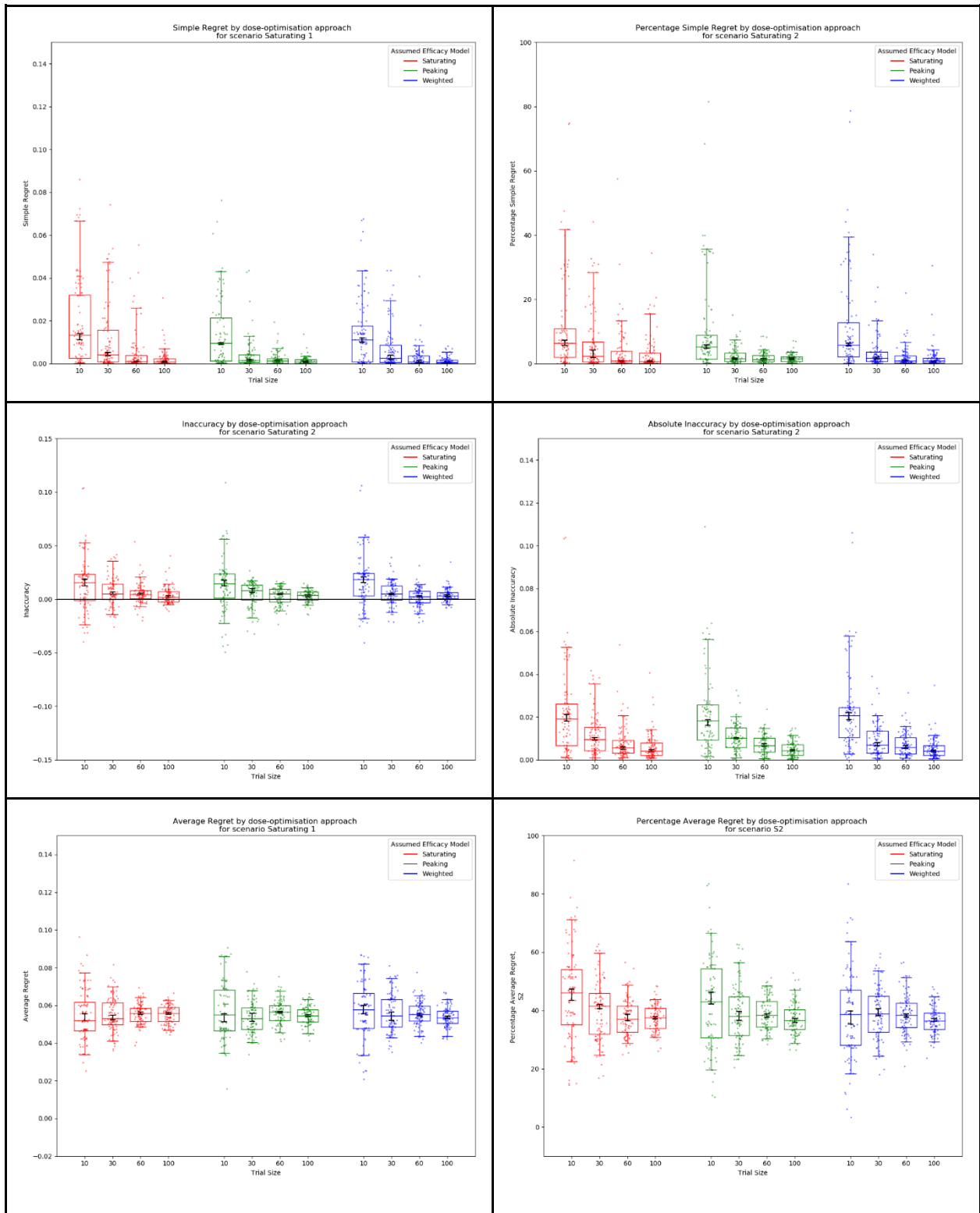
Figure S.11.2. Metrics by dose-optimisation approach for objective 1 for scenario Saturating 2.
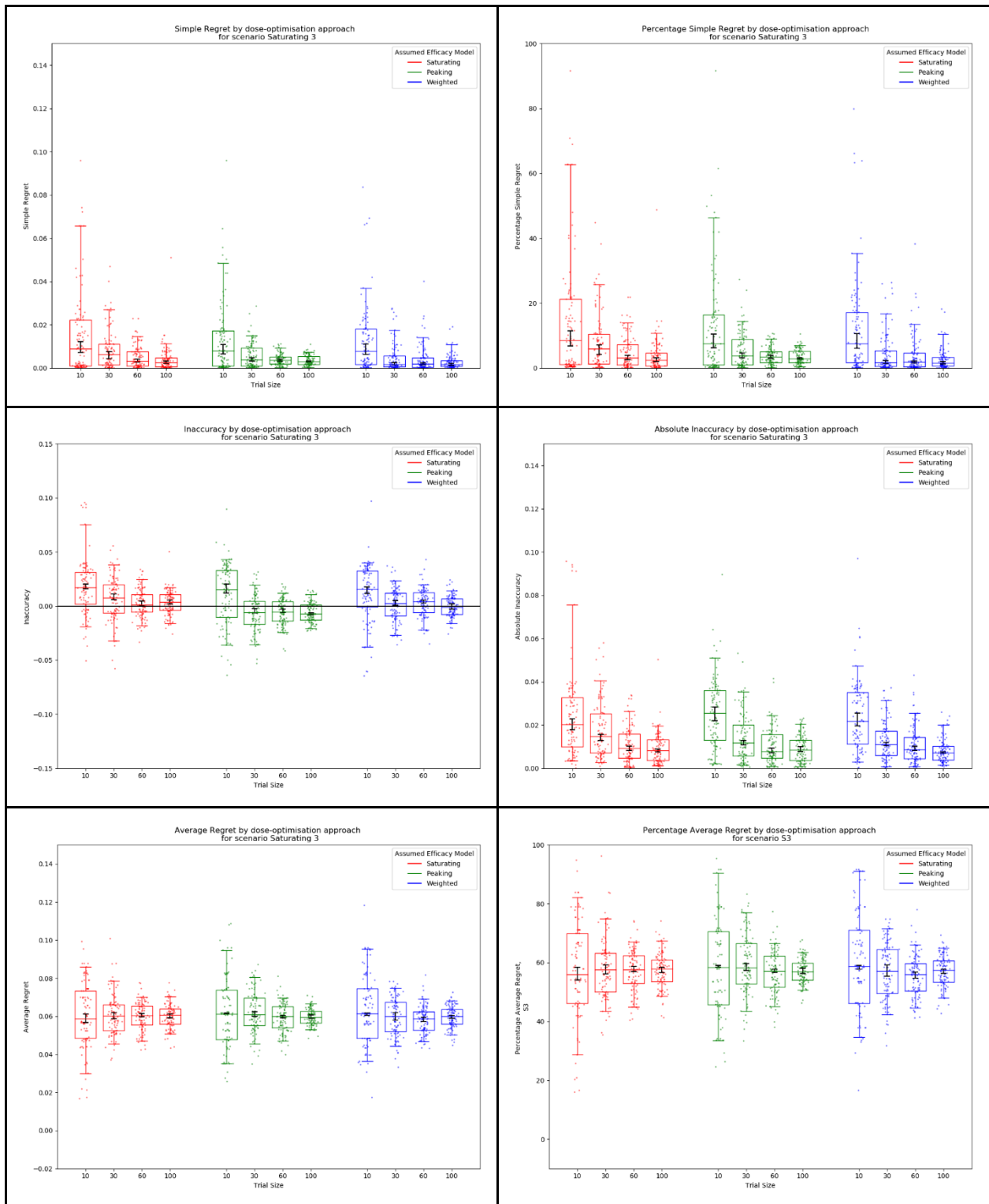
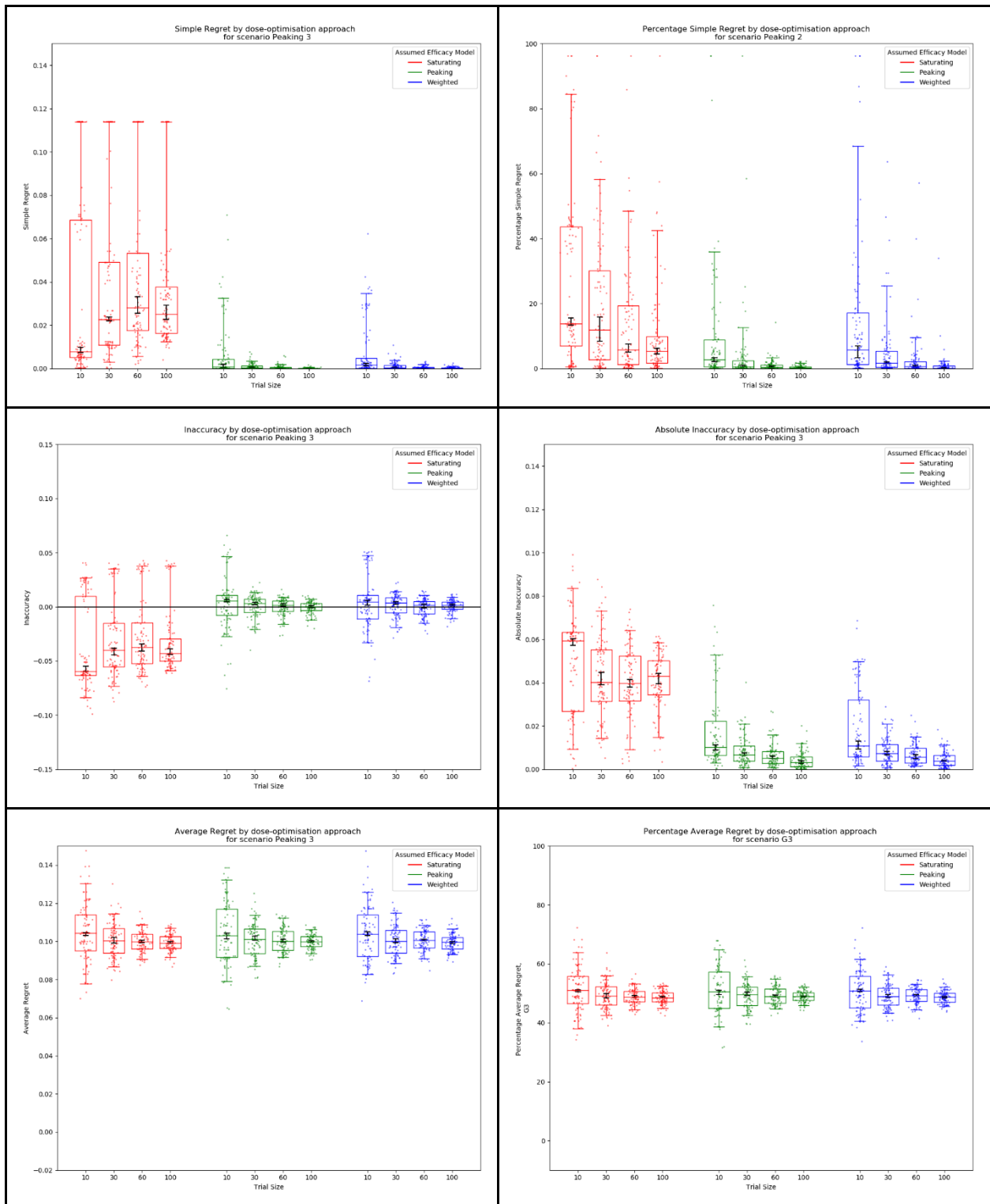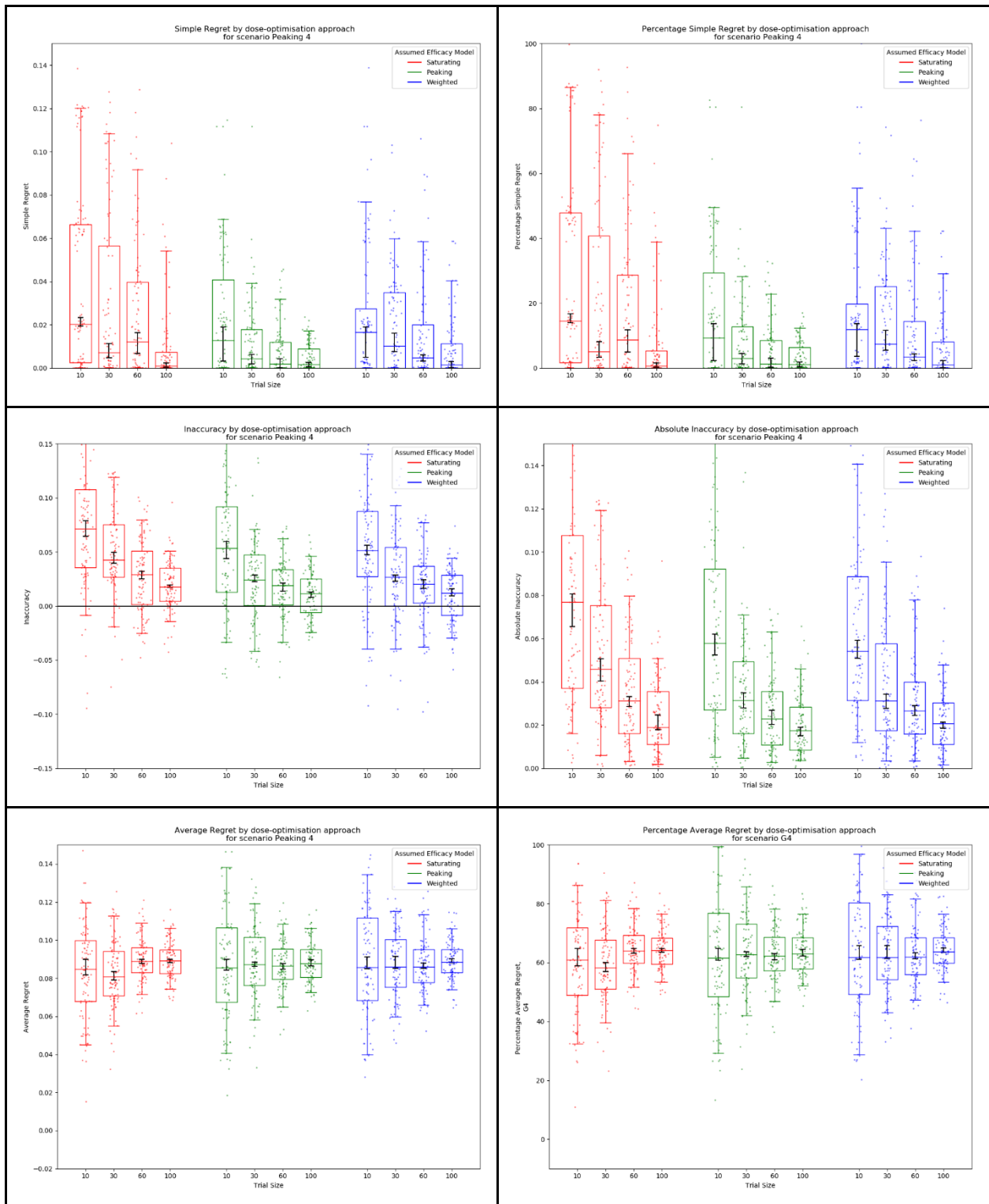**Figure S.11.3. Metrics by dose-optimisation approach for objective 1 for scenario Saturating 3**
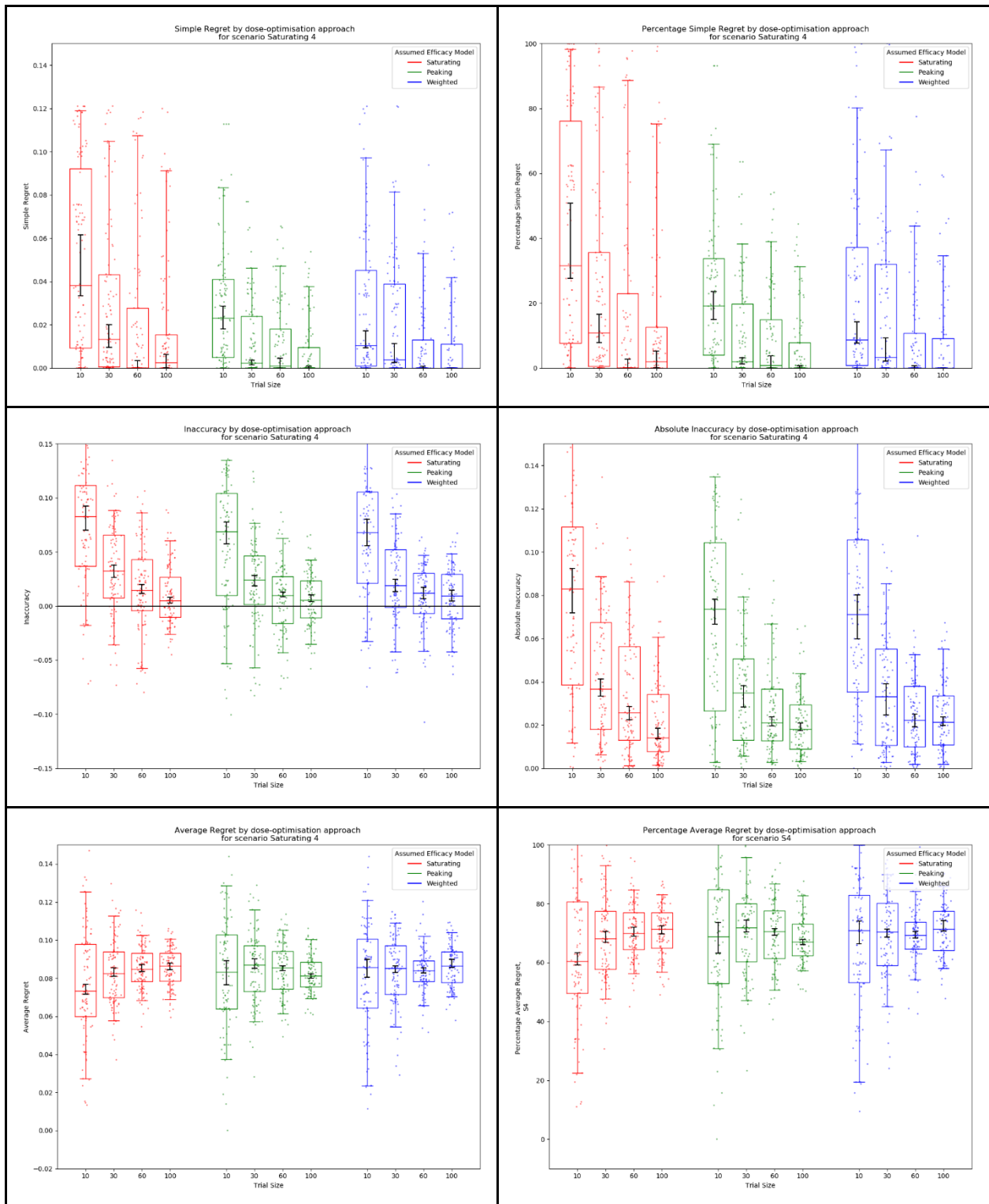
Figure S.11.4. Metrics by dose-optimisation approach for objective 1 for scenario Saturating 4

Scenario Saturating 5
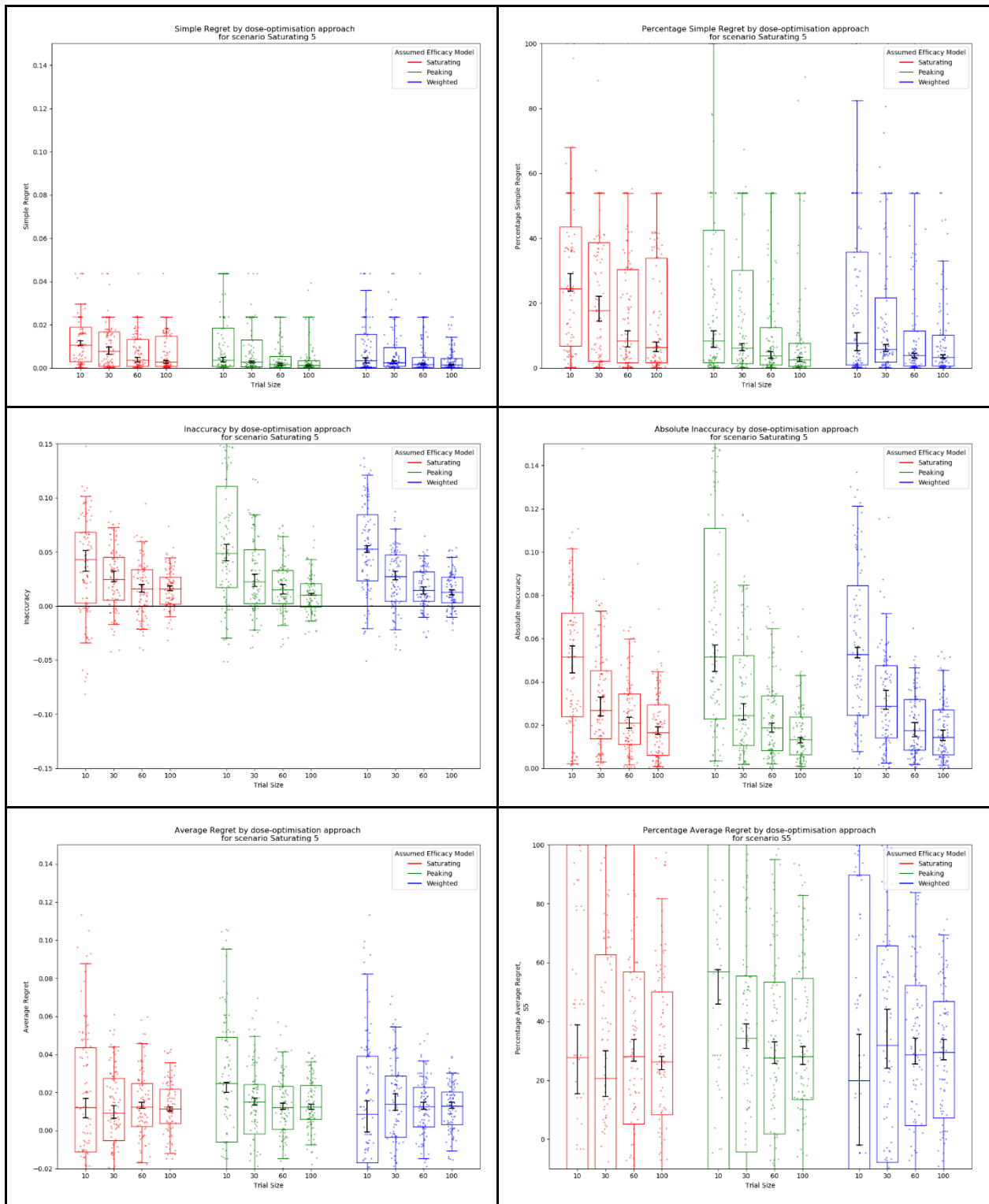
Figure S.11.5. Metrics by dose-optimisation approach for objective 1 for scenario Saturating 5
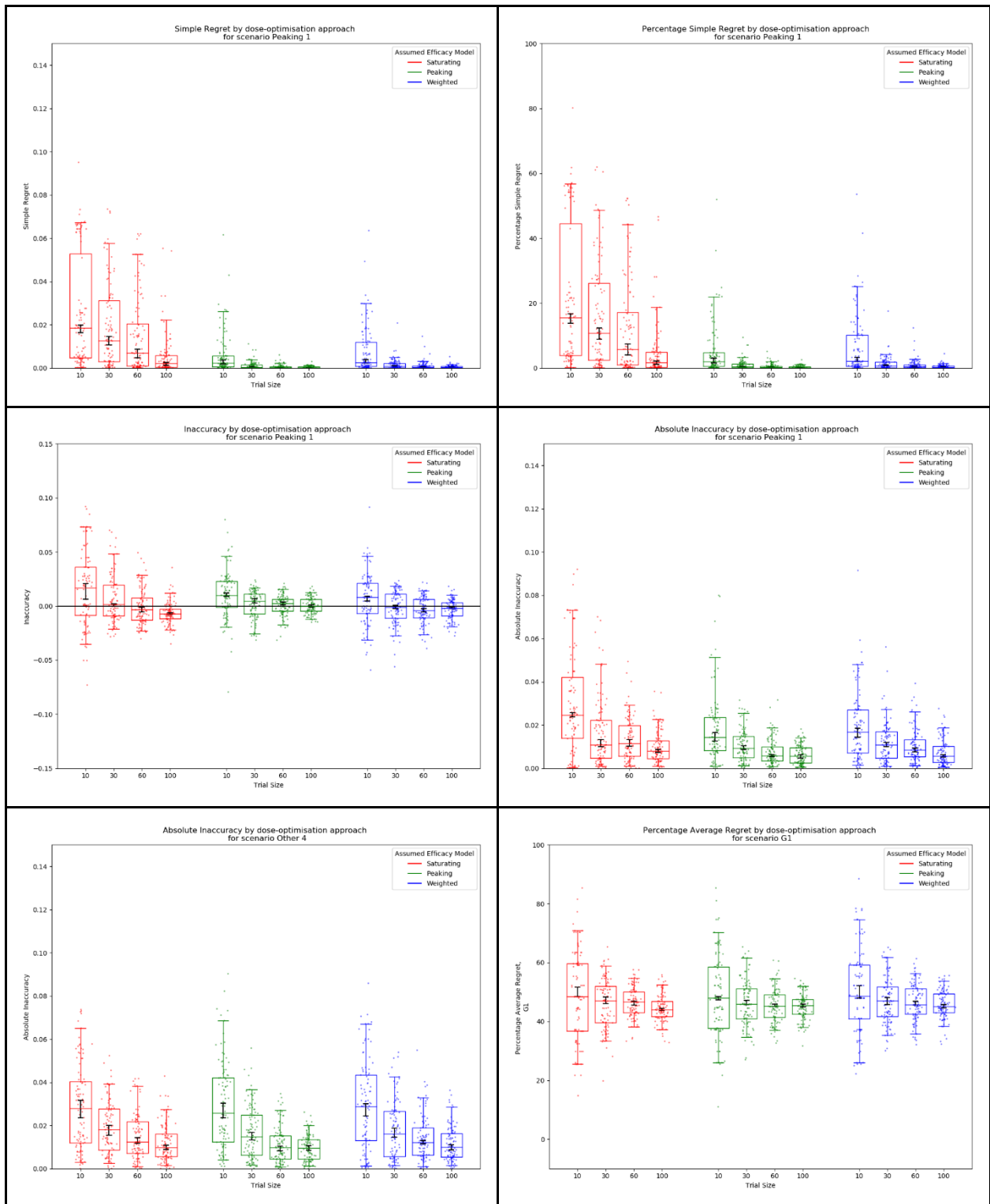
Figure S.11.6. Metrics by dose-optimisation approach for objective 1 for scenario Peaking 1
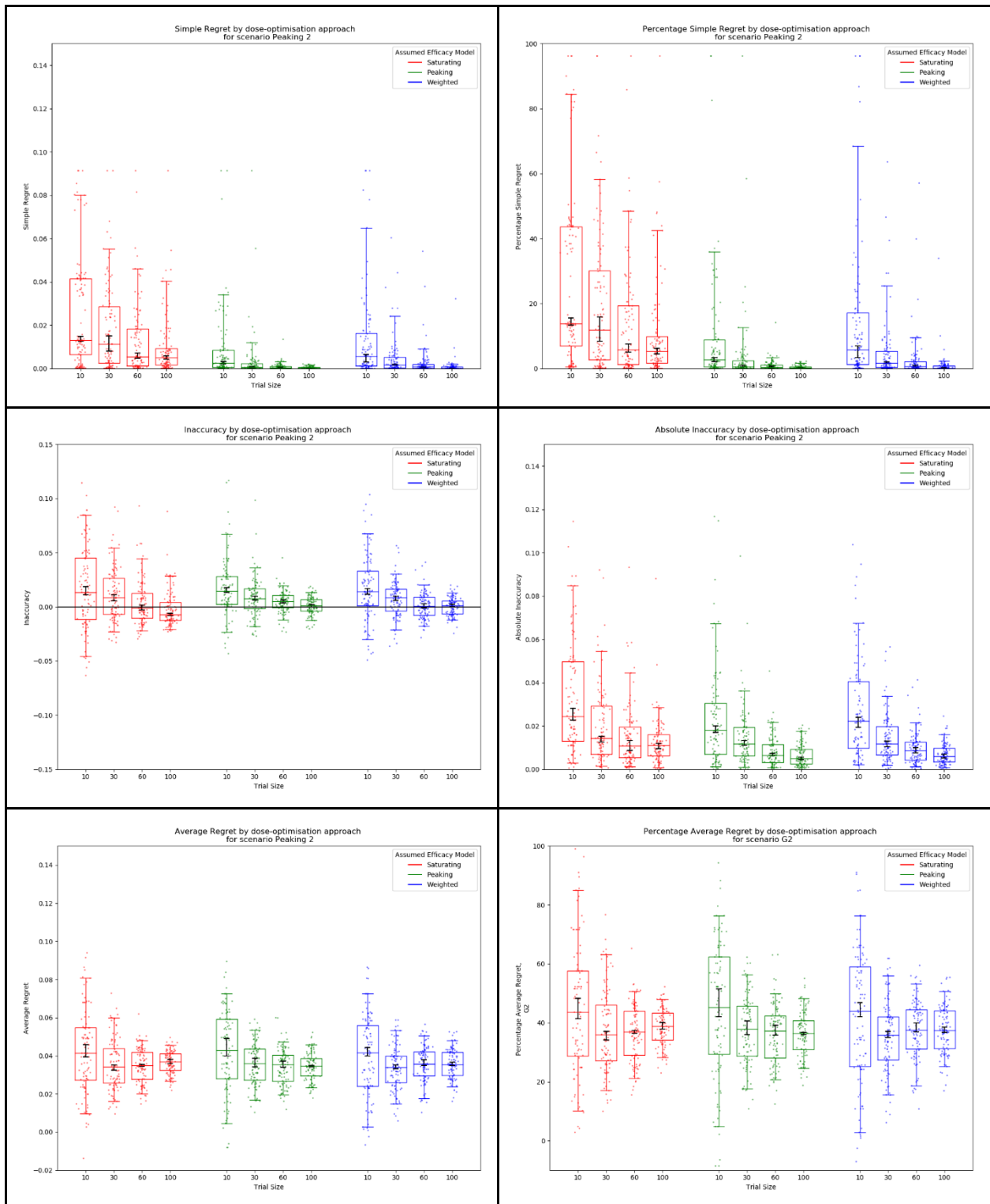
Scenario Peaking 2

Figure S.11.7. Metrics by dose-optimisation approach for objective 1 for scenario Peaking 2

Scenario Peaking 3

Figure S.11.8. Metrics by dose-optimisation approach for objective 1 for scenario Peaking 3
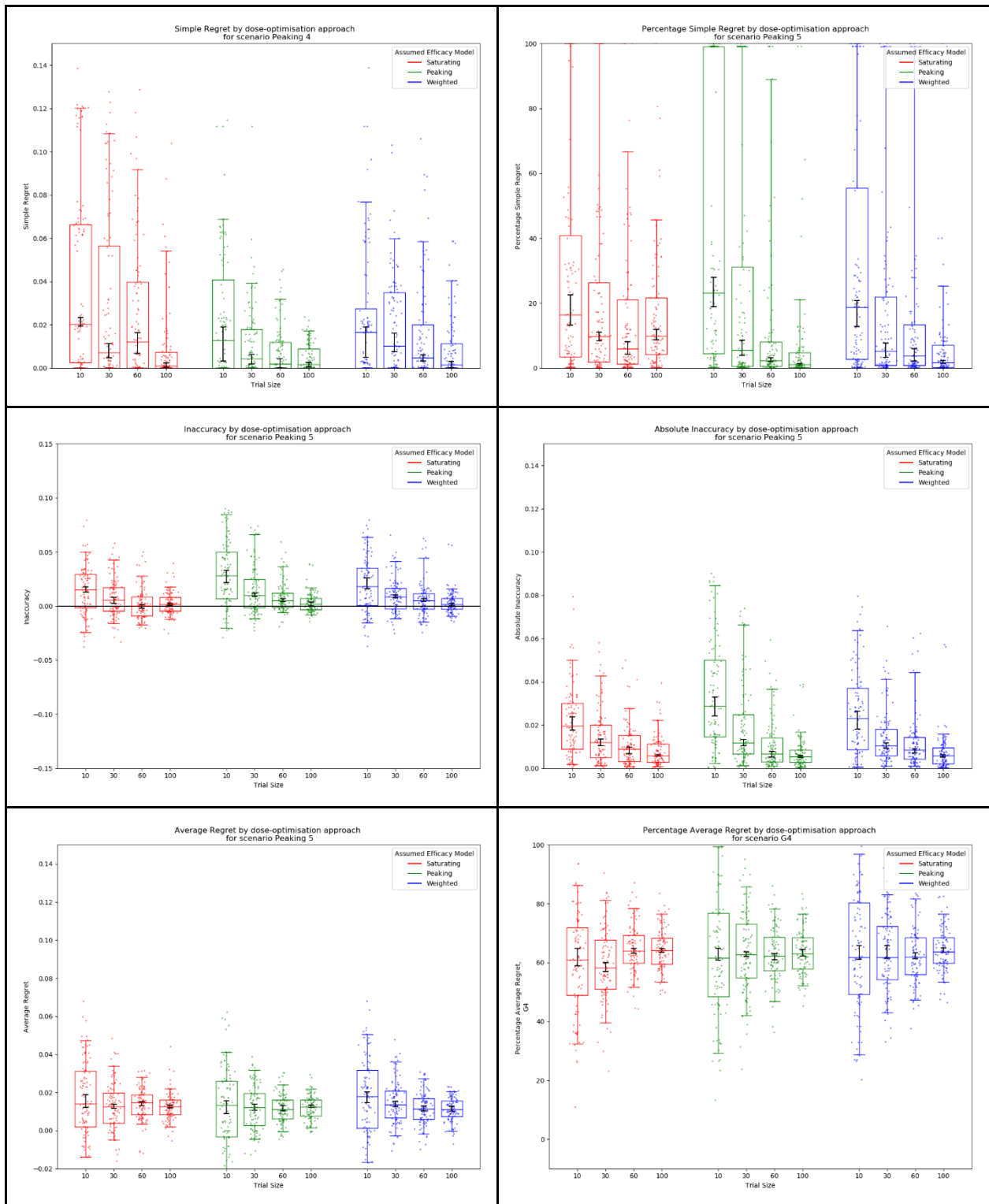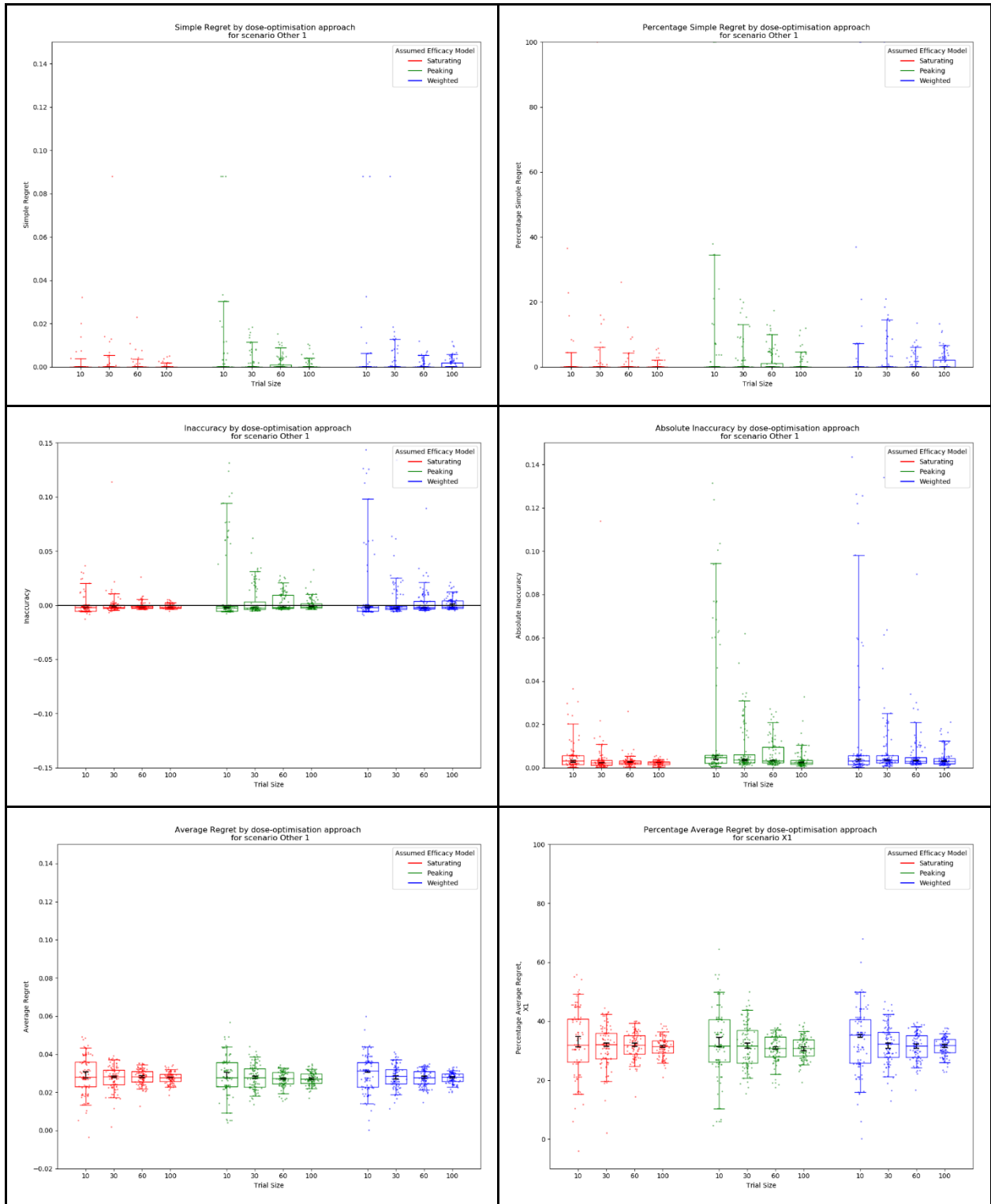
Scenario Peaking 4

Figure S.11.9. Metrics by dose-optimisation approach for objective 1 for scenario Peaking 4

Scenario Peaking 5

Figure S.11.10. Metrics by dose-optimisation approach for objective 1 for scenario Peaking 5

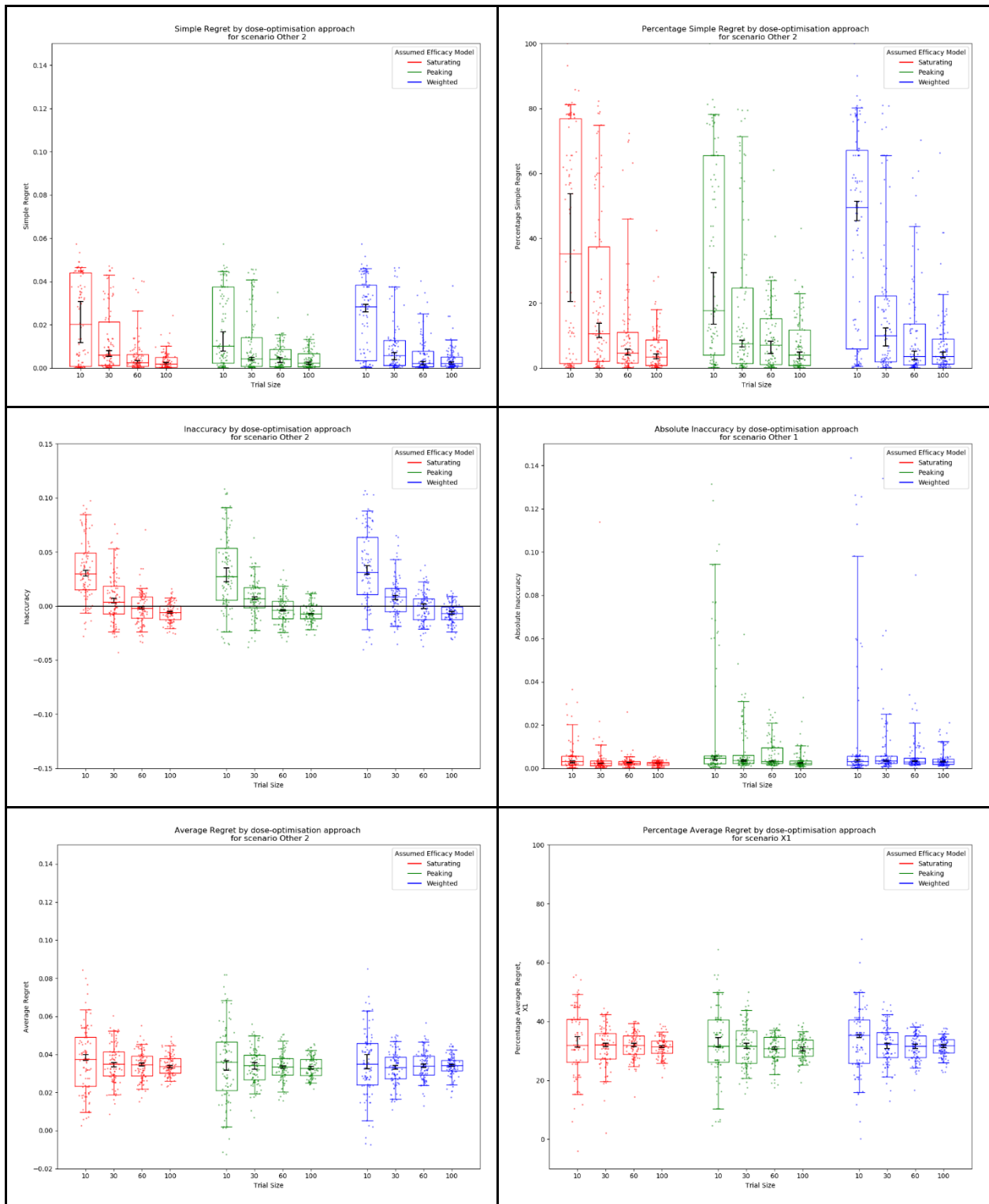**Figure S.11.11. Metrics by dose-optimisation approach for objective 1 for scenario Other 1**
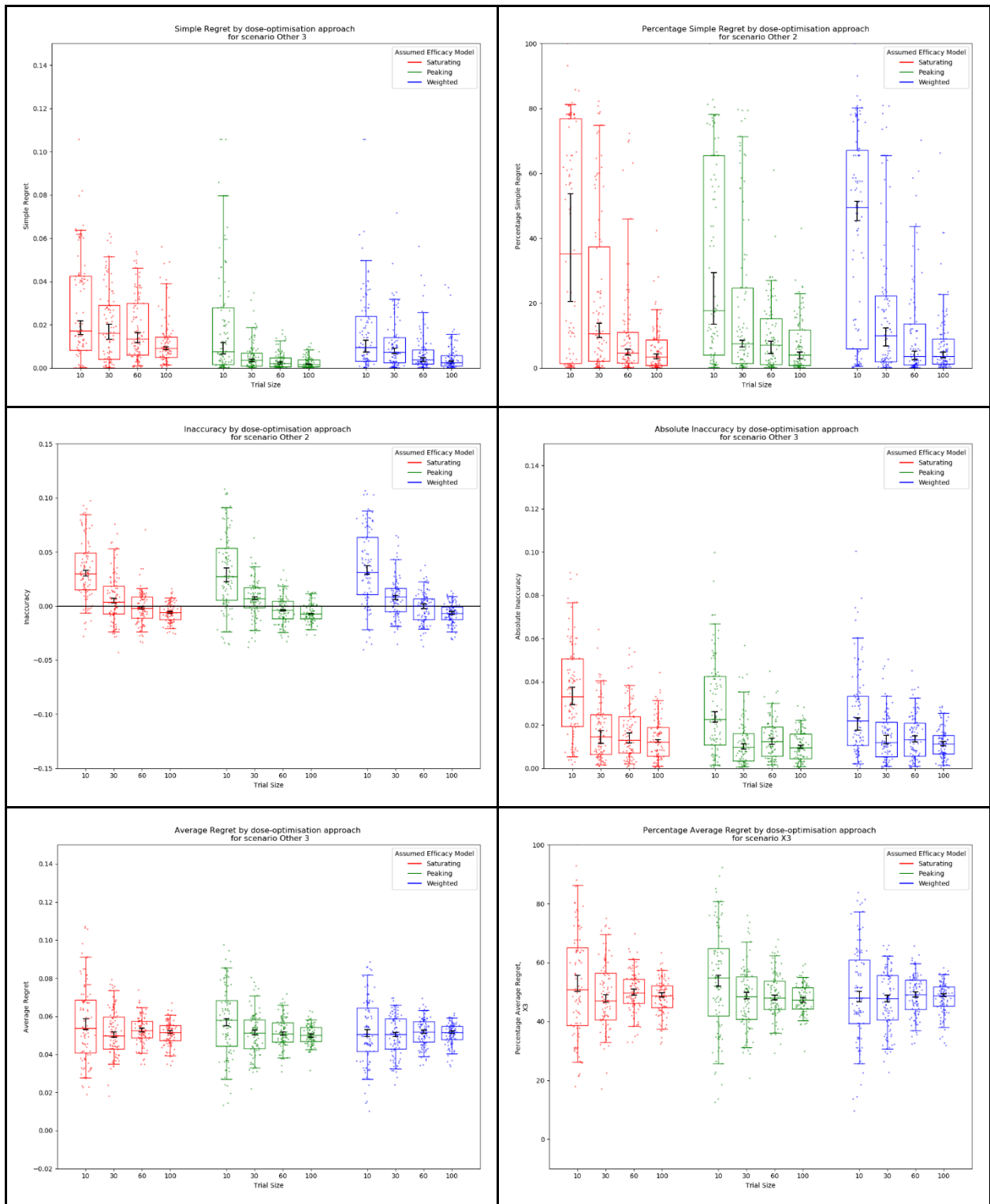
Figure S.11.12. Metrics by dose-optimisation approach for objective 1 for scenario Other 2

Scenario Other 3

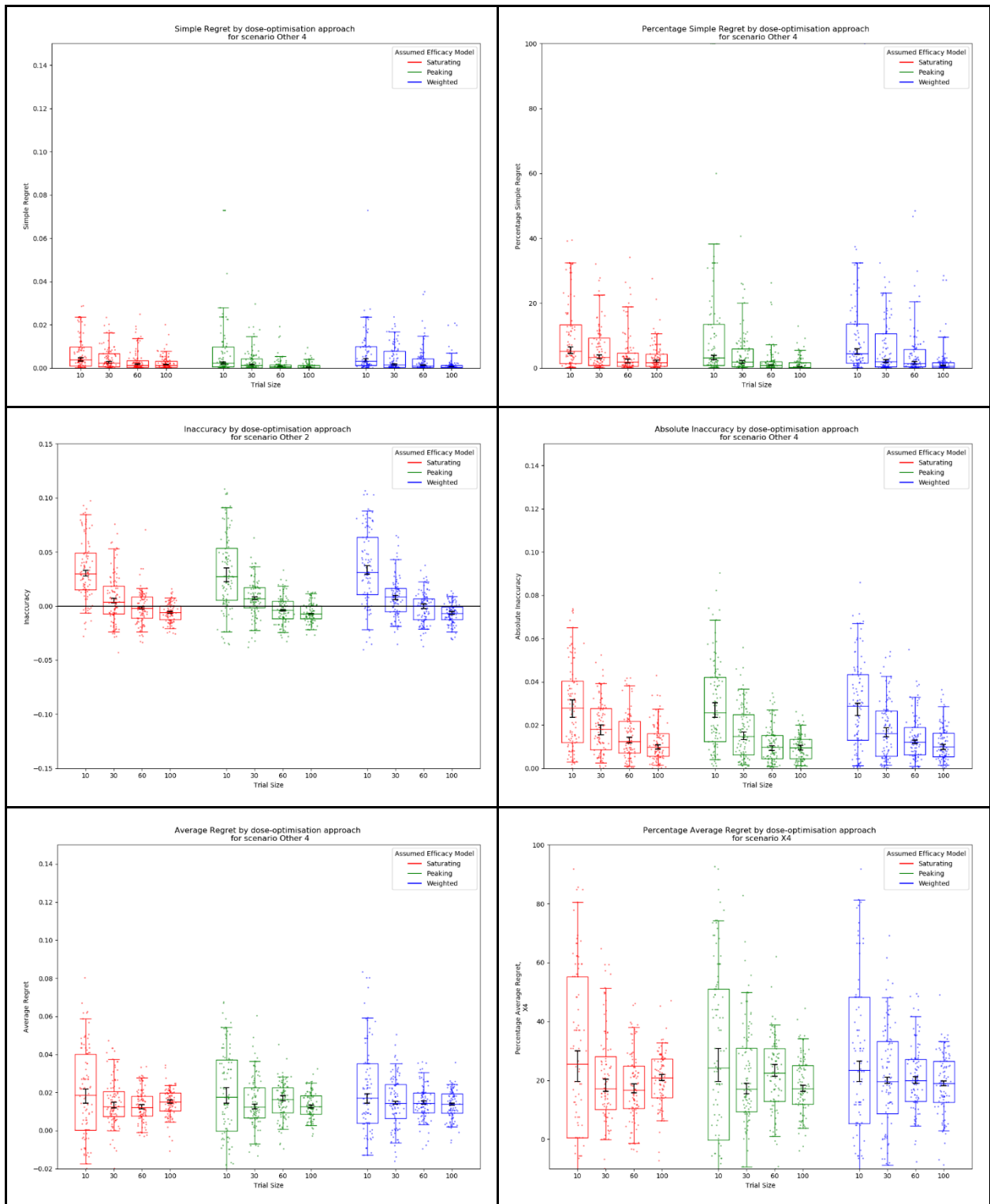**Figure S.11.13. Metrics by dose-optimisation approach for objective 1 for scenario Other 3**

Figure S.11.14. Metrics by dose-optimisation approach for objective 1 for scenario
Other 4

S12.2. Objective 1

Here I show the p-values for objective 1 for the metrics of PSR, Absolute Inaccuracy, and PAR. These are the metrics for the data stratified on scenario. This is the data in Supplementary 11.

For interpretation, the Kolmogorov–Smirnov heatmaps are symmetric, with significance representing evidence that the true distribution for the approach-scenario test metrics of PSR, Absolute Inaccuracy, and PAR differ between the two dose-optimisation approaches across all scenarios. The One-sided Mann-Whitney U test heatmaps are not symmetric, with significance for the cell in row A and column B representing 'statistically significant' evidence that approach A was preferable to approach B with regards to that metric (e.g. lower PSR, lower Absolute Inaccuracy, Lower PAR).

Each of the below figures shows the Kolmogorov–Smirnov (left) and Mann-Whitney U (right) heatmaps of p-values for objective 1 for each scenario. These are for the metrics of PSR (top), Absolute Inaccuracy (middle) and PAR (bottom). Cells with a light pink hue represent the test statistic for that comparison would be significant under the threshold p<0.05, cells with a red hue represent the test statistic for that comparison would be significant under the threshold p<0.05 with Bonferroni multiple comparison correction.
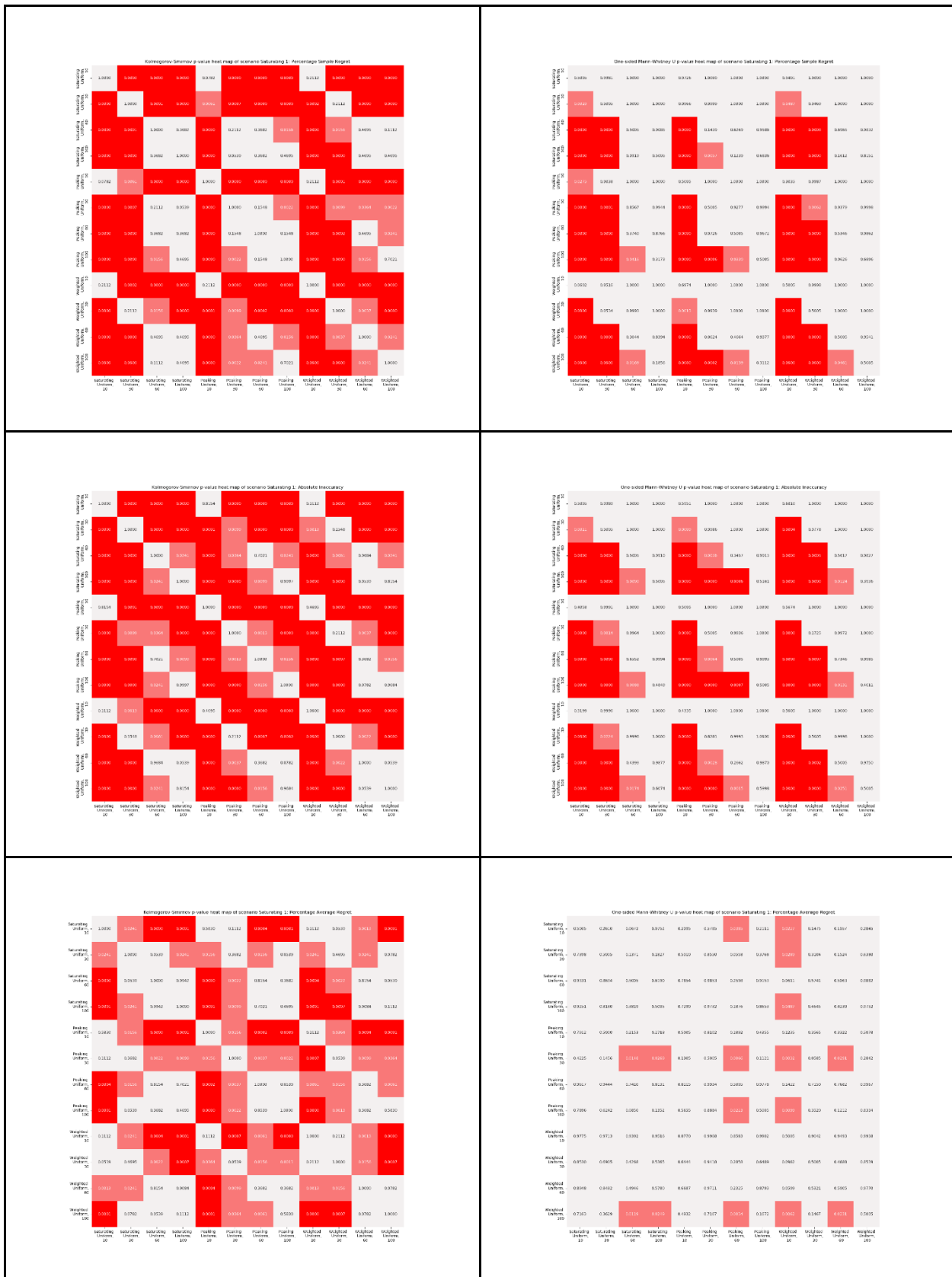
Figure S12.2.1. Kolmogorov–Smirnov (left) and Mann-Whitney U (right) heatmaps of p-values for objective 1 for scenario Saturating 1

Figure S12.2.2. Kolmogorov–Smirnov (left) and Mann-Whitney U (right) heatmaps of p-values for objective 1 for scenario Saturating 2

Figure S12.2.3. Kolmogorov–Smirnov (left) and Mann-Whitney U (right) heatmaps of p-values for objective 1 for scenario Saturating 3

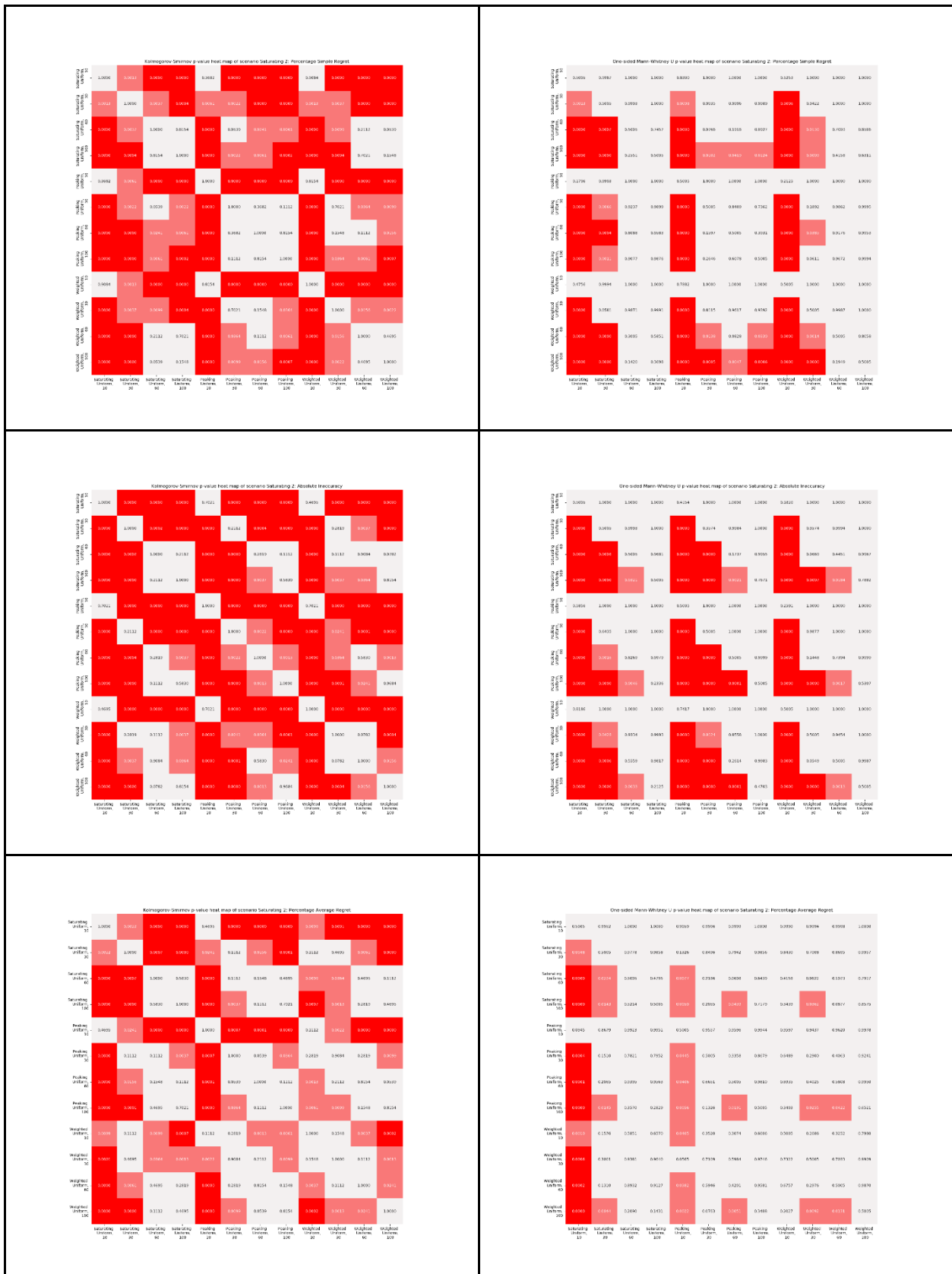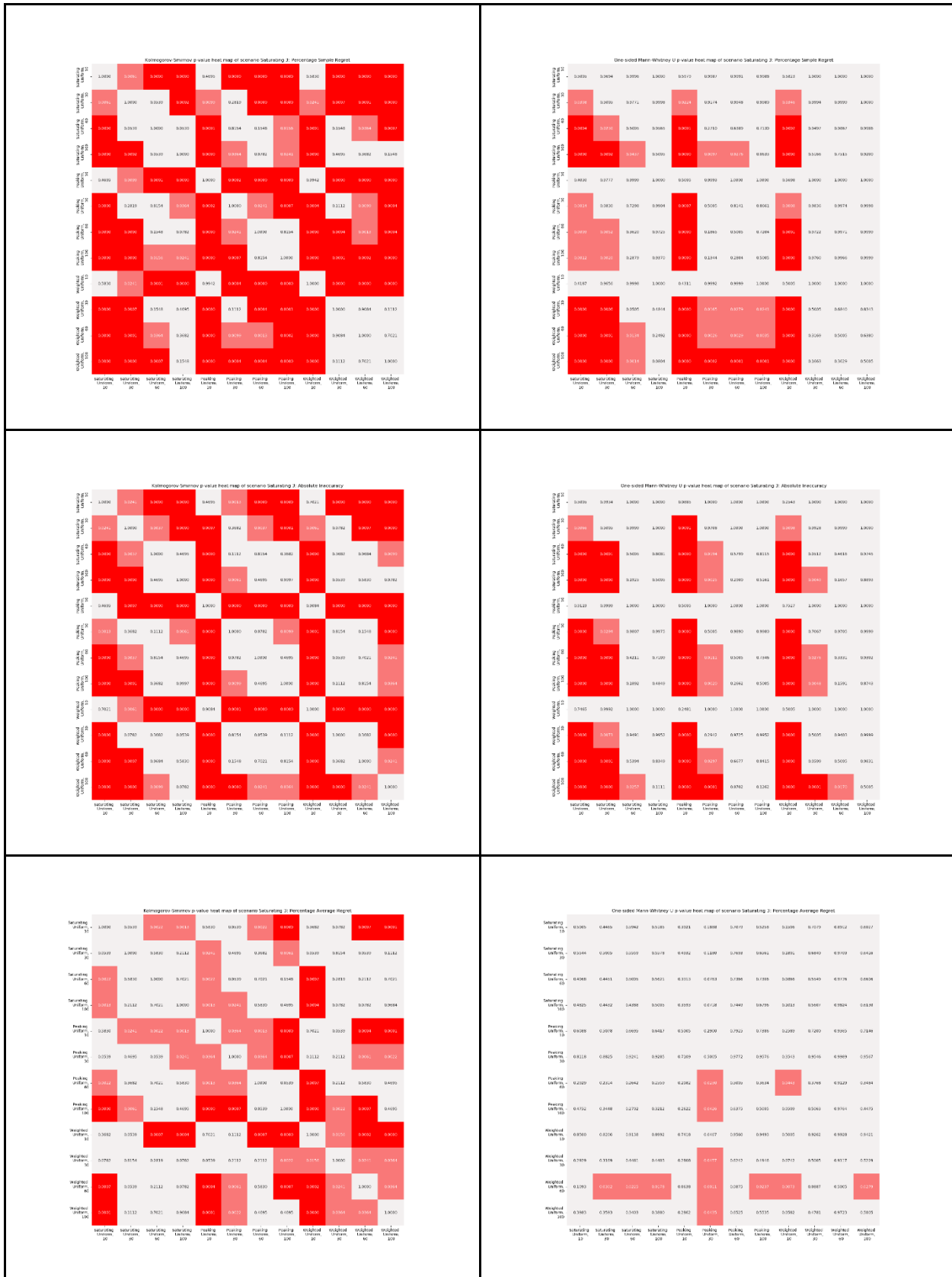**Figure S12.2.4. Kolmogorov–Smirnov (left) and Mann-Whitney U (right) heatmaps of p-values for objective 1 for scenario Saturating 4**

Figure S12.2.5. Kolmogorov–Smirnov (left) and Mann-Whitney U (right) heatmaps of p-values for objective 1 for scenario Saturating 5

Figure S12.2.6. Kolmogorov–Smirnov (left) and Mann-Whitney U (right) heatmaps of p-values for objective 1 for scenario Peaking 1

Figure S12.2.7. Kolmogorov–Smirnov (left) and Mann-Whitney U (right) heatmaps of p-values for objective 1 for scenario Peaking 2

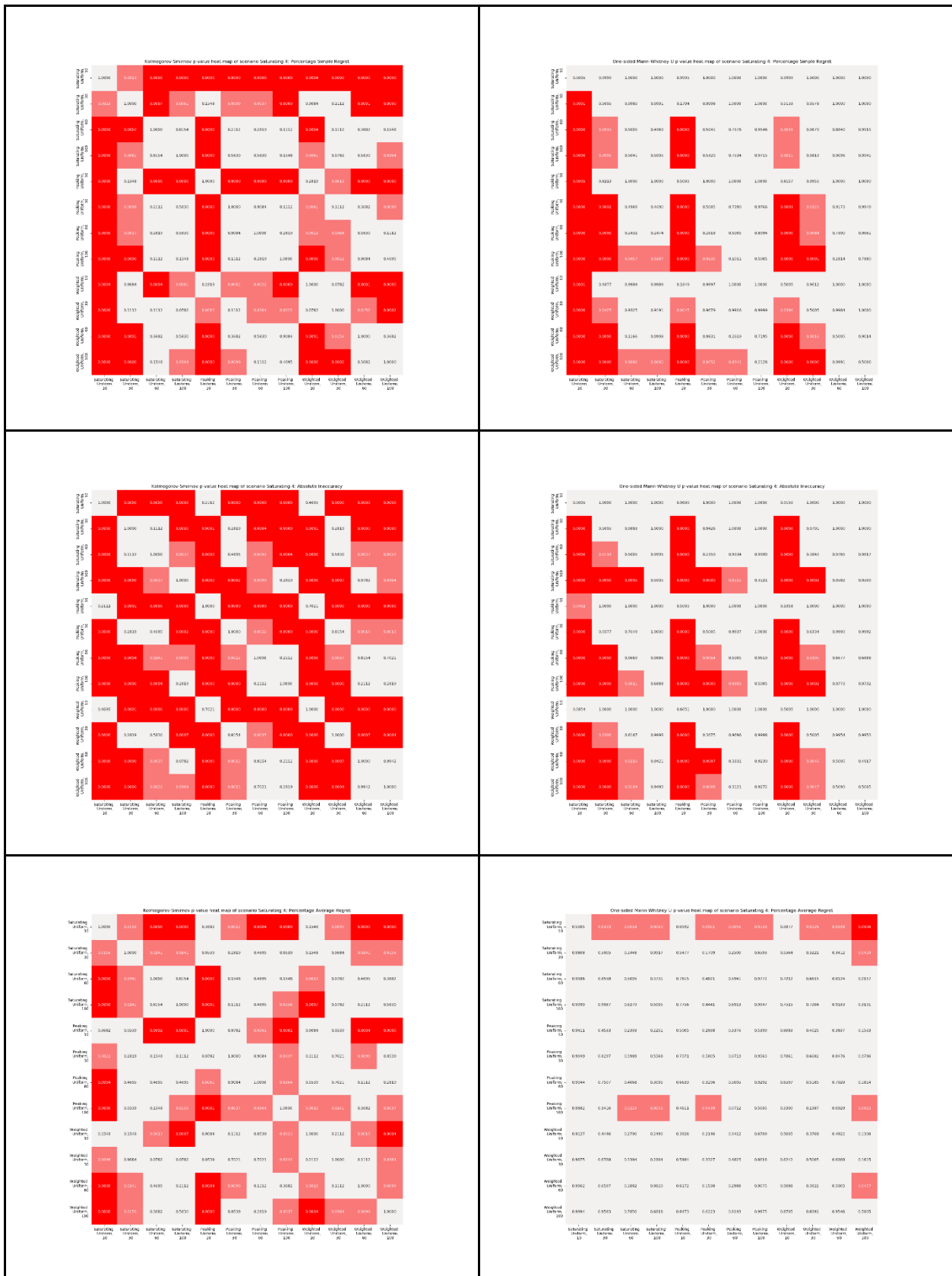**Figure S12.2.8. Kolmogorov–Smirnov (left) and Mann-Whitney U (right) heatmaps of p-values for objective 1 for scenario Peaking 3**

Figure S12.2.9. Kolmogorov–Smirnov (left) and Mann-Whitney U (right) heatmaps of p-values for objective 1 for scenario Peaking 4

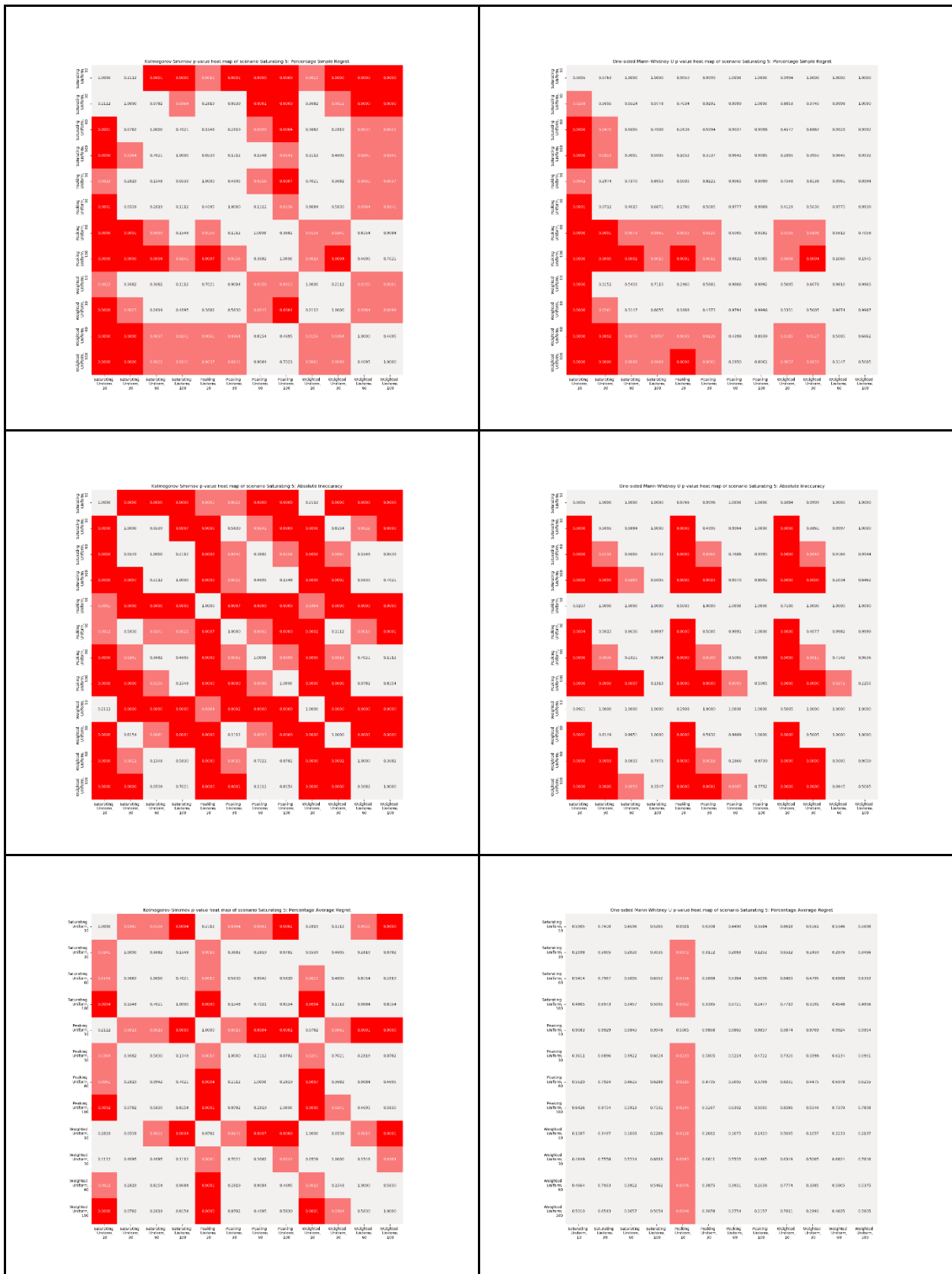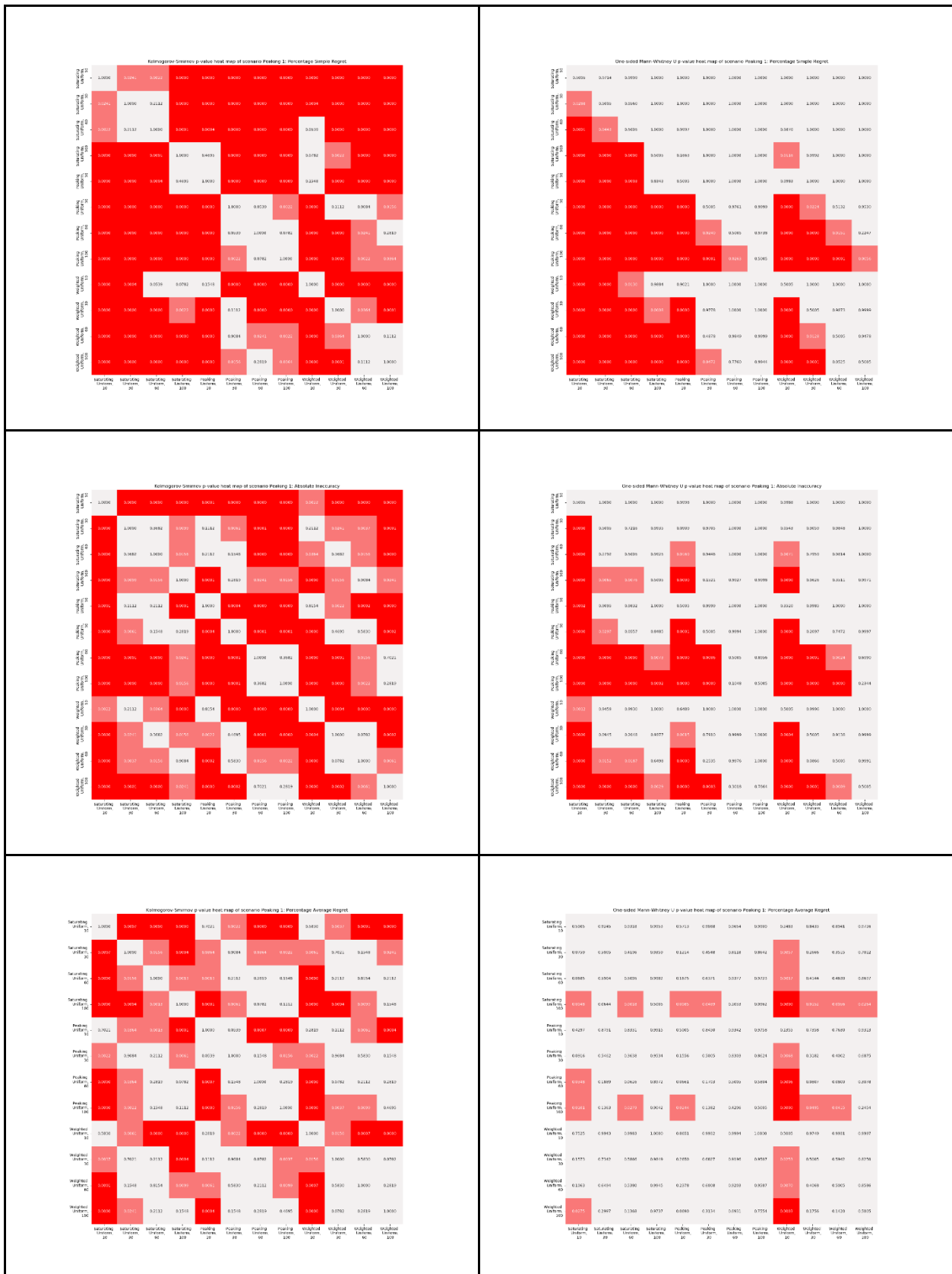**Figure S12.2.10. Kolmogorov–Smirnov (left) and Mann-Whitney U (right) heatmaps of p-values for objective 1 for scenario Peaking 5**

Figure S12.2.11. Kolmogorov–Smirnov (left) and Mann-Whitney U (right) heatmaps of p-values for objective 1 for scenario Other 1

Figure S12.2.12. Kolmogorov–Smirnov (left) and Mann-Whitney U (right) heatmaps of p-values for objective 1 for scenario Other 2

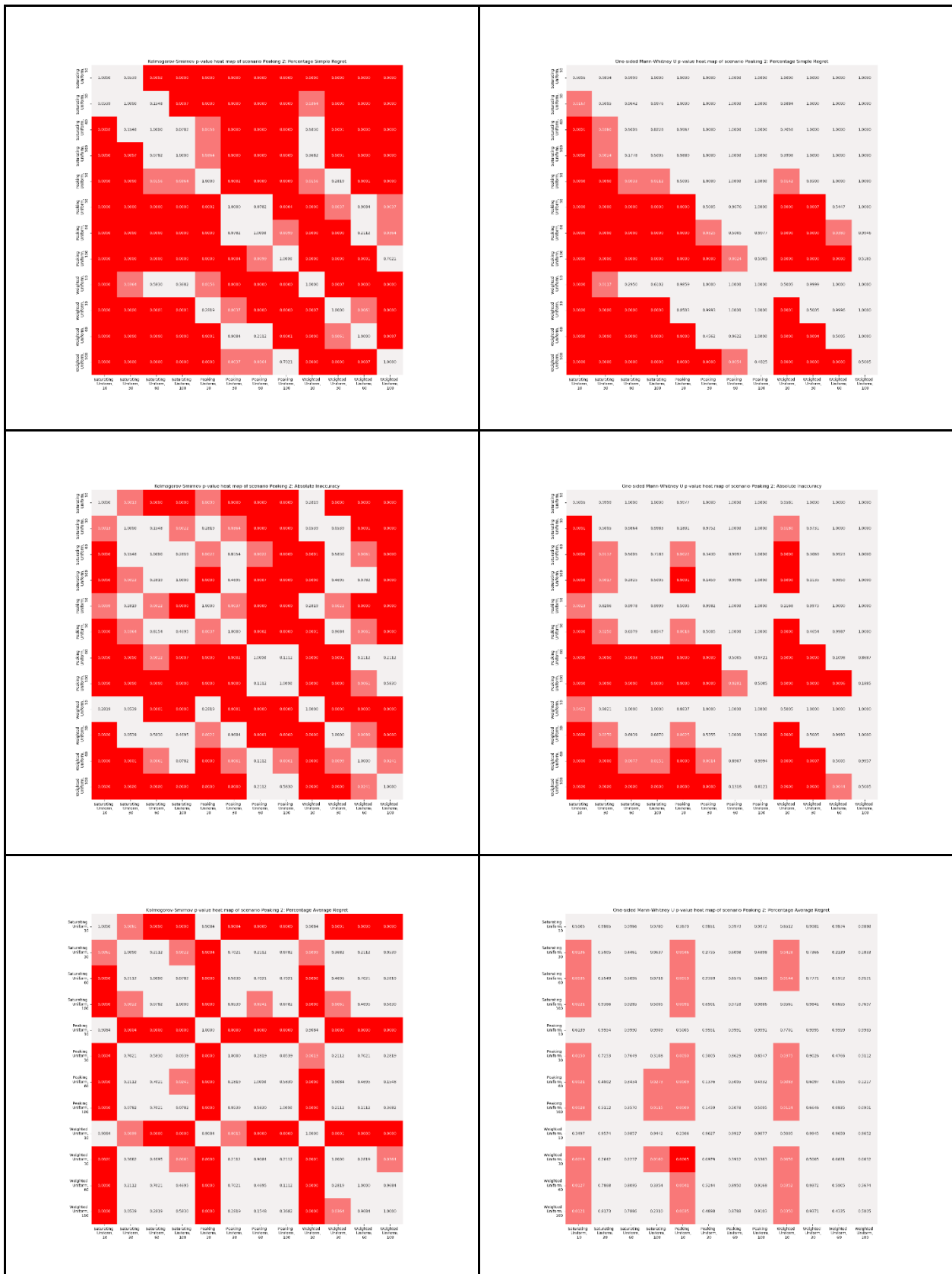**Figure S12.2.13. Kolmogorov–Smirnov (left) and Mann-Whitney U (right) heatmaps of p-values for objective 1 for scenario Other 3**

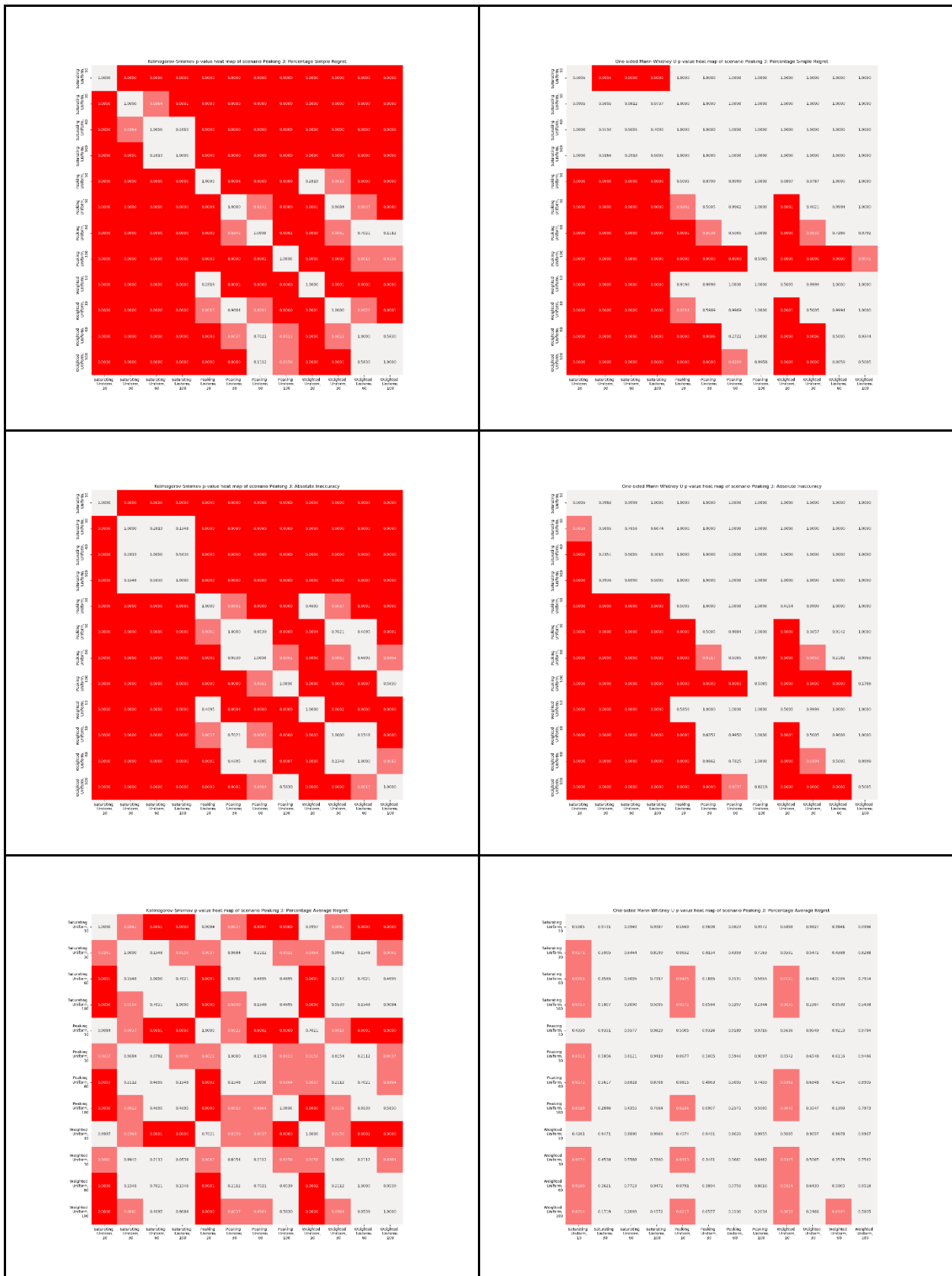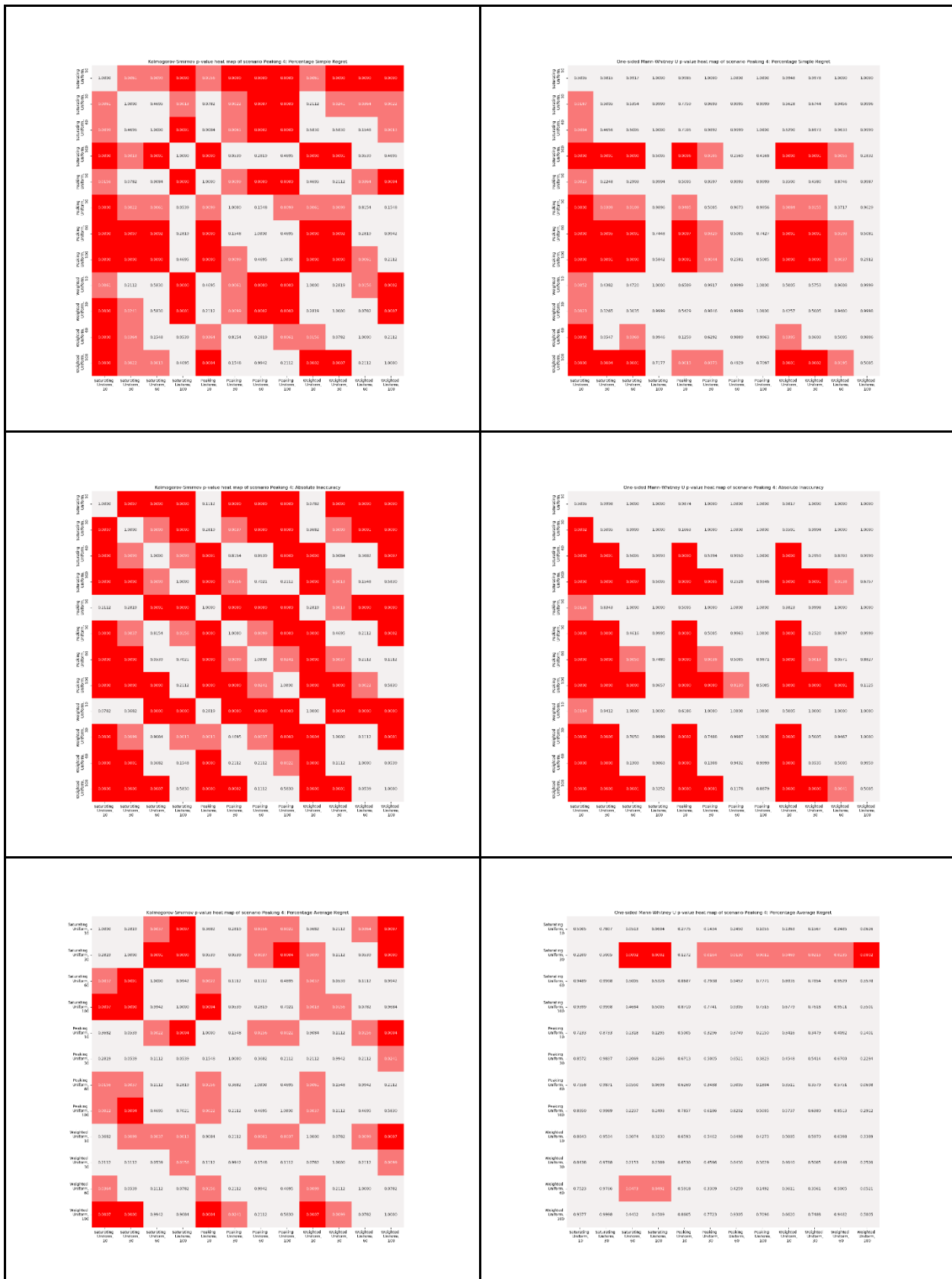**Figure S12.2.14. Kolmogorov–Smirnov (left) and Mann-Whitney U (right) heatmaps of p-values for objective 1 for scenario Other 4**

Here I show the p-values for objective 2 for the metrics of PSR, Absolute Inaccuracy, and PAR. These are the metrics for the data stratified on scenario. This is the data in Supplementary 12.

For interpretation, the Kolmogorov–Smirnov heatmaps are symmetric, with significance representing evidence that the true distribution for the approach-scenario test metrics of PSR, Absolute Inaccuracy, and PAR differ between the two dose-optimisation approaches across all scenarios. The One-sided Mann-Whitney U test heatmaps are not symmetric, with significance for the cell in row A and column B representing 'statistically significant' evidence that approach A was preferable to approach B with regards to that metric (e.g. lower PSR, lower Absolute Inaccuracy, Lower PAR).

Each of the below figures shows the Kolmogorov–Smirnov (left) and Mann-Whitney U (right) heatmaps of p-values for objective 2 for each scenario. These are for the metrics of PSR (top), Absolute Inaccuracy (middle) and PAR (bottom). Cells with a light pink hue represent the test statistic for that comparison would be significant under the threshold p<0.05, cells with a red hue represent the test statistic for that comparison would be significant under the threshold p<0.05 with Bonferroni multiple comparison correction.

Figure S12.4.1. Kolmogorov–Smirnov (left) and Mann-Whitney U (right) heatmaps of p-values for objective 2 for scenario Saturating 1

Figure S12.4.2. Kolmogorov–Smirnov (left) and Mann-Whitney U (right) heatmaps of p-values for objective 2 for scenario Saturating 2
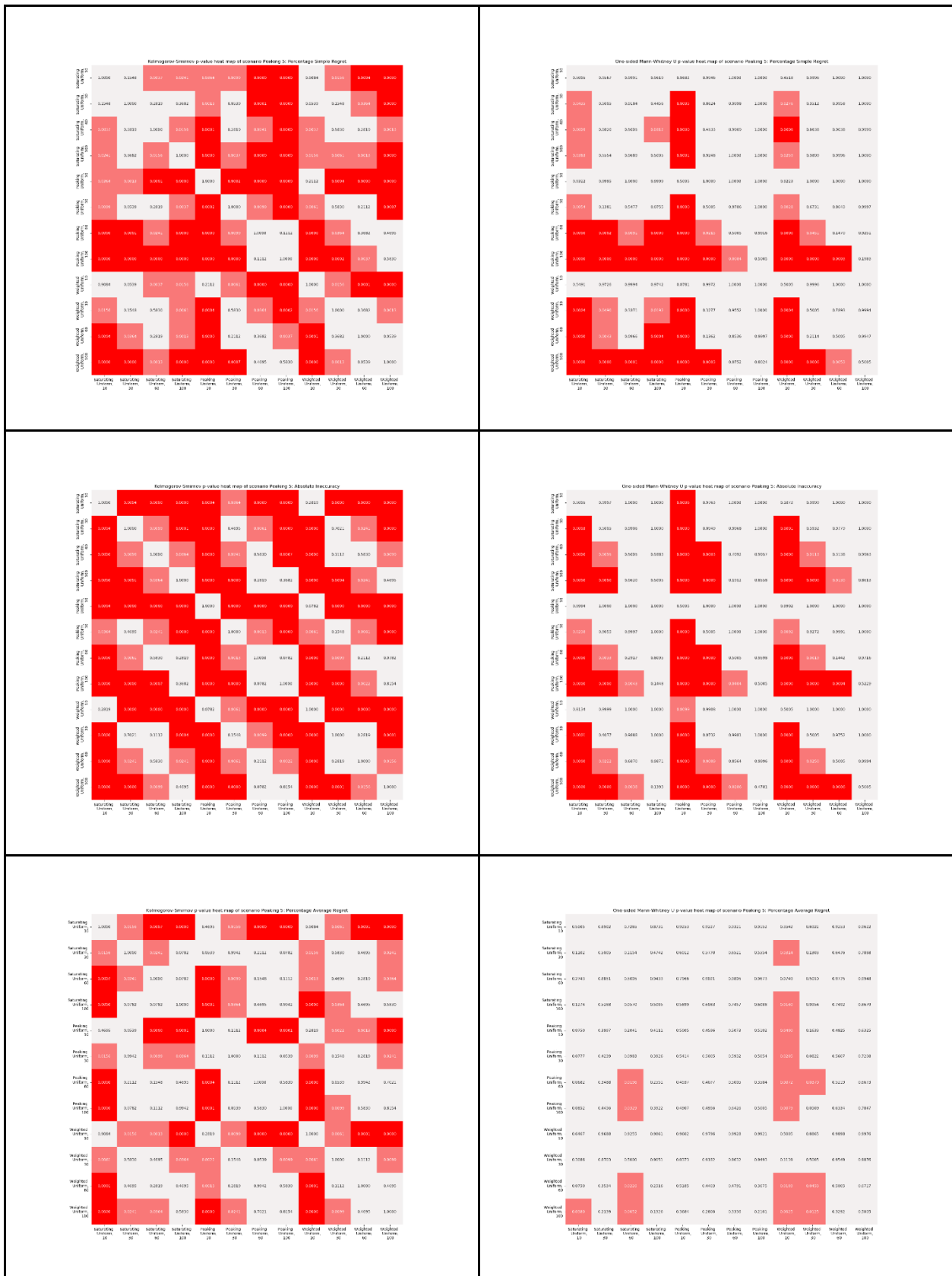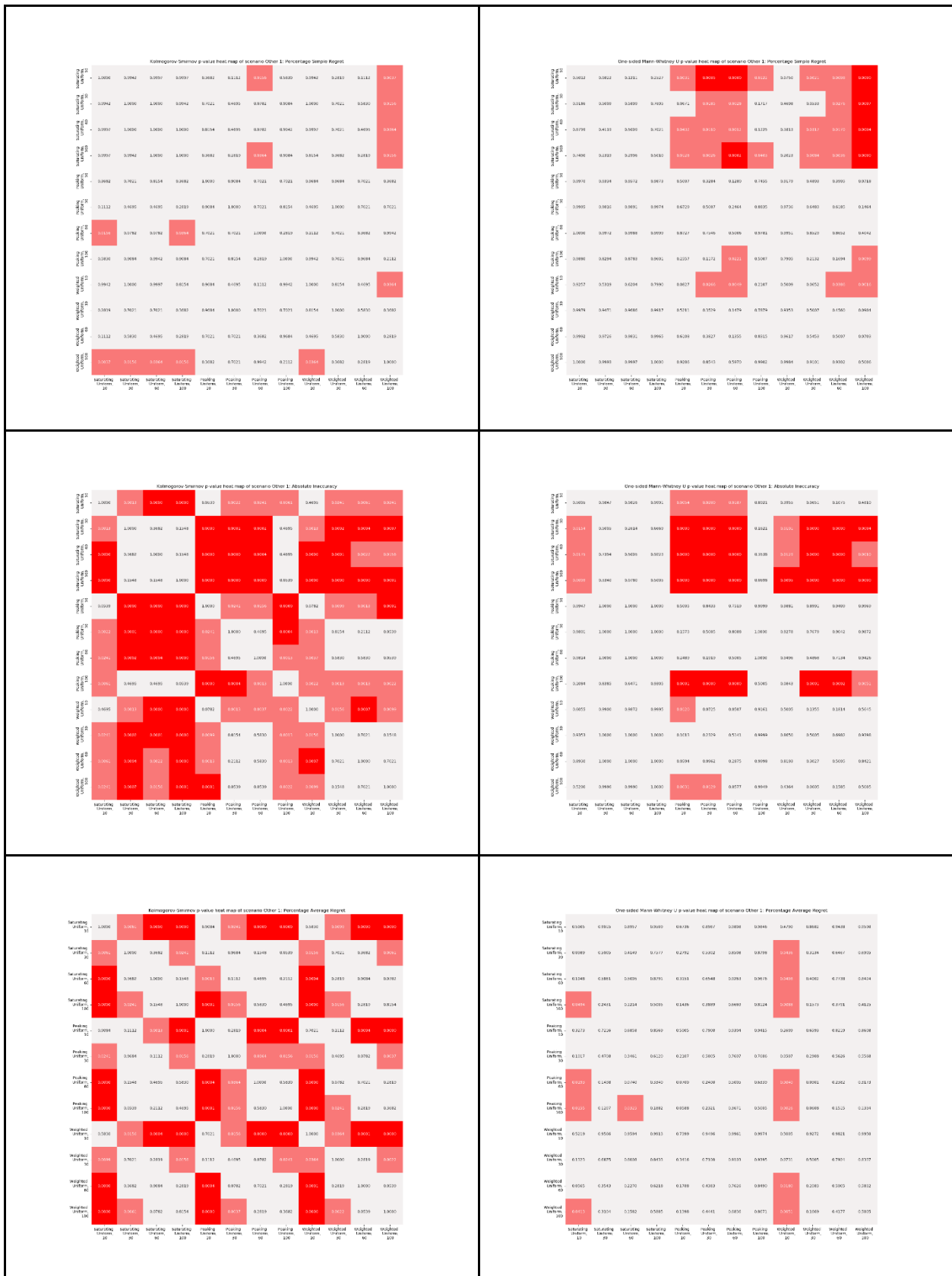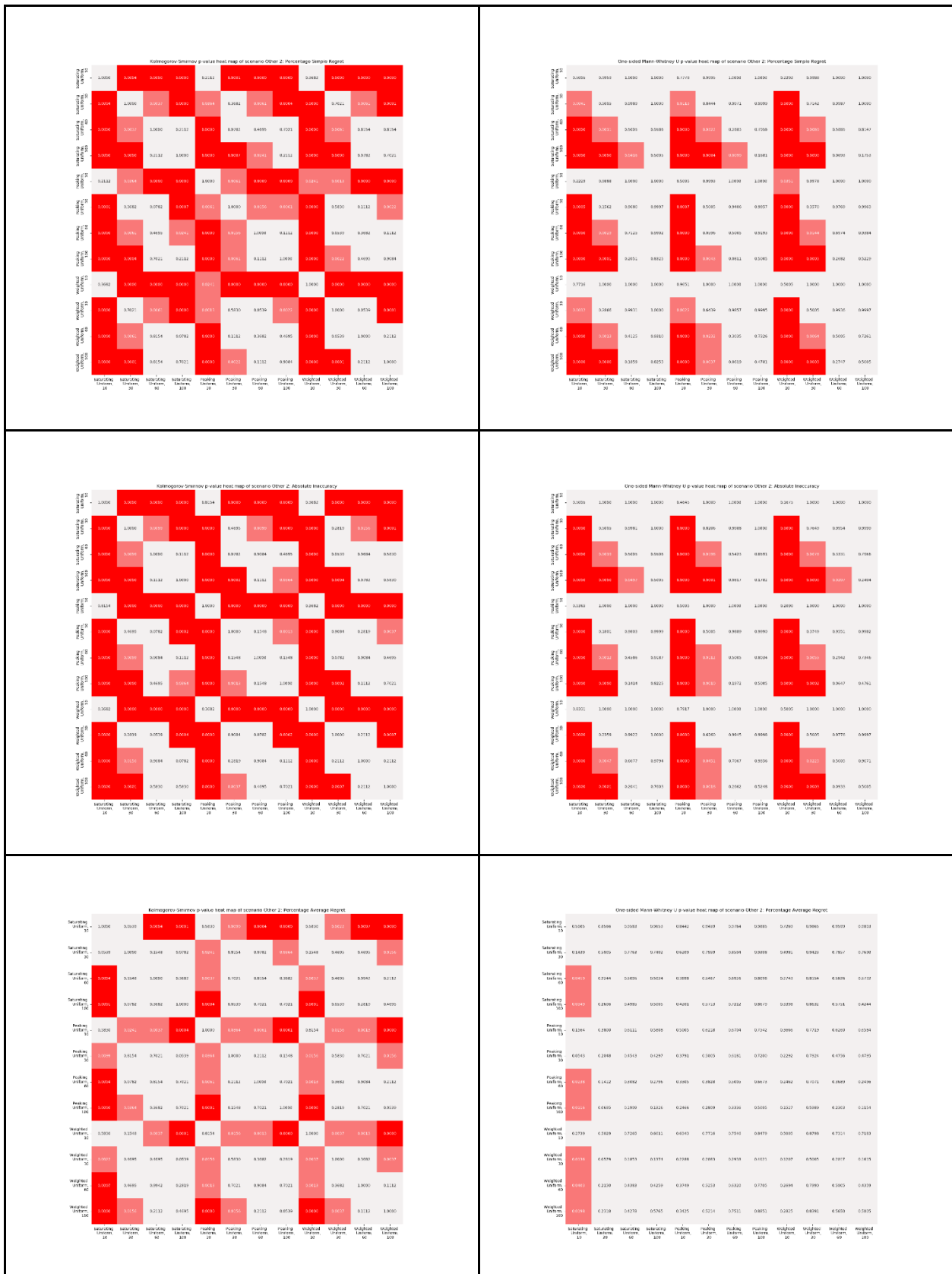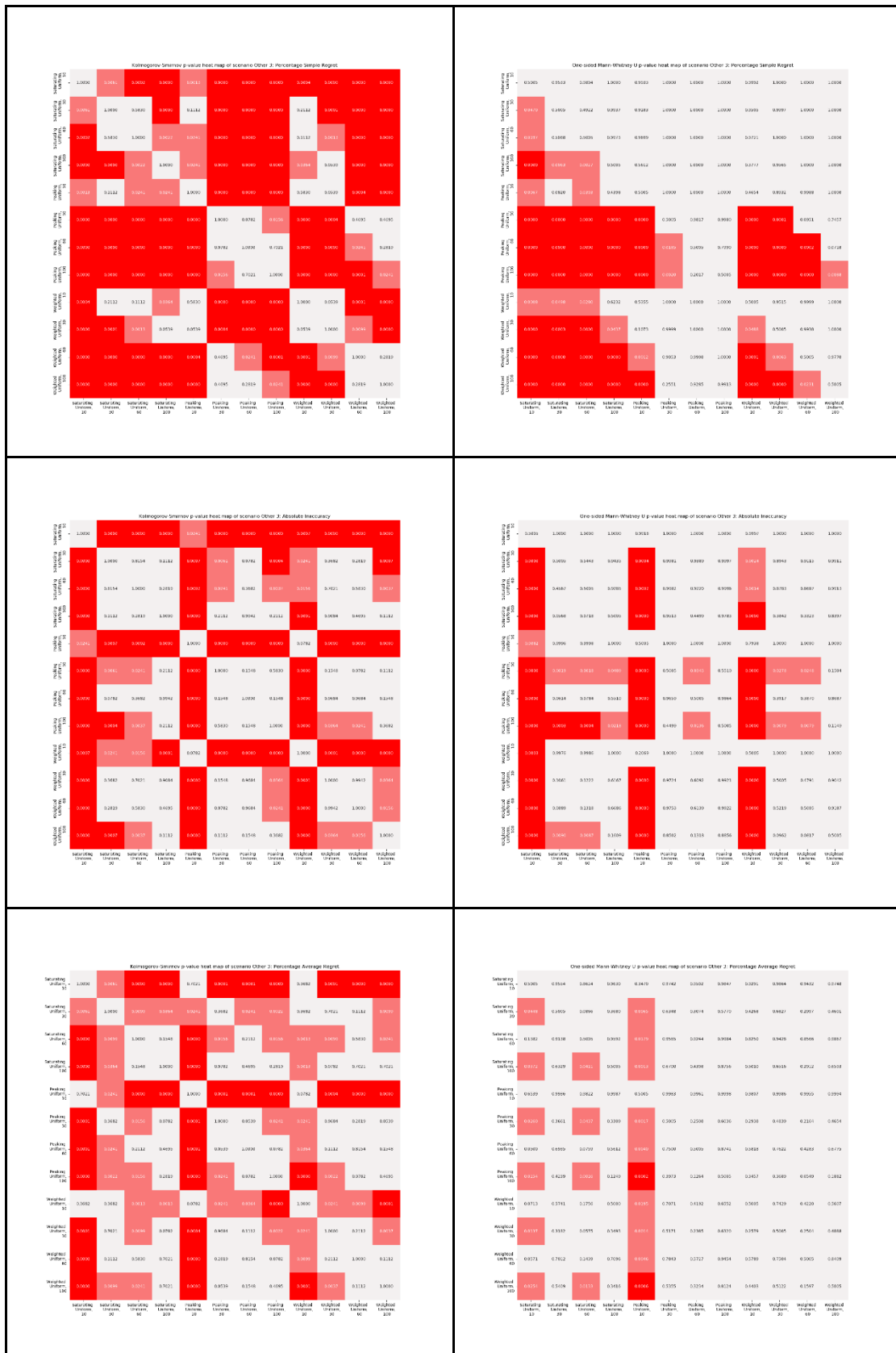
Scenario Saturating 3
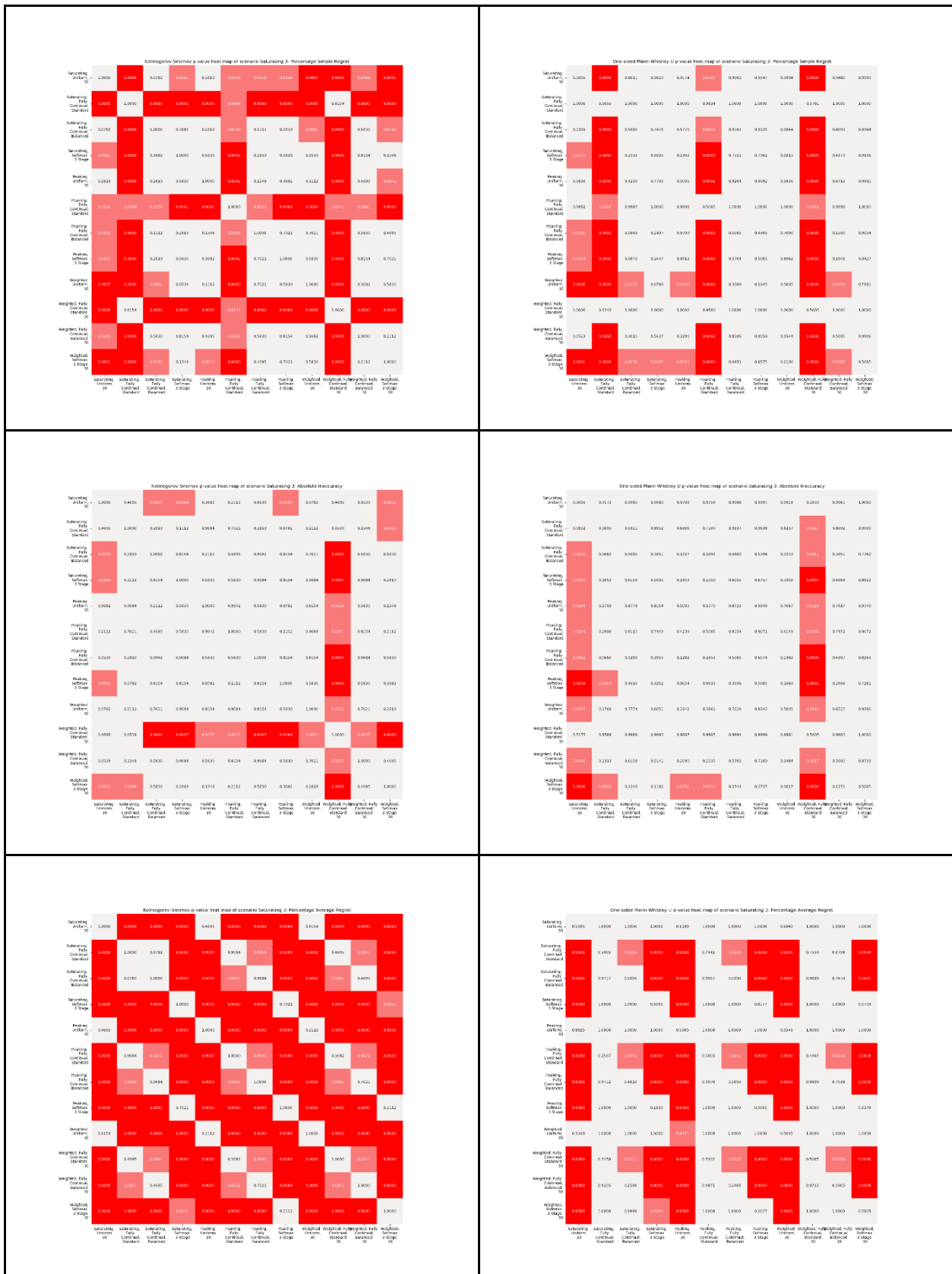
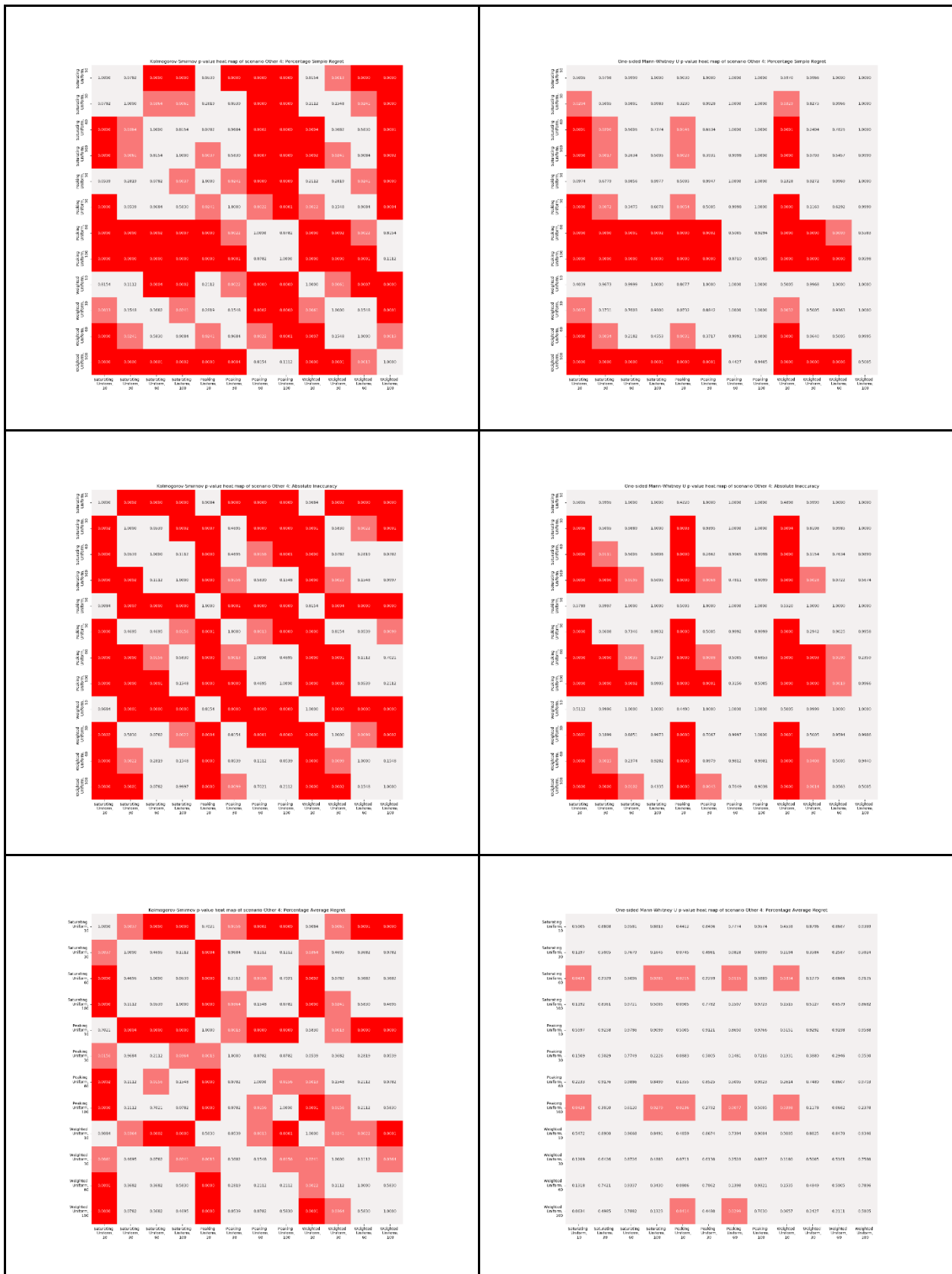**Figure S12.4.3. Kolmogorov–Smirnov (left) and Mann-Whitney U (right) heatmaps of p-values for objective 2 for scenario Saturating 3**
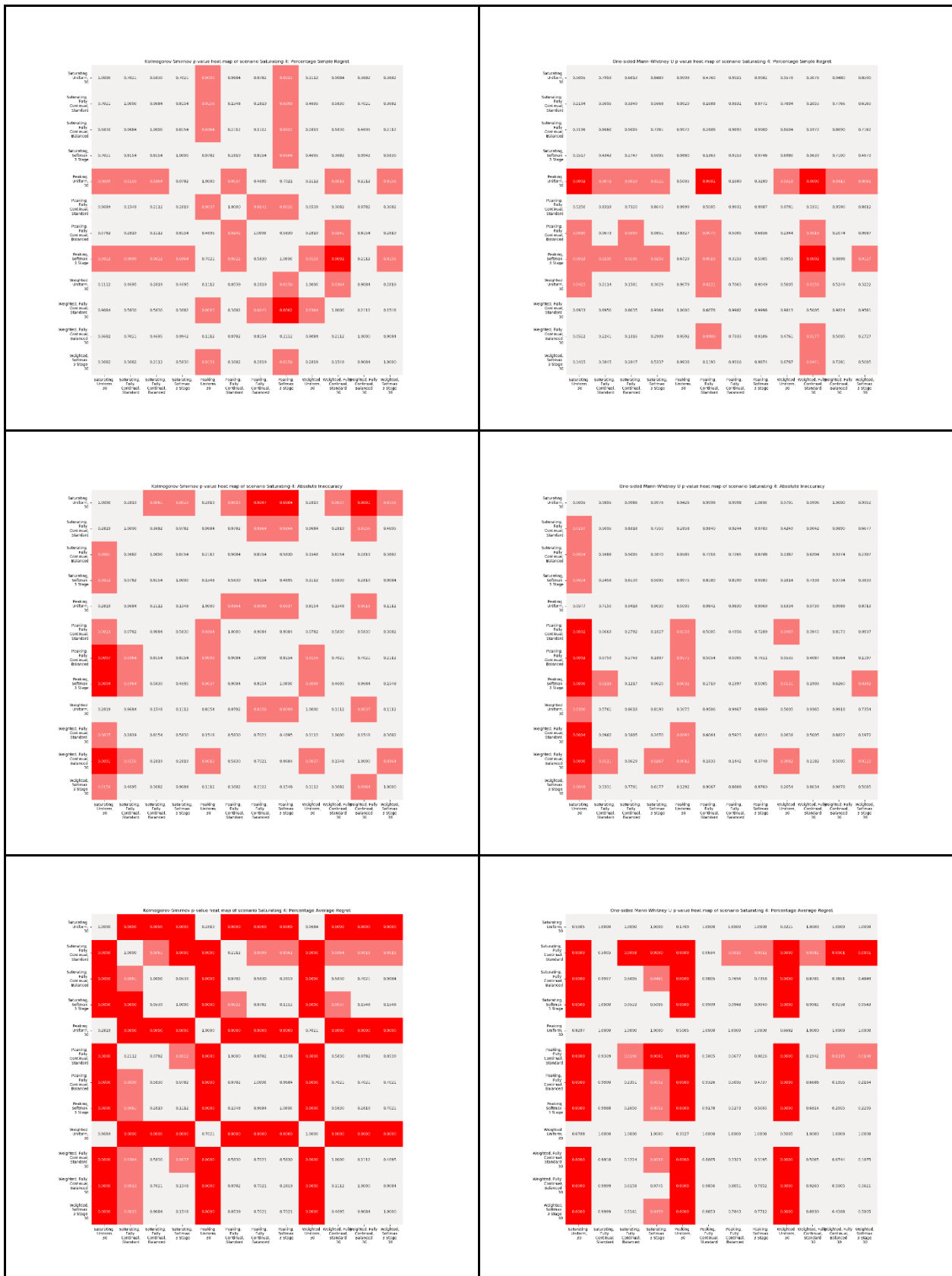
Figure S12.4.4. Kolmogorov–Smirnov (left) and Mann-Whitney U (right) heatmaps of p-values for objective 2 for scenario Saturating 4
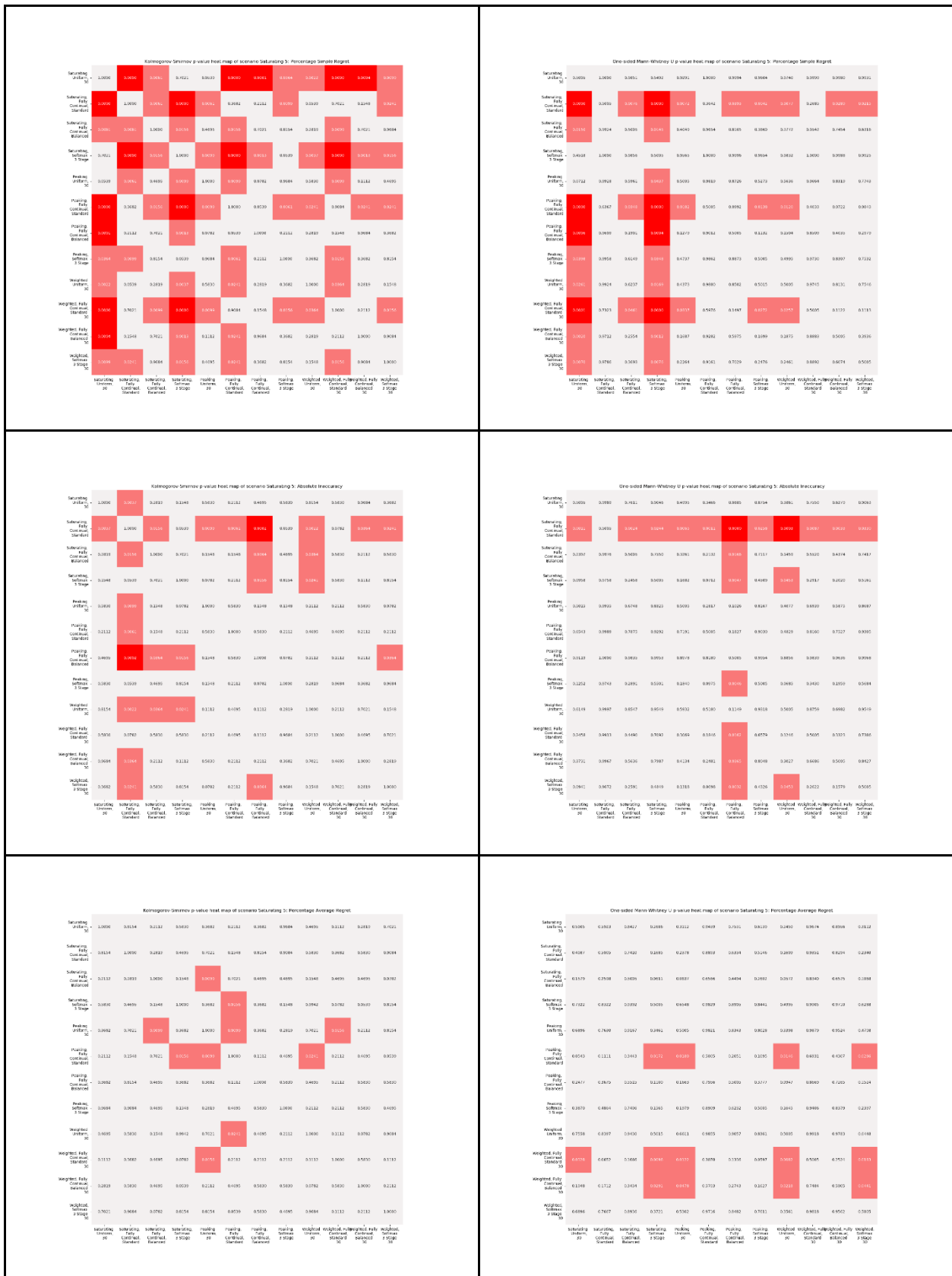
Figure S12.4.5. Kolmogorov–Smirnov (left) and Mann-Whitney U (right) heatmaps of p-values for objective 2 for scenario Saturating 5
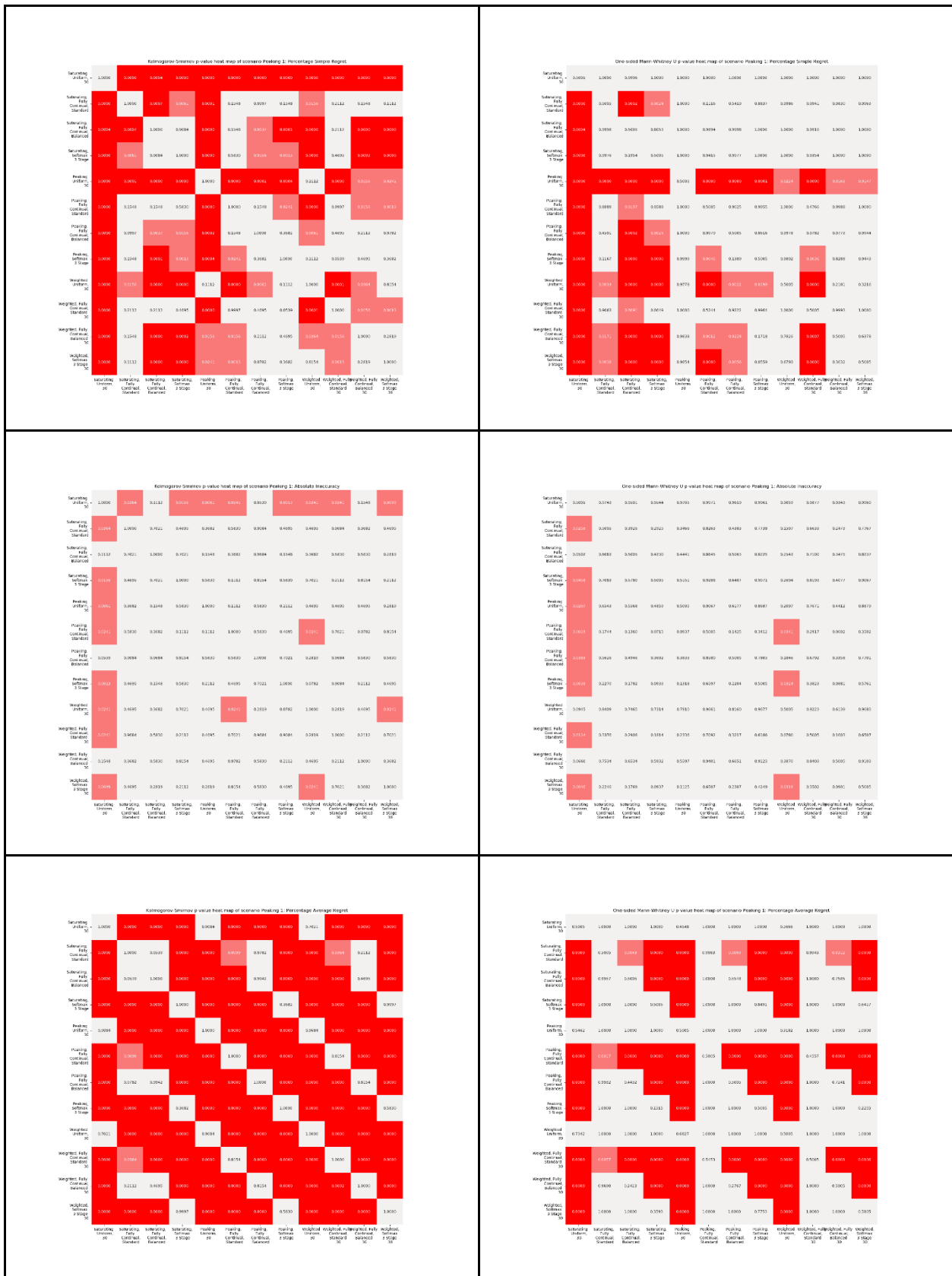
Scenario Peaking 1

Figure S12.4.6. Kolmogorov–Smirnov (left) and Mann-Whitney U (right) heatmaps of p-values for objective 2 for scenario Peaking 1
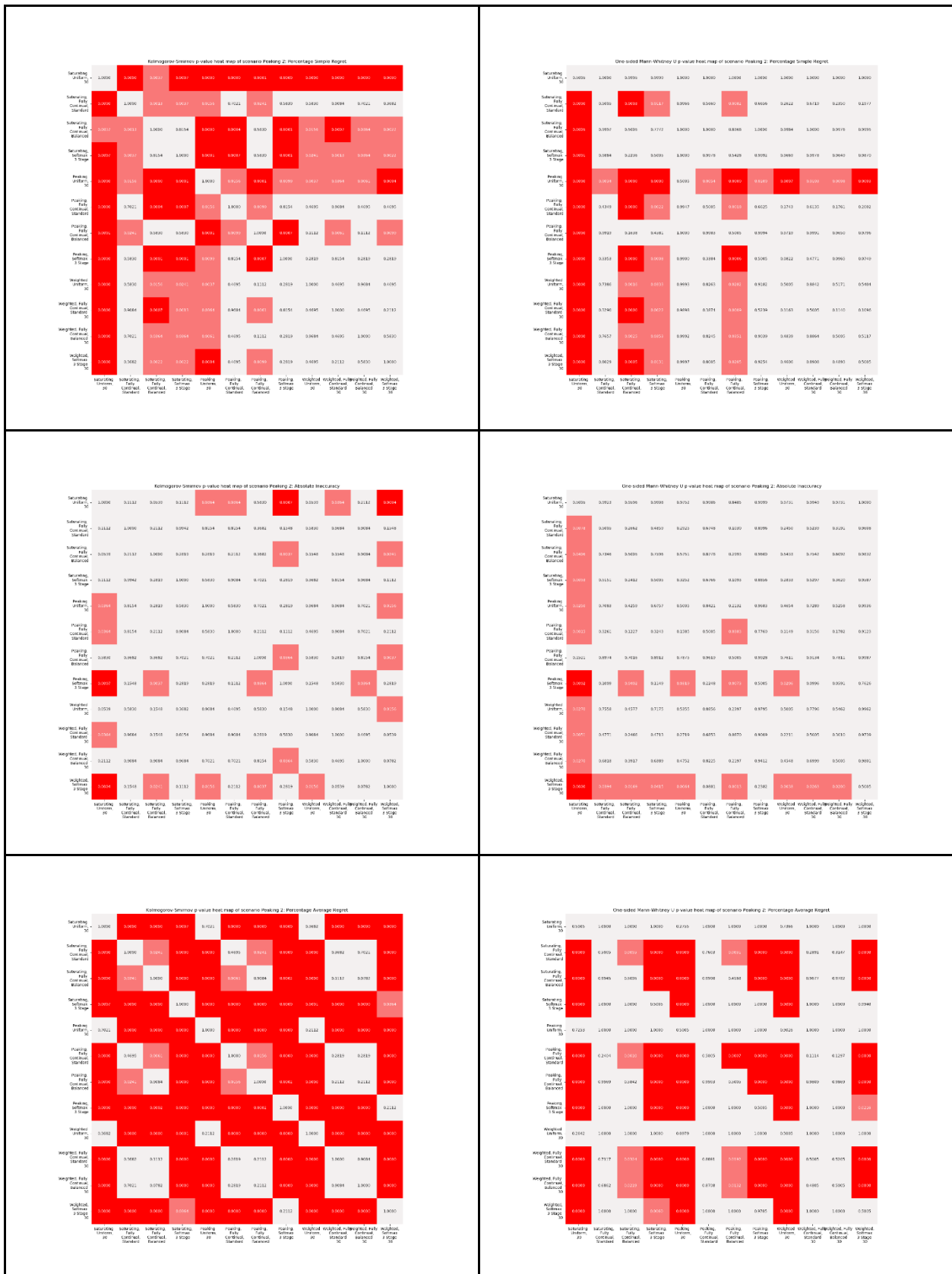
Figure S12.4.7. Kolmogorov–Smirnov (left) and Mann-Whitney U (right) heatmaps of p-values for objective 2 for scenario Peaking 2
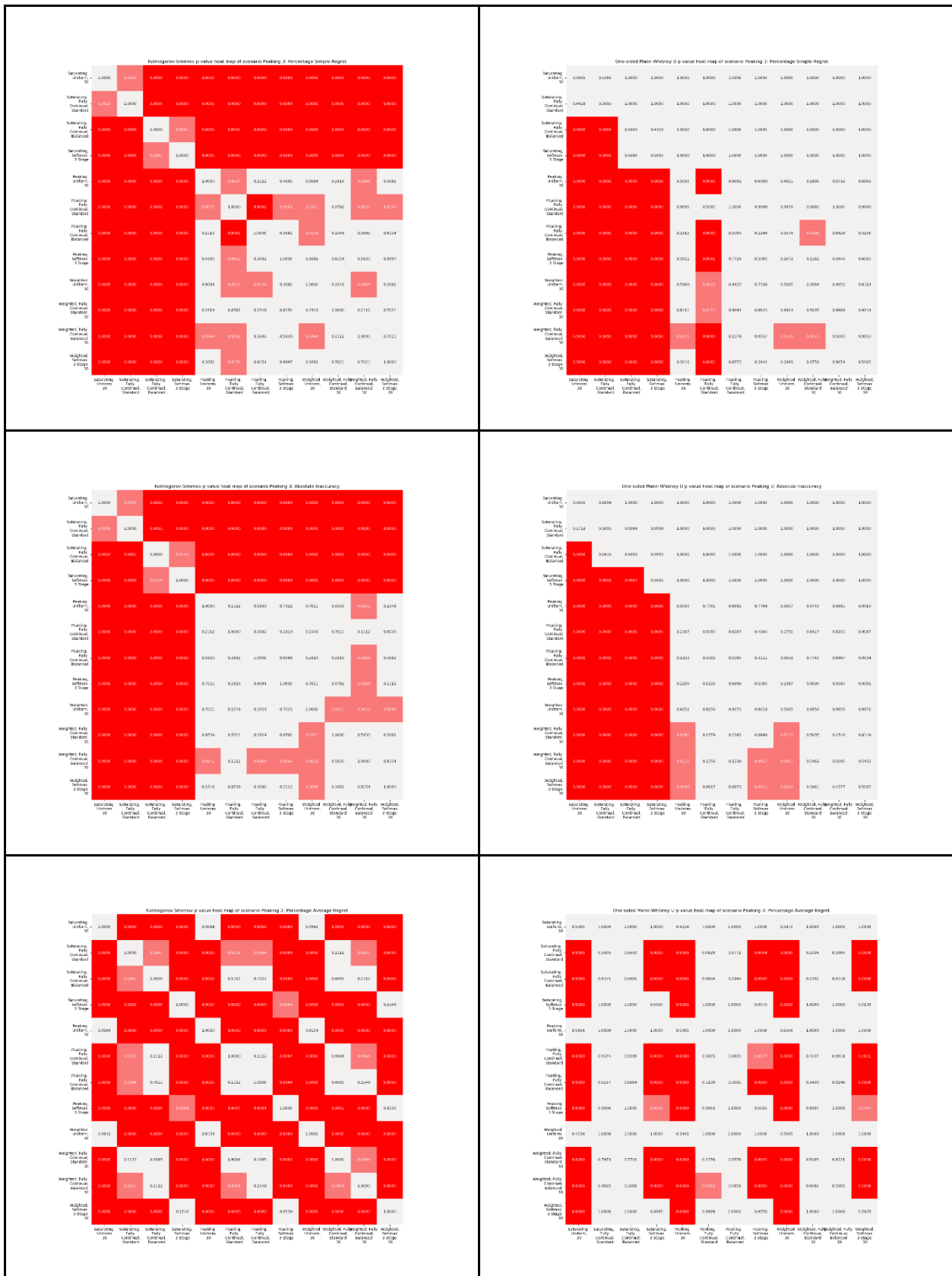
Figure S12.4.8. Kolmogorov–Smirnov (left) and Mann-Whitney U (right) heatmaps of p-values for objective 2 for scenario Peaking 3
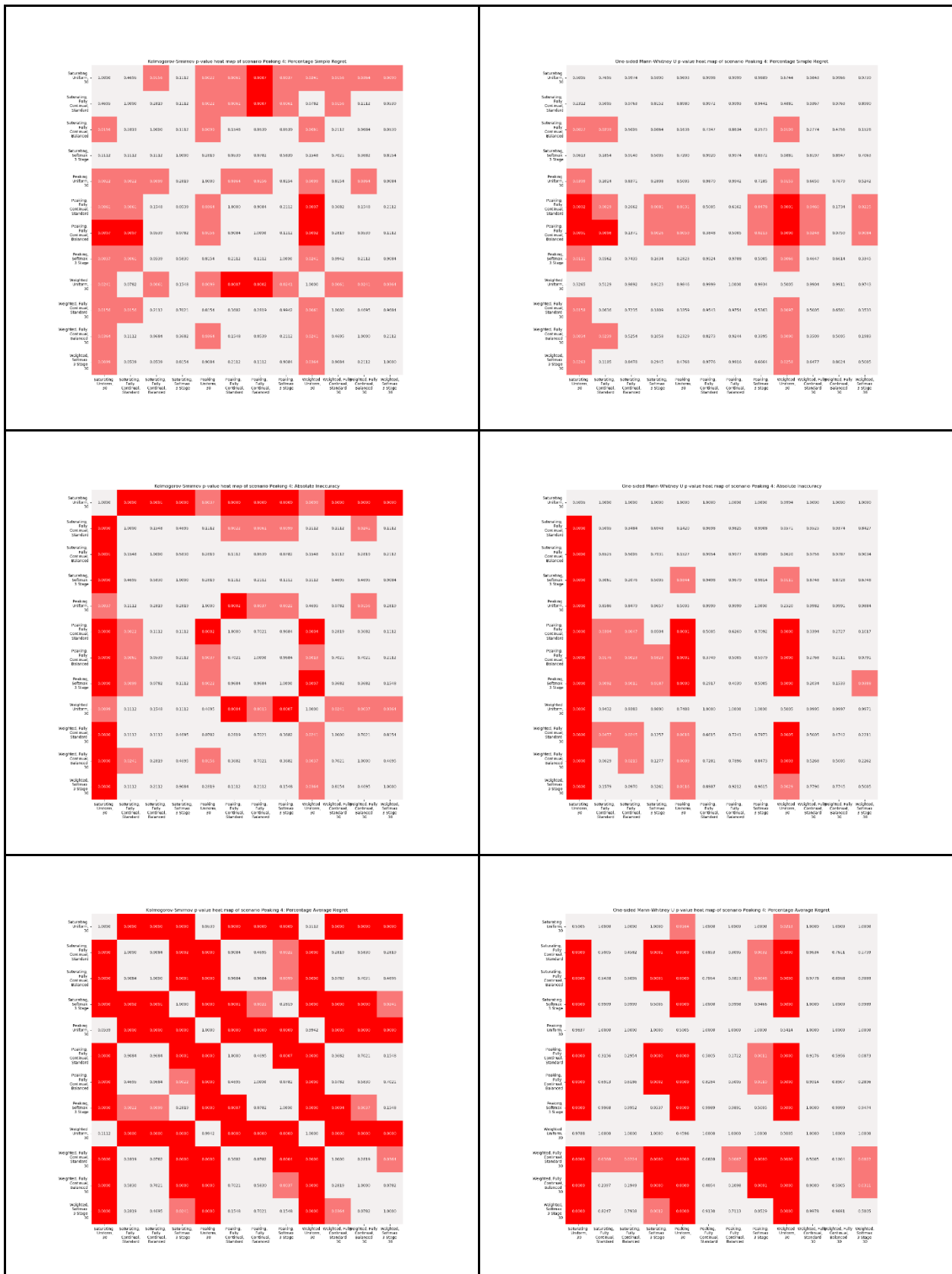
Figure S12.4.9. Kolmogorov–Smirnov (left) and Mann-Whitney U (right) heatmaps of p-values for objective 2 for scenario Peaking 4
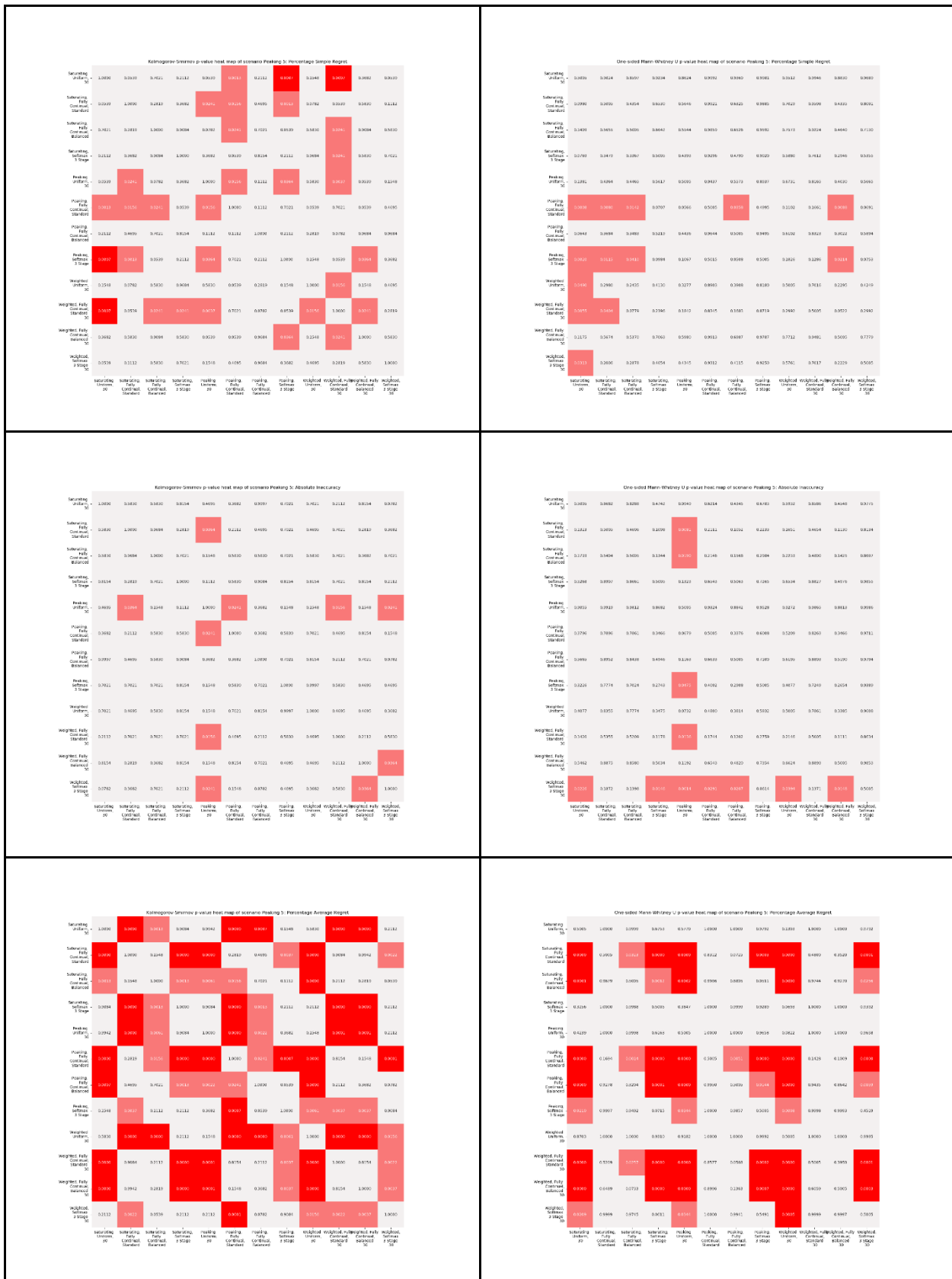
Figure S12.4.10. Kolmogorov–Smirnov (left) and Mann-Whitney U (right) heatmaps of p-values for objective 2 for scenario Peaking 5
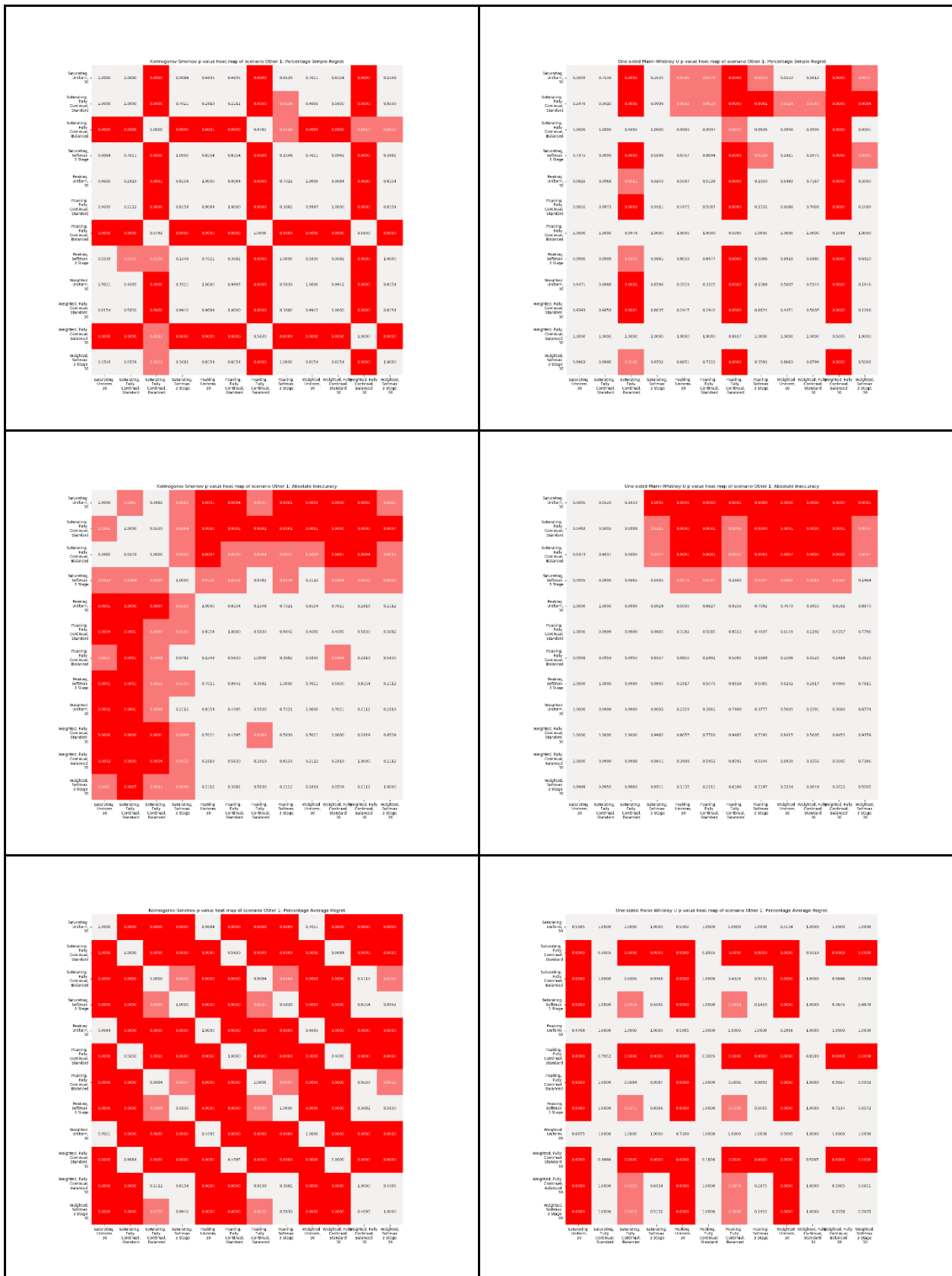
Figure S12.4.11. Kolmogorov–Smirnov (left) and Mann-Whitney U (right) heatmaps of p-values for objective 2 for scenario Other 1

Scenario Other 2
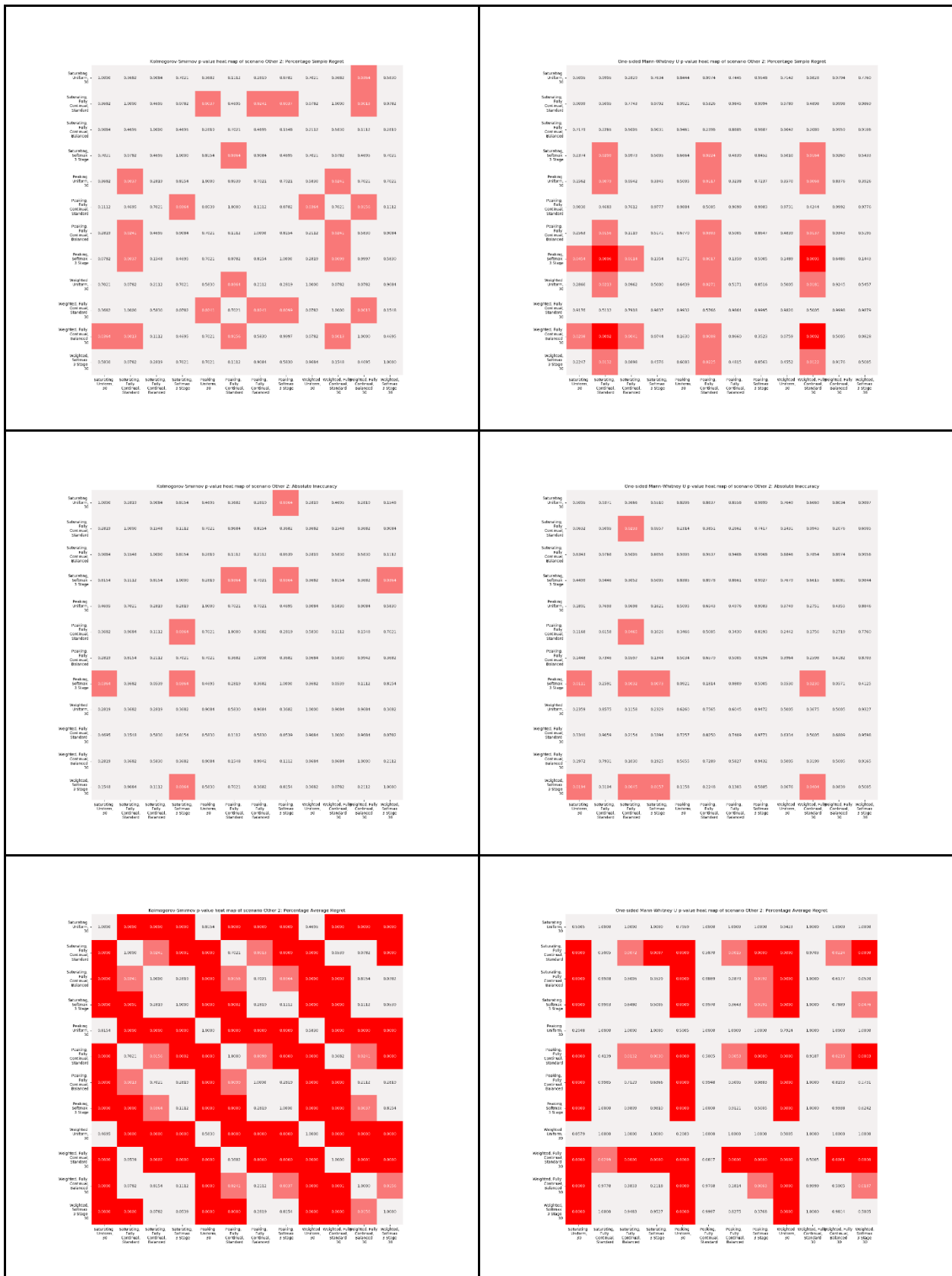
Figure S12.4.12. Kolmogorov–Smirnov (left) and Mann-Whitney U (right) heatmaps of p-values for objective 2 for scenario Other 2
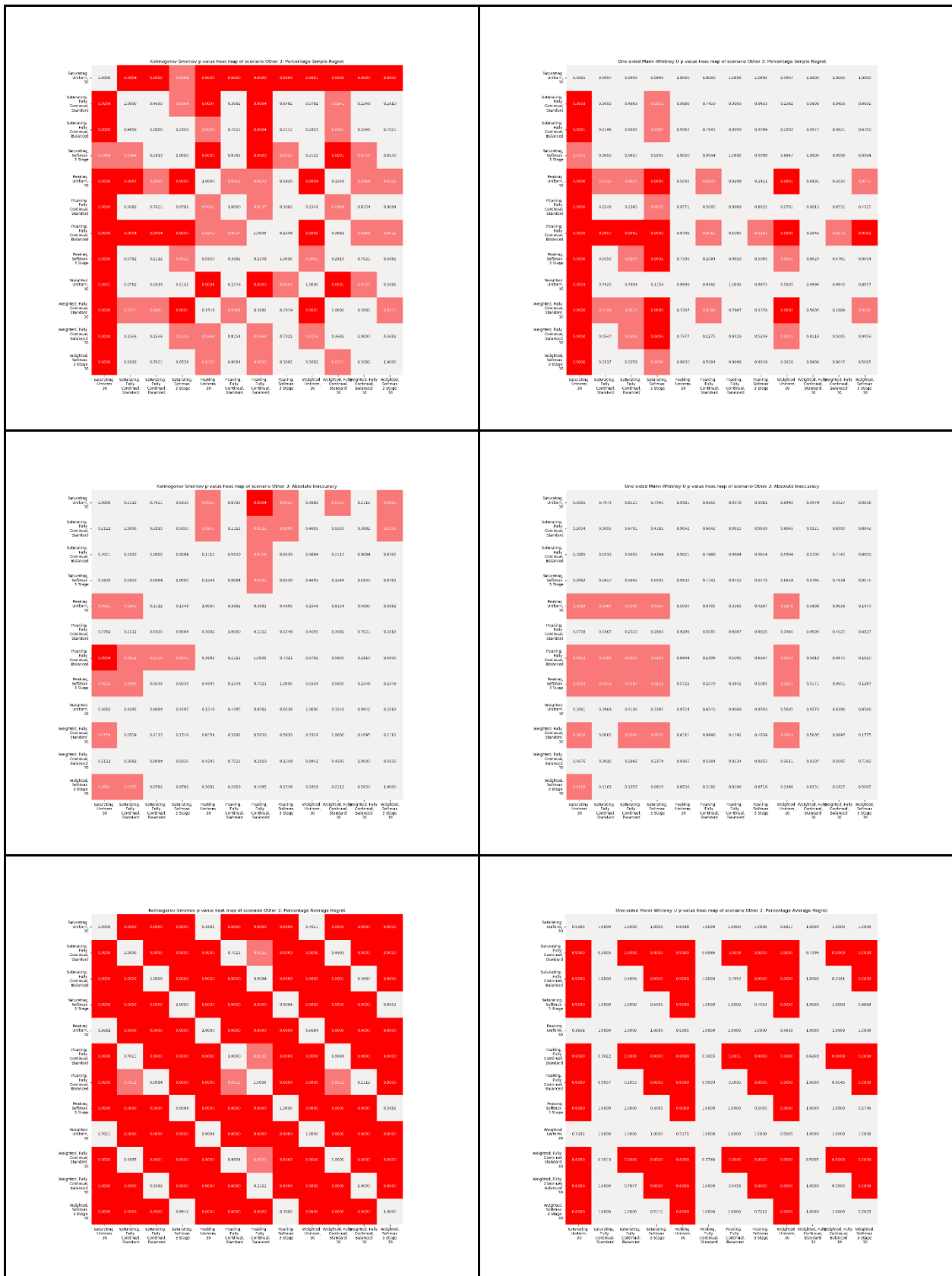
Figure S12.4.13. Kolmogorov–Smirnov (left) and Mann-Whitney U (right) heatmaps of p-values for objective 2 for scenario Other 3
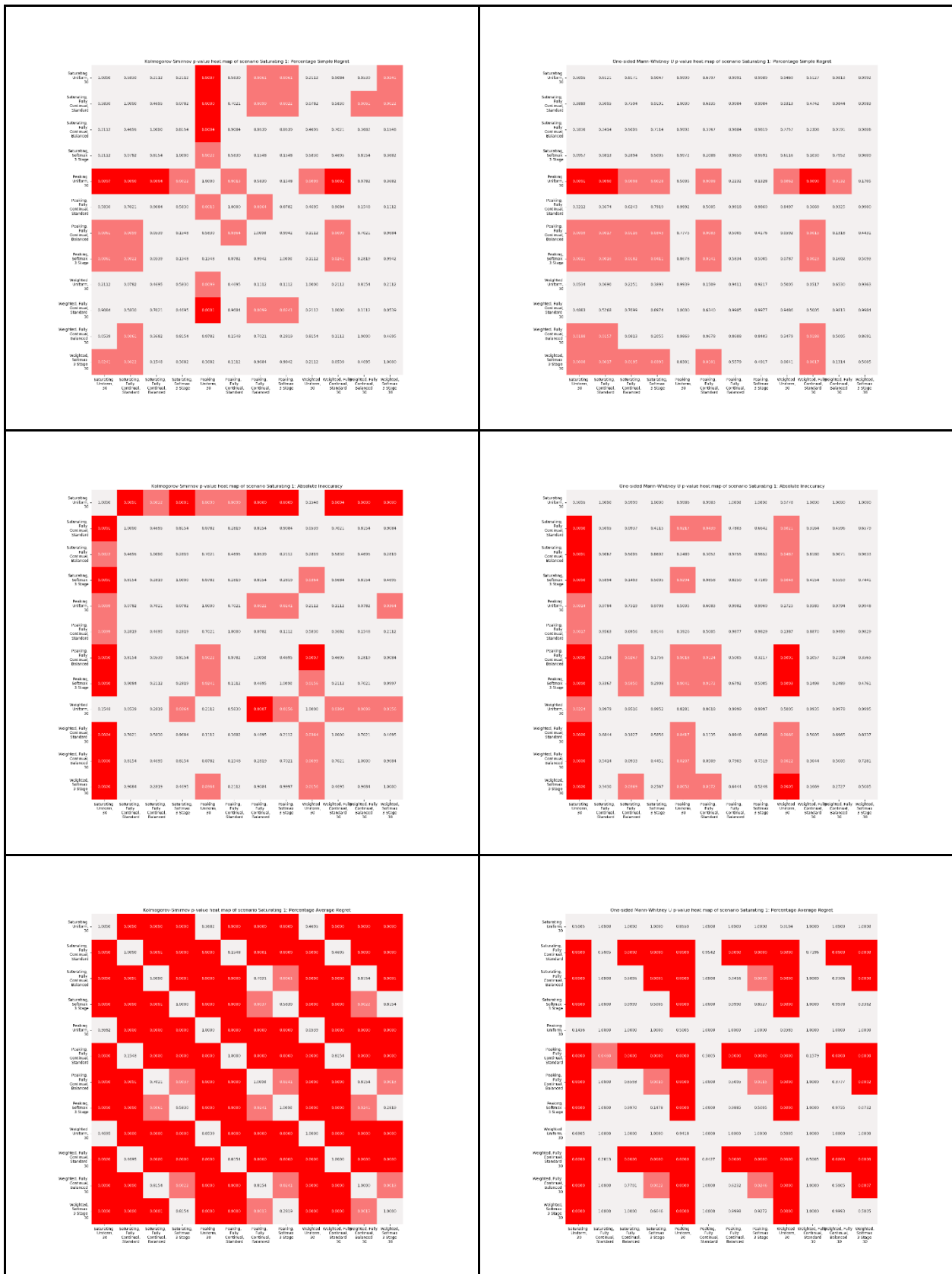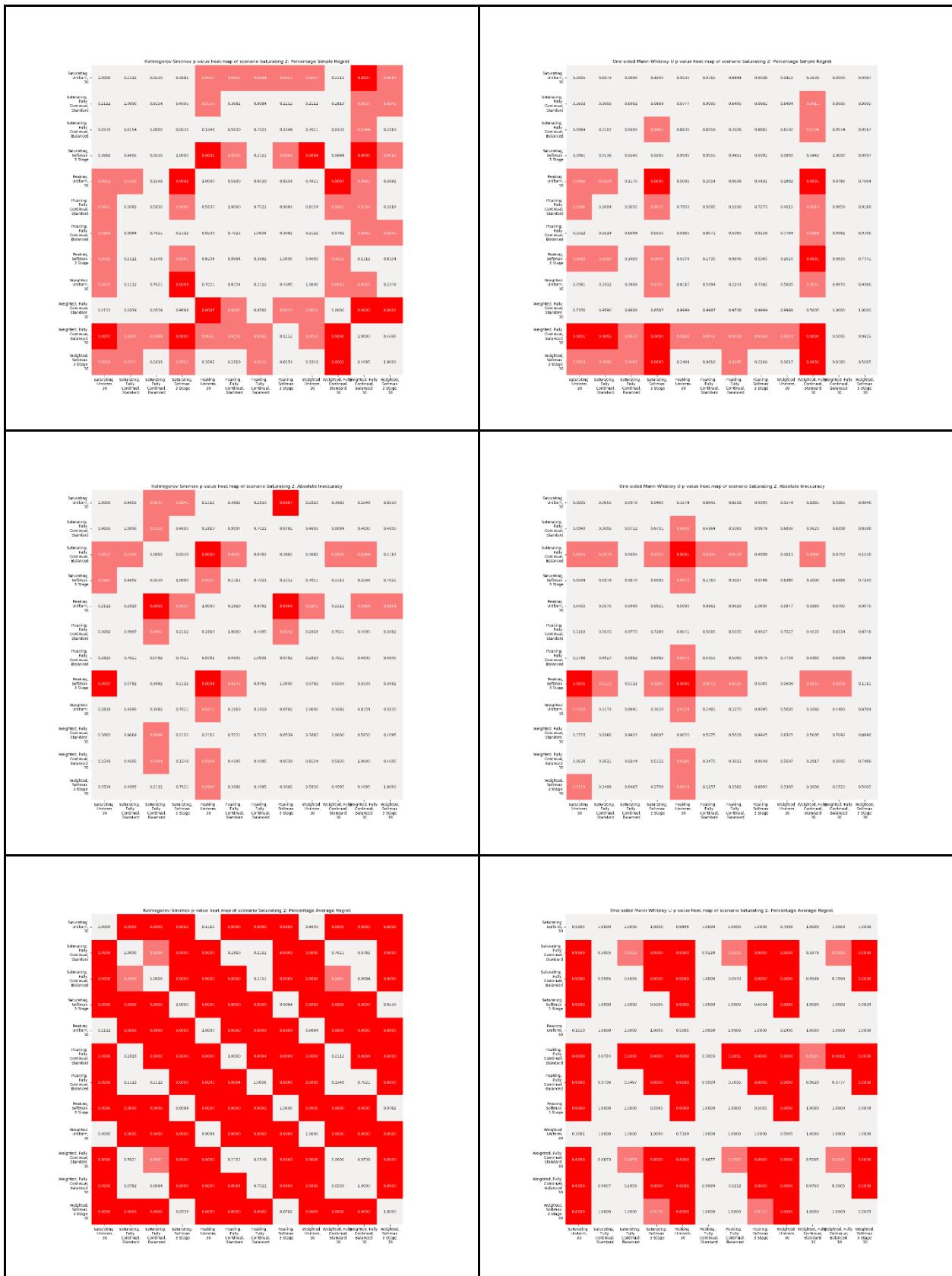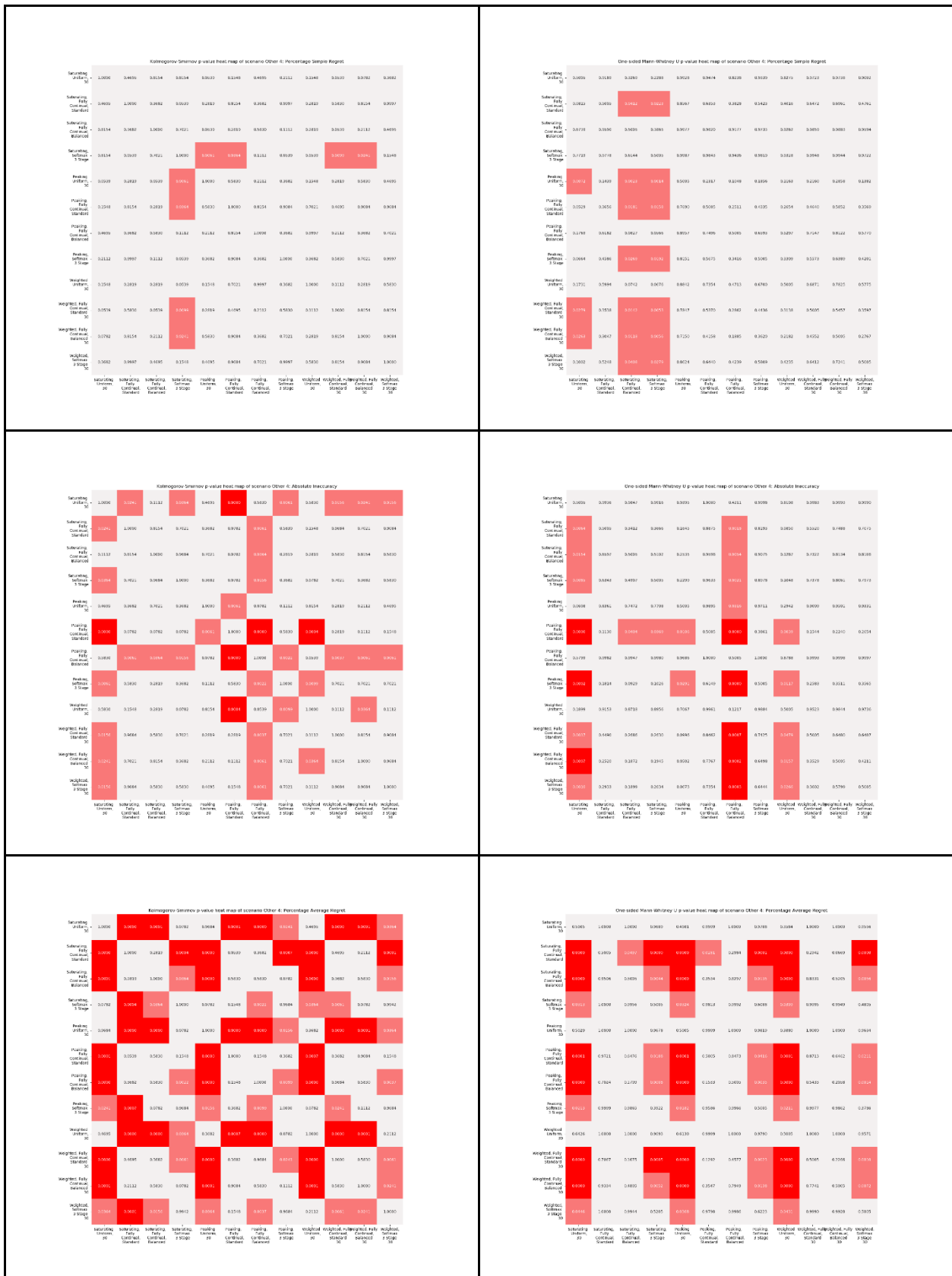
Figure S12.4.14. Kolmogorov–Smirnov (left) and Mann-Whitney U (right) heatmaps of p-values for objective 2 for scenario Other 4

Scenario Saturating 2
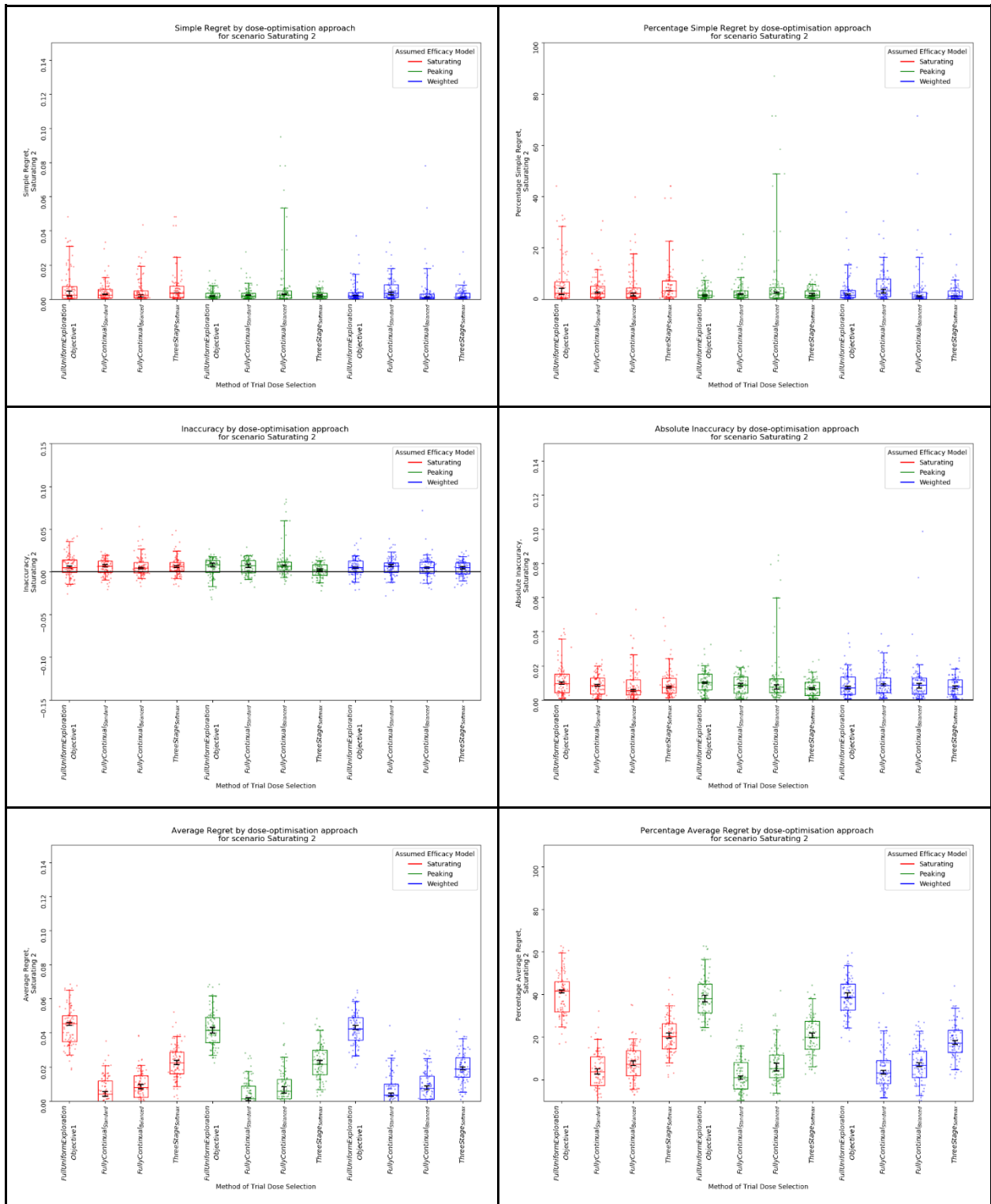
Figure S13.2. Metrics by dose-optimisation approach for objective 2 for scenario Saturating 2
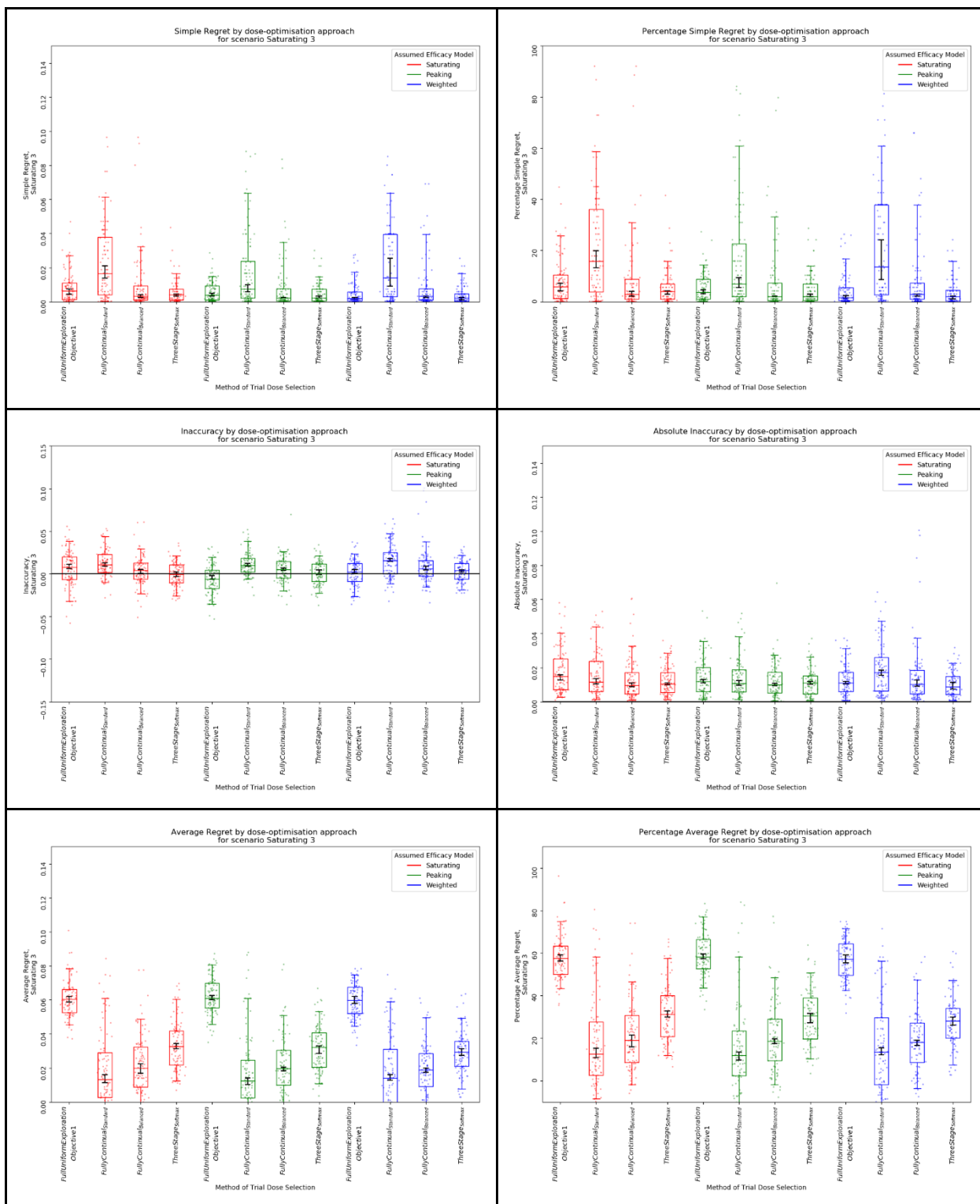
Figure S13.3. Metrics by dose-optimisation approach for objective 2 for scenario Saturating 3
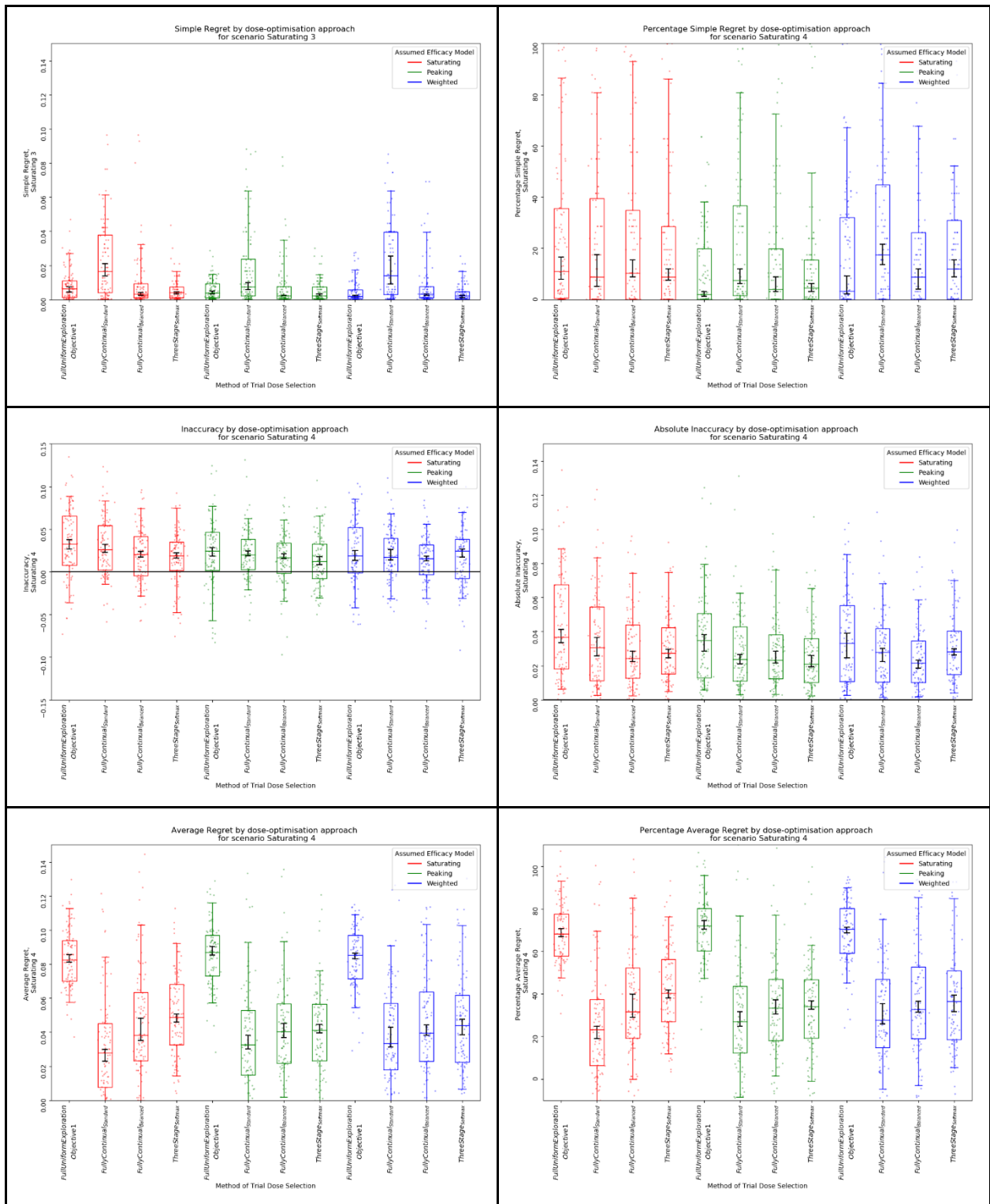
Figure S13.4. Metrics by dose-optimisation approach for objective 2 for scenario Saturating 4
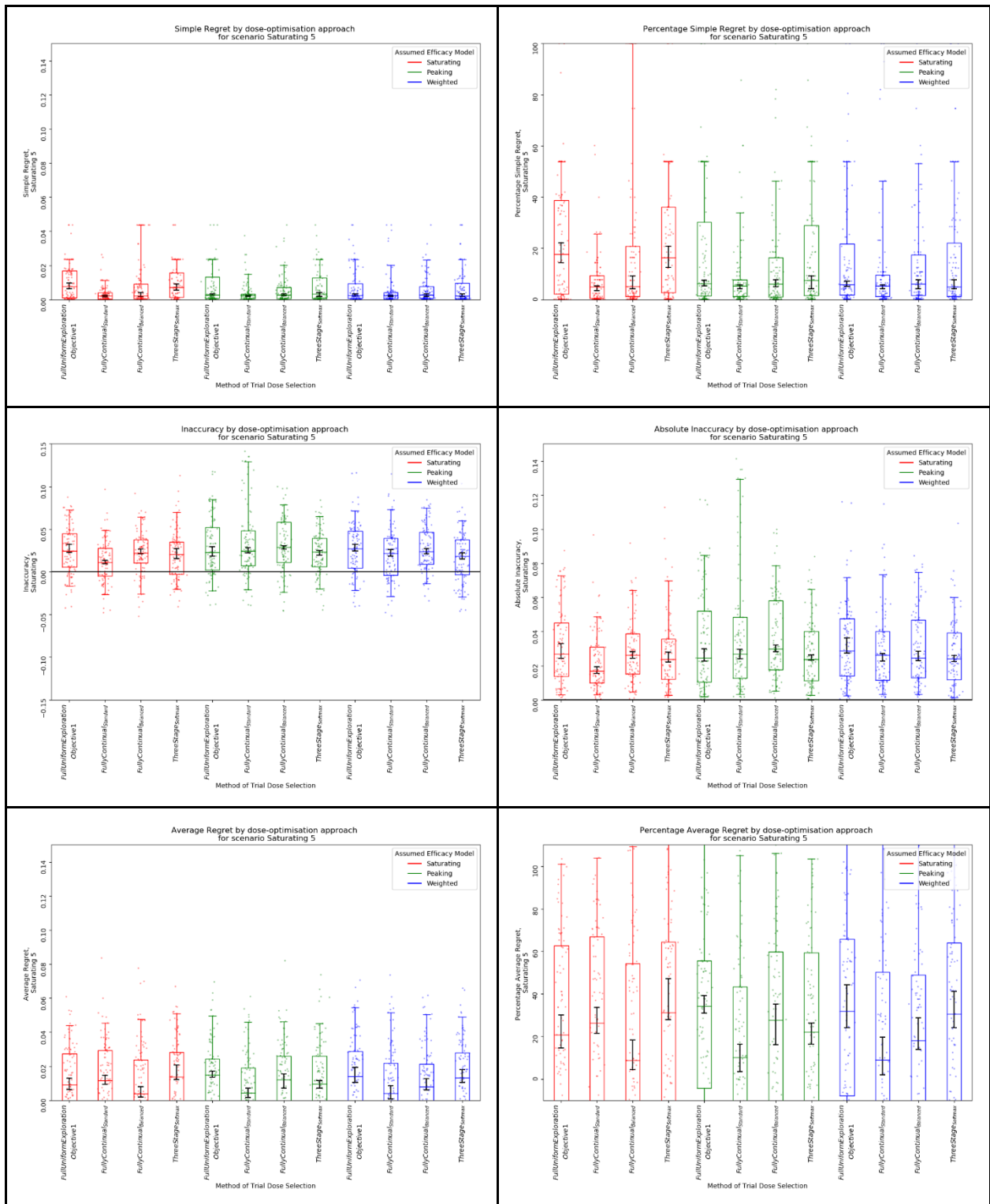
Figure S13.5. Metrics by dose-optimisation approach for objective 2 for scenario Saturating 5
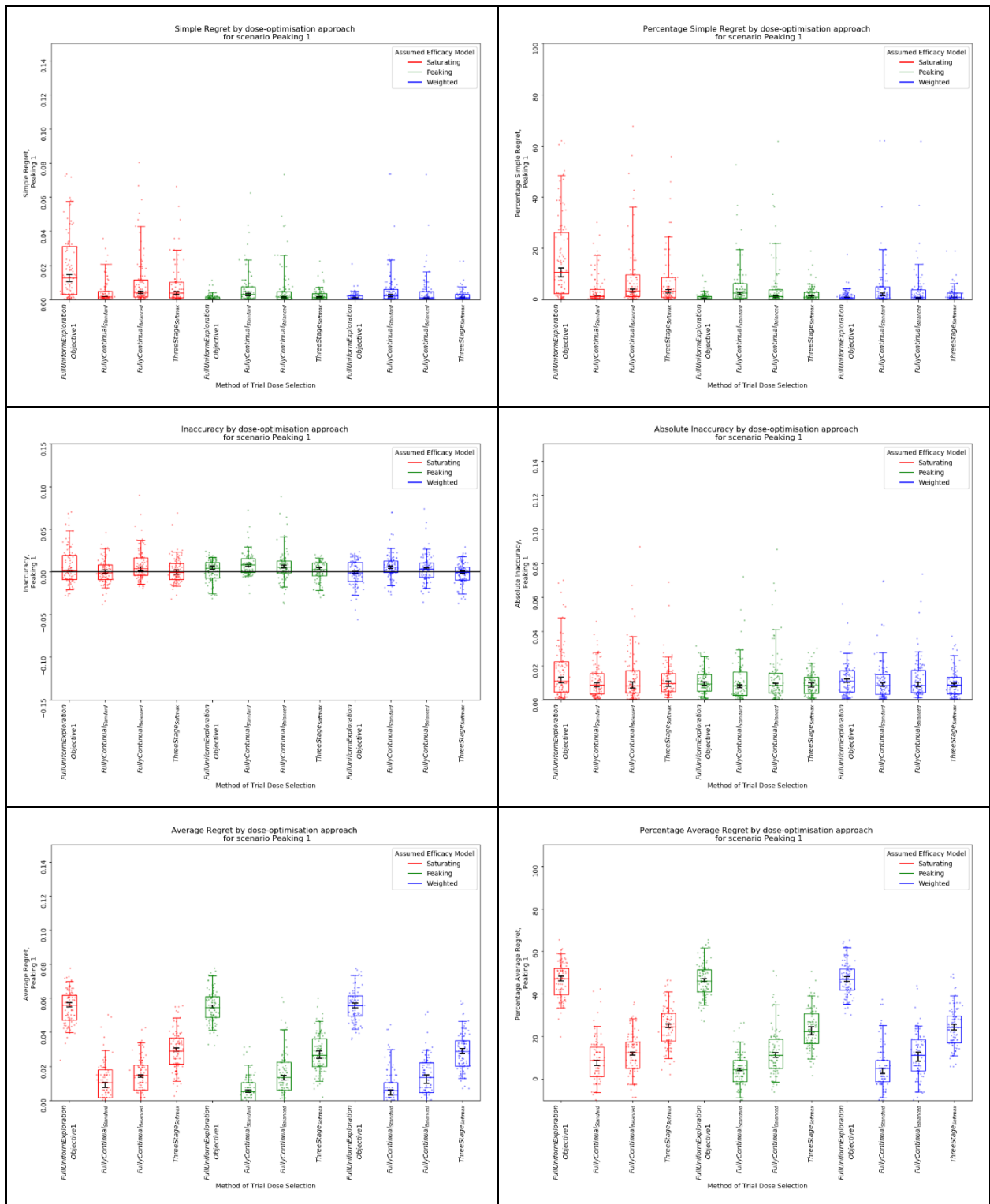
Figure S13.6. Metrics by dose-optimisation approach for objective 2 for scenario Peaking 1
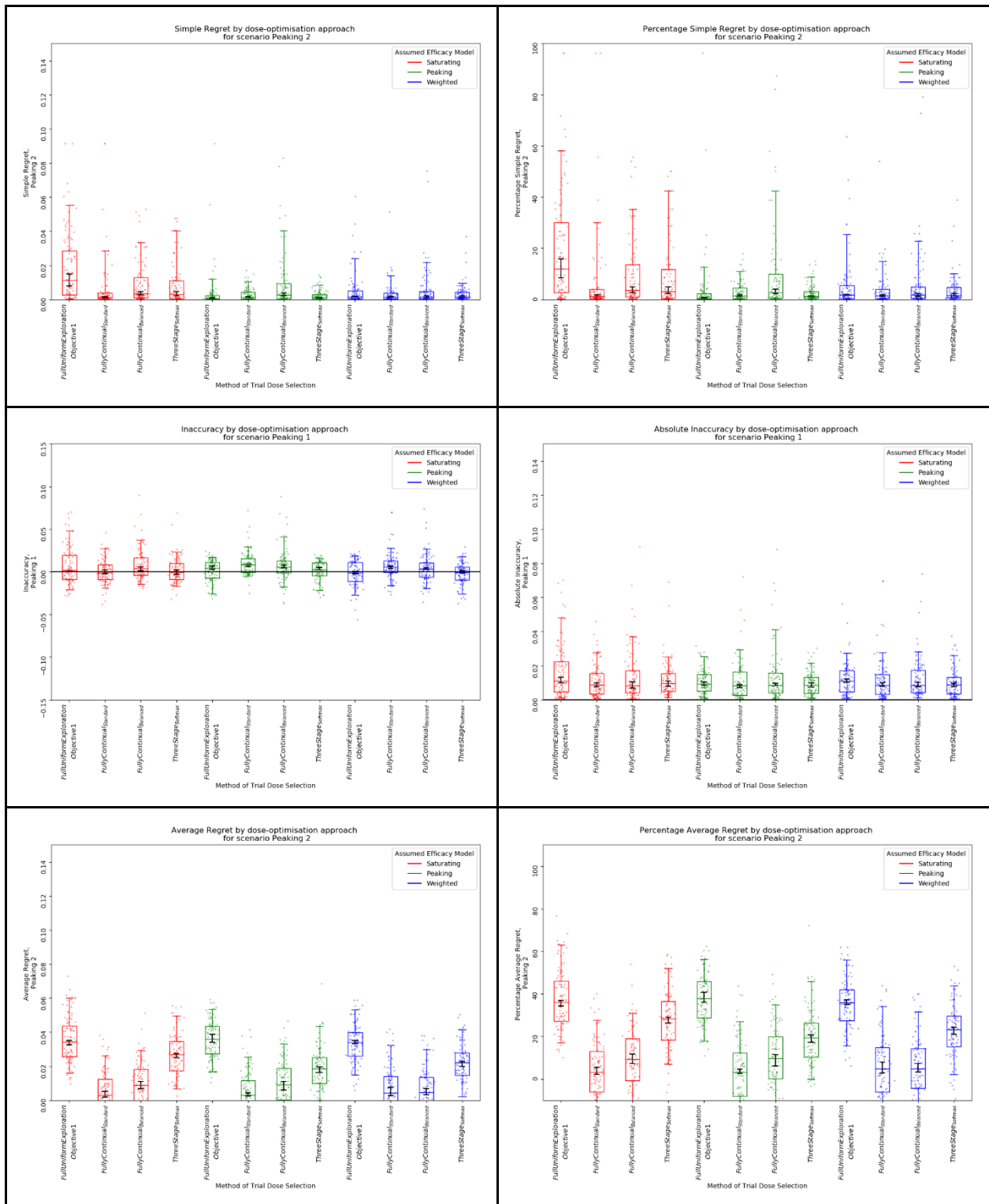
Figure S13.7. Metrics by dose-optimisation approach for objective 2 for scenario Peaking 2
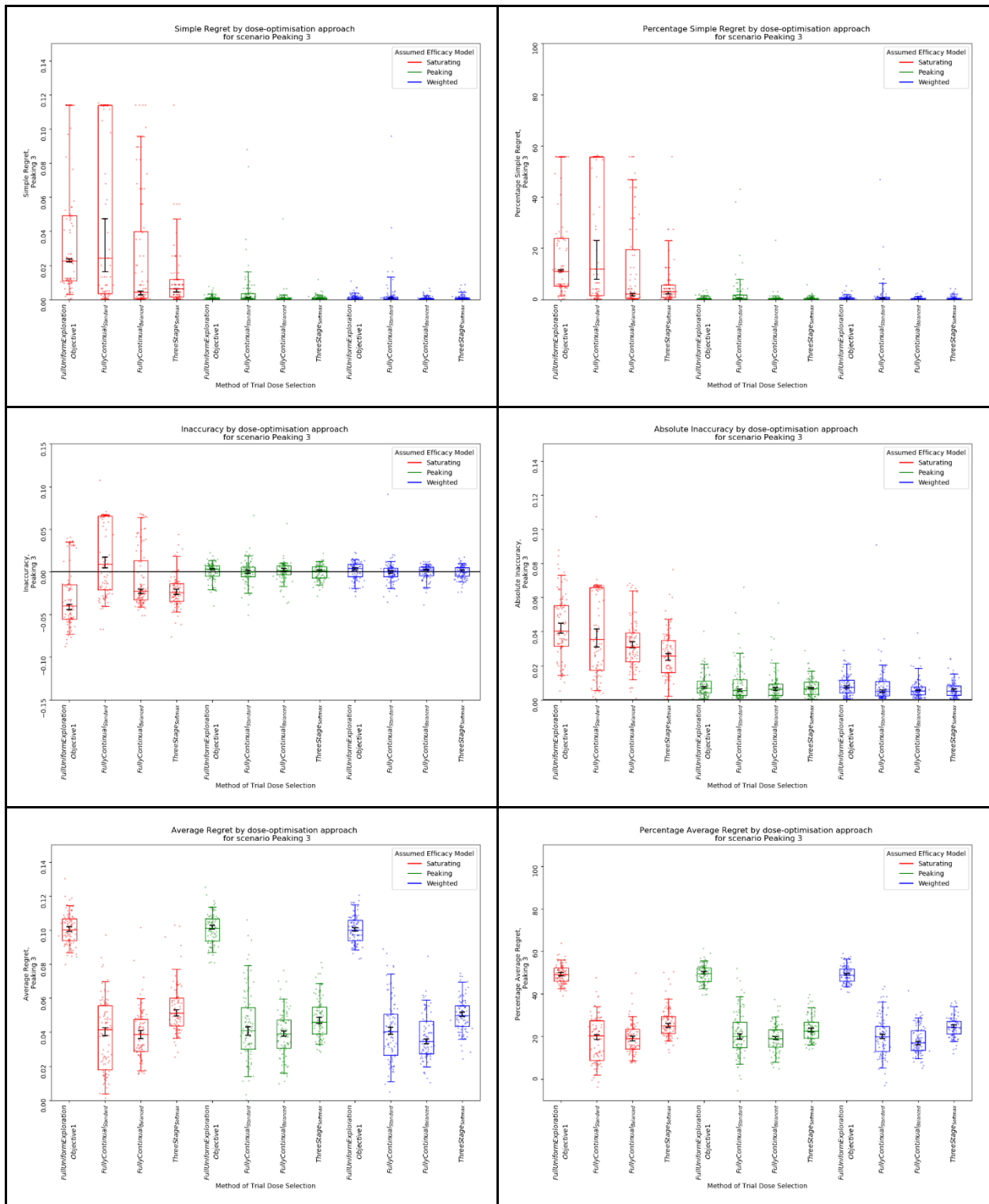
Figure S13.8. Metrics by dose-optimisation approach for objective 2 for scenario Peaking 3
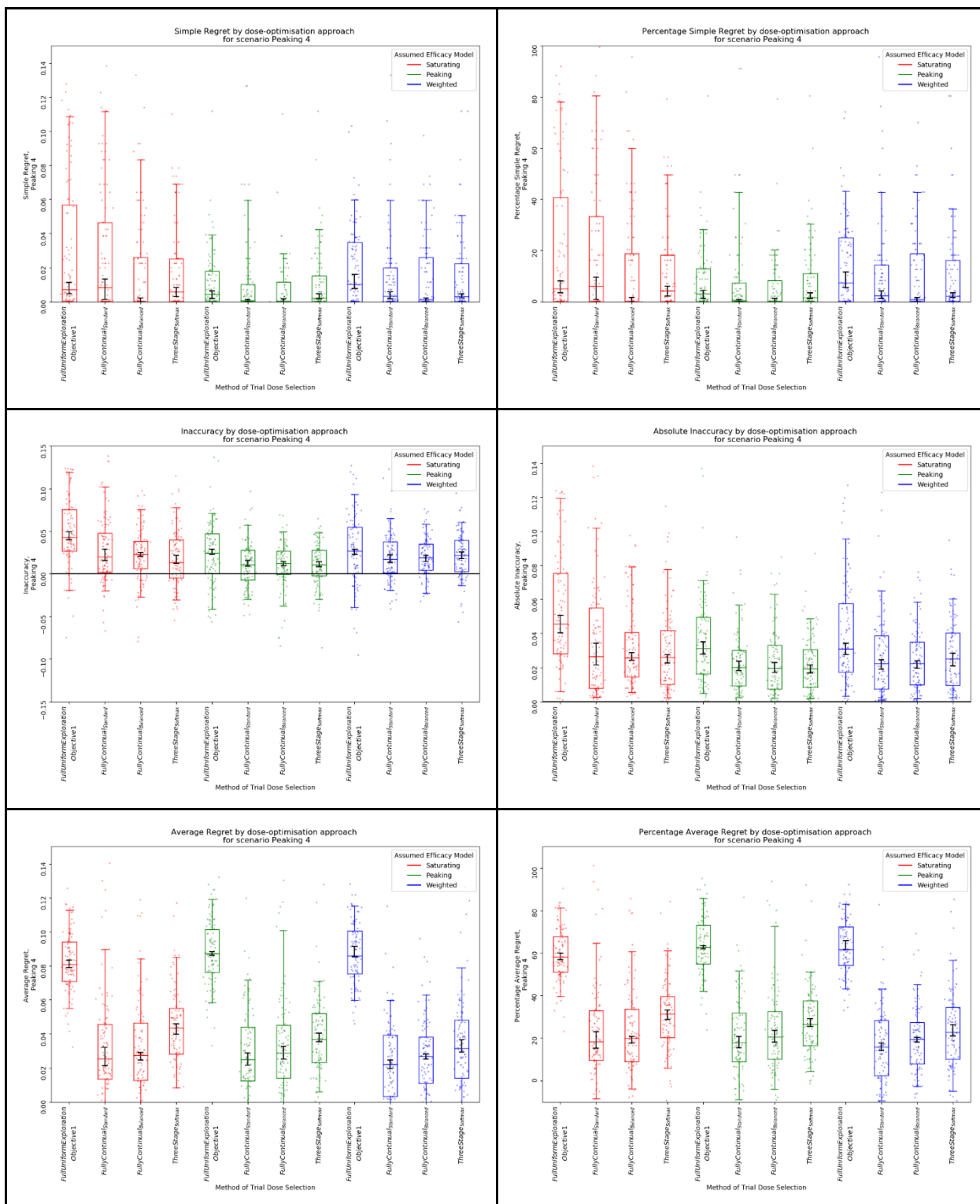
Figure S13.9. Metrics by dose-optimisation approach for objective 2 for scenario Peaking 4
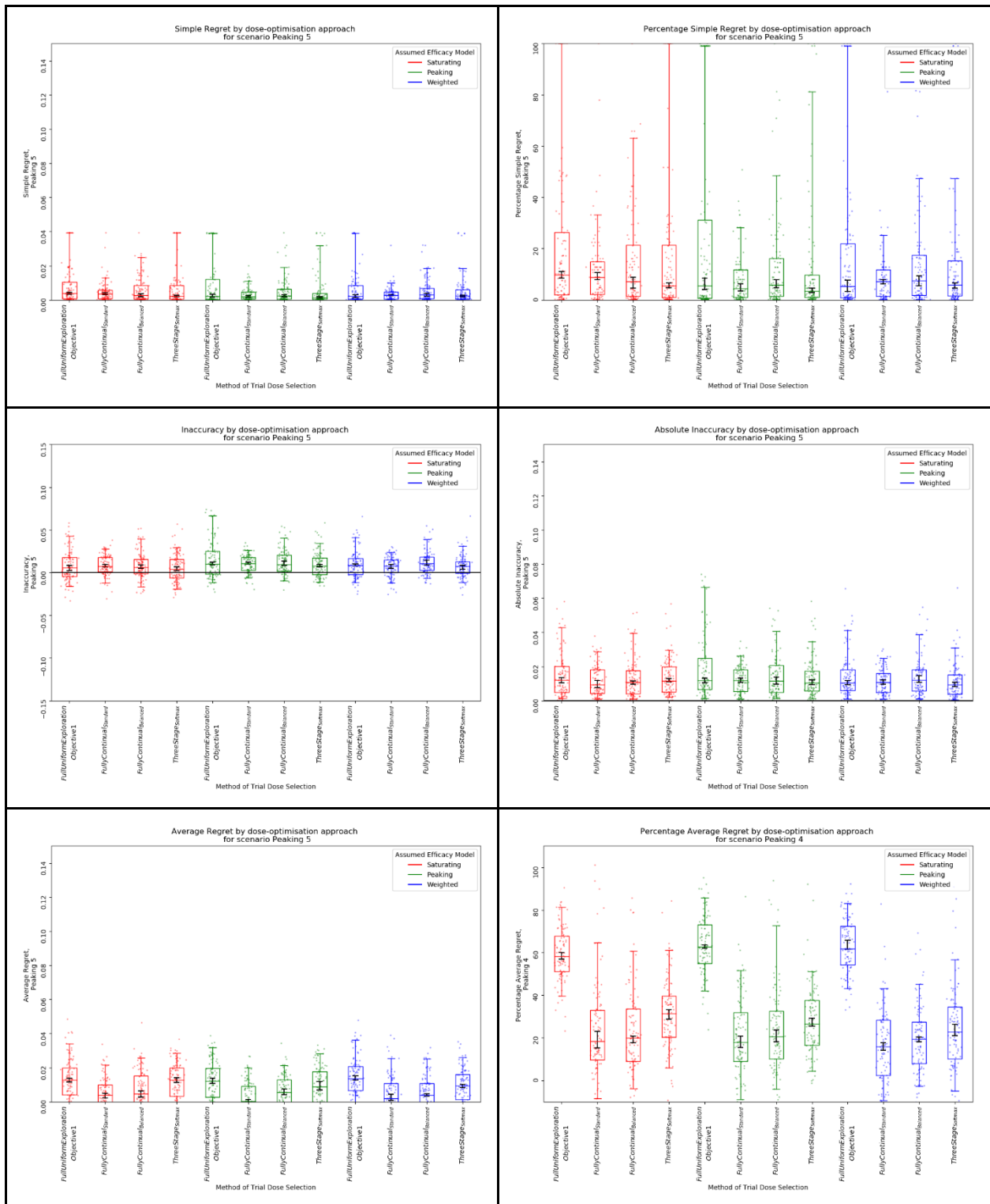
Figure S13.10. Metrics by dose-optimisation approach for objective 2 for scenario Peaking 5
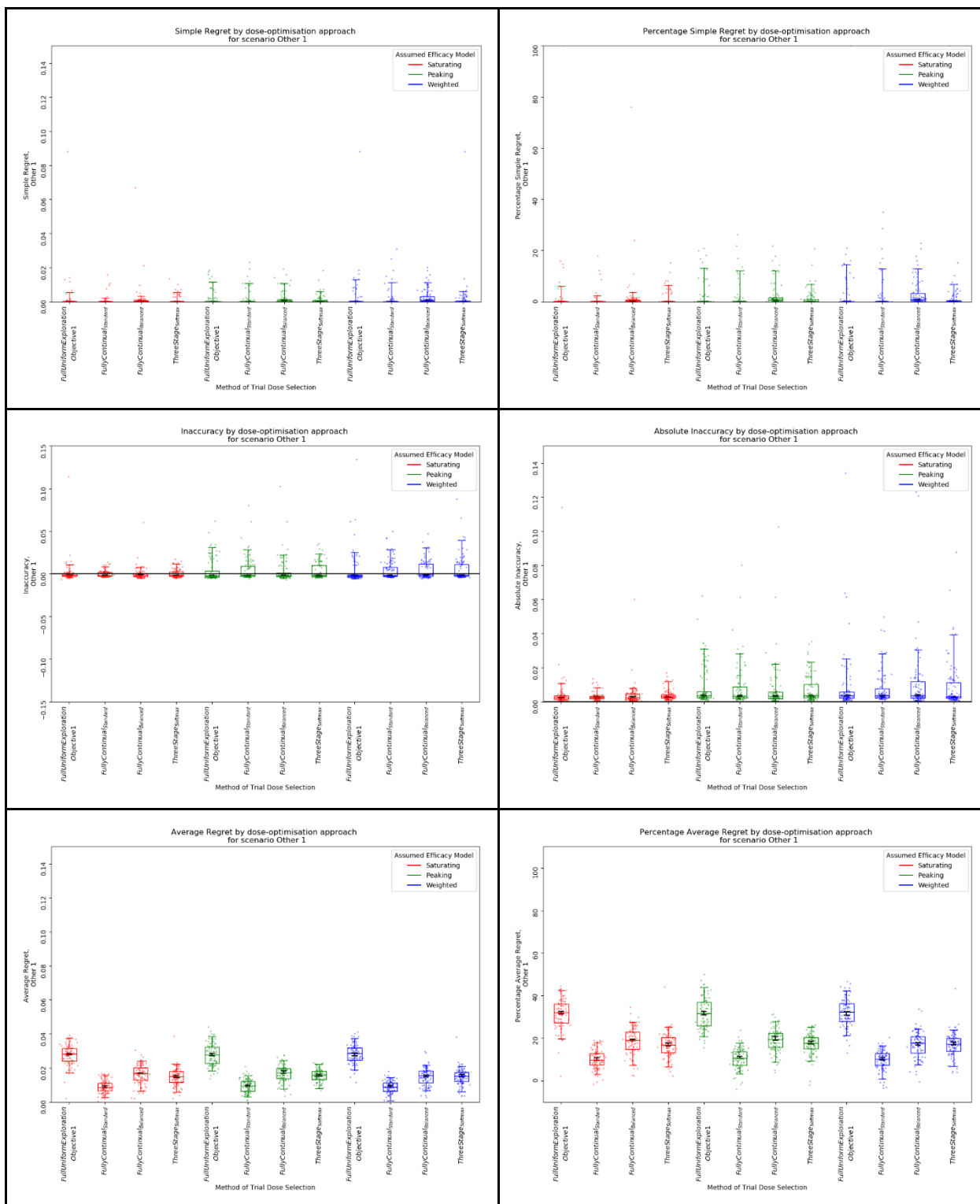
Figure S13.11. Metrics by dose-optimisation approach for objective 2 for scenario Other 1
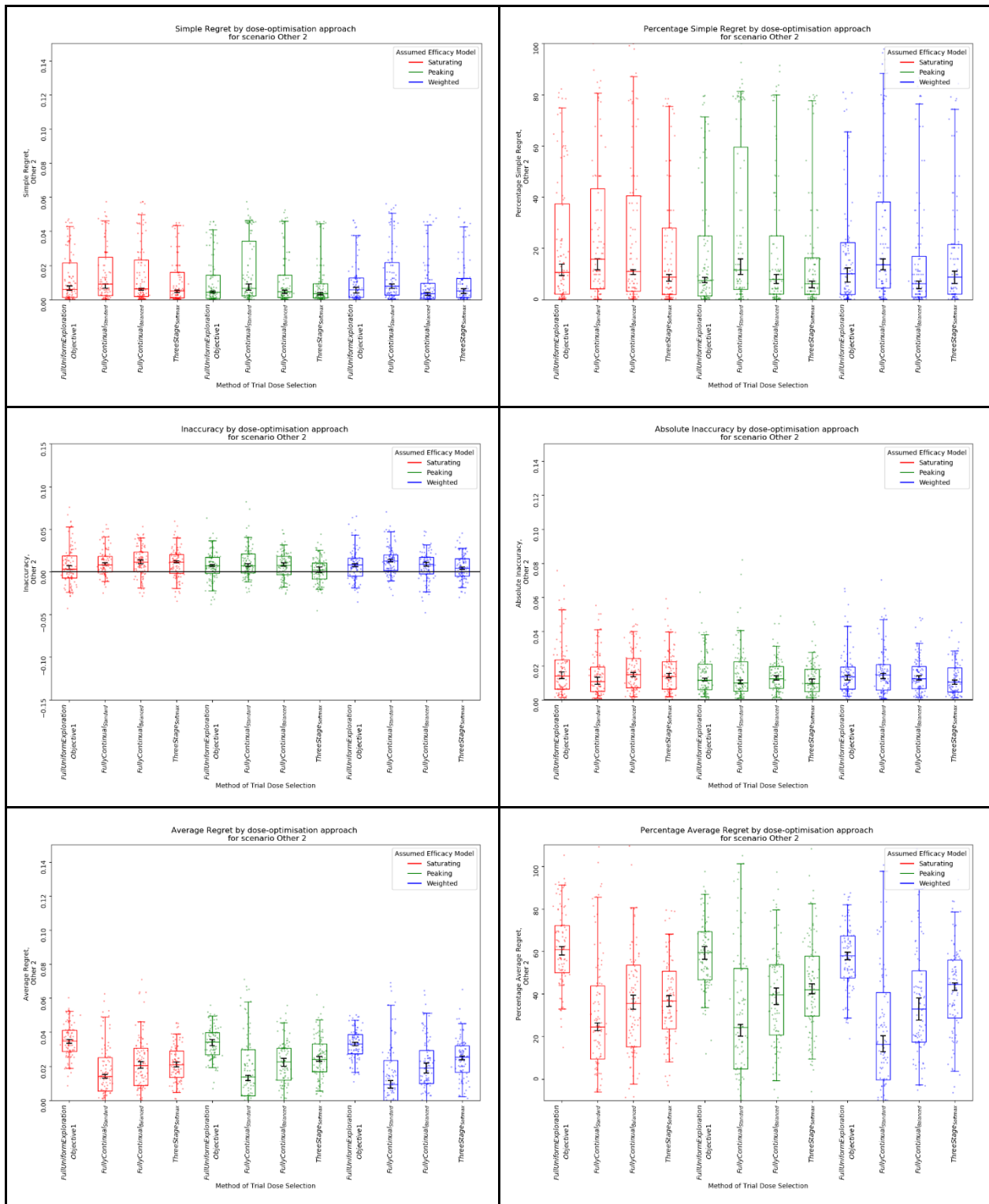
Figure S13.12. Metrics by dose-optimisation approach for objective 2 for scenario Other 2
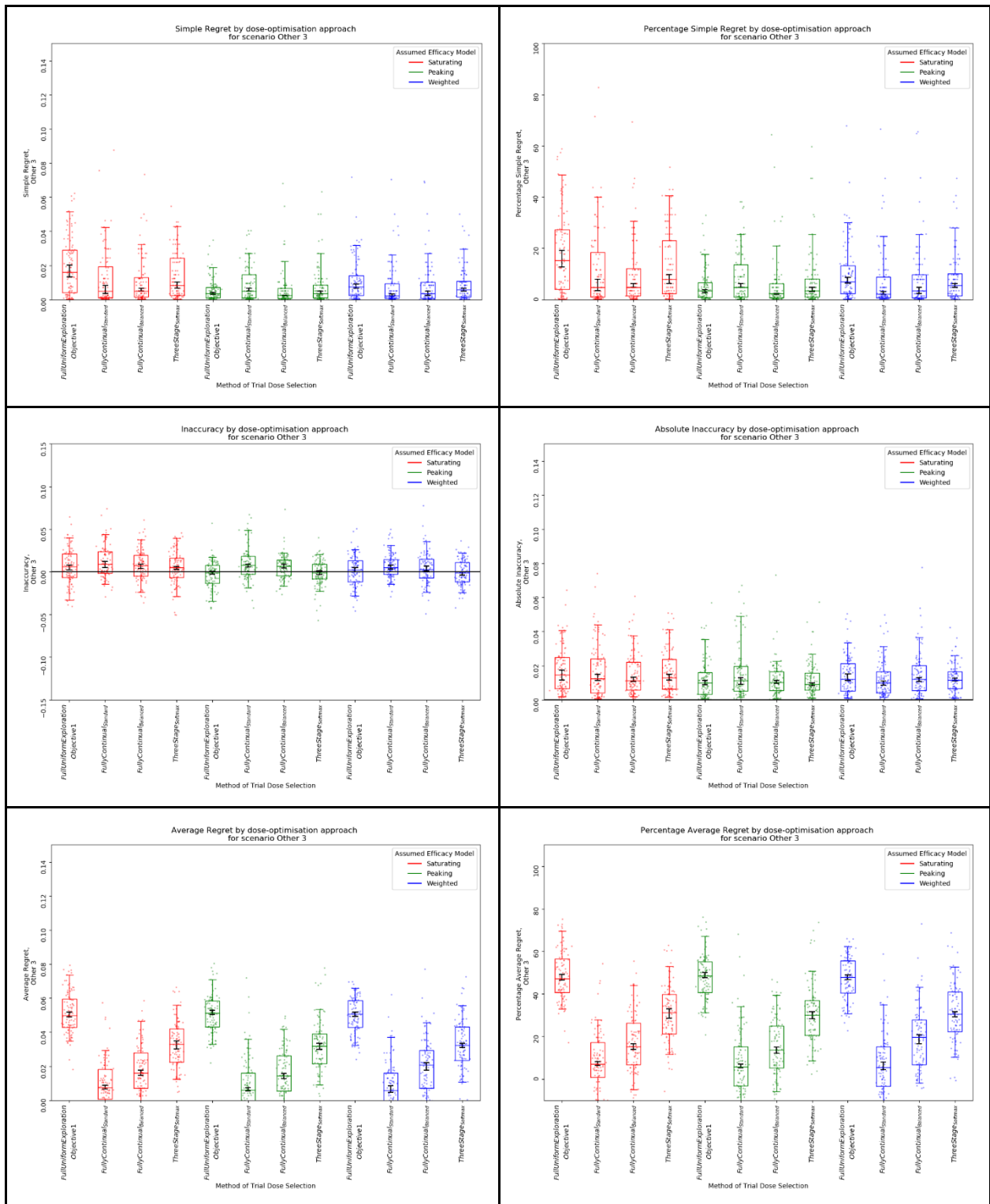
Figure S13.13. Metrics by dose-optimisation approach for objective 2 for scenario Other 3
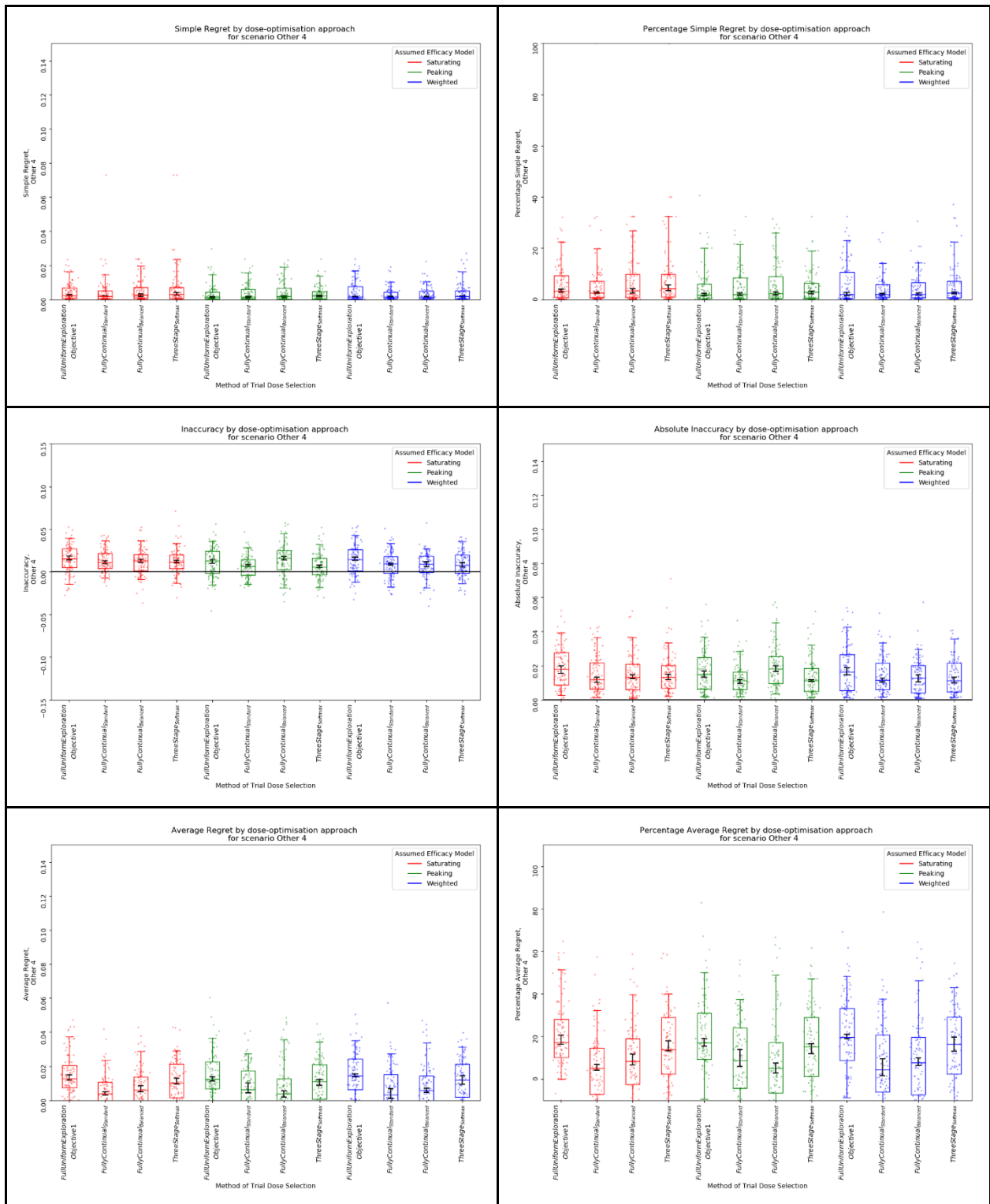
Figure S13.14. Metrics by dose-optimisation approach for objective 2 for scenario Other 4

S14. Copeland Tables

Scenario Saturating 2

Approach	AR		IA		SR	
	Rank	Score	Rank	Score	Rank	Score
Saturating Uniform, 30	12	0.111	11	0.44	10	0.434
Peaking Uniform, 30	10	0.123	12	0.403	3	0.55
Weighted Uniform, 30	11	0.116	4	0.519	6	0.51
Peaking, SoftMax 3 Stage	9	0.403	1	0.591	4	0.544
Saturating, SoftMax 3 Stage	8	0.406	5	0.51	11	0.412
Weighted, SoftMax 3 Stage	7	0.46	3	0.537	2	0.573
Saturating CRM, Fully Continual, Standard	2	0.751	10	0.471	9	0.468
Peaking CRM, Fully Continual, Standard	1	0.806	9	0.475	5	0.518
Weighted CRM, Fully Continual, Standard	3	0.741	8	0.479	12	0.393
Saturating CRM, Fully Continual, Balanced	6	0.681	2	0.58	7	0.494
Peaking CRM, Fully Continual, Balanced	4	0.708	7	0.489	8	0.483
Weighted CRM, Fully Continual, Balanced	5	0.692	6	0.508	1	0.622

Table S.14.2. Copeland metrics for scenario Saturating 2

Approach	AR		IA		SR	
	Rank	Score	Rank	Score	Rank	Score
Saturating Uniform, 30	11	0.407	11	0.407	9	0.473
Peaking Uniform, 30	9	0.49	9	0.49	7	0.525
Weighted Uniform, 30	7	0.512	7	0.512	2	0.608
Peaking, SoftMax 3 Stage	2	0.543	2	0.543	3	0.577
Saturating, SoftMax 3 Stage	4	0.53	4	0.53	5	0.553
Weighted, SoftMax 3 Stage	1	0.58	1	0.58	1	0.637
Saturating CRM, Fully Continual, Standard	10	0.462	10	0.462	12	0.297
Peaking CRM, Fully Continual, Standard	8	0.498	8	0.498	10	0.383
Weighted CRM, Fully Continual, Standard	12	0.38	12	0.38	11	0.317
Saturating CRM, Fully Continual, Balanced	3	0.541	3	0.541	8	0.517
Peaking CRM, Fully Continual, Balanced	5	0.529	5	0.529	4	0.576
Weighted CRM, Fully Continual, Balanced	6	0.527	6	0.527	6	0.537

Table S.14.3. Copeland metrics for scenario Saturating 3

Scenario Saturating 4

Approach	AR		IA		SR	
	Rank	Score	Rank	Score	Rank	Score
Saturating Uniform, 30	10	0.192	12	0.379	10	0.445
Peaking Uniform, 30	12	0.178	11	0.444	1	0.595
Weighted Uniform, 30	11	0.189	10	0.464	4	0.518
Peaking, SoftMax 3 Stage	4	0.614	2	0.566	2	0.576
Saturating, SoftMax 3 Stage	9	0.516	7	0.5	6	0.493
Weighted, SoftMax 3 Stage	6	0.575	8	0.487	7	0.491
Saturating CRM, Fully Continual, Standard	1	0.71	9	0.472	8	0.484
Peaking CRM, Fully Continual, Standard	2	0.654	3	0.54	11	0.439
Weighted CRM, Fully Continual, Standard	3	0.623	4	0.528	12	0.423
Saturating CRM, Fully Continual, Balanced	7	0.574	6	0.514	9	0.467
Peaking CRM, Fully Continual, Balanced	5	0.61	5	0.527	3	0.552
Weighted CRM, Fully Continual, Balanced	8	0.565	1	0.58	5	0.517

Table S.14.4. Copeland metrics for scenario Saturating 4

Approach	AR		IA		SR	
	Rank	Score	Rank	Score	Rank	Score
Saturating Uniform, 30	10	0.445	9	0.478	11	0.394
Peaking Uniform, 30	1	0.595	8	0.485	10	0.476
Weighted Uniform, 30	4	0.518	11	0.463	8	0.48
Peaking, SoftMax 3 Stage	2	0.576	4	0.517	9	0.477
Saturating, SoftMax 3 Stage	6	0.493	2	0.532	12	0.389
Weighted, SoftMax 3 Stage	7	0.491	3	0.524	6	0.507
Saturating CRM, Fully Continual, Standard	8	0.484	1	0.614	1	0.596
Peaking CRM, Fully Continual, Standard	11	0.439	10	0.466	2	0.578
Weighted CRM, Fully Continual, Standard	12	0.423	5	0.509	3	0.567
Saturating CRM, Fully Continual, Balanced	9	0.467	6	0.504	7	0.49
Peaking CRM, Fully Continual, Balanced	3	0.552	12	0.415	4	0.528
Weighted CRM, Fully Continual, Balanced	5	0.517	7	0.493	5	0.518

Table S.14.5. Copeland metrics for scenario Saturating 5

Scenario Peaking 1

Approach	AR		IA		SR	
	Rank	Score	Rank	Score	Rank	Score
Saturating Uniform, 30	10	0.119	12	0.412	12	0.244
Peaking Uniform, 30	11	0.115	8	0.497	1	0.701
Weighted Uniform, 30	12	0.108	11	0.469	2	0.636
Peaking, SoftMax 3 Stage	7	0.451	2	0.544	5	0.558
Saturating, SoftMax 3 Stage	9	0.427	9	0.492	10	0.384
Weighted, SoftMax 3 Stage	8	0.433	1	0.546	3	0.618
Saturating CRM, Fully Continual, Standard	3	0.727	4	0.513	7	0.505
Peaking CRM, Fully Continual, Standard	1	0.816	3	0.522	8	0.452
Weighted CRM, Fully Continual, Standard	2	0.765	5	0.508	9	0.449
Saturating CRM, Fully Continual, Balanced	5	0.676	7	0.503	11	0.352
Peaking CRM, Fully Continual, Balanced	6	0.676	6	0.508	6	0.508
Weighted CRM, Fully Continual, Balanced	4	0.686	10	0.486	4	0.592

Table S.14.6. Copeland metrics for scenario Peaking 1

Approach	AR		IA		SR	
	Rank	Score	Rank	Score	Rank	Score
Saturating Uniform, 30	11	0.19	12	0.406	12	0.269
Peaking Uniform, 30	12	0.169	9	0.48	1	0.655
Weighted Uniform, 30	10	0.203	7	0.485	7	0.518
Peaking, SoftMax 3 Stage	7	0.486	2	0.569	2	0.576
Saturating, SoftMax 3 Stage	9	0.338	6	0.514	9	0.437
Weighted, SoftMax 3 Stage	8	0.42	1	0.595	6	0.522
Saturating CRM, Fully Continual, Standard	1	0.739	4	0.515	5	0.546
Peaking CRM, Fully Continual, Standard	2	0.733	5	0.515	4	0.559
Weighted CRM, Fully Continual, Standard	4	0.706	3	0.516	3	0.569
Saturating CRM, Fully Continual, Balanced	5	0.652	10	0.475	11	0.393
Peaking CRM, Fully Continual, Balanced	6	0.644	11	0.447	10	0.437
Weighted CRM, Fully Continual, Balanced	3	0.721	8	0.484	8	0.517

Table S.14.7. Copeland metrics for scenario Peaking 2

Approach	AR		IA		SR	
	Rank	Score	Rank	Score	Rank	Score
Saturating Uniform, 30	11	0.095	12	0.117	12	0.118
Peaking Uniform, 30	12	0.092	7	0.622	5	0.636
Weighted Uniform, 30	10	0.097	8	0.603	6	0.623
Peaking, SoftMax 3 Stage	7	0.559	6	0.637	4	0.647
Saturating, SoftMax 3 Stage	9	0.482	9	0.272	10	0.298
Weighted, SoftMax 3 Stage	8	0.516	1	0.703	3	0.654
Saturating CRM, Fully Continual, Standard	4	0.693	10	0.194	11	0.206
Peaking CRM, Fully Continual, Standard	6	0.646	5	0.64	8	0.511
Weighted CRM, Fully Continual, Standard	5	0.675	3	0.682	7	0.597
Saturating CRM, Fully Continual, Balanced	2	0.708	11	0.188	9	0.339
Peaking CRM, Fully Continual, Balanced	3	0.699	4	0.647	2	0.671
Weighted CRM, Fully Continual, Balanced	1	0.738	2	0.695	1	0.7

Table S.14.8. Copeland metrics for scenario Peaking 3

Approach	AR		IA		SR	
	Rank	Score	Rank	Score	Rank	Score
Saturating Uniform, 30	10	0.159	12	0.27	12	0.413
Peaking Uniform, 30	11	0.127	10	0.431	8	0.494
Weighted Uniform, 30	12	0.127	11	0.41	11	0.418
Peaking, SoftMax 3 Stage	8	0.552	1	0.605	5	0.514
Saturating, SoftMax 3 Stage	9	0.5	7	0.5	9	0.475
Weighted, SoftMax 3 Stage	7	0.605	6	0.519	7	0.495
Saturating CRM, Fully Continual, Standard	5	0.633	8	0.488	10	0.444
Peaking CRM, Fully Continual, Standard	3	0.658	3	0.585	2	0.577
Weighted CRM, Fully Continual, Standard	1	0.711	4	0.562	6	0.511
Saturating CRM, Fully Continual, Balanced	4	0.633	9	0.474	3	0.536
Peaking CRM, Fully Continual, Balanced	6	0.62	2	0.593	1	0.593
Weighted CRM, Fully Continual, Balanced	2	0.677	5	0.562	4	0.53

Table S.14.9. Copeland metrics for scenario Peaking 4

Approach	AR Rank	AR Score	IA Rank	IA Score	SR Rank	SR Score
Saturating Uniform, 30	11	0.389	8	0.488	12	0.427
Peaking Uniform, 30	9	0.405	12	0.434	8	0.489
Weighted Uniform, 30	12	0.334	6	0.498	4	0.508
Peaking, SoftMax 3 Stage	7	0.47	5	0.508	2	0.562
Saturating, SoftMax 3 Stage	10	0.396	10	0.482	6	0.497
Weighted, SoftMax 3 Stage	8	0.464	1	0.556	5	0.503
Saturating CRM, Fully Continual, Standard	3	0.596	4	0.516	10	0.474
Peaking CRM, Fully Continual, Standard	1	0.65	7	0.498	1	0.57
Weighted CRM, Fully Continual, Standard	2	0.615	2	0.534	3	0.533
Saturating CRM, Fully Continual, Balanced	6	0.546	3	0.531	9	0.476
Peaking CRM, Fully Continual, Balanced	5	0.564	11	0.472	7	0.492
Weighted CRM, Fully Continual, Balanced	4	0.571	9	0.482	11	0.471

Table S.14.10. Copeland metrics for scenario Peaking 5

Scenario Other 1

Approach	AR		IA		SR	
	Rank	Score	Rank	Score	Rank	Score
Saturating Uniform, 30	10	0.129	1	0.658	11	0.402
Peaking Uniform, 30	11	0.124	11	0.427	6	0.47
Weighted Uniform, 30	12	0.113	7	0.46	8	0.456
Peaking, SoftMax 3 Stage	7	0.53	9	0.445	4	0.503
Saturating, SoftMax 3 Stage	4	0.562	4	0.531	10	0.424
Weighted, SoftMax 3 Stage	5	0.561	5	0.482	5	0.488
Saturating CRM, Fully Continual, Standard	2	0.83	2	0.613	12	0.385
Peaking CRM, Fully Continual, Standard	3	0.81	8	0.446	7	0.465
Weighted CRM, Fully Continual, Standard	1	0.835	12	0.413	9	0.449
Saturating CRM, Fully Continual, Balanced	8	0.481	3	0.598	3	0.584
Peaking CRM, Fully Continual, Balanced	9	0.469	6	0.481	2	0.675
Weighted CRM, Fully Continual, Balanced	6	0.556	10	0.445	1	0.7

Table S.14.11. Copeland metrics for scenario Other 1

Approach	AR		IA		SR	
	Rank	Score	Rank	Score	Rank	Score
Saturating Uniform, 30	12	0.253	10	0.464	8	0.489
Peaking Uniform, 30	11	0.274	6	0.508	3	0.536
Weighted Uniform, 30	10	0.294	8	0.484	7	0.518
Peaking, SoftMax 3 Stage	9	0.481	1	0.568	2	0.566
Saturating, SoftMax 3 Stage	5	0.557	11	0.464	5	0.519
Weighted, SoftMax 3 Stage	8	0.5	2	0.549	4	0.523
Saturating CRM, Fully Continual, Standard	2	0.657	3	0.526	11	0.427
Peaking CRM, Fully Continual, Standard	3	0.634	4	0.524	10	0.433
Weighted CRM, Fully Continual, Standard	1	0.697	9	0.47	12	0.424
Saturating CRM, Fully Continual, Balanced	6	0.551	12	0.444	9	0.463
Peaking CRM, Fully Continual, Balanced	7	0.539	5	0.509	6	0.519
Weighted CRM, Fully Continual, Balanced	4	0.563	7	0.489	1	0.582

Table S.14.12. Copeland metrics for scenario Other 2

Approach	AR		IA		SR	
	Rank	Score	Rank	Score	Rank	Score
Saturating Uniform, 30	12	0.148	12	0.429	12	0.31
Peaking Uniform, 30	11	0.154	1	0.563	3	0.587
Weighted Uniform, 30	10	0.157	8	0.478	10	0.438
Peaking, SoftMax 3 Stage	7	0.429	3	0.548	4	0.553
Saturating, SoftMax 3 Stage	8	0.414	10	0.465	11	0.393
Weighted, SoftMax 3 Stage	9	0.413	5	0.519	7	0.491
Saturating CRM, Fully Continual, Standard	3	0.771	11	0.452	8	0.477
Peaking CRM, Fully Continual, Standard	2	0.78	7	0.486	6	0.502
Weighted CRM, Fully Continual, Standard	1	0.788	2	0.554	2	0.593
Saturating CRM, Fully Continual, Balanced	5	0.645	9	0.469	9	0.471
Peaking CRM, Fully Continual, Balanced	4	0.678	4	0.541	1	0.636
Weighted CRM, Fully Continual, Balanced	6	0.624	6	0.494	5	0.549

Table S.14.13. Copeland metrics for scenario Other 3

Approach	AR		IA		SR	
	Rank	Score	Rank	Score	Rank	Score
Saturating Uniform, 30	12	0.355	11	0.399	10	0.452
Peaking Uniform, 30	11	0.364	9	0.481	1	0.559
Weighted Uniform, 30	10	0.374	10	0.448	8	0.5
Peaking, SoftMax 3 Stage	7	0.473	2	0.568	5	0.519
Saturating, SoftMax 3 Stage	8	0.471	6	0.513	12	0.424
Weighted, SoftMax 3 Stage	9	0.463	4	0.538	7	0.51
Saturating CRM, Fully Continual, Standard	1	0.63	7	0.504	6	0.512
Peaking CRM, Fully Continual, Standard	6	0.552	1	0.575	4	0.527
Weighted CRM, Fully Continual, Standard	3	0.591	5	0.526	3	0.528
Saturating CRM, Fully Continual, Balanced	4	0.571	8	0.499	11	0.434
Peaking CRM, Fully Continual, Balanced	2	0.592	12	0.394	9	0.498
Weighted CRM, Fully Continual, Balanced	5	0.564	3	0.553	2	0.536

Table S.14.14. Copeland metrics for scenario Other 4

S6. Scenario Creation

Objective 1, Scenario 2 Efficacy

Inputs	Values
Prime doses	[0.00, 0.01, …, 0.99, 1.00]
Anchor doses	[-0.0], [1.0]
Anchor probabilities	[0.05], [0.9]
K	9
Iterations	21

Objective 1, Scenario 3 Efficacy

Inputs	Values
Prime doses	[-0.50, -0.49, …, 1.49, 1.50]
Anchor doses	[-0.2], [0.1], [0.5], [0.9], [1.2]
Anchor probabilities	[0.1, 0.5, 0.8, 0.55, 0.2]
K	15
Iterations	21

Objective 1, Scenario 4 Efficacy

Inputs	Values
Prime doses	[0.00, 0.01, …, 0.99, 1.00]
Anchor doses	[0], [0.2], [0.4], [0.6], [0.7], [1]
Anchor probabilities	[0.05], [0.1], [0.15], [0.3], [0.7], [0.5]
K	5
Iterations	21

Objective 1, Scenario 5 Efficacy

Inputs	Values
Prime doses	[0.00, 0.01, …, 0.99, 1.00]
Anchor doses	[0], [0.2], [0.4], [0.6], [0.8], [1]
Anchor probabilities	[0.8], [0.7], [0.3], [0.15], [0.1], [0.05]
K	5
Iterations	21

Objective 1, Scenario 6 Efficacy

Inputs	Values
Prime doses	[0.00, 0.01, …, 0.99, 1.00]
Anchor doses	[0.0], [0.2], [0.55], [0.65], [1.0]
Anchor probabilities	[0.3], [0.5], [0.7], [0.5], [0.7]
K	5
Iterations	21

Objective 1, Scenario 7 Efficacy

Inputs	Values
Prime doses	[0.00, 0.01, …, 0.99, 1.00]
Anchor doses	[0.0], [0.5], [1.0]
Anchor probabilities	[0.75], [0.85], [0.75]
K	11
Iterations	11

Objective 2, Scenario 1 Efficacy

Inputs	Values
Prime doses	[0.00, 0.05, …, 0.95, 1.00]
Boost doses	[0.00, 0.05, …, 0.95, 1.00]
Anchor doses	[1,1],[0,0],[0,1], [1,0], [.6,.6]
Anchor probabilities	[.5],[.1], [.4], [.4], [.9]
K	15
Iterations	9

Objective 2, Scenario 2 Efficacy

Inputs	Values
Prime doses	[0.00, 0.05, …, 0.95, 1.00]
Boost doses	[0.00, 0.05, …, 0.95, 1.00]
Anchor doses	[1,1], [0,0], [0,1], [1,0]
Anchor probabilities	[.5], [.1], [.8], [.9]
K	9
Iterations	11

Objective 2, Scenario 3 Efficacy

Inputs	Values
Prime doses	[0.00, 0.05, …, 0.95, 1.00]
Boost doses	[0.00, 0.05, …, 0.95, 1.00]
Anchor doses	[1,1],[0,0],[0,1], [1,0]
Anchor probabilities	[.95],[.05], [.6], [.6]
K	9
Iterations	11

Objective 2, Scenario 4 Efficacy

Inputs	Values
Prime doses	[0.00, 0.05, …, 0.95, 1.00]
Boost doses	[0.00, 0.05, …, 0.95, 1.00]
Anchor doses	[1,1],[.9,.9],[0,0],[0,1], [1,0]
Anchor probabilities	[.6],[.95],[.05], [.4], [.6]
K	9
Iterations	11

Objective 2, Scenario 5 Efficacy

Inputs	Values
Prime doses	[0.00, 0.05, …, 0.95, 1.00]
Boost doses	[0.00, 0.05, …, 0.95, 1.00]
Anchor doses	[1,1], [1,0.5], [1,0], [.9,1], [.9,0.5], [.9,0], [0,1], [0,0.5], [0,0]
Anchor probabilities	[.8],[.7],[.7],[.9],[.8],[.7],[.1],[.05],[.05]
K	9
Iterations	10

Objective 2, Scenario 6 Efficacy

Inputs	Values
Prime doses	[0.00, 0.10, …, 0.90, 1.00]
Boost doses	[0.00, 0.10, …, 0.90, 1.00]
Second-boost doses	[0.00, 0.10, …, 0.90, 1.00]
Anchor doses	[0,0,0],[0,0,1],[0,1,0],[0,1,1], [1,0,0],[1,0,1],[1,1,0],[1,1,1], [.7,.1,.3], [.7,.1,.3], [.7,.1,.3]
Anchor probabilities	[0],[0],[0],[0], [0],[0],[0],[0], [.9], [.9], [.9]
K	7
Iterations	5

Objective 2, Scenario 7 Efficacy

Inputs	Values
Prime doses	[0.00, 0.10, …, 0.90, 1.00]
Boost doses	[0.00, 0.10, …, 0.90, 1.00]
Second-boost doses	[0.00, 0.10, …, 0.90, 1.00]
Anchor doses	[0,0,0],[0,0,1],[0,1,0], [0,1,1] ,[1,0,0],[1,0,1], [1,1,0],[1,1,1], [.5,0,0]
Anchor probabilities	[0.1],[0.5],[0.5], [0.8], [0.1],[0.9], [0.9],[0.4],[0.1]
K	27
Iterations	3

Objective 3, Scenario 1 Efficacy

As objective 1 scenario1.

Objective 3, Scenario 1 Toxicity

Inputs	Values
Prime doses	[-1.00, -0.99, …, 1.99, 2.00]
Anchor doses	[0.2], [1.5]
Anchor probabilities	[0.05], [0.65]
K	21
Iterations	11

Objective 3, Scenario 2 Efficacy

As objective 1 scenario 4.

Objective 3, Scenario 2 Toxicity

Inputs	Values
Prime doses	[-1.00, -0.99, …, 1.99, 2.00]
Anchor doses	[0.7], [0.8]
Anchor probabilities	[0.05], [0.95]
K	3
Iterations	11

Objective 3, Scenario 3 Efficacy

As objective 1 scenario 1.

Objective 3, Scenario 3 Toxicity

As objective 3 scenario 1.

Objective 3, Scenario 4 Efficacy

As objective 1 scenario 4.

Objective 3, Scenario 4 Toxicity

As objective 3 scenario 2.

Objective 3, Scenario 5 Efficacy

As objective 2 scenario 3.

Objective 3, Scenario 5 Toxicity

Inputs	Values
Prime doses	[0.00, 0.05, …, 0.95, 1.00]
Boost doses	[0.00, 0.05, …, 0.95, 1.00]
Anchor doses	[1,1],[0,0],[0,1], [1,0],[.5,.5], [0,.5], [.5,0],[1,.5], [.5,1]
Anchor probabilities	[.9], [.1], [.7], [.8], [.2], [.23], [.2], [.85], [.75]
K	13
Iterations	9

Objective 3, Scenario 6 Efficacy

As objective 2 scenario 2.

Objective 3, Scenario 6 Toxicity

As objective 3 scenario 5.

Objective 4

Objective 4 reused scenarios from the previous sections as detailed in the main body. Please refer to the relevant scenarios.

Appendix E: D-Optimal Design Theory

Throughout this work I have described various methods of trial dose selection or clinical trial design, in both adaptive and non-adaptive settings. One method in particular that was discussed was the 'Uniform' method in chapter 5. In that work I suggested that it may have been a reasonable method of trial dose selection because it samples the entire dosing domain, and hence should be effective for estimating model parameters and therefore useful for selecting optimal dose.

Whilst the above statement is likely to be reasonable if there is no prior understanding of model structure ('curve shape') or if there is little initial understanding of the likely model parameters, in the case where both a likely model form and expected model parameters are known in advance it is possible to select a method of trial dose selection ('trial design') that is better able to provide more accurate estimates of model parameters or model prediction than the 'Uniform' method given the same trial size. This is referred to as optimal experimental design theory.

Here I describe 'optimal design' theory, and in particular here I discuss D-optimal design, which focuses on maximising accuracy in the estimates of model parameters. C-optimal, E-optimal, and various other optimal design topics have also been discussed but will not be detailed here [37]. I also compare D-optimal designs to the 'Uniform' design discussed in chapter 5, and also to a 'Three near optimal' design which would involve testing three doses near the true optimal dose.

This is meant to be demonstrative of the ideas behind the field of optimal design in the context of the topics of this thesis. In this section I do not intend to provide recommendations.

A.E.1. D-Optimal Design: Methodology

D-optimal design aims to find a trial design that minimises the variance of the estimates of model parameters. Formally a 'design' is given by

$$\xi = \{(x_i, p_i), i = i, \dots k\}$$

Where K is the number of 'support points' which is to say the number of dosing levels that will be tested under that design. x_i is the ith dosing level and p_i is the proportion of trial individuals that should receive that dose.

A design is said to be optimal if it maximises the determinant of the information matrix $\det\big(M(\xi,\theta)\big)$ under that design for model $f(x,\theta)$ with parameter vector θ. The information matrix $M(\xi,\theta)$ is given by

$$M(\xi,\theta)) = \sum_{i=1}^{K} p_i \times \mu(x_i,\theta)$$

Where

$$\mu(x_i,\theta) = -E\left[\frac{\delta}{\delta\theta}\log\big(f(x_i,\theta)\big)\frac{\delta}{\delta\theta}\log(f(x_i,\theta))^T]\right]$$

is the information matrix of a single measurement taken at dose x_i.

A.E.2 D-Optimal design in chapter 5

Here I discuss application of D-optimal design to an example model that was used in chapter 5. In particular I use the latent-quadratic model for probability of binary efficacy used in the scenario 'Peaking 1'. This was given by

$$f(x_i,\theta) = peaking(x_i) = \frac{1}{1 + e^{base+gradient1\,x_i+gradient2\,x_i^2}}$$

With

$$\theta = (base, gradient1, gradient2) = (-9.0, 3.0, -0.214)$$

Following [38–40] I find that

$$\mu(x_i,\theta) = \frac{H'^{(z_i)^2}}{H(z_i)\big(1 - H(z_i)\big)}(1, x_i, x_i^2)^T(1, x_i, x_i^2)$$

Which is a 3x3 matrix with

$$z_i = e^{base + gradient1 \, x_i + gradient2 \, x_i^2}$$

$$H(z_i) = \frac{e^{base + gradient1 \, x_i + gradient2 \, x_i^2}}{1 + e^{base + gradient1 \, x_i + gradient2 \, x_i^2}}$$

$$H'(z_i) = \frac{e^{base + gradient1 \, x_i + gradient2 \, x_i^2}}{(1 + e^{base + gradient1 \, x_i + gradient2 \, x_i^2})^2}$$

A design that maximises $\det(M(\xi, \theta))$ can be found using the Particle Swarm Optimisation algorithm as described in [41]. Given that there are three parameters in this model, Carathéodory's theorem guarantees that the number of support points K for the optimal design will be at least three and no greater than seven[42]. The works of Hyun et al. and Yang et al. suggest that K=3 is likely to be the true optimal [43,44], so I set K=3 in this example.

Using a particle swarm optimisation algorithm with

- Objective function = maximise $\det(M(\xi, \theta))$
- K=3
- Swarm size = 40
- Length of the particle position vector = 5, namely (x1,x2,x3,p1,p2) with p3 = 1 - p1 -p2 then being calculated
- Inertia = 0.8
- Personal Best Weight = Global Best Weight = 1
- Alpha = 1
- Iterations = 1000
- Bounds 0<xi<10, 0 <pi<1, p1 + p2 < 1

This algorithm predicted that the optimal design was ξ_{opt}= (3.20, 0.33),(5.89, 0.33),(10.00, 0.33). A visualisation of this design and the efficacy curve $f(x_i, \theta)$ that was assumed in order to generate it are given in figure A.E.2.1

Figure A.E.2.1. Optimal design for Chapter 5's Scenario Peaking 1. The black line is the a-priori efficacy probability predicted by the model for each dose given the assumed model parameters. The red dots are the doses and proportion of doses that would be chosen by the D-optimal design method. The blue dots are the doses and proportion of doses that would be chosen by the uniform design method discussed in the chapter with n=30 trial individuals. The green dots are the doses and proportion of doses that would be chosen by choosing 3 doses near the optimal predicted dose, as is the standard for vaccine clinical trial design.

The efficiency of a design can be calculated as

$$efficiency(\xi, \theta) = \left(\frac{\det(M(\xi, \theta))}{\det\left(M\left(\xi_{opt}, \theta\right)\right)} \right)^{\frac{1}{J}}$$

where J is the number of parameters in the model (here J=3). The efficiencies of the three designs (d-optimal, uniform, Three near optimal) can thus be calculated, see table A.E.2.1. These values can be interpreted as requiring approximately 1.6 (= 1.000/0.624) times the number of trial participants under the uniform design to get the same degree of accuracy in parameter estimation relative to the D-Optimal

design. Approximately 53.2 times the number of trial participants would be required under the 'three near optimal' design.

Design	Efficiency
D-Optimal	1.000
Uniform	0.624
Three near optimal	0.0187

Table A.E.2.1. Efficiencies of three designs for chapter 5's Scenario Peaking 1.

This highlights that the uniform method of trial dose selection may be suboptimal relative to the D-optimal design, however the uniform design was still more informative for the purposes of modelling than selecting a small number of doses near the predicted optimal for this dose-response curve.

A.E.3. Balancing efficiency with benefit to trial population

Note that above the only optimisation criteria for choosing a design is the determinant of the information matrix. Whilst maximising for this is clearly important for estimating model parameters, it is also possible that benefit to trial participants needs to be considered. This leads to the concept of 'penalised D-optimal design'. Here I choose a trial design that maximises some function $\varphi(\xi, \theta)$ which should include not only the determinant of the information matrix but also some function $d(\xi, \theta)$ called the desirability function. Desirability functions are described in more detail in [A.C.4], but here are used to represent desirability of an optimal design as opposed to desirability/utility of a dose.

Using the 'bigger is better' desirability function from [35] I define the desirability function with respect to trial efficacy as

$$d_m(\xi, \theta) = \left(\frac{meanefficacy(\xi, \theta) - minimumefficacy(\theta)}{maximumefficacy(\theta) - minimumefficacy(\theta)} \right)^m$$

Where minimumefficacy() and maximumefficacy() are respectively the efficacy probabilities for the least and most efficacious dose in the [0,10] range as predicted by $f(x_i, \theta)$. $meanefficacy(\xi, \theta)$ over the k dosing groups is given by

$$meanefficacy(\xi, \theta) = \sum_{i=1}^{K} p_i \times f(x_i, \theta)$$

and m is a parameter that is used to determine the steepness of the desirability function [figure A.E.3.1].
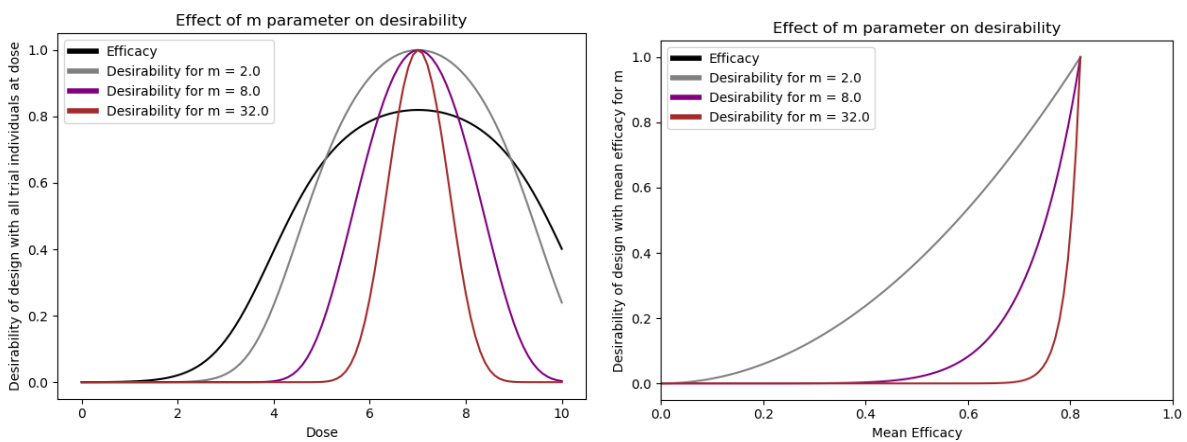
Figure A.E.3.1. The effect of m of the desirability function, with (left) showing the efficacy desirability of designs that place all doses at a specific dose and (right) showing the desirability of a design given some mean efficacy. The black line shows the 'Scenario Peaking 1' efficacy model that is assumed for this analysis.

Here I chose to define a penalised D-optimal design as that maximised $\varphi_m(\xi, \theta)$ with specific form

$$\varphi_m(\xi, \theta) = \det\big(M(\xi, \theta)\big) d_m(\xi, \theta)$$

Many forms $\varphi(\xi, \theta)$ have been suggested, but this simple one was used to highlight the trade-off between efficiency and benefit to trial participants.

I investigated six designs. 'D-optimal', 'Uniform', and ' Three near optimal' were as discussed in the previous section. 'Penalised m = 2.0', 'Penalised m = 8.0', and 'Penalised m = 32.0', instead each optimised $\varphi_m(\xi, \theta)$ for their respective values of

m. Optimal designs were again found through the particle swarm optimisation algorithm described above. Visualisations of theses designs are shown in figure A.E.3.2. The $d_m(\xi, \theta)$ scores and meanefficacy(ξ, θ) for each design were calculated for each of m = 2.0, 8.0, and 32.0 [Table A.E.3.1]
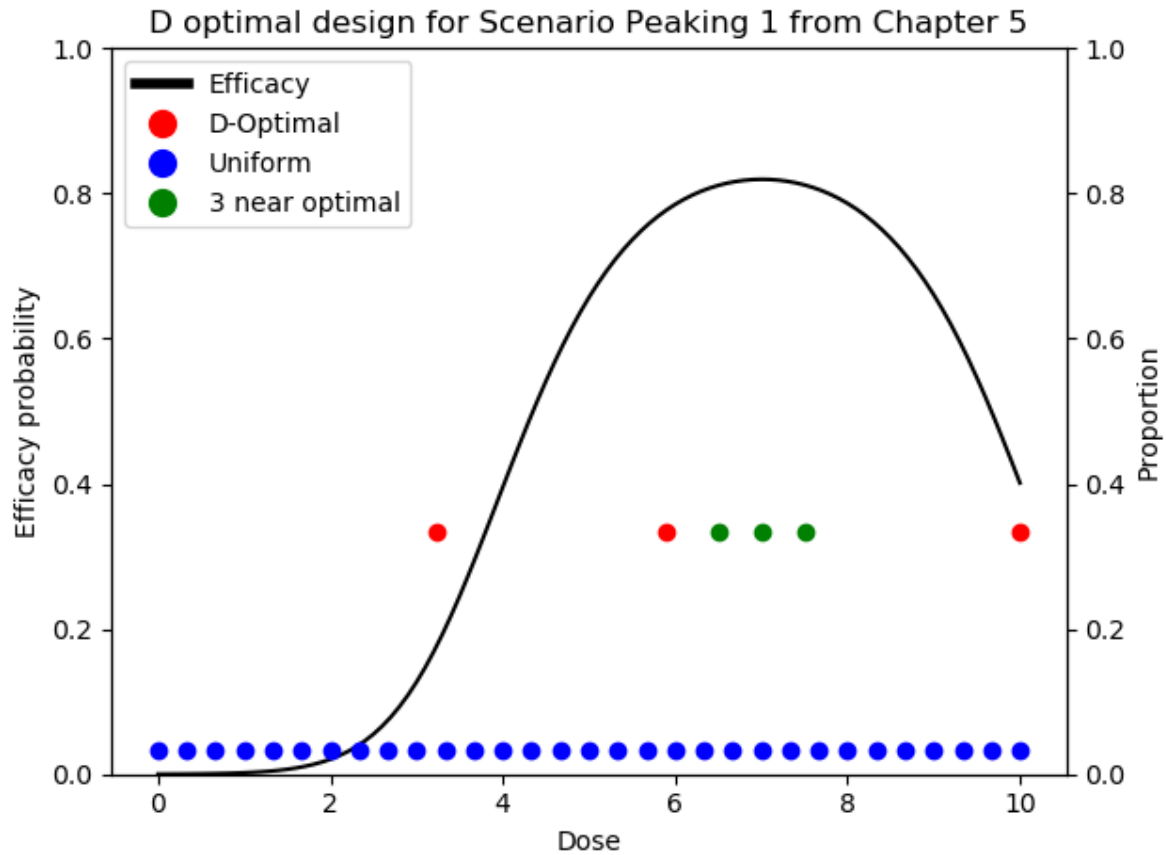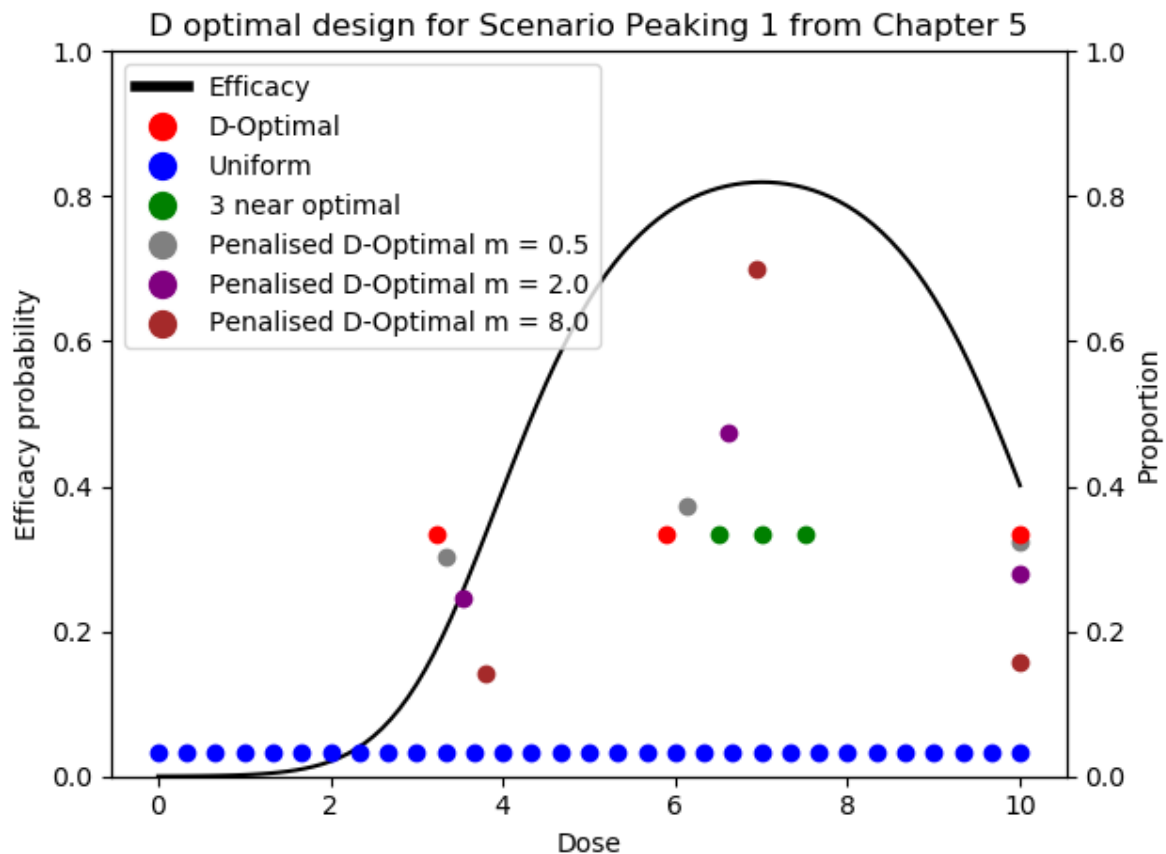

Figure A.E.3.2. Penalised optimal design for Chapter 5's Scenario Peaking 1. The black line is the efficacy probability predicted by the model for each dose given the assumed model parameters. The red dots are the doses and proportion of doses that would be chosen by the D-optimal design method. The blue dots are the doses and proportion of doses that would be chosen by the uniform design method discussed in the chapter with n=30 trial individuals. The green dots are the doses and proportion of doses that would be chosen by choosing a 3 doses near the optimal predicted dose, as is the standard for vaccine clinical trial design. The grey, brown, and purple dots represent the doses and proportion of doses that would be chosen by the penalised optimal design for m = 2.0, 8.0, and 32.0 respectively, with increasing m representing an increased desire for trial doses to be highly efficacious.

This showed that each of these designs were each 'optimal' for the metric for which they were optimised. As the value of m that was optimised for increased, the efficiency of the design decreased but the mean efficacy increased. The uniform

design was dominated by the D-optimal and penalised designs, but again did not require any assumptions on model form or parameterisation. The 'Three near optimal' design had the highest mean efficacy, but again reduced efficiency and would require a-priori knowledge of the true optimal dose.

Design Name	Efficiency	Score m = 2.0	Score m = 8.0	Score m = 32.0	Mean Efficacy
D-optimal	1.00	0.91	0.37	0.01	0.45
Penalised m = 2.0	0.99	0.93	0.43	0.02	0.49
Penalised m = 8.0	0.94	0.84	0.48	0.05	0.56
Penalised m = 32.0	0.71	0.40	0.31	0.10	0.68
Uniform	0.62	0.22	0.09	0.00	0.44
Three near optimal	0.02	0.00	0.00	0.00	0.81

Table A.E.3.1. Efficiencies, $\varphi_m(\xi, \theta)$ scores, and mean efficacy for the six designs.

A.E.4. Bayesian D-optimal Design

So far, I have assumed that there was a model $f(x, \theta)$ where θ was the 1xJ vector $\theta = (\theta_1, \dots, \theta_J)$ with specific known values for $\theta_1, \dots, \theta_J$. This is referred to as 'locally D-optimal design', as the designs that are generated are only optimal for if the parameter values $\theta_1, \dots, \theta_J$ are accurate. If there is uncertainty in the parameter values, then 'Bayesian D-optimal Design' is instead preferable.

In Bayesian D-Optimal Design, a prior can be placed over the model parameters . Therefore the model becomes $f(x, \bar{\theta})$ where $\bar{\theta}$ is an LxJ matrix of L possible values for the J parameters. The information matrix for which the determinant is to be maximised then becomes

$$M(\xi, \theta)) = \frac{1}{L} \sum_{l=1}^{L} \sum_{i=1}^{K} p_i \times \mu(x_i, \theta_l)$$

Where θ_l is the lth row of $\bar{\bar{\theta}}$ and again there are K dosing groups in the trial.

For a specific demonstration, I assumed that all 5 of the parameter value sets for the 5 'peaking' scenarios in chapter 5 where possible, and thus had $\bar{\bar{\theta}}$ as the 5x3 matrix of 5 possible values for the 3 parameters. The dose-efficacy of these 5 curves and the mean of the 5 efficacy predictions for each dose are shown in figure A.E.4.1
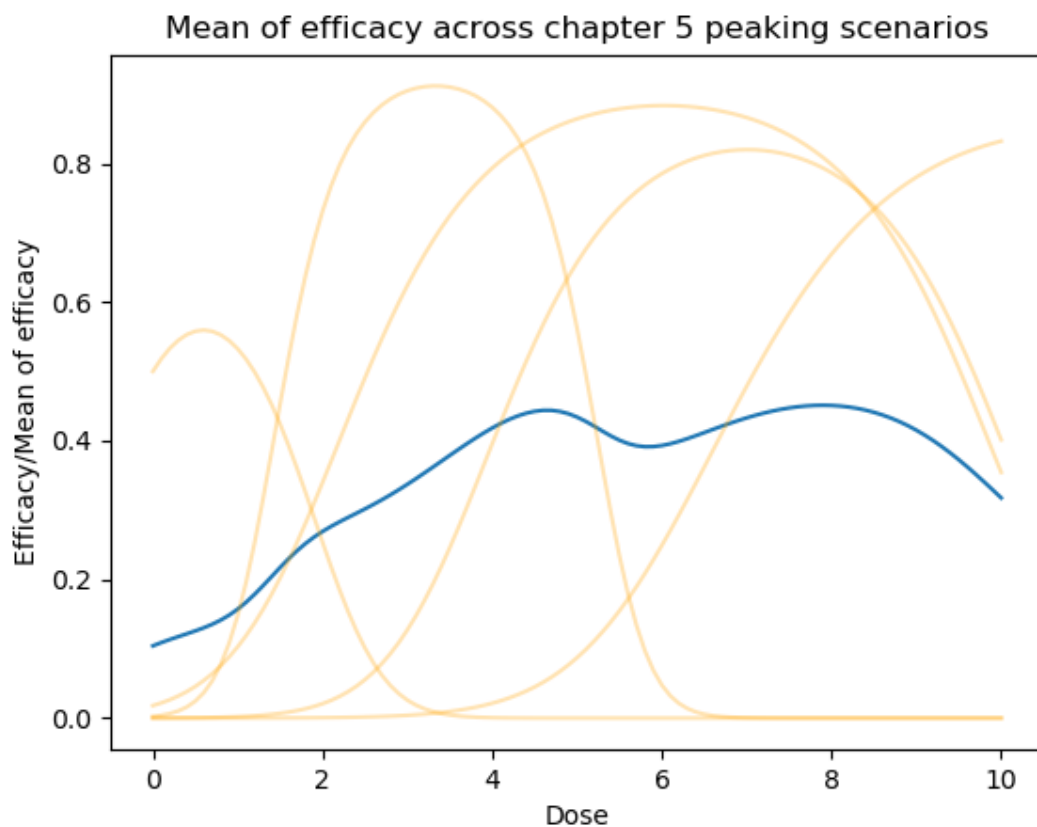
Figure A.E.4.1. The five efficacy curves that are assumed to be possible for demonstrating Bayesian D-optimal design. The 5 possible efficacy curves are shown in orange, and the mean of these curves is shown blue.

Bayesian D-optimal design was calculated by maximising the determinant M(,).
Bayesian Penalised D-optimal designs was calculated by maximising

$$\varphi_m(\xi, \bar{\theta}) = \det\left(M(\xi, \bar{\theta})\right) d_m(\xi, \bar{\theta})$$

With

594

$$d_m(\xi, \theta) = \left(\frac{meanefficacy(\xi, \bar{\theta}) - minimumefficacy(\bar{\theta})}{maximumefficacy(\bar{\theta}) - minimumefficacy(\bar{\theta})} \right)^m$$

$$meanefficacy(\xi, \bar{\theta}) = \frac{1}{L} \sum_{l=1}^{L} \sum_{i=1}^{K} p_i \times f(x_i, \bar{\theta})$$

$$maximumefficacy(\bar{\theta}) = max_{x_x} \sum_{l=1}^{L} f(x_i, \bar{\theta_l})$$

$$minimuumefficacy(\bar{\theta}) = min_{x_x} \sum_{l=1}^{L} f(x_i, \bar{\theta_l})$$

Which is to say that desirability is determined based on the mean prior prediction of efficacy at each dose over the 5 parameter sets.

I investigated the same six trial design as before, again using particle swarm optimisation. Visualisations of these designs are shown in figure A.E.4.2. The $\varphi_m(\xi, \bar{\theta})$s and $meanefficacy(\xi, \bar{\theta})$ for each design were calculated [Table A.E.4.1]
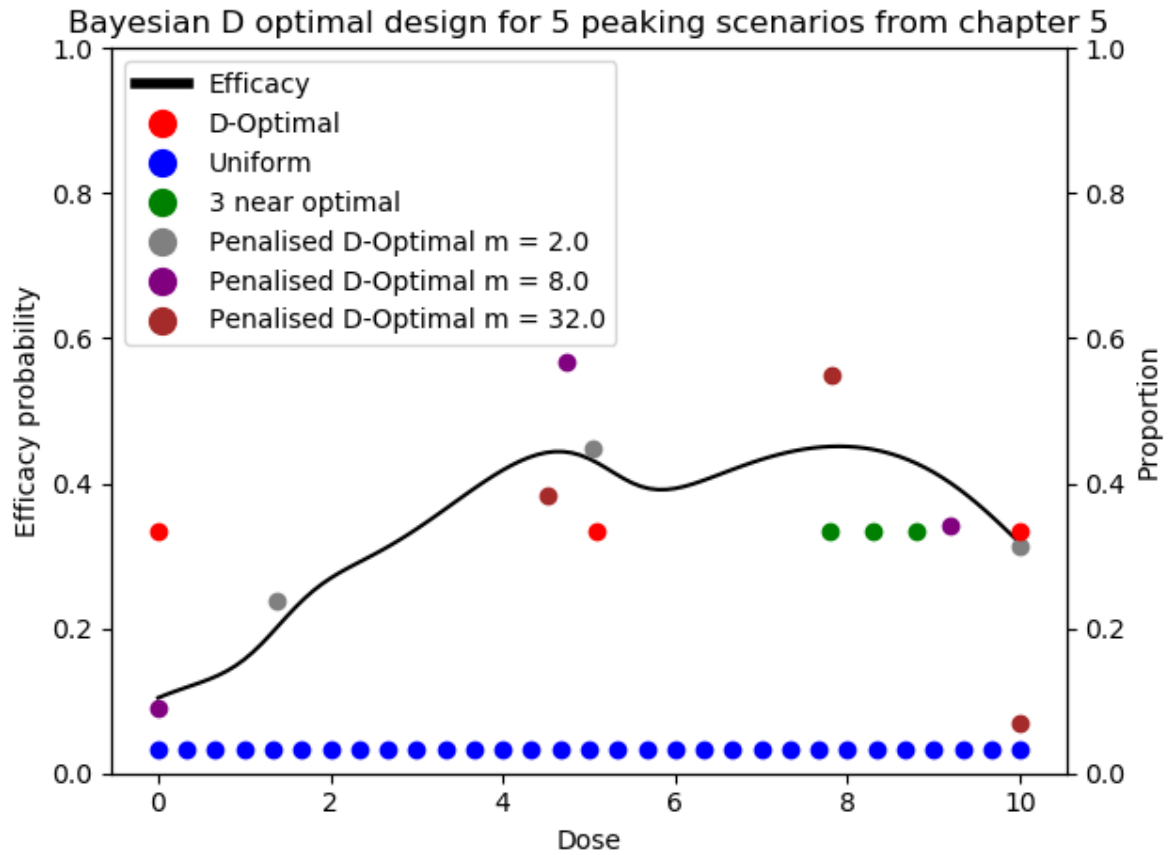
Figure A.E.4.2. Bayesian optimal designs for Chapter 5's Scenario Peaking 1. The black line is the mean of the efficacy probabilities predicted by the 5 model for each dose given the assumed 5 sets of model parameters. The red dots are the doses and proportion of doses that would be chosen by the D-optimal design method. The blue dots are the doses and proportion of doses that would be chosen by the uniform design method discussed in the chapter with n=30 trial individuals. The green dots are the doses and proportion of doses that would be chosen by choosing three doses near the optimal predicted dose, as is the standard for vaccine clinical trial design. The grey, brown, and purple dots represent the doses and proportion of doses that would be chosen by the penalised Bayesian optimal designs for m = 2.0, 8.0, and 32.0 respectively, with increasing m representing an increased desire for trial doses to be highly efficacious.

Design Name	Efficiency	Score m = 2.0	Score m = 8.0	Score m = 32.0	Mean Efficacy
D-optimal	1	0.55	0.01	0	0.28
Penalised m = 2.0	0.94	0.79	0.08	0	0.34
Penalised m = 8.0	0.66	0.42	0.15	0	0.4
Penalised m = 32.0	0.26	0.03	0.03	0.01	0.44
Uniform 30	0.67	0.3	0.04	0	0.35
Three near optimal	0.01	0	0	0	0.44

Table A.E.4.1. Efficiencies, $\varphi_m(\xi, \overline{\theta})$s and $meanefficacy(\xi, \overline{\theta})$ for the six designs on the Bayesian optimal design problem.

Again, this showed that each of the penalised d-optimal designs were 'optimal' for the metric they were optimised for, and the same impact of increasing m was observed. The efficiency of the uniform design (0.67) was increased relative to the efficiency in the locally D-optimal setting (0.62). This was expected and is consistent with the hypothesis that a uniform design may be reasonable if there is little a-priori certainty in model form and parameters. Again the 'three near optimal' design was not as efficient as the uniform trial design but had higher mean efficacy.

AE.5 Potential limitations with D-optimal design with regards to vaccine clinical trial design

Here I have demonstrated how D-optimal design and penalised D-optimal design can be used to maximise the amount of information regarding model parameters that can be gathered in a clinical trial. Whilst this may be highly beneficial in reducing trial costs, improving accuracy in estimations of model parameters, and hence choosing optimal dose, there are a number of limitations to implementation.

Firstly, the D-optimal design process is complicated relative to some other methods of choosing clinical trial doses designs, and clinicians may not be willing to invest the time or resources to implement such a complex methodology. Further to the topics discussed here, it may be necessary to find a design that is D-optimal with regards to modelling multiple responses, for example both an efficacy model and a toxicity model. Whilst methods for doing so exist [45], this would even further increase complexity.

Secondly, these methods rely on assuming a model structure and parameters. As I showed in chapter 3, it may not be reasonable to assume that a 'correct' model is known a-priori. Methods of D-optimal design with model uncertainty exist [46,47] but again increase complexity in calculating the optimal design and also require some a-priori estimate for the probability that each model is correct.

A further argument against D-optimal design is that it assumes that the parametric model that is used (or at least one of the models investigated in the case of model uncertainty) is for some parameter set a completely accurate representation of the true underlying dose-response. Whilst the models that are used to describe dose-response may be good approximations of dose-response, it is unlikely that they are perfectly accurate to the truth.

Finally, it must be remembered that the overall goal of the mathematical modelling that I discuss in this work is to optimise dose. Before D-optimal design is implemented mathematical models would have to be accepted as effective for the purpose of optimising dose, and that there would be significant benefit in designing trials specifically around accurate parameter estimation of these models. For example, the benefits of choosing trial doses proportionally to the probability that they are optimal based on the present model ('Thompson Sampling', chapter 6) may be more intuitively reasonable to clinicians than choosing trial doses that will aid in model parameterisation.

I note here that, at least with regards to efficiency, the uniform method of trial dose selection may be a reasonable approximation of D-optimal design without the complexity in implementation or requirement of a-priori model knowledge. I also found that in this investigation, using three doses near the optimal dose was not efficient for parameter estimation for this assumed parameterisation of the latent quadratic model, but was often the most efficacious design. Further work would be needed to validate if this is the case for other models and parameterisation, but it may imply that clinicians should consider whether such designs are optimal if mathematical modelling for selecting optimal vaccine dose is to be used.

Appendix F: Dosing Space Density

Throughout this work I at many points refer to the concept that the number of dosing groups that are considered when attempting to select optimal vaccine may affect how effectively dose may be optimise. This is the concept that, although vaccine dose is technically a continuous measure, it is often discretised to a finite (and small) number of potential doses during dose finding studies (I call this dosing-domain discretisation). For example, the dose-finding study from which data were extracted for chapter 4 considered only 3 doses [48]. These were namely 0.5mL, 1.0mL and 1.5mL. However, other than practicalities of vial size, there is no reason that for example 0.75mL or 1.28mL could not be possible optimal doses. Any volume in the continuous range between 0 and 1.5mL could have been considered, but the dose-optimisation approach used in that work considered only three discrete doses.

I here use the term dosing-domain density to refer to the number of discretised doses that are considered within a given continuous dosing domain. A clinical trial that considers 3 equally spaced discretised doses would be considered to be sparse relative to a clinical trial considering 12 equidistantly spaced doses over the same possible range (same maximum and minimum doses). It would be however dense relative to a clinical trial considering only 2 equidistantly spaced doses.

A.F.2. The effect of dose domain density in the uniform naive setting

In this section I ignore the possibility of modelling to discuss the effect of increasing the discretisation density of a dosing domain using the 'uniform naive'/'direct comparison' dose optimisation approach in chapter 6. I conducted a small simulation study to investigate this, which is not large enough to provide meaningful implications but may be of interest.

Consider the case of trying to locate the optimal single-administration dose d_{opt}, defined as the maximum efficacy dose, for a vaccine. Based on past data and input from clinicians I assume that there is known a dosing-range which is guaranteed to contain the optimal dose. However, I assume that there is a uniform prior over this

range for where the true optimal dose is. Knowing that there are some total N individuals with which to conduct a dose ranging trial, and deciding to use the uniform naive trial design, the question posed is to choose some K such that there are K dosing groups each containing N/K individuals. I assume that these K potential doses are equidistantly distributed over the range that is known to contain the optimal dose. An increase in K therefore represents an increase in the density of the discretised dosing domain.

I hypothesised that as K increases there should be an increase in the efficacy of the optimal of the K discretised doses. I also hypothesised that the probability of selecting this optimal dose would be decreased, as an increase in K would decrease the N/K individuals per dosing group, decreasing the statistical power to determine which of the groups was most optimal. I hypothesised that these two above statements would combine to mean there was some K that would be optimal for maximising the true efficacy of the dose that is chosen based on the clinical trial.

To investigate this hypothesis, I conducted a simulation study. I used the efficacy curve from chapter 5's 'Scenario Peaking 1', but with the dosing range extended. See Figure A.F.2.1. The optimal/maximum efficacy dose d_{opt} = 7.0.
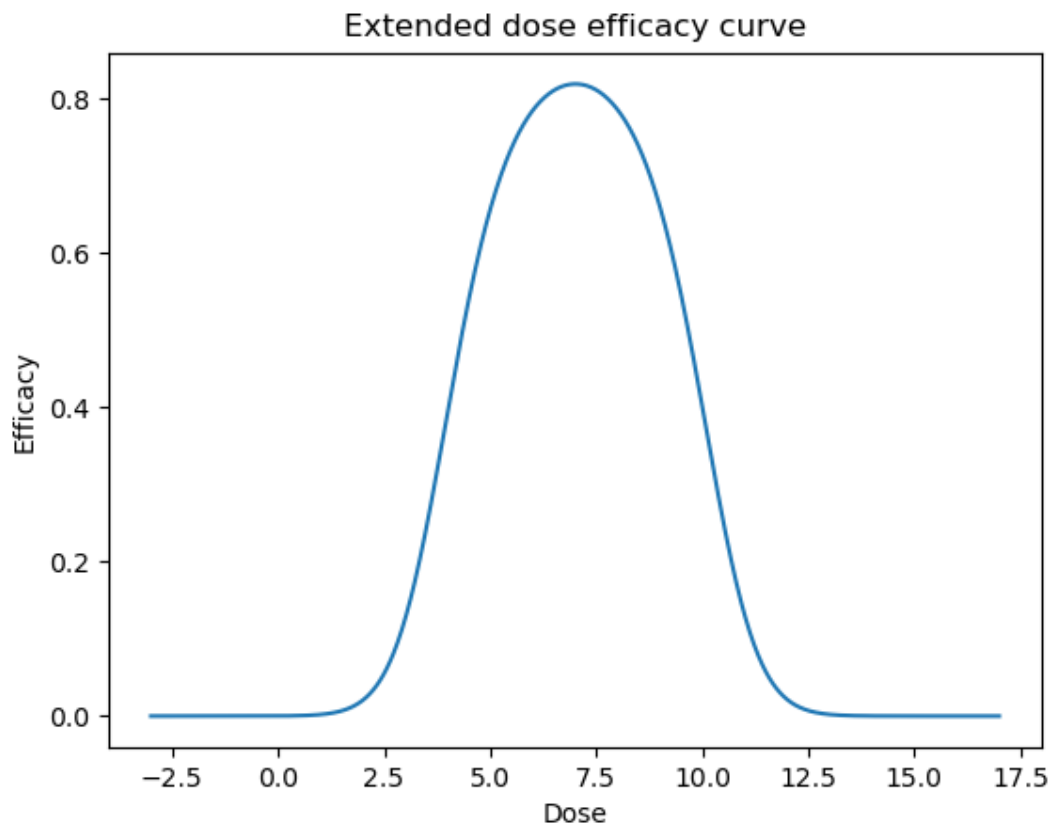
Figure A.F.2.1. True dose-response curve

The range that would be investigated to locate optimal dose was calculated as [min, max] where

$$\text{min} = d_{opt} - r \times size_{interval}$$

$$\text{max} = \text{min} + size_{interval}$$

Where r was uniformly sampled in the range [0,1] and $size_{interval}$ is the size of the interval that is known to contain the optimal dose. This guaranteed that d_{opt} was in the range [min, max] as required. Examples of potential ranges for $size_{interval}$ = 2, 5, 10 are shown [Figure A.F.2.2]. Note that an increase in $size_{interval}$ reflects a greater variance in efficacy between doses in the resulting interval, as for $size_{interval}$ 2 all doses have between 72% and 82% efficacy, whereas for $size_{interval}$ =10 efficacy could vary between 0% and 82% for some intervals.
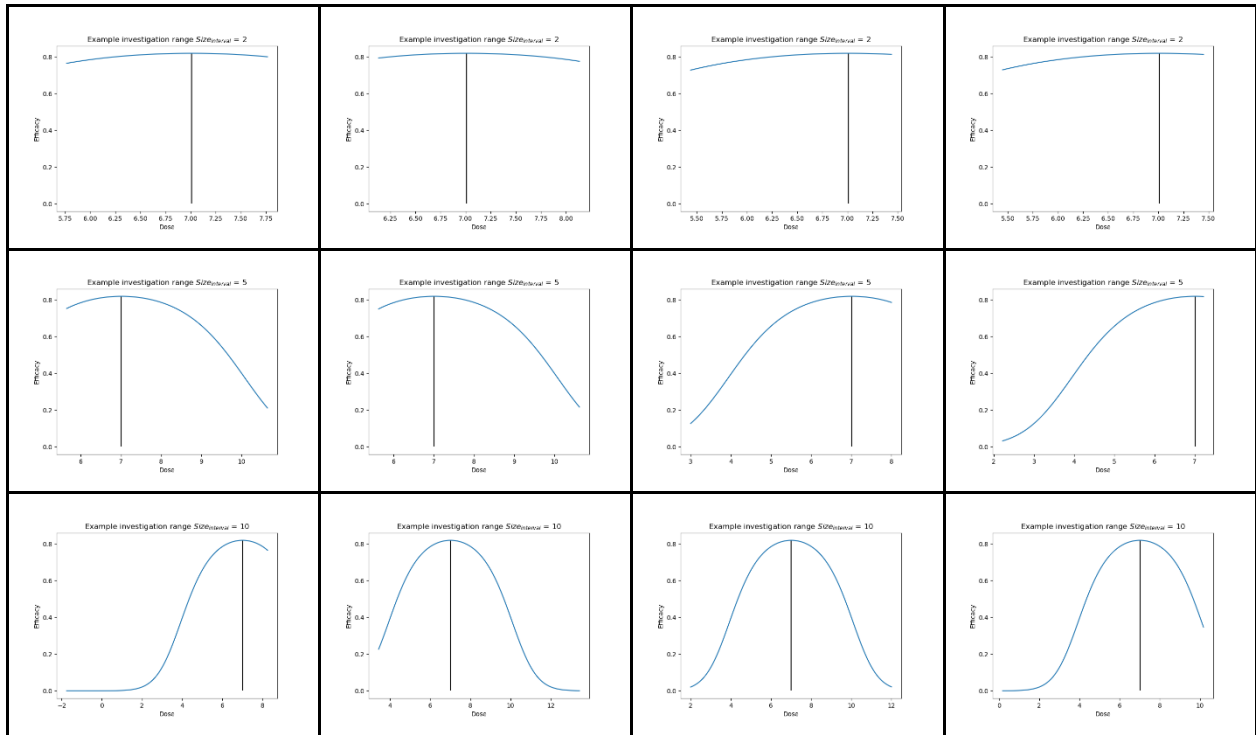
Figure A.F.2.2. Example dosing ranges for different interval sizes. Top $size_{interval}$ 2, middle $size_{interval} = 5$, bottom $size_{interval} = 10$.

The doses for the K dosing groups are then chosen to be equidistant from each other. If K = 1 the dose is in the middle of the dosing range. If K = 2 then the doses are respectively at the ⅓ and ⅔ points along the dosing range. For K>2 the doses are chosen such that the smallest dose is at the bottom of the dosing range, the largest is at the top of the dosing range, and the rest divide the range into K-1 equal segments. Examples for K = 1, 2, 3, 4, 5 and 10 are shown [Figure A.F.2.3.].

603
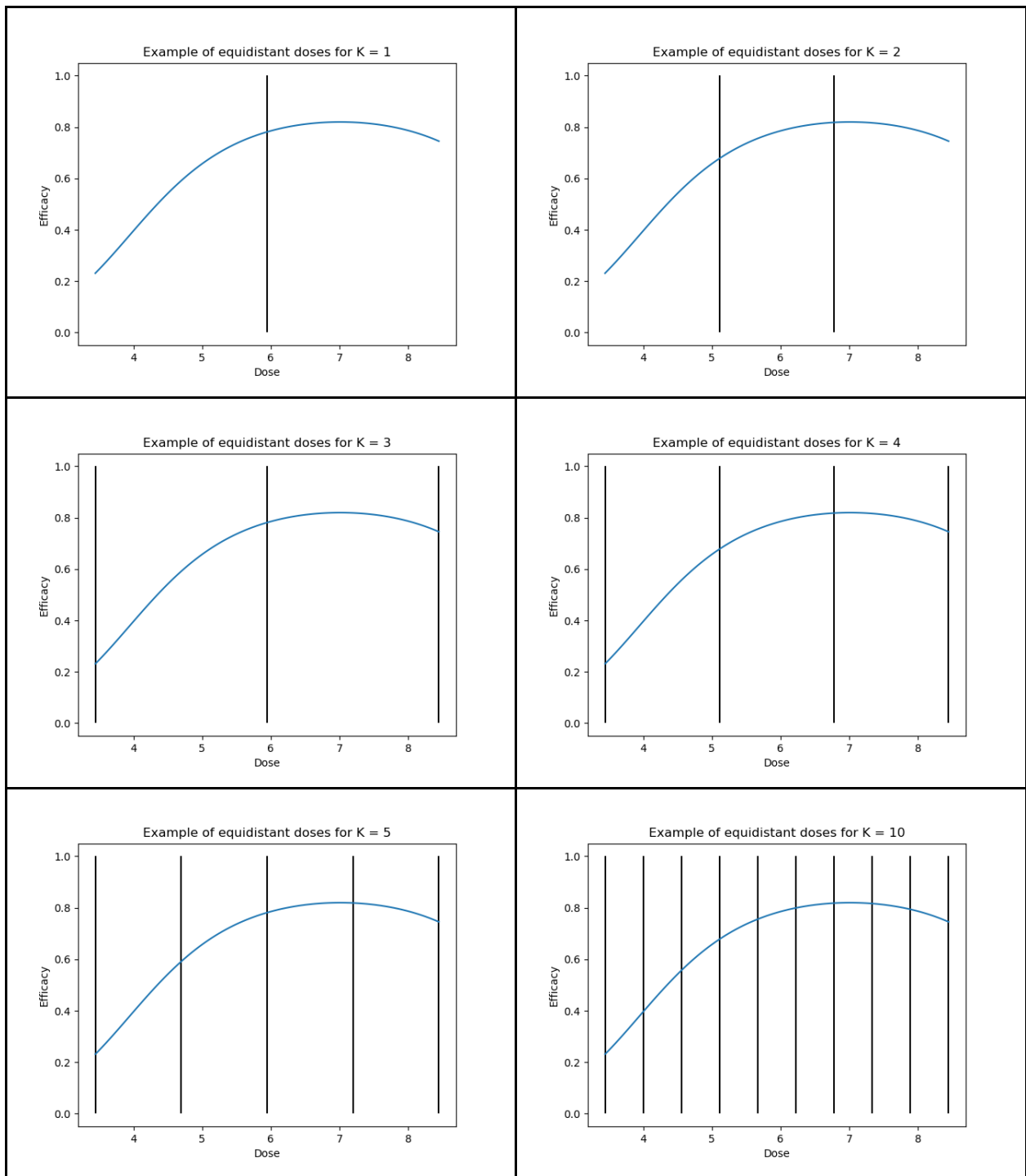
Figure A.F.2.3. Example dosing groups for different K values over the same interval.

Clinical trials could then be simulated as per the uniform naive dose optimisation approach in chapter 6. Briefly, this involves splitting the N potential trial participants equally amongst the K dosing groups. The number of efficacious responses in each group is observed, and the dose for which the group with the greatest number of efficacious responses was observed is selected as optimal. If there was a tie, the smallest dose was selected. For each clinical trial I recorded the metrics

- K,
- the efficacy of the K potential doses that was truly optimal
- whether the truly optimal dose out of the K doses was predicted to be the optimal dose
- the true efficacy of the predicted optimal dose.
- the predicted efficacy for the predicted optimal dose

I simulated clinical trials for

- $\{size\}_{\{interval\}}$ =2,5,10
- N = 30, 60, 120 and 240,
- K = 1, 2, 3, 5, 6, 10, 15, 30, 60, and 120.

I conducted 1000 simulations for each trio of $\{size\}_{\{interval\}}$, K and N, each with a different randomly sampled r, and the results are plotted below [A.F.2.4. If K were greater than N a clinical trial was not simulated as this would mean that the group size N/K would be less than 1. All N/K were integers by design. The mean of each of these metrics stratified by N, $\{size\}_{\{interval\}}$ and K were calculated. The K that maximised the mean of the true efficacy of the predicted optimal dose for each N and $size_{interval}$ was said to be optimal for that N and $size_{interval}$ for this dose-efficacy curve.

Figure A.F.2.4. Metrics of simulated naïve uniform clinical trials for different K values. N increases from top to bottom. $size_{interval}$ increases from left to right.

These results support my hypotheses. Increasing K led to an increase in the true best efficacy amongst the K possible dose, particularly for large $\{size\}_{\{interval\}}$. Increasing K decr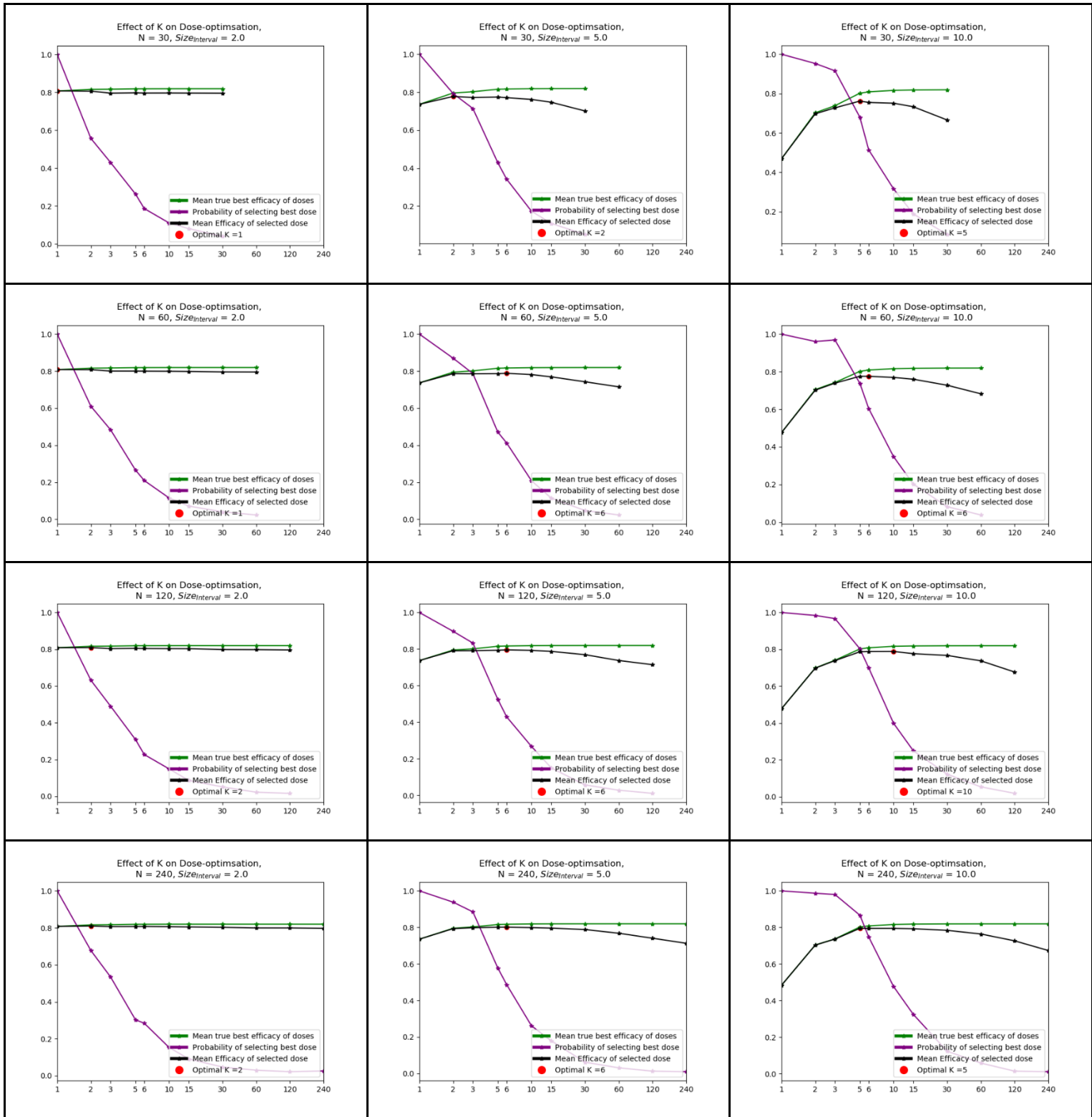eased the probability that the true optimal dose would be selected. The relationship between K and the efficacy of the dose that was predicted to be optimal based on the clinical trials was non-monotonic, initially increasing but decreasing past a certain value of K.

The optimal K for selecting maximum efficacy dose was dependent on N and $size_{interval}$. An increase in N generally led to an increase in the value for the optimal K. This is likely because an increase in N would counterbalance the increased value of K with regards to the value of N/K. Increasing $size_{interval}$ also led to an increase in the optimal value of K. This is likely because when $size_{interval}$ was small there was little difference in efficacy between any of the potential doses, and so a greater value of N/K was needed to determine a difference in efficacy between them.

These results suggest that, if the 'uniform naive' dose-optimisation approach is being used, then the number of dosing groups that should be used (which is to say the density of the dosing domain), may be dependent on a number of factors. If there was expected to be a large variance in efficacy between different doses in the range that is being investigated, represented here by a large $size_{interval}$, then a larger number of dosing groups should have been considered. If it was believed that there would be minor difference in efficacy between doses in the range that is being considered, represented here by a small $size_{interval}$ then there was little benefit in using more than 1 dosing group. This was particularly true for small N, as it was unlikely that any distinction in efficacy between multiple doses can be determined given the small group size.

Further, in this example increasing K in the case where there was slight difference in efficacy maximised the risk of optimistic bias and overestimating vaccine efficacy. Results for this are shown below [A.F.2.5.], with the mean inaccuracy (blue) increasing with K for all N and $size_{interval}$
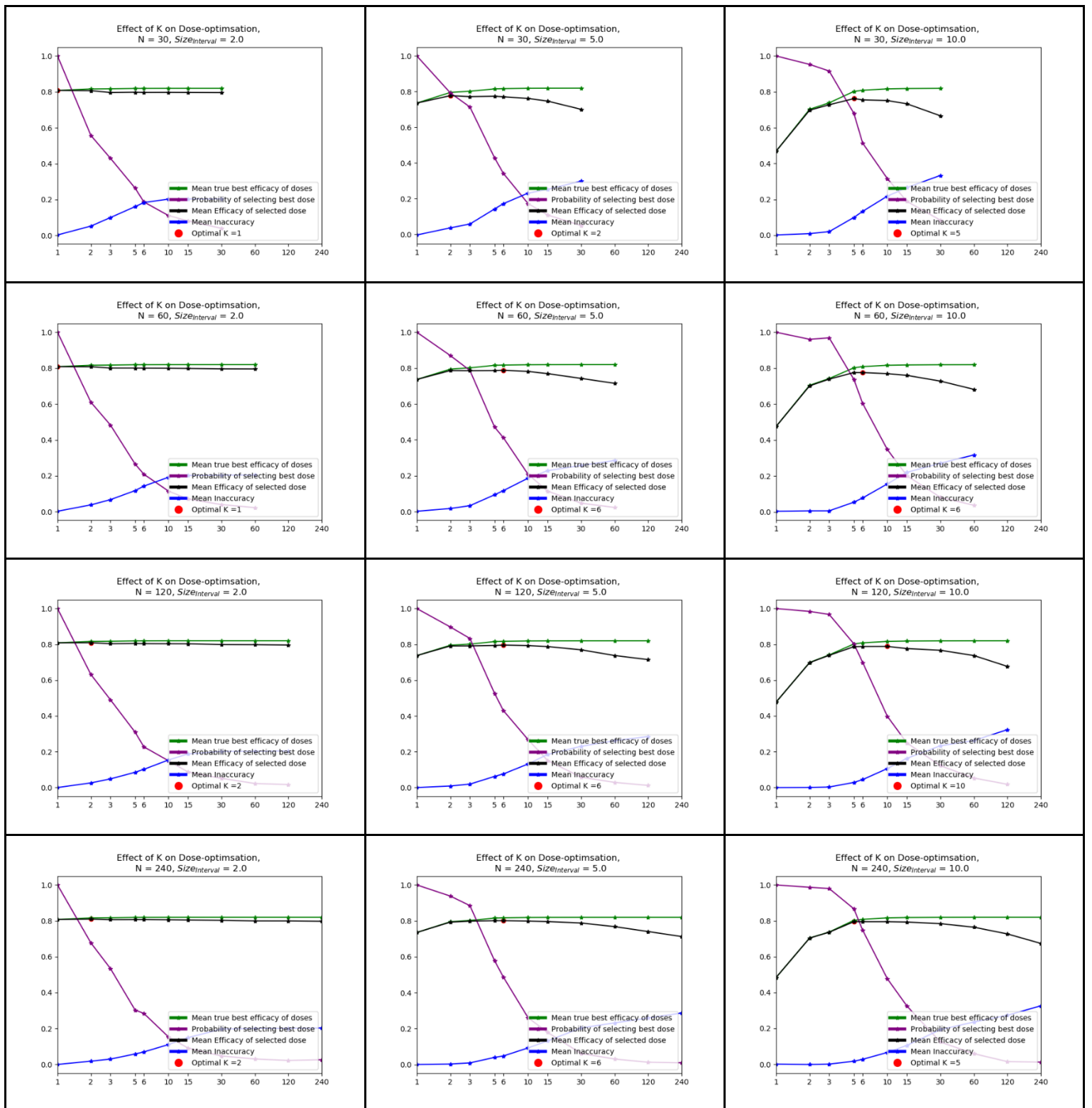
Figure A.F.2.5. Metrics of simulated naïve uniform clinical trials for different K values, including inaccuracy. N increases from top to bottom. $size_{interval}$ increases from left to right.

I note that, as this was an informal simulation study, this is more a demonstration of the potential impact that the decision of dosing space density could have for some dose-efficacy curve under certain assumptions (peaking dose-efficacy, optimal dose guaranteed to be within the dosing domain and with uniform distribution). These findings are presented purely for the interested reader. A larger simulation study using more scenarios and with a sensitivity analysis of these assumptions would be needed to thoroughly investigate these hypotheses or to make recommendations regarding optimal dosing space density.

Appendix G: Objective 3 for chapter 5

In addition to the objectives i and ii which investigated the effect of changing i) efficacy model and trial size and ii) efficacy model and trial dose selection method in chapter 5, I also considered the logical third objective, investigating the effect of changing iii) trial size and ii) method of trial dose selection. This was excluded from the published work to reduce the complexity and length of that paper, but results are included here as they highlight the importance of exploration for the purposes of optimal dose selection.

For this objective I repeated the simulation study on the same 14 scenarios, as in objectives i and ii. I considered dose-optimisation approaches of the form
Efficacy Model: Weighted

    i.   Trial size: 10, 30, 60, or 100
    ii.   Trial dose selection method: Full uniform exploration, standard fully continual modelling, balanced exploration (SoftMax) fully continual modelling, or three stage (SoftMax).

I chose to fix the assumed efficacy model to the 'Weighted' efficacy model, as dose-optimisation approaches that used this efficacy model were found to be effective in the first two objectives.

With the efficacy model being the 'weighted' model, I found that the importance of including exploration in the dose-optimisation approach increased with trial size. In general, across all 14 scenarios, increasing trial size above 30 did not decrease PSR when the standard fully continual method of trial dose selection was used [Figure AG.1.]. This is in contrast to the other three methods of trial dose selection, for which an increase in trial size always decreased PSR. This is in line with the findings of the main body of the work, that not allowing for exploration of doses that are predicted to be suboptimal reduces the quality of the final selected dose. These findings suggest that this effect increases as trial size dose.

Figure A.G.1. Percentage Simple Regret (PSR) for all scenarios, by trial size and trial dose selection method. Assumed efficacy model was 'Weighted'. Individual points represent PSR for a single simulated clinical trial using that dose-optimisation approach for one of the 14 scenarios. The middle line of each boxplot is the median value, the box marks the 25th and 75th percentiles, and the whiskers mark the 5th and 95th percentiles of the data. A lower PSR denotes a more optimal final dose.

With the efficacy model being the 'weighted' model there was minimal difference in Inaccuracy and Absolute Inaccuracy across the approaches [figure AG.2]. This further suggests that accuracy of utility predictions at the model predicted optimal vaccine dose was not dramatically improved by using continual modelling method of dose selection. There was still an optimistic bias observed. I again found that an increased trial size reduced Inaccuracy and Absolute Inaccuracy.

611

(a)   **(b)**

Figure A.G.2. Inaccuracy (a) and Absolute Inaccuracy (b) for all scenarios, by trial size and trial dose selection method. Assumed efficacy model was 'Weighted'. Individual points represent Inaccuracy/Absolute Inaccuracy for a single simulated clinical trial using that dose-optimisation approach for one of the 14 scenarios. The middle line of each boxplot is the median value, the box marks the 25th and 75th percentiles, and the whiskers mark the 5th and 95th percentiles of the data. The closer Inaccuracy/Absolute Accuracy were to 0 the more accurate that the prediction of utility was at the predicted optimal dose.

With the efficacy model being the 'weighted' model, the results suggest that fully continual modelling (both standard and balanced) and three stage approaches identify optimal dose with a greater net benefit to trial participants than the retrospective full uniform exploration approaches (as shown by decreased Average Regret) [Fig A.G.3]. The balanced exploration variant of the fully continual modelling dose-selection method appeared to have a marginally increased Percentage Average Regret compared to approaches with standard fully continual modelling dose-selection, but Average Regret was still significantly reduced relative to approaches using the three stage SoftMax or full uniform exploration methods of trial dose selection. The three stage SoftMax approaches showed a reduced Average Regret relative to full uniform exploration but a greater Average Regret relative to the fully continual approaches. As trial size increased Percentage Average Regret decreased for all methods of trial dose selection other than the full uniform method.
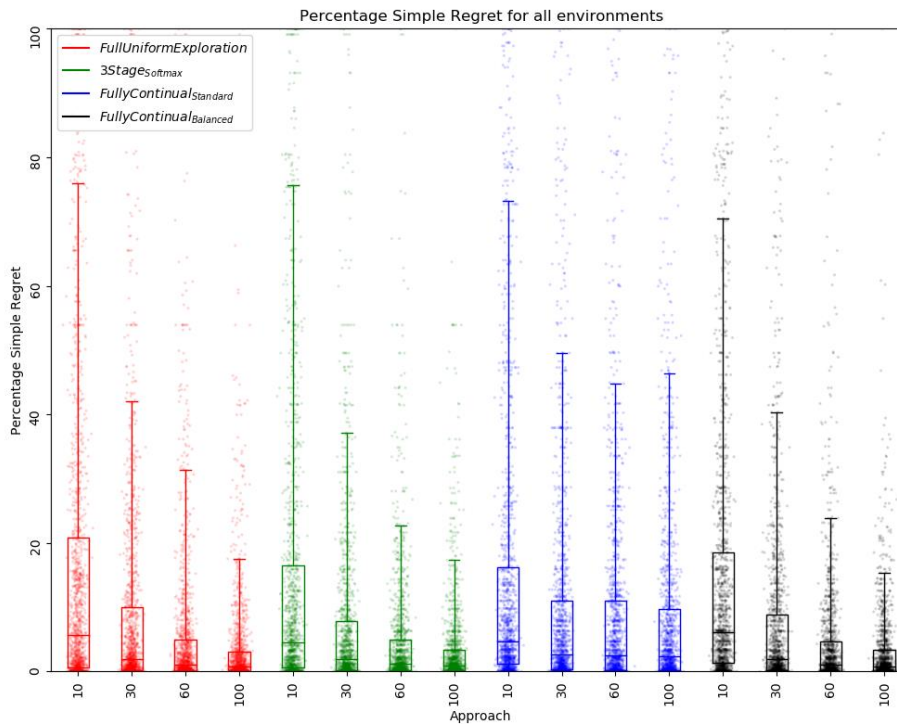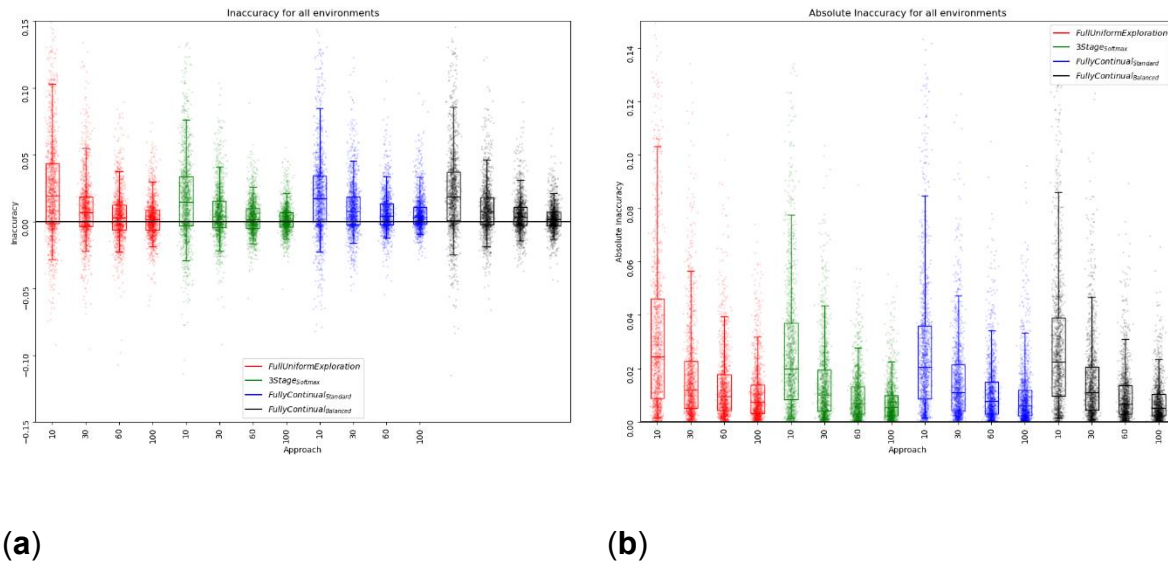
612
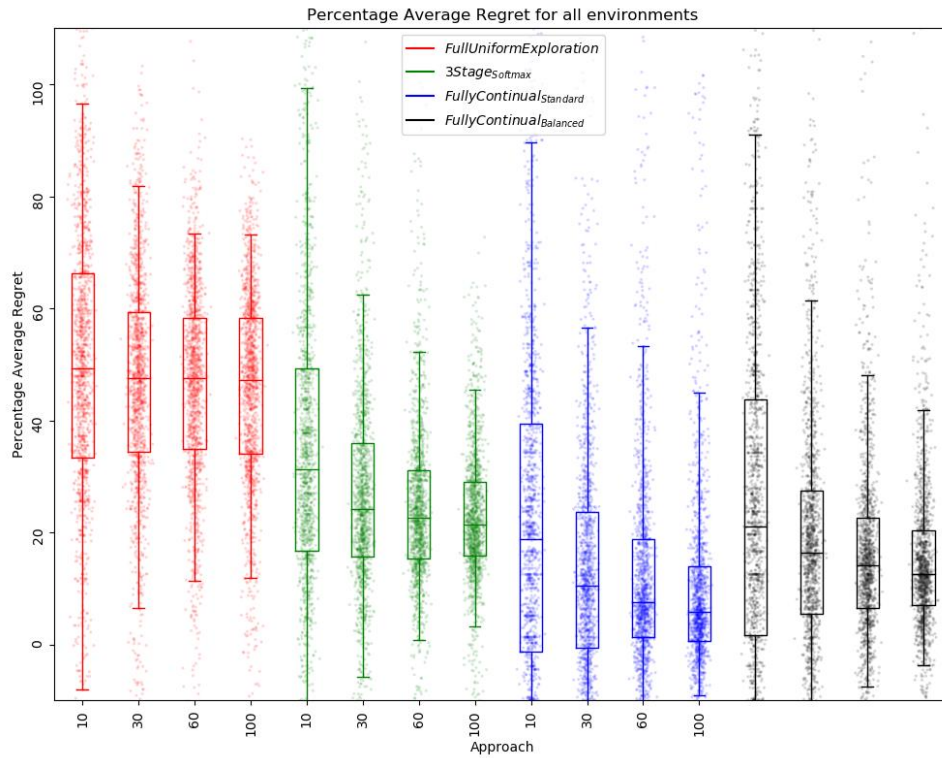
Figure A.G.3. Percentage Average Regret for all scenarios, by trial size and trial dose selection method. Assumed efficacy model was 'Weighted'. Individual points represent Percentage Average Regret for a single simulated clinical trial using that dose-optimisation approach for one of the 14 scenarios. The middle line of each boxplot is the median value, the box marks the 25th and 75th percentiles, and the whiskers mark the 5th and 95th percentiles of the data. A lower percentage Average Regret denotes better outcomes to trial participants.

Appendix References

1. O'Quigley, J.; Iasonos, A.; Bornkamp, B. Dose-Response Functions. In *Handbook of Methods for Designing, Monitoring, and Analyzing Dose-Finding Trials: Handbooks of Modern Statistical Methods*; 2017; p. 199 ISBN 978-1-315-15198-4.

2. Slob, W. Dose-Response Modeling of Continuous Endpoints. *Toxicological Sciences* **2002**, *66*, 298–312, doi:10.1093/toxsci/66.2.298.

3. Thall, P.F. BAYESIAN ADAPTIVE DOSE-FINDING BASED ON EFFICACY AND TOXICITY. *Journal of Statistical Research* **2012**, *46*, 187–202.

4. Rhodes, S.J.; Zelmer, A.; Knight, G.M.; Prabowo, S.A.; Stockdale, L.; Evans, T.G.; Lindenstrøm, T.; White, R.G.; Fletcher, H. The TB Vaccine H56+IC31 Dose-Response Curve Is Peaked Not Saturating: Data Generation for New Mathematical Modelling Methods to Inform Vaccine Dose Decisions. *Vaccine* **2016**, *34*, 6285–6291, doi:10.1016/j.vaccine.2016.10.060.

5. O'Quigley, J.; Iasonos, A.; Bornkamp, B. *Handbook of Methods for Designing, Monitoring, and Analyzing Dose-Finding Trials: Handbooks of Modern Statistical Methods*; 2017; ISBN 978-1-315-15198-4.

6. Brock, K.; Billingham, L.; Copland, M.; Siddique, S.; Sirovica, M.; Yap, C. Implementing the EffTox Dose-Finding Design in the Matchpoint Trial. *BMC Medical Research Methodology* **2017**, *17*, 112, doi:10.1186/s12874-017-0381-x.

7. Fedorov, V.V.; Leonov, S.L. Generalized Regression and Elemental Fisher Information Matrices. In *Optimal Design for Nonlinear Response Models*; CRC Press: Boca Raton, 2013; pp. p23-30 ISBN 978-0-429-10398-8.

8. Gadagkar, S.R.; Call, G.B. Computational Tools for Fitting the Hill Equation to Dose-Response Curves. *J Pharmacol Toxicol Methods* **2015**, *71*, 68–76, doi:10.1016/j.vascn.2014.08.006.

9. O'Quigley, J.; Iasonos, A.; Bornkamp, B. MCP-Mod Methodology. In *Handbook of Methods for Designing, Monitoring, and Analyzing Dose-Finding Trials: Handbooks of Modern Statistical Methods*; 2017; pp. p206-210 ISBN 978-1-315-15198-4.

10. Prentice, R.L. A Generalization of the Probit and Logit Methods for Dose Response Curves. *Biometrics* **1976**, *32*, 761–768, doi:10.2307/2529262.

11. Yuen, B.; Hoang, M.T.; Dong, X.; Lu, T. Universal Activation Function for Machine Learning. *Sci Rep* **2021**, *11*, 18757, doi:10.1038/s41598-021-96723-8.

12. Keshtkar, E.; Kudsk, P.; Mesgaran, M.B. Perspective: Common Errors in Dose–Response Analysis and How to Avoid Them. *Pest Management Science* **2021**, *77*, 2599–2608, doi:10.1002/ps.6268.

13. Cumberland, W.N.; Fong, Y.; Yu, X.; Defawe, O.; Frahm, N.; De Rosa, S. Nonlinear Calibration Model Choice between the Four and Five-Parameter Logistic Models. *J Biopharm Stat* **2015**, *25*, 972–983, doi:10.1080/10543406.2014.920345.

14. Dinse, G.E. An EM Algorithm for Fitting a 4-Parameter Logistic Model to Binary Dose-Response Data. *J Agric Biol Environ Stat* **2011**, *16*, 221–232, doi:10.1007/s13253-010-0045-3.

15. Goutelle, S.; Maurin, M.; Rougier, F.; Barbaut, X.; Bourguignon, L.; Ducher, M.; Maire, P. The Hill Equation: A Review of Its Capabilities in Pharmacological Modelling. *Fundamental & Clinical Pharmacology* **2008**, *22*, 633.

16. Gu, J.; Zhang, X.; Ma, Y.; Li, N.; Luo, F.; Cao, L.; Wang, Z.; Yuan, G.; Chen, L.; Xiao, W.; et al. Quantitative Modeling of Dose–Response and Drug Combination Based on Pathway Network. *Journal of Cheminformatics* **2015**, *7*, 19, doi:10.1186/s13321-015-0066-6.

17. Phanitchat, T.; Ampawong, S.; Yawootti, A.; Denpetkul, T.; Wadmanee, N.; Sompornrattanaphan, M.; Sivakorn, C. Dose-Dependent Blood-Feeding Activity and Ovarian Alterations to PM2.5 in Aedes Aegypti. *Insects* **2021**, *12*, 948, doi:10.3390/insects12100948.

18. Yu, H.; Samuels, D.C.; Zhao, Y.; Guo, Y. Architectures and Accuracy of Artificial Neural Network for Disease Classification from Omics Data. *BMC Genomics* **2019**, *20*, 167, doi:10.1186/s12864-019-5546-z.

19. Yang, S.X. A Neural Network Model for Dose–Response of Foodborne Pathogens. *Applied Soft Computing* **2003**, *3*, 85–96, doi:10.1016/S1568-4946(03)00013-9.
20. Schwab, P.; Linhardt, L.; Bauer, S.; Buhmann, J.M.; Karlen, W. Learning Counterfactual Representations for Estimating Individual Dose-Response Curves. *Proceedings of the AAAI Conference on Artificial Intelligence* **2020**, *34*, 5612–5619, doi:10.1609/aaai.v34i04.6014.
21. Takahashi, A.; Suzuki, T. Bayesian Optimization Design for Dose-Finding Based on Toxicity and Efficacy Outcomes in Phase I/II Clinical Trials. *Pharmaceutical Statistics* **2021**, *20*, 422–439, doi:10.1002/pst.2085.
22. Zang, Y.; Lee, J.J.; Yuan, Y. Adaptive Designs for Identifying Optimal Biological Dose for Molecularly Targeted Agents. *Clin Trials* **2014**, *11*, 319–327, doi:10.1177/1740774514529848.
23. Van Meter, E.M.; Garrett-Mayer, E.; Bandyopadhyay, D. Dose-Finding Clinical Trial Design for Ordinal Toxicity Grades Using the Continuation Ratio Model: An Extension of the Continual Reassessment Method. *Clin Trials* **2012**, *9*, 303–313, doi:10.1177/1740774512443593.
24. Pan, H.; Yuan, Y. A Default Method to Specify Skeletons for Bayesian Model Averaging Continual Reassessment Method for Phase I Clinical Trials. *Stat Med* **2017**, *36*, 266–279, doi:10.1002/sim.6941.
25. Zitzler, E.; Laumanns, M.; Bleuler, S. A Tutorial on Evolutionary Multiobjective Optimization. In Proceedings of the Metaheuristics for Multiobjective Optimisation; Gandibleux, X., Sevaux, M., Sörensen, K., T'kindt, V., Eds.; Springer: Berlin, Heidelberg, 2004; pp. 3–37.
26. Martinez, N.; Bertran, M.; Sapiro, G. Minimax Pareto Fairness: A Multi Objective Perspective. In Proceedings of the Proceedings of the 37th International Conference on Machine Learning; PMLR, November 21 2020; pp. 6755–6764.
27. Li, H.; Shui, Y.; Sun, J.; Zhang, Q. Approximating Pareto Fronts in Evolutionary Multiobjective Optimization with Large Population Size. In Proceedings of the Evolutionary Multi-Criterion Optimization; Ishibuchi, H., Zhang, Q., Cheng, R., Li, K., Li, H., Wang, H., Zhou, A., Eds.; Springer International Publishing: Cham, 2021; pp. 65–76.
28. Emmerich, M.T.M.; Deutz, A.H. A Tutorial on Multiobjective Optimization: Fundamentals and Evolutionary Methods. *Nat Comput* **2018**, *17*, 585–609, doi:10.1007/s11047-018-9685-y.
29. Patterson, S.; Francis, S.; Ireson, M.; Webber, D.; Whitehead, J. A Novel Bayesian Decision Procedure for Early-Phase Dose-Finding Studies. *Journal of Biopharmaceutical Statistics* **1999**, *9*, 583–597, doi:10.1081/BIP-100101197.
30. Murphy, K.P. Bayesian Decision Theory. In *Machine learning: a probabilistic perspective*; Adaptive computation and machine learning series; MIT Press: Cambridge, MA, 2012; pp. p176-189 ISBN 978-0-262-01802-9.
31. O'Quigley, J.; Iasonos, A.; Bornkamp, B. A Bayesian Formulation. In *Handbook of Methods for Designing, Monitoring, and Analyzing Dose-Finding Trials: Handbooks of Modern Statistical Methods*; 2017; pp. p84-85 ISBN 978-1-315-15198-4.
32. Talbi, E.-G. Scalar Approaches. In *Metaheuristics: from design to implementation*; John Wiley & Sons: Hoboken, N.J, 2009; pp. p324-326 ISBN 978-0-470-27858-1.
33. Gibb, R. Optimal Treatment Combination Estimation for Univariate and Multivariate Response Surface Applications, Virginia Commonwealth University, 1998.
34. Shih, M.; Gennings, C.; Chinchilli, V.M.; Carter Jr., W.H. Titrating and Evaluating Multi-Drug Regimens within Subjects. *Statistics in Medicine* **2003**, *22*, 2257–2279, doi:10.1002/sim.1440.
35. Parker, S.M.; Gennings, C. Penalized Locally Optimal Experimental Designs for Nonlinear Models. *Journal of Agricultural, Biological, and Environmental Statistics* **2008**, *13*, 334–354.
36. Chen, H.-W.; Wong, W.K.; Xu, H. Data-Driven Desirability Function to Measure Patients' Disease Progression in a Longitudinal Study. *J Appl Stat* **2016**, *43*, 783–795, doi:10.1080/02664763.2015.1077378.
37. Fedorov, V.V.; Leonov, S.L. Optimality Criteria. In *Optimal Design for Nonlinear Response Models*; CRC Press: Boca Raton, 2013; pp. p52-61 ISBN 978-0-429-10398-8.

38. Biedermann, S.; Dette, H.; Zhu, W. Optimal Designs for Dose–Response Models With Restricted Design Spaces. *Journal of the American Statistical Association* **2006**, *101*, 747–759, doi:10.1198/016214505000001087.

39. Huang, S.-H.; Lo Huang, M.-N.; Lin, C.-W. Optimal Designs for Binary Response Models with Multiple Nonnegative Variables. *Journal of Statistical Planning and Inference* **2020**, *206*, 75–83, doi:10.1016/j.jspi.2019.09.006.

40. Khan, M.K.; Yazdi, A.A. On D-Optimal Designs for Binary Data. *Journal of Statistical Planning and Inference* **1988**, *18*, 83–91, doi:10.1016/0378-3758(88)90018-3.

41. Shi, Y.; Zhang, Z.; Wong, W.K. Particle Swarm Based Algorithms for Finding Locally and Bayesian D-Optimal Designs. *Journal of Statistical Distributions and Applications* **2019**, *6*, 3, doi:10.1186/s40488-019-0092-4.

42. Silvey, S. Carathéodory's Theorem. In *Optimal Design: An Introduction to the Theory for Parameter Estimation*; Springer Netherlands, 1980; p. p72 ISBN 978-0-412-22910-7.

43. Hyun, S.W.; Yang, M.; Flournoy, N. A Procedure for Finding an Improved Upper Bound on the Number of Optimal Design Points. *Computational Statistics & Data Analysis* **2013**, *58*, 276–282, doi:10.1016/j.csda.2012.08.012.

44. Yang, M. On the de La Garza Phenomenon. *The Annals of Statistics* **2010**, *38*, 2499–2524, doi:10.1214/09-AOS787.

45. Fedorov, V.V.; Leonov, S.L. Dose-Finding for Efficacy-Toxicity Response. In *Optimal Design for Nonlinear Response Models*; CRC Press: Boca Raton, 2013; p. p161 ISBN 978-0-429-10398-8.

46. Fedorov, V.V.; Leonov, S.L. Model Discrimination. In *Optimal Design for Nonlinear Response Models*; CRC Press: Boca Raton, 2013; p. p141 ISBN 978-0-429-10398-8.

47. Aouni, J.; Bacro, J.N.; Toulemonde, G.; Colin, P.; Darchy, L.; Sebastien, B. Design Optimization for Dose-Finding Trials: A Review. *J Biopharm Stat* **2020**, *30*, 662–673, doi:10.1080/10543406.2020.1730874.

48. Zhu, F.-C.; Li, Y.-H.; Guan, X.-H.; Hou, L.-H.; Wang, W.-J.; Li, J.-X.; Wu, S.-P.; Wang, B.-S.; Wang, Z.; Wang, L.; et al. Safety, Tolerability, and Immunogenicity of a Recombinant Adenovirus Type-5 Vectored COVID-19 Vaccine: A Dose-Escalation, Open-Label, Non-Randomised, First-in-Human Trial. *The Lancet* **2020**, *395*, 1845–1854, doi:10.1016/S0140-6736(20)31208-3.