2007-08-06

An Adaptive Bayesian Approach to Bernoulli-Response Clinical Trials

Andrew W. Stacey
Brigham Young University - Provo

AN ADAPTIVE BAYESIAN APPROACH TO BERNOULLI-RESPONSE CLINICAL TRIALS

by

Andrew W. Stacey

A Thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

Department of Statistics
Brigham Young University
December 2007

BRIGHAM YOUNG UNIVERSITY

GRADUATE COMMITTEE APPROVAL

of a Thesis submitted by

Andrew W. Stacey

This Thesis has been read by each member of the following graduate committee and by majority vote has been found to be satisfactory.

Date	C. Shane Reese, Chair
Date	Scott D Grimshaw
Date	Gilbert W. Fellingham

BRIGHAM YOUNG UNIVERSITY

As chair of the candidate's graduate committee, I have read the Thesis of Andrew W. Stacey in its final form and have found that (1) its format, citations, and bibliographical style are consistent and acceptable and fulfill university and department style requirements; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the graduate committee and is ready for submission to the university library.

Date

C. Shane Reese
Chair, Graduate Committee

Accepted for the Department

Scott D Grimshaw
Graduate Coordinator

Accepted for the College

Thomas W. Sederberg
Associate Dean, College of Physical and Mathematical Sciences

ABSTRACT

AN ADAPTIVE BAYESIAN APPROACH
TO BERNOULLI-RESPONSE CLINICAL TRIALS

Andrew W. Stacey
Department of Statistics
Master of Science

Traditional clinical trials have been inefficient in their methods of dose finding and dose allocation. In this paper a four-parameter logistic equation is used to model the outcome of Bernoulli-response clinical trials. A Bayesian adaptive design is used to fit the logistic equation to the dose-response curve of Phase II and Phase III clinical trials. Because of inherent restrictions in the logistic model, symmetric candidate densities cannot be used, thereby creating asymmetric jumping rules inside the Markov chain Monte Carlo algorithm. An order restricted Metropolis-Hastings algorithm is implemented to account for these limitations.

Modeling clinical trials in a Bayesian framework allows the experiment to be adaptive. In this adaptive design batches of subjects are assigned to doses based on the posterior probability of success for each dose, thereby increasing the probability of receiving advantageous doses. Good posterior fitting is demonstrated for typical dose-response curves and the Bayesian design is shown to properly stop drug trials for clinical futility or clinical success. In this paper we demonstrate that an adaptive Bayesian approach to dose-response studies increases both the statistical and medicinal effectiveness of clinical research.

ACKNOWLEDGEMENTS

Most importantly, I would like to thank my wife, Lauren, and my parents for their continuing support. I would also like to thank the BYU Department of Statistics for being willing to work with a graduate-student-athlete. Notwithstanding odd hours, interviews across the country, and frequent weekend commitments taking me away from campus I always found the department to be remarkably flexible.

CONTENTS

CHAPTER

1 Introduction 1

2 Literature Review 4

2.1 Dose-Response Experiments 4

2.1.1 Continuous Versus Binary Response 4

2.1.2 Experimentation 5

2.2 Logistic Models 6

2.2.1 Breadth of Logistic Models 7

2.2.2 The Four-Parameter Model 8

2.2.3 Interpretation of the Model 9

2.3 Bayesian Approach to Clinical Trials 10

2.3.1 Why a Bayesian Approach? 10

2.3.2 History 10

2.3.3 Two Schools of Bayesian Thought 11

2.3.4 Methods of Simulation 12

2.4 Adaptive Clinical Trial Design 13

3 Methods 16

3.1 The Model 16

3.1.1 Parameter Interpretability 16

3.1.2 Model Restrictions 19

3.2 Likelihood and Priors 19

3.2.1 Derivation of the Likelihood 19

3.2.2 Prior Specification 21

3.3 Simulation Techniques 22
3.3.1 Complete Conditionals 22
3.3.2 Candidate Densities 23
3.4 Asymmetric Jumping Rule 25
3.4.1 Implications of Using Truncated Normal Density 25
3.4.2 Acceptance Probabilities 27
3.4.3 Application in Simulation 28
3.5 Adaptive Trial Design 29
3.5.1 Purpose of an Adaptive Trial 29
3.5.2 Methods of Calculation ED95 30
3.5.3 Design: Allocation by Batch 31
3.6 Computational Implementation 32
3.6.1 Starting Values 32
3.6.2 Logistic Regression and ED95 Threshold 33
3.6.3 ED95 Probabilities 34
3.6.4 Stopping Rules 34
3.6.5 Reasons for Failure 35
3.6.6 Dose Allocation 38
3.7 Model Assessment, Operating Characteristics of Bayesian Adaptive Design 39
3.8 Situations under Consideration and Simulation 40

4 Results 44
4.1 Success and Failure Probabilities 44
4.2 Examining One Simulated Clinical Trial 46
4.3 Examining Multiple Simulations for each Situation 51
4.3.1 Data from Multiple Clinical Simulations 51
4.3.2 Interpretation of Results 51

5 Summary and Conclusions 54

BIBLIOGRAPHY 55

APPENDIX

A Appendix 61

A.1 Case 1: Data from Simulated Clinical Trial of Situation "A" 61

A.2 Model Variations, Maximizing Success Probabilities 63

A.3 R Code 65

A.4 Tables 73

TABLES

Table

3.1 Summary of parameter values used in each of the 7 clinical situations tested in this study. 43

4.1 Type I error rate and cap realization probability of 5 different success probabilities 45

4.2 Operating characteristics of Bayesian adaptive approach to clinical trials. 51

5.1 Average dose allocation and success rates for a simulated clinical trial of type "A" ($n \approx 590$) 55

5.2 Dose allocation and success rates for a generic, traditional clinical trial of type "A" with 590 subjects 56

A.1 Case 1: Simulated data and results for each dose in a simulated Situation "A" clinical trial 62

A.2 Alterations to Bayesian stopping rules: 3 ways to maximize success rates and minimize cap realization rates 63

A.3 Situation "A": Mean number of subjects at each dose and successes at each dose for one simulated clinical trial 74

A.4 Situation "B": Mean number of subjects at each dose and successes at each dose for one simulated clinical trial 75

A.5 Situation "C": Mean number of subjects at each dose and successes at each dose for one simulated clinical trial 76

A.6 Situation "D": Mean number of subjects at each dose and successes at each dose for one simulated clinical trial 77

A.7 Situation "E": Mean number of subjects at each dose and successes at each dose for one simulated clinical trial 78

A.8 Situation "F": Mean number of subjects at each dose and successes at each dose for one simulated clinical trial 79
A.9 Situation "G": Mean number of subjects at each dose and successes at each dose for one simulated clinical trial 80

FIGURES

Figure

1.1 Ordinary Dose-Response Curve 1

3.1 Four Parameters 18

3.2 Truncated Normal Densities 26

3.3 Situation 1 36

3.4 Situation 2 37

3.5 Situation 3 37

3.6 Basic Sigmoid Curve 41

3.7 Horizontally Shifted Curves 41

3.8 Underrepresented Curves 42

3.9 Control Curves. Curve 'G' has no true ED95 dose 42

4.1 Type I error rates of 5 different success probabilities 45

4.2 Final response values of a complete clinical trial simulation of type "A" 47

4.3 Final response values of a complete clinical trial simulation of type "B" 47

4.4 Final response values of a complete clinical trial simulation of type "C" 48

4.5 Final response values of a complete clinical trial simulation of type "D" 48

4.6 Final response values of a complete clinical trial simulation of type "E" 49

4.7 Final response values of a complete clinical trial simulation of type "F" 49

4.8 Final response values of a complete clinical trial simulation of type "G" 50

A.1 Case 1: Posterior fit to theoretical response values 61

1. INTRODUCTION

Inherent in a traditional clinical trial is the necessity to randomly allocate an equal number of subjects to each dose in question. While an experiment of this design will provide adequate power to reject a null hypothesis, it may not be the most efficient or medically beneficial way to conduct clinical research. In order to understand why the traditional methodology is inefficient it is important to understand the characteristics of an ordinary dose-response curve.

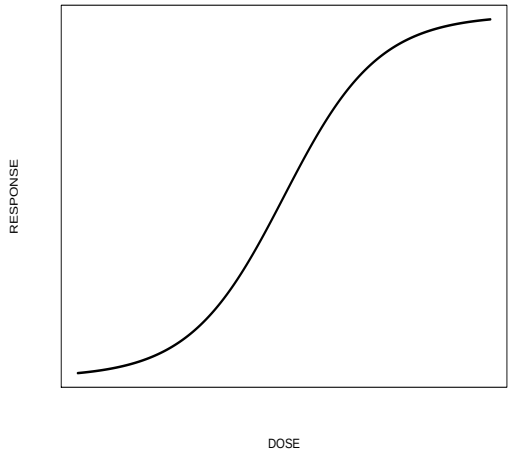

Figure 1.1: Ordinary Dose-Response Curve

In most clinical studies there is a positive, nonlinear relationship between dose and response. Figure 1.1 demonstrates this relationship as a sigmoid with a shallow slope at the ends of the dose range and a steep curve in the middle.

The dose-response relationship is well understood but often ignored in clinical studies because the dose-response curve for a drug is never known until after hundreds of subjects have been tested. In order for the curve to be approximated, subjects are allocated to a broad range of doses. Every part of the dose-response curve must be tested in order to understand its characteristics and dynamics. Consequently, traditional clinical trials only provide good medicine to a fraction of the patients

involved.

In this paper a four-parameter logistic equation is proposed to model the dose-response relationship of a drug during clinical trials. Researchers will then have the ability to assign a large percentage of subjects to beneficial doses on the dose-response curve. Logistic models have been used in biological studies since 1941. They are praised for their ease of use and straightforward interpretability. The four-parameter logistic equation used in this paper,

$$\beta + \frac{\delta}{1 + e^{\frac{\theta - x_i}{\tau}}}, \tag{1.1}$$

contains parameters that can be easily interpreted and understood as characteristics of the dose-response curve. The parameters, β, δ, θ, and τ each control a different dynamic of the sigmoid and can be used to provide good posterior fit in most dose-response studies.

Markov chain Monte Carlo (MCMC) techniques are used to estimate the values of the four parameters in our model. Inherent conditions in the model prevent the MCMC process from being carried out in the conventional manner. To adjust for these conditions, the Metropolis-Hastings algorithm is implemented with an asymmetric jumping rule. The simulation methodology allows posterior distributions to be fluid and shift appropriately to describe most dose-response curves.

The Bayesian clinical design produces a drug trial that is completely adaptive. Rather than lumping all subjects into one clinical trial, subjects are divided into batches of 20 individuals. After the data from the first batch are recorded, the Metropolis-Hastings algorithm is used to approximate the probability of success at each dose. We can then allocate the next 20 subjects to the doses that demonstrate the best success rates. The process of simulation, patient allocation, and data recording is repeated until the trial is stopped due to clinical success or clinical futility. The goal of the Bayesian adaptive design is to quickly identify and terminate clinically

futile dose-response studies.

This paper demonstrates how the four-parameter logistic model and an adaptive Bayesian design provide more precision and more success to clinical research. Implementing these methods in clinical trials can increase the proportion of successful treatments during clinical trials and can decrease the time and resources necessary to make statistical inference about a dose-response curve.

2. LITERATURE REVIEW

2.1 Dose-Response Experiments

Dose-response clinical research is a long-standing and well understood field. The dose-response curve is well documented and is used in a wide variety of research areas. There have been different approaches attempting to interpret and analyze the dose-response relationship. In 1941, DeBeer (1941) proposed a method of graphical analysis and possible calculation of these relationships. Today, we use software and graphical tools to understand dose-response curves, but the ideas behind the characteristics of each graph are the same as those DeBeer studied.

A general overview of modern statistical methods used in dose-response research is given by Kodell and Chen (1991). A broad range of topics are promoted in their paper. This review will only discuss those arguments that specifically relate to this model. An important debate in *current* dose-response research is not how to understand the data, as in DeBeer's day, but how to model it. The first question an experimenter needs to ask when modeling a clinical study is whether the response will be continuous or quantal in nature. This turns out to be an important question because it can change the entire makeup of a study.

2.1.1 Continuous Versus Binary Response

Sand and Rosen (2003) state that if a study only deals with the presence or absence of a particular response, it should be modeled in a binary manner. Most other forms of data will be continuous; that is, there will be varying levels of response. There is, however, a natural desire to model response as binary data because continuous data are more difficult to model. In fact, there is an abundance of research currently devoted to creating workable models for continuous data. Crump (2002) discussed

the best methods to use while conducting dose-response studies with continuous data. Interestingly, one of the examples in this study is a method of converting a continuous response into binary data. This is done by setting a threshold on the response; if the subject is beyond the threshold he or she is considered part of the study, if below the threshold the subject is not considered.

The gains from using binary data are numerous. Modeling and interpretation in binary dose-response studies is straightforward. In addition, with binary data the response can be represented as a probability. This is a feature that makes both prior specification and posterior interpretation more clear-cut.

Sometimes an optimal dose is better explored using two variables. Whitehead et al. (2004) conducted a series of experiments using models similar to those used in this paper but collected two binary responses for each subject. Multiple variables, when used correctly, can provide more information than univariate models. However, multivariate modeling is not relevant to all studies. In addition, the analysis of multivariate models may become intense and difficult to explain. Karen Han's work (Han et al., 2004) offers some interesting insights into binary, multivariate analysis in biological assays.

This paper focuses on models that only relate to one-variable, binary, dose-response studies. It is believed that these data can be both better represented and better understood than multivariate or continuous models. The equation used in this study, however, can be used to model continuous data if the guidelines set out by Crump (2002) are followed. In addition, following the format proposed by Whitehead (2004), these methods could also be used in bivariate studies.

2.1.2 Experimentation

Perhaps the most important question in modern dose-response studies is how to carry out the experiment. Traditional studies use a cohort of numerous subjects and

a broad range of doses. In Phase I clinical studies the dose given to the patients is escalated until the response either levels off or begins to decrease. The ineffectiveness of these studies is widely known but little has been done to change their methods. Specific improvements to these trials are proposed by Whitehead et al. (2006). Their work focuses on proper dose escalation techniques and stopping rules. Another solution has been proposed by Kalish (1990). This paper focuses on using tolerance experimentation to more quickly find effective drugs and doses. While Kalish also attempts to resolve the problem of inefficient dose finding, the methods in this paper will rely more heavily on simulation and Bayesian prior distributions. These methods are discussed in the Bayesian methods portion of this chapter.

One of the most important issues with the ineffectiveness of traditional dose-response clinical studies is the well-being of the subject. In the traditional method described above, many patients will not be assigned to a dose that will benefit them; only a small proportion of subjects will actually receive a beneficial dose.

It has been suggested (Simon, 1999) that generic dose-response studies are unethical, especially in the case of patients who have life threatening illnesses. The suggested solution to this issue is called "Active Control." In such a study, experimenters test a new treatment by only comparing it to another treatment that has already been proven effective. This answers the question of whether the person receives a good drug, but does not answer the issue of whether a person receives a good dose. There is not much literature exploring beneficial dose allocation in clinical studies.

2.2 Logistic Models

The previous section suggests that an important step in designing better dose-response studies is a better understanding of the underlying characteristics of dose-response data. The current understanding of dose-response data has lead to a number

of proposed models to represent the data. Whitehead et al. (2004), as well as many current researchers and most of Whitehead's collaborators, use logistic models to represent these types of data. While their approach is successful, it is not new.

2.2.1 Breadth of Logistic Models

Logistic models have been used in biological studies for almost one hundred years; one of the benchmark papers was published by C. W. Emmens (1941). He proposed a variety of three- and four-parameter logistic models that could be used to represent bioassay data from the pituitary gland. Though he was not the first to use these models in the literature, he began using such models in specific biological studies.

Today, logistic models and corresponding logistic regression are widely used tools in binary data analysis. The breadth of the current research using logistic models in biological and dose-response studies is enormous. While this paper will focus more on the Bayesian side of the issue, there is not a lack of Frequentist literature on logistic dose-response research. Some recent applications include a modified χ^2 analysis using logistic models (Tang, 2000), a method of splining the dose-response curve to obtain better fit (Li and Hunt, 2004), and the use of logistic models to build better confidence intervals for estimating effective doses (Huang et al., 2002).

The breadth of these applications is partially due to the abundance of functional logistic models. In the early part of the nineteenth century, Pierre Francois Verhulst proposed the first logistic model for use in studying human populations (Verhulst, 1838). His original equation, where m_0 is the Malthusian parameter (maximum growth rate) and k is the maximum capacity parameter, is written as follows:

$$m = m_0(1 - \frac{n}{k}). \tag{2.1}$$

While the theory behind this model is still used, it has been modified so that it

can be applied to many different fields of study. As is seen in Verhulst's early equation, one of the benefits of using logistic models is the ease of parameter interpretation. This interpretability has lead to a plethora of logistic parameterizations.

2.2.2 The Four-Parameter Model

Aage Volund (1978) showed the relationship between a number of different logistic models used in biological studies. He proposed the use of a four-parameter model to decide whether a slope ratio or parallel line method should be used to interpret biological data. The use of the four-parameter logistic model is not new, but has exploded into the literature in recent years. In a dose-escalation paper by Whitehead et al. (2006), a clue is offered as to why this model has become so widely used. They state that the four parameters allowed an informative evaluation of dose ranges while at the same time minimizing the consequences from using too many parameters in the model.

The four-parameter model optimizes the amount of information that can be inferred from an equation. In the case of dose-response studies, fewer parameters would provide less information, and more parameters are only needed if the data do not conform to hypotheses of symmetry (Streibig and Kudsk, 1993). However, even when the four-parameter model is chosen, there still needs to be a decision of parameterization. Because of the model's flexibility, there are innumerable ways to parameterize the four-parameter logistic model. The one used in this paper was chosen because of its interpretability and ease of implementation. The model is similar to many models in the literature, but has not been used explicitly in dose-response research. A more in-depth analysis of the parameterization choice will follow in the next section.

Ratkowsky and Reedy have studied which of the infinite number of four-parameter models should be used in research (Ratkowsky and Reedy, 1986). In their paper they

directly study six of the most commonly used parameterizations. The parameterization used in this thesis is not directly found in their article, but is one algebraic step away from one of their parameterizations. Interestingly, this model was found to be stable and the authors suggested that it can be very useful in biological studies.

2.2.3 Interpretation of the Model

As mentioned earlier, one of the most convincing reasons to incorporate logistic models into biological studies is their inherent interpretability. Most of the literature already cited contains an explanation of the specific interpretation for each model. In order to understand how one model can be applied to so many different fields, it is important to understand the generic interpretations. In Streibig's 1993 book on bioassays (Streibig and Kudsk, 1993), he explains the interpretation of each parameter in detail. Because the interpretations used in this study are similar to his, parameter explanations are limited to those used in our model. Definitions from both Streibig and Ratkowsky will be used to explain each variable. These explanations can be found in the methods section of this thesis, Chapter 3.

$$Y_i = \beta + \frac{\delta}{1 + e^{\frac{\theta - x_i}{\tau}}} \tag{2.2}$$

In equation 2.2, Y_i represents the proportion of successes and x_i represents the dose for all possible doses in the study. The parameter, β, represents the minimum possible value of efficacy. The maximum efficacy value is represented by $\beta + \delta$. Therefore, δ is the possible change from the minimum to the maximum efficacy. From this interpretation, two things are evident. First, when $\beta = 0$, δ alone is the maximum possible efficacy. Second, because β is the minimum efficacy we know that $\beta \leq \delta$. This relationship, which will become very interesting during MCMC simulation, is discussed further in the section on Bayesian methods.

The other parameters can be interpreted as well. The parameter θ is understood in many papers as a halfway point between the minimum and maximum parameters (Streibig and Kudsk, 1993). It can also be thought of as the point of inflection in the dose-response curve. For our purposes, we understand θ as the value of the i^{th} dose, (x_i), that achieves 50% of the change from the minimum dose (β) to the maximum dose $(\beta + \delta)$. The parameter τ is a scaling factor that can change the characteristics of the dose-response curve.

2.3 Bayesian Approach to Clinical Trials

2.3.1 Why a Bayesian Approach?

It must be understood that the Bayesian paradigm includes a few things that other statistical approaches do not include. Firstly, the incorporation of prior information is important in dose-response studies because as more is known about the drug, the prior information can become more and more precise. Secondly, a point estimate of dose efficacy is not logical; there is too much inherent variation in modern drugs to report a point estimate. A probability distribution would be a more reasonable summary of a dose's effectiveness. Thirdly, the variation from drugs and subjects can be more adequately summarized using distributions on an interpretable model such as the logistic model. Lastly, Bayesian modeling provides the possibility of simulation that can be used to fully understand the distribution of the response (Dodds and Vicini, 2004); (Whitehead et al., 2006).

2.3.2 History

The Bayesian approach to experimental design is a relatively new field. In 1995 Chaloner and Verdinelli published a Bayesian stepwise design process that could be applied to multiple fields of study (Chaloner and Verdinelli, 1995). This was an

attempt to convince the experimental world of the benefits of Bayesian methods in research.

For years, researchers have attempted to find the most effective way of analyzing dose-response data. Although a Bayesian approach to this problem is relatively new, the literature is full of evidence that it is highly preferred to other methods. D. A. Jones (1996) was the first proponent of Bayesian modeling in clinical dose-response studies. Since then there has been an explosion of such models in the medical literature. One widely accepted benefit of Bayesian modeling in clinical studies is the insight it gives into dose-efficient experimentation. O'Hagan and Stevens (2002) published a review of the breakthrough Bayesian models in medical experimentation. Although their goal was to show how Bayesian methods can save money in experimentation, the research they cite also reveals how these methods can be used to improve modeling, interpretation, and experimentation itself. The advantages of Bayesian methods mentioned by O'Hagan and Stevens include intuitive and meaningful inferences, powerful computational tools, and the use of genuine prior information.

2.3.3 Two Schools of Bayesian Thought

There are two main schools of thought in Bayesian experimental design (Mukhopadhyay, 2000). In the first, the nonparametric approach, a prior distribution is used to represent the full space of the response variable. The second recognizes a trend in the relationship between response and dose. This relationship can then be modeled in terms of explanatory equations, like the four-parameter logistic equations. The parameters of these equations are represented by prior distributions and are used to predict a posterior distribution. This paper will follow the standards of the second methodology. While most of the literature recommends the use of the second approach, Mukhopadhyay suggests that a parametric approach may be too sensitive for dose-response studies (Mukhopadhyay, 2000). This may be a valid concern, but we

believe the benefits of the model used in this paper, such as interpretation and ease of simulation, far outweigh the possibiility of over-sensitivity.

The parameter-based Bayesian approach to dose-response modeling is widely used in dose-response studies. The earliest such methods involved Dirichlet prior distributions (Ramsey, 1972), but were not very informative. Sun and Tsutakawa researched the inherent variability that arises when prior knowledge is used in dose-response studies (Sun and Tsutakawa, 1997). In 1998 the Bayesian bootstrap was applied to dose-response bioassays (Bowman, 1998). While these results are still used, most current dose-response research relies heavily computer simulation, something that was not the focus of earlier papers.

2.3.4 Methods of Simulation

Bayesian statistics is the practice of finding the posterior distribution of a parameter. An important tool in accomplishing this task is the Markov chain Monte Carlo (MCMC) technique used in computer simulation. This type of simulation gives the researcher the ability to find the posterior distribution by finding the convergence point of a long Markov chain. These techniques have been shown to be useful in recent dose-response studies (Dodds and Vicini, 2004). One specific form of MCMC uses the Metropolis-Hastings algorithm published in 1970 by W. K. Hastings (Hastings, 1970). Hastings refers to two characteristics that will be very important to this thesis. First, MCMC can draw samples from any probability distribution. This is imperative when the distribution in question cannot be written in closed form. Second, the Metropolis-Hastings algorithm, unlike its predecessor, the Metropolis algorithm, does not require a symmetric jumping rule. In other words, given state a and state b; the probability of going from state a to state b does not have to equal the probability of going from state b to state a.

$$P(a, b) \neq P(b, a)$$

Because the jumping rule in this paper is asymmetrical, this feature will prove to be invaluable. Englehardt and Swartout (2006) have recently published research on a dose-response model in which the Metropolis-Hastings algorithm was lightly used, but the literature relating the Metropolis-Hastings algorithm to dose-response data is relatively sparse.

The four-parameter logistic model proposed in this paper contains restrictions that will make the MCMC process interesting. As stated earlier, we know that $\beta \leq \delta$ and $\beta, \delta > 0$. Because the Metropolis-Hastings algorithm uses random draws from a candidate distribution, this inequality will make a traditional distribution to be very inefficient. To overcome this inefficiency the use of a truncated normal distribution is used as the candidate distribution for β and δ in our Markov chain. Random truncated normal generation can be done in a number of different ways. The documentation for the random truncated normal generator can be found in the R documentation (R Development Core Team, 2003).

2.4 Adaptive Clinical Trial Design

A number of advantages have already been demonstrated when using a Bayesian approach and a logistic model in dose-response studies. The most convincing argument for the use of this model, however, lies in its ability to change and adapt to new information. The motivating force behind the model presented in this thesis is that it allows a researcher to create a clinically sound adaptive trial.

Thall and Russell (Thall and Russell, 1998) attempted one of the first Bayesian adaptive trial designs. The goals listed in their paper are now a standard for adaptive Bayesian trial design. These goals are:

> (1) To find a dose of a new agent that satisfies specific safety and efficacy requirements, (2) to stop the trial early if it is likely that no dose is both safe and efficacious, and otherwise (3) to treat as

many subjects as possible at the optimal doses, so participants receive maximal benefit.

These three goals have been the aim of many recent clinical studies (Biswas et al., 2006); (Chang and Chow, 2005); (Cheng and Shen, 2006). Each of these papers, published in the last year, establishes new techniques that can aid in achieving the goals listed by Thall and Russel.

In this thesis, an adaptive model will utilize logistic models and Bayesian design in order to find an optimal dose. Unlike other designs that use the proportion odds model (Thall and Russell, 1998) or the continuation-ration model (Zhang et al., 2005), the model in this study will apply prior distributions to a simple binomial response variable. In this way, the design can be straighforward and can investigate each dose individually while applying the same prior information to all doses. This is possible using logistic regression methods and the Metropolis-Hastings algorithm previously mentioned.

In this way, every time data are collected, the priors can be updated and the model can be used to predict the response for each dose in the trial. Then, as in the trials cited above, the subjects can be assigned to those doses that have simulated the best responses. The trial will continue until an optimal dose is agreed upon.

The decision of when a trial should stop is an interesting one. The Bayesian adaptive design will return probability distributions on the response to each dose in a trial. It will not, however, tell a researcher when a dose is found to be optimal. This is done at the discretion of the researcher. Many papers have been written about the best stopping rules. Some interesting ideas are recommended by Zhou and Whitehead. One idea is to "never use a dose for which the probability of toxicity exceeds η, for some value $\eta < \gamma$." Zhou and Whitehead also mention that very pessimistic prior distributions can be used to overestimate the harmful effects of a drug. This may slow down the process of dose-finding, but it will ensure that the harmful effects of a

drug are never overlooked (Zhou and Whitehead, 2003).

The Bayesian design used in this thesis effectively answers each of the points made by Thall and Russel. In particular, the model focuses on ensuring that subjects will be given doses that have the greatest probability of success. The stopping rules and mechanics of this study will be discussed in the methods chapter of this paper.

3. METHODS

3.1 The Model

This study uses a fully interpretable, four-parameter logistic model as a means of representing Bernoulli-response clinical trial data. The model is not well known, but it has many strengths: its mathematical stability (Ratkowsky and Reedy, 1986), breadth of explanatory mechanisms, and ease of interpretation make it an excellent choice for dose-response data. The logistic model,

$$\beta + \frac{\delta}{1 + e^{\frac{\theta - x_i}{\tau}}}, \tag{3.1}$$

contains four parameters $(\beta, \delta, \theta, \tau)$, which are constant for all possible doses in a trial and one variable (x_i) that indicates which dose is being considered. Choosing this type of model facilitates the use of logistic regression to analyze each dose in a trial. By varying only the dose indication variable, drug performance can be properly compared across dose ranges.

3.1.1 Parameter Interpretability

The four parameters in our model are interpretable and well documented in the literature. Parameter interpretability is essential when applying mathematical models to biological systems; without it, models would be meaningless and it would be impossible to make biological conclusions about the data.

Because our model has not been used before in a clinical trial setting, our parameter interpretations are derivations of other biological interpretations, especially those used by Ratkowsky and Reedy (1986). Converting our model into a simple equation, we have

$$Y_i = \beta + \frac{\delta}{1 + e^{\frac{\theta - x_i}{\tau}}}, \tag{3.2}$$

where Y_i represents the probability that the i^{th} dose, denoted as x_i, will be successful. In the model, β represents the minimum response to a given dose, reported as a proportion. The function $\beta + \delta$ represents the maximum response, which means that δ represents the range of the response. Or, if $\beta = 0$, then δ would represent the maximum possible response to the drug dose. These explanations reveal an interesting relationship between β and δ, as they relate to clinical trials.

The parameter θ can be thought of as the point of inflection on the dose response curve; if θ is increased, the curve's concavity will change later in the dose range and vice versa. In addition, θ can be thought of as the dose at which 50% of the change from minimum to maximum dose is achieved; if the curve is steep early in the dose range and then levels off, the value of θ will be relatively small. τ changes the peak dynamics of the curve; as τ increases the curve is less peaked and begins to have very sharp angles.

A graphical representation may aid in understanding the influence of each of these parameters. In the following graphs (Figure 3.1), the logistic model is presented in black using constant but generic parameter estimates. The influence of each parameter on the dose-response curve is shown by either increasing or decreasing the value of the parameter in question and plotting those lines on the same graph as the model with constant parameters.

The change in parameter values shown in these graphs is never more than 20% of the original value, demonstrating how distinct and important each parameter is to the logistic model. The characteristics of these variables allow the model to adapt and mimic most dose response curves in medicine.

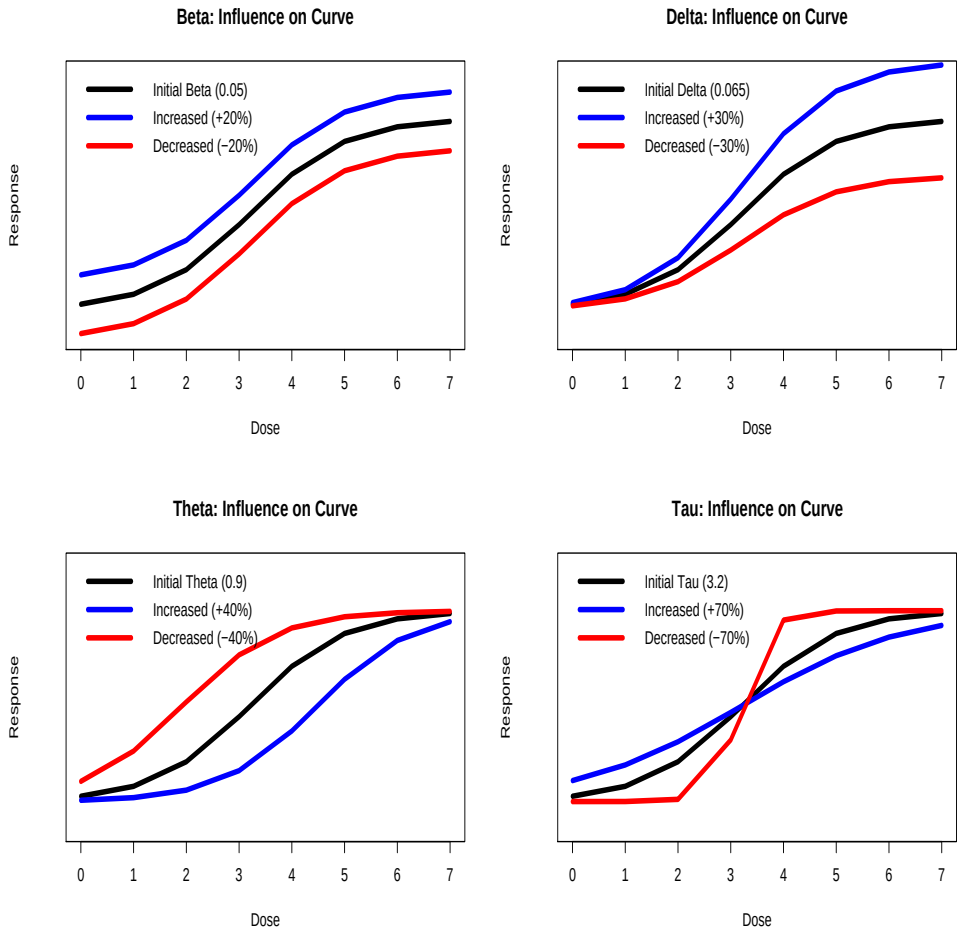

Figure 3.1: Four Parameters

3.1.2 Model Restrictions

In this study we are interested in later-phase (II and III) clinical trials in which initial success rates are low and drug intervention has a positive effect. If drugs are not effective or minimally effective, the issue is not statistical, but pharmaceutical and the drug effect should be explained to the drug company. Under these assumptions, all clinical trials will follow the restriction that $\beta \leq \delta$. In other words, the effect of a drug is larger than the effect of a placebo. This restriction can be thought of as an indicator function placed at the end of the original logistic equation,

$$\left(\beta + \frac{\delta}{1 + e^{\frac{\theta - x_i}{\tau}}}\right) * I(\beta \leq \delta).$$

3.2 Likelihood and Priors

For notational purposes, it is important to note that in a Bayesian computational setting we represent a posterior density as

$$\Pi(\Theta|y) = \frac{f(y|\Theta)p(\Theta)}{\int f(y|\Theta)p(\Theta)d\Theta},$$

where y represents the data and Θ represents the parameters of the model, in this case $\beta, \delta, \theta, \tau$. $\Pi(\cdot)$ represents the posterior distribution, $f(\cdot)$ represents the likelihood function, and $p(\cdot)$ represents the prior distribution. This notation will be used throughout the remainder of the text.

3.2.1 Derivation of the Likelihood

A typical Bernoulli-response clinical trial will allocate n number of subjects to drug dose x_i. Of these subjects, there will be m number of successes. Therefore, a Bernoulli trial can be represented by a binomial variable, Λ, where

$$\Lambda \sim \text{BIN}(n, \frac{m}{n}).$$

The response variable in the logistic equation, Y_i, is the probability that a drug dose, x_i, is successful. Because this study models Bernoulli clinical trials, Λ can be written as a function of the response variable,

$$\Lambda \sim \text{BIN}(n, Y_i).$$

The four-parameter logistic equation can now be substituted in the place of Y_i in the equation for Λ. Because the logistic model represents the proportion of successes for a particular dose, subscripts must be placed on all the variables to denote that the equation for Λ_i only refers to the i^{th} dose:

$$\Lambda_i \sim \text{BIN}(n_i, \left(\beta + \frac{\delta}{1 + e^{\frac{\theta - x_i}{\tau}}}\right)).$$

This function can also be represented by the corresponding binomial density function, demonstrated below. Additionally, it was mentioned that the likelihood under a Bayesian framework can be represented by $f(\text{Data}|\text{Parameters})$, or $f(y|\Theta)$. This notation will be substituted in as follows:

$$f(y|\Theta) = \binom{n_i}{m_i} \left(\beta + \frac{\delta}{1 + e^{\frac{\theta - x_i}{\tau}}}\right)^{m_i} \left(1 - \beta - \frac{\delta}{1 + e^{\frac{\theta - x_i}{\tau}}}\right)^{(n_i - m_i)}.$$

The MCMC algorithm that we will use is invariant to proportionality constants in the likelihood. Thus, it is possible to drop those expressions that are not functions of the four parameters in the model. The density can then be written as a proportional,

$$f(y|\Theta) \propto \left(\beta + \frac{\delta}{1 + e^{\frac{\theta - x_i}{\tau}}}\right)^{m_i} \left(1 - \beta - \frac{\delta}{1 + e^{\frac{\theta - x_i}{\tau}}}\right)^{(n_i - m_i)}.$$

The likelihood is calculated from the binomial density, where r is the total number of doses,

$$L(y|\Theta) \propto \prod_{i=1}^{r} \left\{ \left(\beta + \frac{\delta}{1 + e^{\frac{\theta - x_i}{\tau}}}\right)^{m_i} \left(1 - \beta - \frac{\delta}{1 + e^{\frac{\theta - x_i}{\tau}}}\right)^{(n_i - m_i)} \right\},$$

And the log-likelihood follows simply,

$$\log(L(y|\Theta)) \propto \sum_{i=1}^{r} \left(m_i \log \left(\beta + \frac{\delta}{1 + e^{\frac{\theta - x_i}{\tau}}} \right) + (n_i - m_i) \log \left(1 - \beta - \frac{\delta}{1 + e^{\frac{\theta - x_i}{\tau}}} \right) \right). \tag{3.3}$$

The log likelihood function contains three variables whose values will be known through experimentation: x_i represents each individual dose, m_i represents the number of successes for each dose, and n_i represents the total number of subjects for each dose. The other four parameters will be estimated using prior distributions and calculated using MCMC techniques. Prior distributions for the four parameters in the model are chosen based on each parameter's inherent interpretation from the model.

3.2.2 Prior Specification

It has been stated that β is the minimum success rate of a Bernoulli-response clinical trial and that δ is the change in minimum to maximum response. Therefore, it can be deduced that both β and δ must be between 0 and 1. This relationship indicates that both β and δ are distributed as Beta random variables.

The parameter θ can be any value between the minimum and maximum dose range. Therefore, θ is symmetrical around the values of β and $\beta + \delta$ and can easily be represented by a Normal prior distribution because of its symmetric characteristics. The parameter τ is a scaling factor that must be positive in order for the function to be increasing. A negative value of τ describes a decreasing sigmoid, which is not a characteristic of dose-response studies. To ensure an increasing sigmoid function, τ is represented by a Gamma prior distribution.

In summary, the prior distributions for our model and their notation are,

$$\begin{aligned} \beta &\sim Beta(a_\beta, b_\beta), \\ \delta &\sim Beta(a_\delta, b_\delta), \\ \theta &\sim Normal(\mu, \sigma^2), \end{aligned}$$

$$\tau \sim Gamma(\alpha, \kappa).$$

In the initial stages of the research, these prior distributions will be less informative in nature, ensuring that the MCMC algorithm can converge. As a trial proceeds through time, information about a drug will periodically be added to the priors in the form of dose allocation probabilities, or the percentage of patients allocated to each drug dose. This process, explained thoroughly in the Adaptive Trial section of the text (Chapter 4), will update priors until they become informative enough to assist in making conclusions about a drug.

3.3 Simulation Techniques

3.3.1 Complete Conditionals

In order to simulate posterior distributions of each parameter, complete conditional densities for each variable must be used in the Markov chain Monte Carlo process. Complete conditionals are found by combining the prior density functions for the parameters and the likelihood function. In the interest of simulation and computation speed, the density functions for each of the prior distributions, along with the likelihood function listed above, will be used only in their log forms. The complete conditionals that are used in this study are listed below. Let $\log(f(y|\Theta))$ be denoted by $l(y|\Theta)$;

$$\begin{aligned}
\beta &\propto l(y|\Theta) + (a_\beta - 1)\log\beta + (b_\beta - 1)\log(1-\beta), \\
\delta &\propto l(y|\Theta) + (a_\delta - 1)\log\delta + (b_\delta - 1)\log(1-\delta), \\
\theta &\propto l(y|\Theta) + \left(\frac{-1}{2\sigma^2}(\theta - \mu)^2\right), \\
\tau &\propto l(y|\Theta) + (\alpha - 1)\log(\tau) - \left(\frac{\tau}{\kappa}\right).
\end{aligned}$$

The conditional distributions allow draws to be generated from the posterior distribution for each of the four unknown parameters. This is done using the Metropolis-

Hastings algorithm. Using this algorithm, the value of each parameter will be updated using only the most recent value of each of the other parameters in the likelihood equation. One loop of the Metropolis-Hastings algorithm is complete when each of the unknown parameters has been updated and properly saved. The looping procedure can be repeated until the posterior density of each parameter has been properly explored and demonstrates convergence.

3.3.2 Candidate Densities

Inherent in the Metropolis-Hastings algorithm is the use of a candidate density. Each parameter is updated randomly using a density, usually the Normal density, centered around the previous value of that parameter. Therefore, a model with restrictions on the parameters will present an interesting problem. In order to properly cycle through the Markov chain, the generated value of each parameter must strictly follow the constraints on the model. In this case, the issue of constrained parameters applies directly to the parameters β and δ. Remember the constraints on the model,

$$\begin{aligned} \beta &< \delta, \\ 0 < \beta &< 1, \\ 0 < \delta &< 1. \end{aligned}$$

Furthermore, we know that $\beta + \delta$ is the maximum response value, which implies

$$\beta + \delta < 1.$$

These inequalities lead to a fourth restriction on the model,

$$\beta < 0.5.$$

We can algebraically restructure this information into two sets of formulas. The first set is true when $\delta > 0.5$ and the second set is true when $\delta < 0.5$.

$$1: \quad 0 < \beta < 1 - \delta,$$

$$
\begin{aligned}
& & \beta &< \delta < 1 - \beta. \\
2: & & 0 &< \beta < \delta, \\
& & \beta &< \delta < 1 - \beta.
\end{aligned}
$$

Using a traditional candidate density, like the Normal density, will not ensure that these constraints are met. In order to efficiently generate numbers while still conforming to these constraints, the algorithm will make use of a Truncated Normal Distribution as a candidate density in the Metropolis-Hastings algorithm. This will ensure, even when there is only a small region of possible values, that values for both β and δ will be properly and efficiently simulated to follow these constraints.

The truncation bounds for β and δ can be understood from the above inequalities. It should be noted that correct truncation bounds depend on whether δ is greater than or less than 0.05. If δ is greater than 0.05, a new value of β must be drawn from a Truncated Normal Distribution centered at the previous value of β with bounds of 0 and $1-\delta$ and a new value of δ must be drawn from a Truncated Normal Distribution centered at the previous value of δ with bounds of β and $1-\beta$. If δ is less than 0.05, the bound for β change to $0 < \beta < \delta$ and the bound for δ remain the same. The two candidate densities, assuming $\delta < 0.05$, can be represented as follows:

$$\beta \sim \text{TN}(\mu_\beta, \sigma_\beta^2, 0, 1-\delta),$$

where 0 is the lower bound and $(1-\delta)$ is the upper bound, and

$$\delta \sim \text{TN}(\mu_\delta, \sigma_\delta^2, \beta, 1-\beta),$$

where β is the lower bound and $(1-\beta)$ is the upper bound.

Because there are no restrictions on the parameters θ and τ, we can utilize the simple normal density as a candidate density for these two parameters. Defining our candidate densities in this way, the indicator function in our likelihood, $I(\beta \leq \delta)$, is always true. Because this inequality is guaranteed to be true, the indicator

function can be disregaurded as being part of the likelihood function. A likelihood free of indicator functions provides the opportunity to simulate without worrying about constraints on the model. Therefore, the candidate density for β and δ allows the simulation process to be carried out as if it used a traditional likelihood with no restrictions.

3.4 Asymmetric Jumping Rule

3.4.1 Implications of Using Truncated Normal Density

Random number generation from a Truncated Normal Distribution can be done in a number of different ways. This study uses the rtnorm function in the msm package for R. Documentation for this generator can be found in the R Documentation (R Development Core Team, 2003). Distributions from random, doubly truncated Normal Distributions look just like Normal Distributions without the tails. Figure 3.1, generated using the rtnorm function, shows how a truncated Normal generator can produce very different distributions using the same mean and variance. Two different truncations of the standard normal density are contrasted with the simple standard normal density.

Figure 3.2 demonstrates an important characteristic of the truncated Normal Density when it is used as a candidate density. If the densities are truncated differently for two different parameters, or even if the truncation changes for a single parameter, the jumping probability for the Metropolis-Hastings algorithm is asymmetrical. In other words, the probability of going from state a to state b is not the same as going from state b to state a;

$$P(a, b) \neq P(b, a).$$

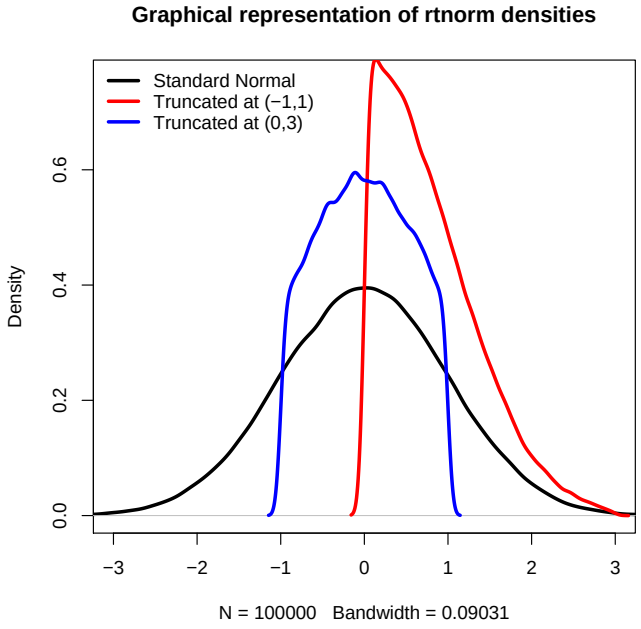

Figure 3.2: Truncated Normal Densities

3.4.2 Acceptance Probabilities

The defining characteristic of the Metropolis-Hastings algorithm, however, is that these asymmetrical jumping rules will not necessary confound the Markov chain Monte Carlo process. In order to create an algorithm that still converges to the posterior distribution, the probability of accepting a newly generated value for β and δ must be changed.

The acceptance probability of going from a to b, denoted as $\alpha(a, b)$, is a simple ratio,

$$\alpha(a, b) = \frac{f(y|b) * p(b) * P(b, a)}{f(y|a) * p(a) * P(a, b)}, \tag{3.4}$$

where $f(y|b)$ is the likelihood using the new value, $p(b)$ is the prior density using the new value and $P(b, a)$ is the probability of moving from b to a, and vice-versa for the denominator.

When the jumping probabilities are symmetric, this equation simplifies to be a ratio of the likelihood times the prior of the new value divided by the likelihood times the prior of the old value. In this study, the equation becomes more involved.

In order to find the probability of moving from one Truncated Normal state to another, $P(b, a)$, the density of the Truncated Normal Distribution must be known. When a distribution is truncated the new density is just the old density divided by that part of the density that has been left off. The denominator in this equation is known as the normalizing constant. Because the density is doubly truncated, both deleted segments must added to the original density in order to divide by the correct normalizing constant. Accordingly, the density of a Normal Distribution truncated between a and b where $a < b$ can be written as

$$\begin{aligned} f(x_{a,b}) &= \frac{(2\pi\sigma^2)^{\frac{-1}{2}} e^{\left(\frac{-1}{2\sigma^2}(x-\mu)^2\right)}}{\int_{-\infty}^{b} f(x)dx + \int_{a}^{\infty} f(x)dx} \\ &= \frac{N(\mu, \sigma^2)}{1 - \int_{b}^{a} f(x)dx} \end{aligned}$$

$$
\begin{aligned}
&= \frac{N(\mu, \sigma^2)}{1 - \{(1 - \Phi\left(\frac{b-\mu}{\sigma}\right)) + \Phi\left(\frac{a-\mu}{\sigma}\right)\}} \\
&= \frac{N(\mu, \sigma^2)}{\Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)}.
\end{aligned}
$$

Applying this result to our two parameters, the density of both truncated distributions can be represented:

$$t(\beta) = \frac{N(\mu_\beta, \sigma_\beta^2)}{\Phi\left(\frac{1-\delta-\beta}{\sigma}\right) - \Phi\left(\frac{0-\beta}{\sigma}\right)}, \tag{3.5}$$

$$t(\delta) = \frac{N(\mu_\delta, \sigma_\delta^2)}{\Phi\left(\frac{1-\beta-\delta}{\sigma}\right) - \Phi\left(\frac{\beta-\delta}{\sigma}\right)}. \tag{3.6}$$

3.4.3 Application in Simulation

A proper acceptance probability equation can now be constructed for simulation purposes. Because the derivation for the acceptance probability of β is similar to the derivation for δ, we will simply list the result for δ after the derivation for β. Note that because the distribution of β depends on δ, the prior density for δ is necessary for the acceptance probability of β^{new}. However, this probability is constant in the equation, and will therefore cancel from the numerator and denominator. The notation, $g(\cdot)$ represents the likelihood function multiplied by the prior function for a given parameter. The probability of accepting a move from β^{old} to β^{new}, denoted by $\alpha(\beta^{old}, \beta^{new})$, can be derived in a number of steps:

$$
\begin{aligned}
\alpha(\beta^{old}, \beta^{new}) &= \frac{f(y|\beta^{new}) * p(\beta^{new}) * P(\beta^{new}, \beta^{old})}{f(y|\beta^{old}) * p(\beta^{old}) * P(\beta^{old}, \beta^{new})} \\
&= \frac{g(\beta^{new}) * t(\beta^{new})}{g(\beta^{old}) * t(\beta^{old})} \\
&= \frac{g(\beta^{new}) * N(\mu_\beta, \sigma_\beta^2) / \left(\Phi\left(\frac{1-\delta-\beta^{new}}{\sigma}\right) - \Phi\left(\frac{0-\beta^{new}}{\sigma}\right)\right)}{g(\beta^{old}) * N(\mu_\beta, \sigma_\beta^2) / \left(\Phi\left(\frac{1-\delta-\beta^{old}}{\sigma}\right) - \Phi\left(\frac{0-\beta^{old}}{\sigma}\right)\right)}
\end{aligned}
$$

$$= \frac{g(\beta^{new})) * \left(\Phi\left(\frac{1-\delta-\beta^{old}}{\sigma}\right) - \Phi\left(\frac{0-\beta^{old}}{\sigma}\right)\right)}{g(\beta^{old}) * \left(\Phi\left(\frac{1-\delta-\beta^{new}}{\sigma}\right) - \Phi\left(\frac{0-\beta^{new}}{\sigma}\right)\right)}.$$

Because calculations are done on the log scale, the acceptance probability for β must also be on the log scale. Thus,

$$\begin{aligned}\log \alpha(\beta^{old}, \beta^{new}) &= l(y|\beta^{new}) + \log p(\beta^{new}) + \log\left[\Phi\left(\frac{1-\delta-\beta^{old}}{\sigma}\right) - \Phi\left(\frac{-\beta^{old}}{\sigma}\right)\right] \\ &- \left\{l(y|\beta^{old}) + \log p(\beta^{old}) + \log\left[\Phi\left(\frac{1-\delta-\beta^{new}}{\sigma}\right) - \Phi\left(\frac{-\beta^{new}}{\sigma}\right)\right]\right\}.\end{aligned}$$

The log of the acceptance probability for δ follows the same pattern,

$$\begin{aligned}\log \alpha(\delta^{old}, \delta^{new}) &= l(y|\delta^{new}) + \log p(\delta^{new}) + \log\left[\Phi\left(\frac{1-\beta-\delta^{old}}{\sigma}\right) - \Phi\left(\frac{\beta-\delta^{old}}{\sigma}\right)\right] \\ &- \left\{l(y|\delta^{old}) + \log p(\delta^{old}) + \log\left[\Phi\left(\frac{1-\beta-\delta^{new}}{\sigma}\right) - \Phi\left(\frac{\beta-\delta^{new}}{\sigma}\right)\right]\right\}.\end{aligned}$$

The acceptance probabilities, together with the conditional distributions of the likelihood for each parameter, provide all the information needed to run the Metropolis-Hastings algorithm. Values for each parameter are drawn from the appropriate candidate density; the likelihood will be calculated and the value will be accepted or denied based on correct acceptance probabilities. This process will continue for as many times as the MCMC simulation is told to run. Once the parameters are simulated, the values are used in an improved clinical trial which utilizes a clever adaptive design.

3.5 Adaptive Trial Design

3.5.1 Purpose of an Adaptive Trial

Through proper simulation, each piece of data from a trial provides a better understanding of the dose-response characteristics of a drug. The intent of modeling

the data as demonstrated and creating effective simulation methods is to be able to predict a drug's most effective dose range. This range can be effectively achieved by applying an adaptive dose-allocation system in the proposed model.

The aim of the adaptive section of this study is to find the minimum dose in a clinical trial that obtains a 95% level of effectiveness. This level of precision is referred to as ED95, or Effective Dose 95. In this study, the ED95 level is a function of dose effectiveness and dose variability. It is important to note that some models include the effects of toxicity in assigning the ED95 level. While other factors can help in better assessing the ED95 level, drug effectiveness and drug variability provide enough information to make proper inference. Developing a protocol that will ensure that the best doses (those in the ED95 range) are allocated to subjects is the focus of the adaptive trial design portion of our research.

3.5.2 Methods of Calculation ED95

Creating a trial that determines which doses fall in the ED95 range hinges on the ability to correctly find the ED95 range. The definition of ED95 can lead to two possible methods of understanding and calculating the value. The two methodologies are labeled $ED95^{\star}$ and $ED95^{\dagger}$, respectively.

The first method of finding ED95 uses the theoretical interpretation of the logistic model. It is assumed that $\beta + \delta$ is the maximum effectiveness of a drug. Because ED95 is 95% of the maximum efficiency, the relationship is simple;

$$ED95^{\star} = .95(\beta + \delta).$$

However, estimating ED95 in this way may result in an invalid estimate. Because of its theoretical nature, it is possible that the $ED95^{*}$ result represents a dose range that is not part of the study. If the trial was focused on one small range of all possible doses, this may be a valid possibility. However, in a study that includes all doses

under consideration, the ED95 level *must* be part of the study. This is a simple corollary from the definition of ED95.

Though it may be more mechanical, the second method of finding ED95 does not have this problem. The second method is a more conservative, less theoretical approach to finding ED95. The four-parameter logistic model is an increasing function. If a trial utilizes n doses, then the effectiveness of the nth dose will be the maximum value for the trial, and the lowest dose, usually a placebo, will be the relative noise level of the trial. ED95 can be calculated as a function of these two levels, where p(0) represents the efficacy of the lowest dose and p(n) represents the efficacy of the nth dose,

$$ED95^{\dagger} = p(0) + .95\,(p(n) - p(0))\,.$$

Theoretical ED95 levels are not useful when identifying the ED95 dose among all those under experimentation. Therefore, the second method, $ED95^{\dagger}$, will be the used to calculate the ED95 level in this study.

3.5.3 Design: Allocation by Batch

Unlike traditional clinical trials, subjects are processed as batches, not all at once or one at a time. After the data from the initial batch come in, the MCMC loop is simulated and the posterior predictive distribution for the parameters in the logistic model are collected. From these distributions, the effectiveness of each dose in the trial can be calculated. By calculating an overall ED95 level for the trial, it is possible to calculate the probability that each dose is the true ED95 dose (the lowest dose that achieves 95% efficacy). This probability is used to randomly allocate the next batch of subjects to doses. In other words,

$$P(\text{Allocation to } X_i) = P(X_i = ED95),$$

where X_i is the ith dose in the trial. In this way, those doses that perform better will have a greater probability of receiving new subjects, allowing more subjects to be assigned to doses that have a large probability of success.

When the data from the second batch are complete, their results are augmented to the results of the first batch and the adaptive process will be repeated in the same way. In other words, if the placebo received 15 patients during the first dose, 3 of which reported successful results, and if the placebo received 5 patients with the second batch, none of which reported success, the numbers n=20 and m=3 would be reported. In essence, augmenting the data is another way of updating the prior information; with each additional batch of subjects the prior information will become more and more informative. The adaptive process continues until the ED95 range is agreed upon and an optimal dose is found.

3.6 Computational Implementation

3.6.1 Starting Values

This thesis uses 8 total doses (0,1,...,7). The simulation begins by assigning success probabilities to each dose in the trial. Section 3.8 describes more fully the probabilities assigned to each dose in this thesis.

An initial batch of 96 subjects is then equally allocated to all doses in the study. This provides a large enough sample size, n_i=12, at each dose to begin to make conclusions. The number of successes at each dose, m_i, is calculated using the probability of success assigned to each dose and a random binomial number generator. The data are then placed in the MCMC loop to find posterior densities for each of the four parameters in the logistic model. The MCMC algorithm contains a total length of 5,500 iterations with a burn-in period of 500 iterations.

The parameter values for the prior densities used in simulation are listed below;

$$\begin{aligned}
\beta &\sim Beta(1,1), \\
\delta &\sim Beta(1,1), \\
\theta &\sim Normal(3.2,\ \text{sd} = 1), \\
\tau &\sim Gamma(1,\ \text{scale} = 1).
\end{aligned}$$

While these are very broad prior densities, it is important to remember that with the addition of data from each successive batch, the variability of the Metropolis-Hastings algorithm will decrease.

3.6.2 Logistic Regression and ED95 Threshold

The Metropolis-Hastings algorithm outputs posterior densities for the unknown parameters. Each iteration of the MCMC loop provides one value for each of the four parameters in the logistic model. However, the logistic model contains one other variable, x_i, the ith dose in the trial. The binomial nature of the likelihood allows the incorporation of logistic regression into the analysis. Therefore, the posterior predictive density of the success probability of each dose can be found by simply imputing the values of each dose, (0,1,....,7), into the logistic model in place of x_i. At the end of a simulation, the logistic model can be uniquely calculated for as many iterations as the loop cycled; in this thesis, the model can be calculated 5,000 times. Through logistic regression, this provides the success probability for each dose (0,1,..,7) a total of 5,000 times.

Under the assumptions of the second method of ED95 calculation each iteration provides enough information to calculate a unique ED95 level. The inherent stochasticity of the MCMC allows the possibility of calculating thousands of vastly different ED95 values after simulation. Consequently, in order to compare the data from each iteration simultaneously it is necessary to calculate an overall ED95 threshold for the

whole simulation. This is done by taking the success probabilities from the logistic regression and calculating the mean probability of success for maximum and minimum doses in the trial, over all iterations. These mean values can be used in in the $ED95$ equation in order to find an overall ED95 threshold for the data. This value can be described in equation format;

$$ED95_{Thsld} = \mu_{x0} + .95(\mu_{x7} - \mu_{x0}),$$

where μ_{x0} is the mean of all values for the minimum dose and μ_{x7} is the mean of all values for the maximum dose.

3.6.3 ED95 Probabilities

The ED95 threshold can be compared to each MCMC iteration. At each iteration, the minimum dose that is larger than the ED95 threshold is designated the ED95 dose for that iteration. At each dose, the proportion of ED95 designations is divided by the total number of MCMC iterations. In essence, this computes the probability that each dose in the trial is the ED95 dose. It is important to note that because the algorithm that calculates the ED95 threshold is an average over all iterations, it is possible that one iteration does not contain any success probability greater than the threshold. Therefore, it is also possible that the sum of all the ED95 probabilities does not equal 1.

ED95 probabilities are used for two calculations. First, they are used to determine if the drug trial will proceed or if it will stop. These guidelines are commonly known as stopping rules. If the algorithm proceeds to the next batch of patients, the ED95 dose probabilities are used in dose allocation.

3.6.4 Stopping Rules

All drug trials need firm stopping rules. In this study there are three ways to stop a trail: Cap Limit, Success, and Failure.

Federal regulations and financial restrictions keep a drug trial from going on forever. At some point, if no decision has been made, the trial must be stopped and reassessed before more patients are used or wasted. In this study, a cap limit of 600 individuals has been set. As soon as the total number of subjects reaches 580 or above, the trial is stopped so that the next batch of 20 individuals does not exceed the 600-individual limit.

In order to stop a trial for success, there must be convincing evidence that there exists a dose that can be labeled the ED95 dose. In the case of trial success, the adaptive model is interested in the success rates of each dose relative to the other doses in the trial. Consequently, the ED95 probabilities calculated for each dose are normalized so that the sum of all probabilities equals one. The trial will stop for success when a dose obtains an ED95 probability greater than 0.90.

Stopping the trial for failure can be interpreted to mean that the model finds no dose in the trail that achieves the ED95 level. In this case, the adaptive model is interested in the true non-normalized ED95 probabilities for each dose in the trial. In this study there must be at least one dose with an non-normalized ED95 probability greater than 0.25 in order to proceed through the trial; otherwise the trial stops for failure.

3.6.5 Reasons for Failure

Failure can occur for two reasons. First, the stochasticity of patient responses does not always allow researchers to see the true dose-response curve. There is always a possibility that, through chance alone, the data may demonstrate a curve very different from the theoretical curve. Precautions, such as large sample sizes and controlled experiments, can be taken to limit the possibility of such an event, but it can never be ruled out. The second cause of a failed experiment can be the choice of doses in the study. The graphs below demonstrate three possible instances where the

choice of doses will not allow our algorithm to properly find the ED95 dose.

In clinical research it is important that the dose-response curve is well represented by the dose range under investigation. The ED95 dose, remember, is the lowest dose that achieves 95% efficacy. Figure 3.3 below demonstrates one instance where the true high and low values of the curve are not explored. In this case, the algorithm will not be able to locate an ED95 dose because all doses have equal responses.

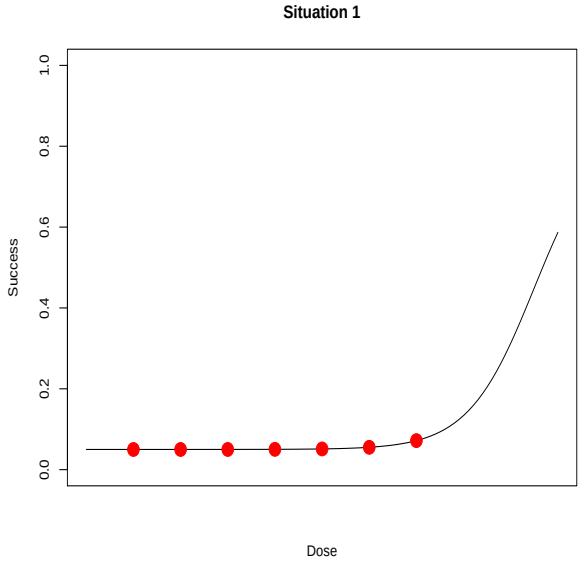

Figure 3.3: Situation 1

Figure 3.4 represents a curve that is completely represented by the range of doses in the study. However, there is only a small change from the lowest dose to the highest dose. With a cap size of 600 subjects, this small difference will be virtually impossible to see. While the parameters in the logistic model will be properly calculated under these circumstances, the algorithm will fail to find the ED95 dose because of the similarity among all doses in the trial.

Figure 3.5 represents a possible and frequent circumstance in clinical research. The dose range captures the correct high and low values on the curve and there is good representation of doses in the sigmoidal section of the curve. However, the

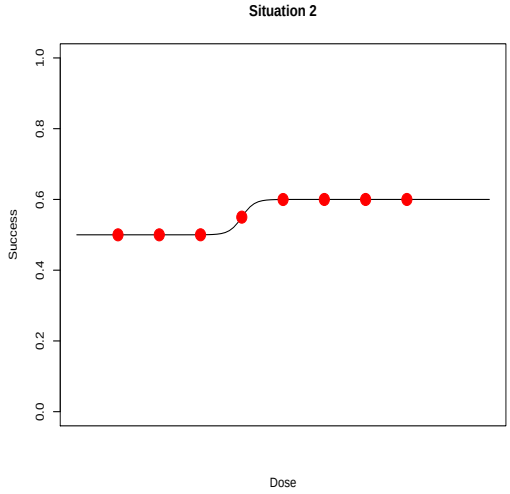

Figure 3.4: Situation 2

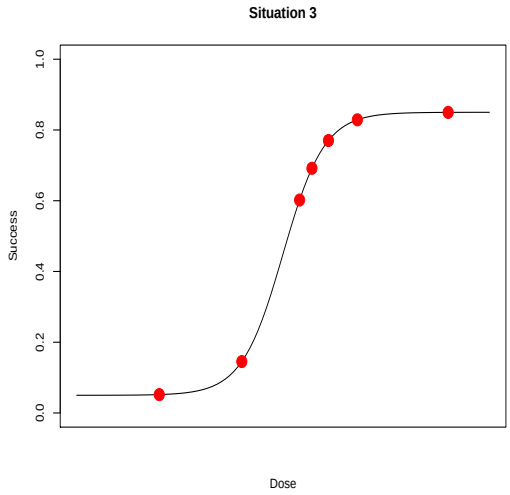

Figure 3.5: Situation 3

doses around the true ED95 range are very close together. The MCMC iterations are random and will not be able to distinguish between the three doses in the middle of the curve. The ED95 probabilities will be distributed among each of these doses and could easily cause the algorithm to fail.

3.6.6 Dose Allocation

After each batch, if the trial meets none of the stopping criteria it can proceed to the next batch. Once the algorithm has determined to proceed to the next batch, the next step is to allocate the next batch of subjects to doses. Two variables are used in the allocation, the ED95 probability and the variability of each dose. In this study, the variation statistic of interest is the variability of the next subject at each dose in the trial. To calculate this value, the draws from the MCMC are used and the variance calculated is multiplied by $\frac{n_i}{n_i+1}$, providing the variance of the next subject.

It should be mentioned that a number of other variables can and have been used to allocate subjects to doses. One of the most obvious deletions from this model is a toxicity variable. Many studies will allocate fewer subjects to a dose that exhibits a high toxicity rating. For simulation purposes, it is not in the best interest of the adaptive design to randomly assign toxicity levels to the doses in the trial. In a true clinical trial, this parameter could be easily added to the model without too much modification.

The probability of allocating one subject to each dose is calculated by multiplying the ED95 probability of each dose by the variability of the next subject at each dose. The allocation probabilities are normalized for sampling purposes;

$$P(\text{allocation to } x_i) = \frac{P(x_i = ED95) \cdot \frac{n_i}{n_i+1} Var(x_i)}{\sum_{i=1}^{r} P(x_i = ED95) \cdot \frac{n_i}{n_i+1} Var(x_i)}.$$

The next batch of 20 subjects is then allocated to doses based on the probabilities calculated, with one small restriction. The FDA requires drug trials to maintain

a placebo effect throughout a trial. In order to comply with FDA regulations, the Bayesian adaptive design automatically allocates one-fourth of the subjects to the placebo, or dose zero. Therefore, a batch of 20 randomly allocated subjects becomes a batch of 15 randomly allocated subjects with 5 automatically assigned to the placebo.

Data from the new batch is augmented to the previous data. In a clinical trial, this would take anywhere from days to months. In this study, these numbers are generated using a random binomial number generator and a true probability of success for each dose included at the beginning of the simulation.

The augmented data is then inserted into the beginning of the algorithm and the process continues until one of the stopping criteria is reached. The augmented data can be thought of as an update to the *apriori* information that was used at the beginning of the simulation. With each update of the data, the MCMC will become more and more precise until it either pinpoints the ED95 dose or determines that no such dose exists.

3.7 Model Assessment, Operating Characteristics of Bayesian Adaptive Design

Clinical trials are inherently stochastic. In this thesis the a Binomial number generator allows the algorithm to capture that stochasticity. Every time a new subject is allocated to a dose a binomial number generator is applied, using the true success probability for that dose, to determine if the treatment is successful. A binomial generator, like human trials, contains a variance and the possibility of a wide range of outcomes.

Because of the inherent variance in the study, it is entirely possible that the Bayesian model returns a false answer. Both Type I and Type II errors are possible in our study. For example, if it is known that the true ED95 dose is the 4th dose in the trial, the model could report that there is no ED95 dose (Type II) or that

the ED95 dose is a different dose (Type I). Consequently, it is necessary to observe the model's performance many times in order to understand how it is functioning. Observing just one simulated clinical trial can only provide a small amount of insight into model performance.

The entire simulated clinical trial, then, will be replicated enough times to obtain the probabilities of success, error, and cap realization. This process allows us to identify the operating characteristics of the Bayesian adaptive design when applied to clinical trials. The simulated clinical trials will provide insight into how the adaptive design performs when used for a variety of clinical situations.

3.8 Situations under Consideration and Simulation

In this study the performance of the adaptive model is tested using data from simulated clinical trials under 7 different clinical situations. Each situation will simulate a different curve dynamic to show how the model responds. In order to minimize the number of variables in the study the parameter values of β and δ are kept constant, which will produce sigmoid curves that range between 0.05 and 0.85.

The curves under consideration are plotted below. Each graph represents the dose-response curve from the placebo dose (0) to the maximum dose (7). The value of each dose is represented by a red dot while the value of the theoretical ED95 dose is represented by a blue dot. The first curve (Figure 3.6) can be considered the basic sigmoid; it is a dose-response curve centered in the middle of the dose range (between doses 3 and 4, or $\theta = 3.5$) with a modest slope (τ just under 1).

The next two curves (Figure 3.7) are similar to the first curve but are shifted to the right and to the left. These two simulations help determine the ability of the model to detect curve dynamics both early and late in the dose range.

The next curves under discussion (Figures 3.8) display few sigmoidal characteristics because the dose range does not capture all the curve dynamics. The first curve

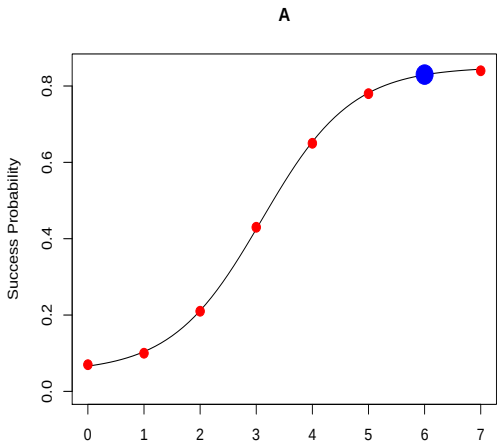

Figure 3.6: Basic Sigmoid Curve

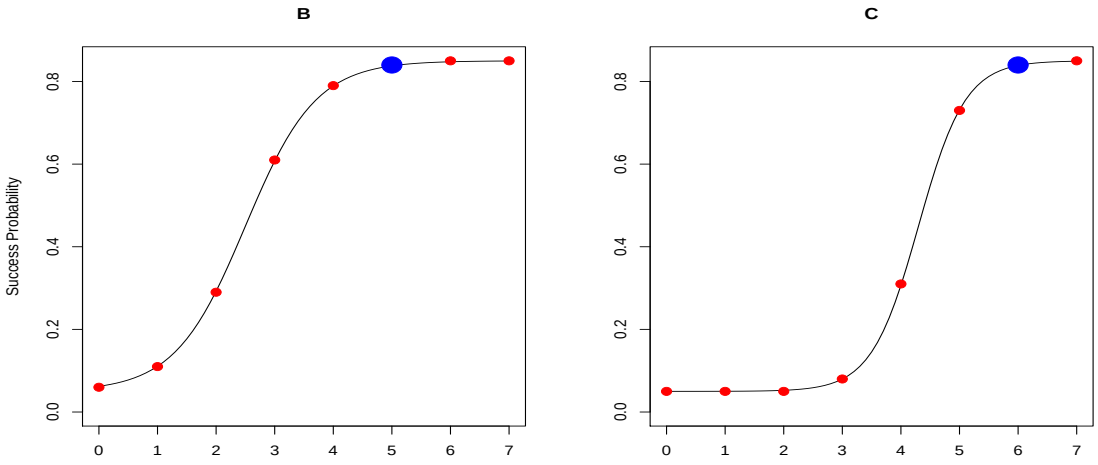

Figure 3.7: Horizontally Shifted Curves

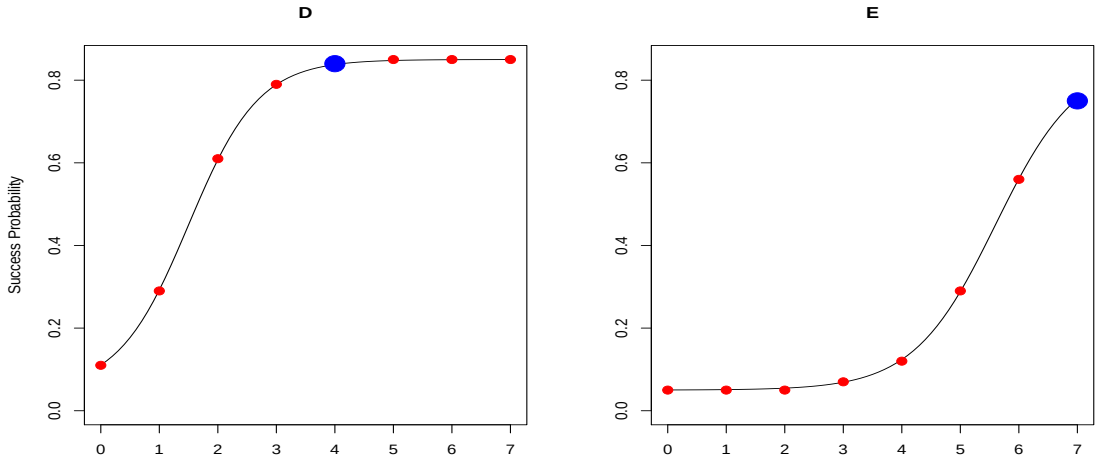

Figure 3.8: Underrepresented Curves

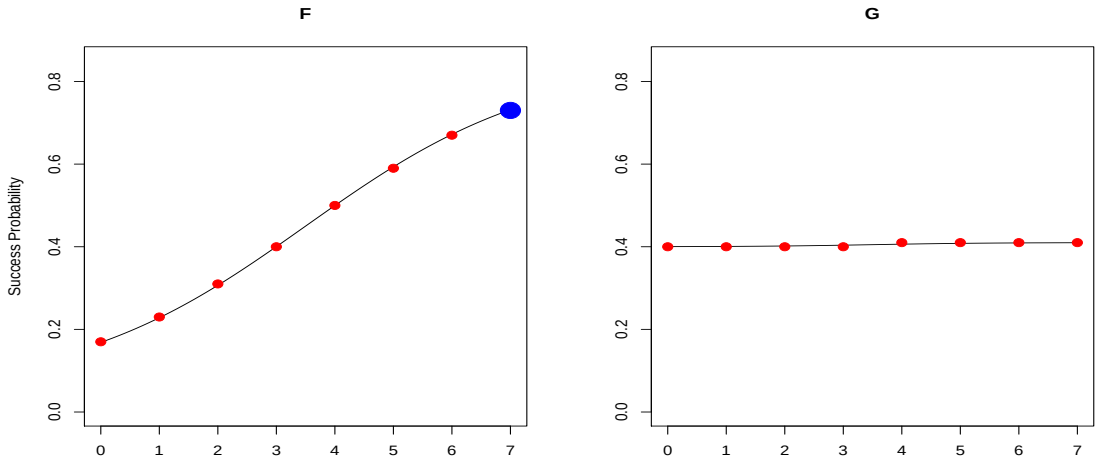

Figure 3.9: Control Curves. Curve 'G' has no true ED95 dose

displays an immediate jump in success probabilities, while the second graph displays a very late jump in success probabilities. These graphs test the ability of the model to find the ED95 dose even when the curve is not properly represented in the dose range.

The next two curves (Figures 3.9) are used as control groups. The first graph has some sigmoidal characteristics but is basically a linearly increasing function from the first dose to the last. The second graph can be thought of as the null case. Here there is essentially no change from the first dose to the last. These curves test how the model responds to functions that do not readily fall into the sigmoid category. While these functions are not probable outcomes of a Phase II or Phase III clinical trial, they are necessary to test.

The parameter values for the seven curves in question are displayed in the following table. For a more detailed explanation of the interpretation and the curvature dynamic effect of each of these parameters, refer to Section 3.1 of the text.

Table 3.1: Summary of parameter values used in each of the 7 clinical situations tested in this study.

	β	δ	θ	τ
Curve A	0.05	0.80	3.5	0.8
Curve B	0.05	0.80	2.5	0.6
Curve C	0.05	0.80	4.3	0.4
Curve D	0.05	0.80	1.5	0.6
Curve E	0.05	0.80	5.6	0.7
Curve F	0.05	0.80	3.5	2.0
Curve G	0.40	0.01	3.5	1.0

4. RESULTS

4.1 Success and Failure Probabilities

Results from simulated clinical trials are inherently dependent on the stopping rules for each trial. Simulation has shown that the values used to stop a trial for success, failure, or cap realization can dramatically alter the results from the simulated clinical trials. Initial simulations have demonstrated that 0.25 is a reasonable value for the stop-for-failure probability. On average, this value consistently fails in trials that have no true ED95 dose (Situation "G") and rarely produces a Type II error. Federal regulations limit the options in choosing an appropriate cap limit. It was found, however, that a cap limit of 600 conforms to protocol regulations and generally provides enough data to make legitimate conclusions.

The stop-for-success probability has been the object of much simulation. This value ultimately controls the probability of Type I errors and the probability that a trial will stop for cap realization. We present the results of simulations that vary the stop-for-success probability between 0.70 and 0.90. Table 4.1 displays the Type I error rate and cap realization rate for possible stop-for-success values; these values are plotted graphically in Figure 4.1. In each simulation, the values of stop-for-failure and cap limit are held constant. Each simulation is averaged over a number of clinical scenarios.

Choosing the best stop-for-success value is also dependant on FDA regulations. Federal mandate states that Type I error rates must be below 0.05. Accordingly, the lowest probability that achieves that regulation, 0.80, was chosen for simulation. This value will be kept constant throughout the clinical trial simulations, allowing the model to compare the performance of the Bayesian adaptive design between all seven clinical situations.

	TYPE I Error Rate	% Reaching Cap
Success Probability = 0.70	0.30	0.01
0.75	0.20	0.15
0.80	0.07	0.34
0.85	0.05	0.60
0.90	0.00	0.90

Table 4.1: Type I error rate and cap realization probability of 5 different success probabilities

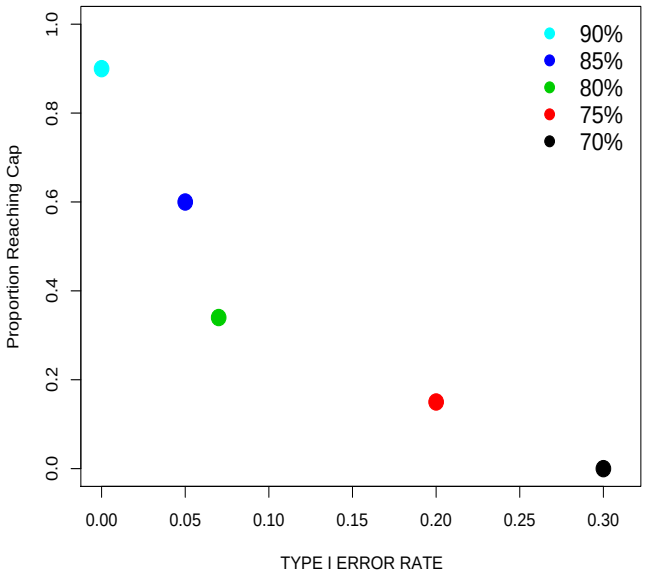

Figure 4.1: Type I error rates of 5 different success probabilities

4.2 Examining One Simulated Clinical Trial

To illustrate the Bayesian adaptive methodology, we present the analysis of the simulated clinical trials of each situation outlined in Section 3.8. For each of the seven situations the posterior mean dose-response curve is superimposed on the theoretical dose-response curve (Figures 4.2-4.9). The fitted curve (red dotted line) shows a good fit to the theoretical curve (black solid line).

The simulated posterior responses demonstrate areas of both good and poor fitting. In the null case, Situation "G", the posterior fit to the theoretical response curve is completely wrong. The likelihood was restricted to produce an increasing function where $\beta < \delta$. Because this is not true of the null case, it follows that the posterior means could not possibly follow the theoretical response curve.

Posterior means fit theoretical response rates better in areas of maximum allocation. In this sense, maximum allocation can refer to doses that receive more patients or clinical trials that require more data. In the first case, Bayesian adaptive design allocates most subjects to the placebo and to doses that are close to the true ED95 level. As demonstrated in Figures 4.2–4.9 the most precise Bayesian estimates are seen at the placebo level and the region around the theoretical ED95 level. In the second case, the data demonstrate that clinical situations that require more data (Figures 4.2, 4.3) produce more precise Bayesian estimates than situations that do not require much data (Figure 4.6). A situation like the clinical trial depicted in Figure 4.6 will not require as much data as other situations because the theoretical dose-response levels are vastly different from one another.

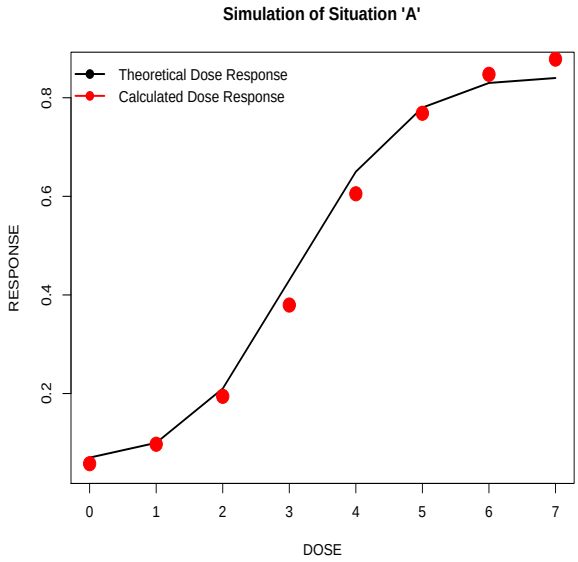

Figure 4.2: Final response values of a complete clinical trial simulation of type “A”

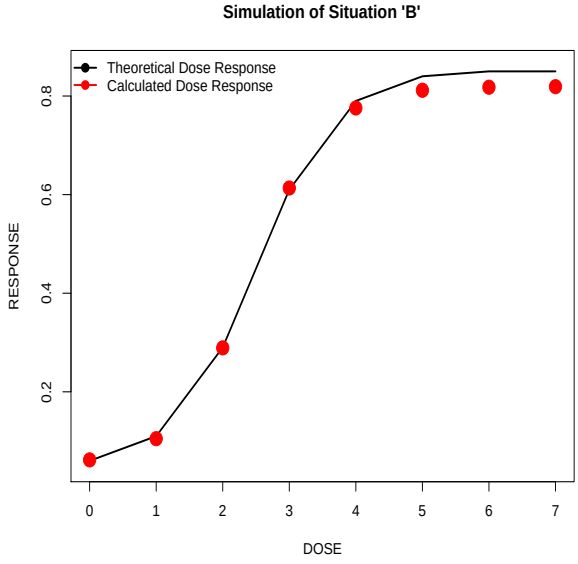

Figure 4.3: Final response values of a complete clinical trial simulation of type “B”

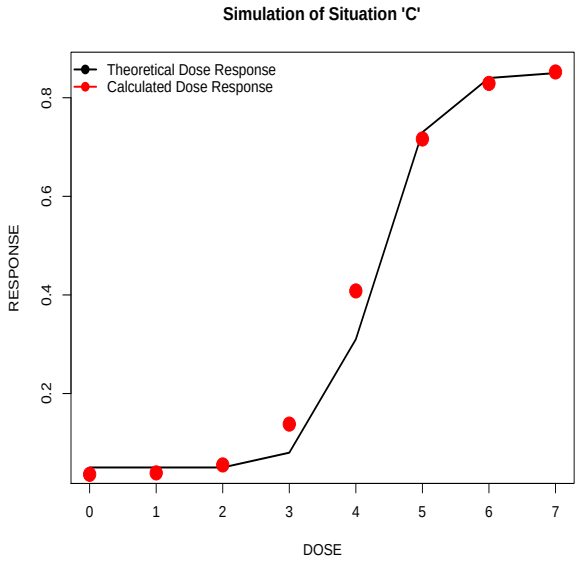

Figure 4.4: Final response values of a complete clinical trial simulation of type “C”

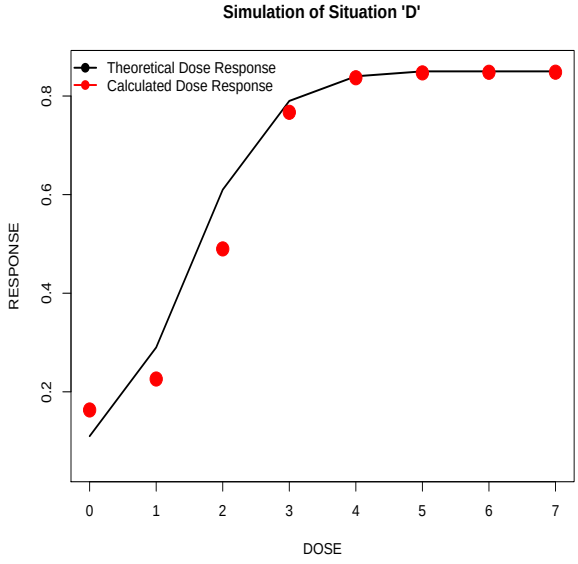

Figure 4.5: Final response values of a complete clinical trial simulation of type “D”

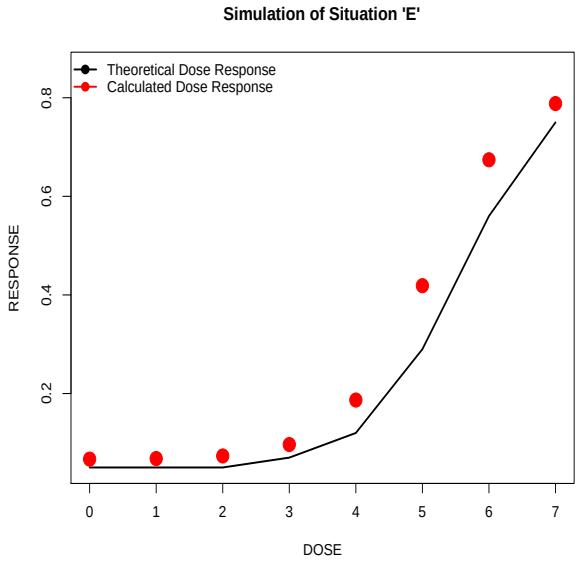

Figure 4.6: Final response values of a complete clinical trial simulation of type “E”

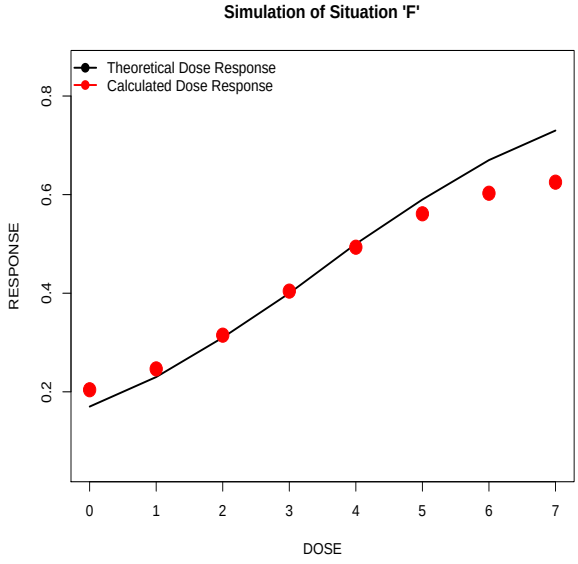

Figure 4.7: Final response values of a complete clinical trial simulation of type “F”

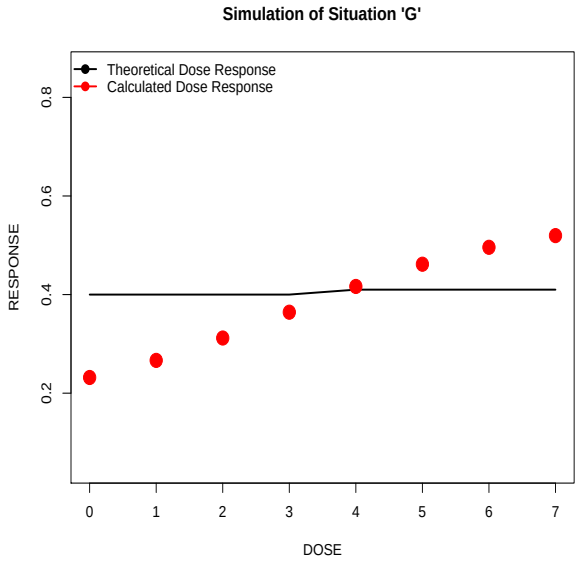

Figure 4.8: Final response values of a complete clinical trial simulation of type “G”

4.3 Examining Multiple Simulations for each Situation

4.3.1 Data from Multiple Clinical Simulations

The graphs depicted above (Figures 4.2–4.8) represent one fully simulated clinical trial. We now present the results of 2,500 fully simulated clinical trials for each of the seven situations identified in Section 3.8. Each simulated clinical trial was stopped for success, failure, or cap realization, as described in Section 3.6. Because the theoretical dose-response rates are known, we can calculate the true ED95 dose (See Figures 3.6–3.9) for each study and calculate Type I error rates. The probability of each result is recorded in Table 4.2.

	Success	Type I Error Rate	Failure	Cap Realization
Situation A	0.114	0.067	0.003	0.816
Situation B	0.209	0.160	0.001	0.630
Situation C	0.711	0.110	0	0.180
Situation D	0.225	0.192	0.021	0.562
Situation E	0.983	0.014	0.001	0.002
Situation F	0.354	0	0.221	0.425
Situation G	0	0	1	0

Table 4.2: Operating characteristics of Bayesian adaptive approach to clinical trials.

In addition, Tables A.2–A.8 in the Appendix report patient allocation data for each of the seven situations. The tables present the mean number of subjects and the mean number of successes at each dose with the standard deviations in parenthesis. They also report the overall mean and standard deviation for the number of subjects and the number of successes for each situation.

4.3.2 Interpretation of Results

The overall Type I error rate for all simulations was 0.06, slightly higher than the goal of 0.05. The remaining results are better understood when analyzed in groups. The data indicate that there are four distinct groups in the study: the null

case, the linear case, obvious ED95 levels, and ambiguous ED95 levels.

All 2,500 simulated clinical trials for Situation "G" failed, demonstrating the consistency of a Bayesian adaptive approach in failing to find the ED95 dose in a situation where no true ED95 dose exists. More importantly, the Bayesian approach allowed these trials to be stopped immediately; the adaptive design never proceeded to the second subject batch.

Data for Situation "F", the linearly increasing control, demonstrate a high failure rate. These rates are explained by the logistic model, which tends to add curvature to the data even when it does not exist (Figure 4.7). The large variances seen in patients allocated to Doses 6 and 7 indicate that a large number of clinical trials ended early. Stopping for failure is a legitimate possibility when small amounts of data are available to an equation that adds curvature to a theoretically straight line.

Situations "A", "B", and "D" demonstrate low success rates and high probabilities of cap realization. Such behavior is explained by the proximity of adjacent doses to the ED95 level. In fact, the difference between the theoretical response rate of the dose just higher than the theoretical ED95 level and the dose just lower than the ED95 level is less than 0.05 in each of the Situations "A", "B", and "D". We will use Situation "A" to illustrate why the Bayesian process arrives at cap realization so often.

Situation "A" contains a theoretical ED95 level of .8015. Dose 5 has a success rate of 0.78, while Dose 6 has a success rate of 0.83. On average, Dose 5 was allotted 140 subjects during clinical simulation. According to the binomial distribution, Dose 5 has a standard deviation of $\sqrt{(140)(.78)(.22)} \approx 4.9$. Theoretically, Dose 5 should see 109.2 successful cases in each simulation. Adding just one standard deviation of successful subjects to the theoretical average, Dose 5 would receive 114.1 successes, a success rate of 0.815. It is noted that within one standard deviation, Dose 5 would

be designated the ED95 dose because $0.815 > 0.8015$.

Stochasticity in the binomial function leads to small deviations from the theoretical response values. It can be demonstrated that even small deviations alter the posterior distribution of the logistic parameters enough to cause a clinical trial to reach cap limit without stopping for success. Case 1 in the Appendix gives a detailed report of a simulated clinical trial of this nature.

Data from the final group, Situations "C" and "E", demonstrate high success rates and relatively low rates of cap realization. Situations "C" and "E" are characterized, in contrast to Situations "A", "B", and "C", by large distances between doses surrounding the theoretical ED95 level, 0.11 and 0.19 respectively. Their high success rates can be explained using the logic described for Situations "A", "B", and "C".

This study is focused on modeling the dose-response curve and minimizing the Type I error rates. Though the results have succeeded in accomplishing these goals, the trade-off is evident in a large proportion of cap realizations. An important feature of our model is its ability to change stopping criteria to maximize the researcher's parameter of choice. In fact, success rates can be increased and the average number of subjects used in clinical trials can be decreased by a simple alteration in the stopping rules. Section 6.2 in the Appendix lays out three stopping rules not used in this study that can be used to maximize success rates at the sacrifice of higher Type I error rates.

5. SUMMARY AND CONCLUSIONS

This thesis was focused on two main goals: to model a clinical trial that can properly estimate the dose-response curve and to create an adaptive trial design that properly stops trials for clinical futility.

In the first case, the Bayesian adaptive approach has demonstrated good fitting characteristics in most Phase II or III clinical scenarios. The Bayesian fit discussed in this paper is robust to curve changes, dose ranges, and even highly variable results (Case 1, Appendix). In the second case, the operating characteristics of the Bayesian adaptive design demonstrate proper protocol in failing clinically futile trials (Situation "G").

The operating characteristics of the Bayesian design dealing with success and cap realization depend on the nature of the dose-response curve. In our simulations the "success" rate for the Phase II or III clinical situations, not including the controls, was only 45%. This number, however, is dependant on the arbitrarily assigned stopping rule for success. In fact, the medicinal effectiveness of the Bayesian adaptive design is not dependant on the percentage of trials that end in success.

The goal of medicine is to provide the best care whenever possible. The adaptive trial design described in this paper provides physicians and researchers with the possibility of providing the best medicine while conducting proper clinical research. Simulations from Situation "A" are used to demonstrate this fact. Table 5.1 reports the average number of subjects, successes, and success rates for each dose in Situation "A". The subsequent table (Table 5.2) provides the same information for a traditional trial in which subjects are equally allotted to doses. Theoretical response rates are used to calculate the success rates of the traditional trial.

The Bayesian adaptive design demonstrated in Table 5.1 reports that 160 of

	Placebo	Dose 1	Dose 2	Dose 3	Dose 4	Dose 5	Dose 6	Dose 7
Subjects	128	20	20	20.5	56.7	140.7	159.6	46.5
Successes	8.7	2.8	4.2	8.9	37.4	108.9	132.7	38.2
Success Rate	.07	0.14	0.21	0.43	0.66	0.77	0.83	0.82

Table 5.1: Average dose allocation and success rates for a simulated clinical trial of type "A" ($n \approx 590$)

the 590 subjects were allocated to the true ED95 dose (Dose 6), and another 140 were allocated to the next closest dose (Dose 5). In all, 342 subjects responded successfully to the clinical study. The traditional clinical trial displayed in Table 5.2 reports that only 66 subjects were allocated to the true ED95 dose and 66 more were allocated to the next closest dose. The entire trial only produced 260 success responses. Comparatively, the adaptive trial allocated 242% more subjects to the ED95 dose and had an overall success rate 32% higher than the traditional clinical trial. Similar numbers could be shown for each of the situations under consideration in this study.

The adaptive Bayesian approach to clinical trials outlined in this paper properly models the dose-response curve of many clinical situations. It allows more patients to be allocated to the most beneficial doses and it increases the overall success rate of patients in clinical studies. The Bayesian adaptive trial design provides better medicine to patients while at the same time more quickly pinpointing the dynamics of the dose-response curve.

	Placebo	Dose 1	Dose 2	Dose 3	Dose 4	Dose 5	Dose 6	Dose 7
Subjects	128	66	66	66	66	66	66	66
Successes	8.7	7	12	28	43	51	55	55
Success Rate	.07	0.10	0.21	0.43	0.65	0.78	0.83	0.84

Table 5.2: Dose allocation and success rates for a generic, traditional clinical trial of type "A" with 590 subjects

BIBLIOGRAPHY

Biswas, A., Hui-Hsin, H., and Wen-Tao, H. (2006), "Covariate-Adjusted Adaptive Designs for Continuous Responses in a Phase III Clinical Trial: Recommendation for Practice," *Journal of Biopharmaceutical Statistics*, 227–239.

Bowman, D. (1998), "Bayesian Bootstrap Procedures in Dose-Response Assessment," *Journal of Agricultural, Biological, and Environmental Statistics*, 370–382.

Chaloner, K. and Verdinelli, I. (1995), "Bayesian Experimental Design: A Review," *Statistical Science*, 273–304.

Chang, M. and Chow, S.-C. (2005), "A Hybrid Bayesian Adaptive design for Dose Response Trials," *Journal of Biopharmaceutical Statistics*, 677–691.

Cheng, T. and Shen, Y. (2006), "Bayesian adaptive designs for clinical trials," *Biometrika*, 633–646.

Crump, K. (2002), "Critical Issues in Benchmark Calculations from Continuous Data," *Critical Revies in Toxicology*, 133–153.

DeBeer (1941), "A Scale for Graphically Determining the Shopes of Dose-Response Curves," *Science*, 521–522.

Dodds, M. and Vicini, P. (2004), "Assessing Convergence of Markov Chain Monte Carlo Simulations in Hierarchical Bayesian Models for Population Pharmacokinetics," *Annals of Biomedical Engineering*, 1300–1313.

Emmens, C. W. (1941), "The dose response relationship for certain principles of the pituitary gland, and of the serum and urine of pregnancy," *Journal of Endocrinology. [J End]*, 194–225.

Englehardt, J. and Swartout, J. (2006), "Predictive Bayesian Microbial Dose-Response Asssessment Based on Suggested Self-Organization in Primary Illness Response," *Risk Analysis*, 543–554.

Han, K., Catalano, P., Senchauduri, A., and Mehta, C. (2004), "Exact Analysis of Dose Response for Multiple Correlated Binary Outcomes," *Biometrics*, 216–224.

Hastings, W. K. (1970), "Monte Carlo Sampling Methods Using Markov Chains and Their Applications," *Biometrika*, 97–109.

Huang, Y., Harris, P., Kirby, S., and Dearden, J. (2002), "Improved approaches to interval estimation of effective doses for a logistic dose-response curve," *Journal of Statistics and Computer Simulation*, 565–584.

Jones, D. A. (1996), *A Bayesian approach to the economic evaluation of health care technologies*, Lippincott-Raven.

Kodell, R. and Chen, J. (1991), "Characterization of Dose-Response Relationships Inferred by Statistally Significant Trend Tests," *Biometrics*, 139–146.

Li, C.-S. and Hunt, D. (2004), "Regression splines for threshold selection with application to a random-effects logistic dose-response model," *Computational Statistics and Data Analysis*, 1–9.

Mukhopadhyay, S. (2000), "Bayesian Nonparametric Inference on the Dose Level with Specified Response Rate," *Biometrics*, 220–226.

O'Hagan, A. and Stevens, J. (2002), "Bayesian methods for design and analysis of cost-effectiveness trials in the evalutation of health care technologies," *Statistical Methods in Medical Research*, 469–490.

R Development Core Team, R. (2003), *R: A Language and Environment for Statistical Computing*, R Development Core Team.

Ramsey, F. (1972), "A Bayesian Appraoch to Bioassay," *Biometrics*, 841–858.

Ratkowsky, D. and Reedy, T. (1986), "Choosing Near-Linear Parameters in the Four-Parameter Logistic Model for Radioligan and Related Assays," *Biometrics*, 575–582.

Sand, S., Rosen, D., and Filipsson, A. (2003), "Benchmard Calculations in Risk Assessment Using Contiuous Dose-Response Information," *Risk Analysis*, 1059–1068.

Simon, R. (1999), "Bayesian Design and Analysis of Active Control Clinical Trials," *Biometrics*, 484–487.

Streibig, J. C. and Kudsk, P. (1993), *Herbicide Bioassays*, CRC Press.

Sun, D. and Tsutakawa, R. (1997), "Bayesian Design for Dose-Response Curves with Penalized Risk," *Biometrics*, 1262–1273.

Tang, M.-L. (2000), "On tests of linearity for dose response data: asymptoic, exact conditinoal and exact unconditional tests," *Journal of Applied Statistics*, 871–880.

Thall, P. and Russell, K. (1998), "A strategy for dose-finding and safety monitoring based on efficacy and adverse outcomes in phase I/II clinical trials," *Biometrics*, 251–264.

Verhulst, P. F. (1838), "Notice sur la loi que la population pursuit dans son accroissement," *Correspondance mathmatique et physique*, 113–121.

Volund, A. (1978), "Application of hte Four-Parameter Logistic Model to Bioassay: Comparison with Slope Ration and Parallel Line Models," *Biometrics*, 357–365.

Whitehead, J., Zhou, Y., Stevens, J., and Blakey, G. (2004), "An Evaluation of a Bayesian Method of Dose Escalation Based on Bivariate Binary Responses," *Journal of Biopharmaceutical Statistics*, 969–983.

Whitehead, J., Zhou, Y., Stevens, J., Blakey, G., Price, J., and Leadbetter, J. (2006), "Bayesian decision procedures for dose-escalation based on evidence of undesirable events and therapeutic benefit," *Statistics in Medicine*, 37–53.

Zhang, W., Sargent, D., and Mandrekar, S. (2005), "An adaptive dose-finding design incorporation both toxicity and efficacy," *Statistics in Medicine*, In Press.

Zhou, Y. and Whitehead, J. (2003), "PRactical Implementation of Bayesian Dose-Excalation Procedures," *Drug Information Journal*, 45–59.

A. APPENDIX

A.1 Case 1: Data from Simulated Clinical Trial of Situation "A"

This section presents all information related to a simulated clinical trial that ended in cap realization. Figure A.1 illustrates the theoretical dose response curve of Situation "A" (black line). The blue diamonds represent the sampled response rate of the patients. The red dots represent the Bayesian posterior fit of the theoretical dose-response curve using the sampled data. The posterior mean results show very good fit of the theoretical response curve despite less-than-perfect data.

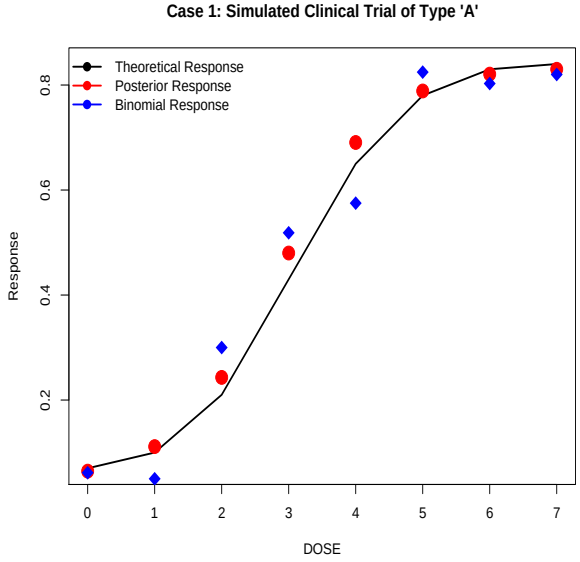

Figure A.1: Case 1: Posterior fit to theoretical response values

Table A.1 presents the data collected for each dose in the simulated trial. The total number of patients and the number of successes at each dose appear in the first two rows. The next row refers to the theoretical response rate of each dose. The Binomial rate is the actual rate calculated from the patients in the clinical trial. The Posterior rates are the response values calculated, using the patient data, by the

Bayesian posterior distributions. The final row refers to the posterior probability that each dose is the ED95 dose.

	Plcbo	Dose1	Dose2	Dose3	Dose4	Dose5	Dose6	Dose7
Patients	130	20	20	27	40	171	142	50
Successes	8	1	6	14	23	141	114	41
Theoretical Rate	.07	.10	.21	.43	.65	.78	.83	.84
Binomial Rate	.06	.05	.30	.52	.58	.82	.80	.82
Posterior Rate	.06	.11	.24	.48	.69	.79	.82	.83
ED95 Probability	0	0	0	0	.01	.41	.54	.05

Table A.1: Case 1: Simulated data and results for each dose in a simulated Situation "A" clinical trial

Table A.1 demonstrates how the Bayesian adaptive approach and the logistic model correctly fit stochastic, binomial data to the theoretical dose-response curve. The patient data in Case 1 create posterior responses that are not exactly equal to theoretical responses. In the theoretical case, the ED95 level is calculated to be .8015, just below Dose 6 (0.83). However, the posterior response rates calculate the ED95 level to be .7915—just .0015 larger than Dose 5 (0.79). The proximity of the calculated ED95 level to the posterior response rate of Dose 5 explains why this clinical trial ended in cap realization. The value, 0.0015, is small enough that Bayesian simulation will often call Dose 5 the ED95 dose. The calculated ED95 probabilities demonstrate this fact. Although 95% of all ED95 designations are split between Doses 5 and 6, the trial stops for cap realization because neither dose has received more than an 80% probability by itself.

Case 1 demonstrates that if a researcher is interested in maximizing the success probability, it is necessary to alter the stopping rule of the Bayesian design. Section A.2 illustrates how such criteria can be changed in the model.

A.2 Model Variations, Maximizing Success Probabilities

A feature of the Bayesian adaptive design is that it can be altered in order to maximize different criteria. In this thesis, efforts have been focused on precise dose-response modeling and on minimizing the Type I Error rate. Because of these goals, some of the clinical situations have resulted in very low success rates and very high cap realization probabilities. In this section we demonstrate how simple alterations of the stopping rules in the Bayesian design can be used to maximize the success rate of a clinical trial and minimize the number of subjects required by the trial.

In this section, three possible alterations to the stopping rule for success are presented. The first variation is a simple change in the probability needed to designate a dose as the ED95 dose; the probability that a dose is ED95 is thus decreased from 0.8 to 0.6. The second variation involves combining the probabilities of adjacent doses. If the combined ED95 probabilities of any two adjacent doses is greater than 0.9, the dose with the largest ED95 probability is designated as the ED95 dose. For the third variation, the algorithm stops for success when less than 10% of the posterior draws demonstrate no ED95 dose. In other words, if the non-normalized ED95 probabilities add to a number greater that 0.9, the dose with the highest ED95 probability is designated as the ED95 dose.

	Success	Type I Error	Futility	Cap Realization	Subjects / Trial
Variation #1	0.587	0.408	0.005	0	256
Variation #2	0.625	0.363	0.005	0.007	244
Variation #3	0.582	0.213	0.001	0.204	483

Table A.2: Alterations to Bayesian stopping rules: 3 ways to maximize success rates and minimize cap realization rates

Table 6.2 reports the results of the three variations to the stopping rule for success. In cases 1 and 2 the algorithm stopped for success more than 99% of the time, though both cases demonstrated high Type I error rates. The average number

of subjects required in clinical simulation is 50% of the number used in previous studies. These simulations successfully demonstrate that the Bayesian design can be altered to maximize various stopping probabilities. Thus, the Bayesian adaptive model demonstrates added versatility and can be applied in studies that have differing clinical objectives.

A.3 R Code

```
## msm library needed for random truncated normal generator ##
install.packages ("msm") library ("msm")

    #~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~]
    # Specifications for Stopping Rules ]
    #~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~]
cap<-600 success<-.80 failure<-.25

################################################################## ##
##  THE 7 SITUATIONS THAT WILL BE TESTED USING THE ALGORITHM      ##

plot(seq(0,7,by=.1),.05+(.8/(1+exp((3.1-seq(0,7,by=.1))/.8)))
    ,ylim=c(0,.85),type="l",ylab="",xlab="",main="A")
a<-round(.05+(.8/(1+exp((3.1-(0:7))/.8))),2)
points(0:7,a,cex=2,col="red",pch=20)
points(6,a[7],cex=4,col="blue",pch=20)
    # ED95: Dose 6

plot(seq(0,7,by=.1),.05+(.8/(1+exp((2.5-seq(0,7,by=.1))/.6)))
    ,ylim=c(0,.85),type="l",ylab="",xlab="",main="B")
b<-round(.05+(.8/(1+exp((2.5-(0:7))/.6))),2)
points(0:7,b,cex=2,col="red",pch=20)
points(5,b[6],cex=4,col="blue",pch=20)
    # ED95: Dose 5

plot(seq(0,7,by=.1),.05+(.8/(1+exp((4.3-seq(0,7,by=.1))/.4)))
    ,ylim=c(0,.85),type="l",ylab="",xlab="",main="C")
c<-round(.05+(.8/(1+exp((4.3-(0:7))/.4))),2)
points(0:7,c,cex=2,col="red",pch=20)
points(6,c[7],cex=4,col="blue",pch=20)
    # ED95: Dose 6

plot(seq(0,7,by=.1),0.05+(.8/(1+exp((1.5-seq(0,7,by=.1))/.6)))
    ,ylim=c(0,.85),type="l",ylab="",xlab="",main="D")
d<-round(0.05+(.8/(1+exp((1.5-(0:7))/.6))),2)
points(0:7,d,cex=2,col="red",pch=20)
points(4,d[5],cex=4,col="blue",pch=20)
    # ED95: Dose 4

plot(seq(0,7,by=.1),.05+(.8/(1+exp((5.6-seq(0,7,by=.1))/.7)))
    ,ylim=c(0,.85),type="l",ylab="",xlab="",main="E")
```

```
e<-round(.05+(.8/(1+exp((5.6-(0:7))/.7))),2)
points(0:7,e,cex=2,col="red",pch=20)
points(7,e[8],cex=4,col="blue",pch=20)
    # ED95: Dose 7

plot(seq(0,7,by=.1),.05+(.8/(1+exp((3.5-seq(0,7,by=.1))/2)))
    ,ylim=c(0,.85),type="l",ylab="",xlab="",main="F")
f<-round(.05+(.8/(1+exp((3.5-(0:7))/2))),2)
points(0:7,f,cex=2,col="red",pch=20)
points(7,f[8],cex=4,col="blue",pch=20)
    # ED95: Dose 7

plot(seq(0,7,by=.1),.4+(.01/(1+exp((3.5-seq(0,7,by=.1))/1)))
    ,ylim=c(0,.85),type="l",ylab="",xlab="",main="G")
g<-round(.4+(.01/(1+exp((3.5-(0:7))/1))),2)
points(0:7,g,cex=2,col="red",pch=20)
    # ED95: DNE


    #~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~]
    # Assign Success Probabilities to Doses for Entire Simulation ]
    #~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~]
dose<-c(0,1,2,3,4,5,6,7) dose.value<-c(0,10,20,40,50,60,80,100)
dose.probs<-c(g) #plot(0:7,dose.probs)
#lines(0:7,mean(beta)+(mean(delta)/
    (1+exp((mean(theta)-(0:7))/mean(tau)))))


        # # # # # # # #
        # MCMC VALUES #
        # # # # # # # #

    ## Logistic model likelihood Function Declared (used in MCMC)
like<-function(beta.func,delta.func,theta.func,tau.func){
    beta.func+(delta.func / (1+ exp((theta.func-dose)/tau.func)))
                                                }

    ### VALUES ###
length<-3000 burn<-50

    ###PRIORS###
tau.a<-1 tau.b<-1

theta.mu<-3.2 theta.var<-1
```

```
beta.a<-1 beta.b<-1

delta.a<-1 delta.b<-1

    ###MCMC SETTINGS###
candsig.tau<-.5 candsig.theta<-.6 candsig.beta<-.1
candsig.delta<-.06



        ############################################################
        ###                                                      ###
        ### OUTER LOOP STARTS HERE                               ###
        ### - mmm is number of outer loop iterations             ###
        ### - why.stop records why each iteration was stopped ###
        ############################################################


mmm<-2500 why.stop<-matrix(0,ncol=3,nrow<-mmm)
colnames(why.stop)<-c("Success","Failure","Cap")



for(j in 1:mmm){



        #Initial batch data found using dose success rates#
n.total<-160 n<-rep(floor(n.total/length(dose)),length(dose))
y<-NULL for(i in 1:length(dose))y<-c(y,rbinom(1,n[i],dose.probs[i]))


        #////////////////////////////////////////////////////#
        ### While loop to stop when a stopping rule is met ###
        #////////////////////////////////////////////////////#

    while(sum(why.stop[j,]) < 1){



    ### Initialize and Clear Parameter Values ###
tau<-numeric(length+burn) theta<-numeric(length+burn)
beta<-numeric(length+burn) delta<-numeric(length+burn)

    ###STARTING VALUES###
```

```
tau[1]<-1 theta[1]<-3.5 beta[1]<-0.05 delta[1]<-0.8

            ########################
            ### INNER LOOP: MCMC ###
            ########################

for (i in 2:(length+burn)){ ###update for beta

beta[i]<-beta[i-1]
old.beta<-beta[i-1]
if(delta[i-1]>=.5)new.beta<-rtnorm(1,old.beta,candsig.beta,0,1-delta[i-1])
    else
if(delta[i-1]<.5)new.beta<-rtnorm(1,old.beta,candsig.beta,0,delta[i-1])

## These calculations disregard the priors because they will =0

    llo<-sum((y * log(like(old.beta,delta[i-1],theta[i-1],tau[i-1]))) +
           ((n-y)* log(1-like(old.beta,delta[i-1],theta[i-1],tau[i-1]))))
   lln<-sum((y * log(like(new.beta,delta[i-1],theta[i-1],tau[i-1]))) +
          ((n-y)* log(1-like(new.beta,delta[i-1],theta[i-1],tau[i-1]))))
    uu<-runif(1,0,1)
        ## These if statements done to keep beta< (delta and 0.5)
 if(delta[i-1]>=.5) accept<-lln+log(ptnorm(1-delta[i-1],old.beta,candsig.beta)
               - ptnorm(0,old.beta,candsig.beta)) - llo -
               log(ptnorm(1-delta[i-1],new.beta,candsig.beta)
                - ptnorm(0,new.beta,candsig.beta)) else
 if(delta[i-1]<.5) accept<-lln + log(ptnorm(delta[i-1],old.beta,candsig.beta)
                   - ptnorm(0,old.beta,candsig.beta))- llo -
                  log(ptnorm(delta[i-1],new.beta,candsig.beta)
                  - ptnorm(0,new.beta,candsig.beta))
    if (log(uu) < accept){beta[i]<-new.beta}

###update for delta
    delta[i]<-delta[i-1]
    old.delta<-delta[i-1]
    new.delta<-rtnorm(1,old.delta,candsig.delta,beta[i],1-beta[i])
    ## These calculations disregard the priors because they will always = 0
        llo<-sum((y * log(like(beta[i],old.delta,theta[i-1],tau[i-1])))
            + ((n-y)* log(1-like(beta[i],old.delta,theta[i-1],tau[i-1]))))
        lln<-sum((y * log(like(beta[i],new.delta,theta[i-1],tau[i-1])))
            + ((n-y)* log(1-like(beta[i],new.delta,theta[i-1],tau[i-1]))))
    uu<-runif(1,0,1)
    acceppt<- lln + log(ptnorm(1-beta[i],old.delta,candsig.delta) -
            ptnorm(beta[i],old.delta,candsig.delta))- llo -
                    log(ptnorm(1-beta[i],new.delta,candsig.delta) -
```

```
                    ptnorm(beta[i],new.delta,candsig.delta))
    if(log(uu) < acceppt){delta[i]<-new.delta}

###update for theta
    theta[i]<-theta[i-1]
    old.theta<-theta[i-1]
    new.theta<-rnorm(1,old.theta,candsig.theta)
        llo<-sum((y * log(like(beta[i],delta[i],old.theta,tau[i-1])))
            + ((n-y) * log(1-like(beta[i],delta[i],old.theta,tau[i-1]))))
            - ((1/(2*theta.var))*((old.theta-theta.mu)^2))
        lln<-sum((y * log(like(beta[i],delta[i],new.theta,tau[i-1])))
            + ((n-y) * log(1-like(beta[i],delta[i],new.theta,tau[i-1]))))
            - ((1/(2*theta.var))*((new.theta-theta.mu)^2))
    uu<-runif(1,0,1)
    if(log(uu) < (lln-llo)){theta[i]<-new.theta}

###update for tau
    tau[i]<-tau[i-1]
    old.tau<-tau[i-1]
    new.tau<-rnorm(1,old.tau,candsig.tau)
    if(new.tau > 0) {
    ### Extra loop needed because it is a gamma rvb and must be > 0 ###
        ## These calculations simplify the prior to be only one variable
        llo<-sum((y * log(like(beta[i],delta[i],theta[i],old.tau)))
            + ((n-y) * log(1-like(beta[i],delta[i],theta[i],old.tau))))
            - old.tau
        lln<-sum((y * log(like(beta[i],delta[i],theta[i],new.tau)))
            + ((n-y) * log(1-like(beta[i],delta[i],theta[i],new.tau))))
            - new.tau
    uu<-runif(1,0,1)
    if(log(uu) < (lln-llo)){tau[i]<-new.tau}
              }
    }



    ##plots to help analyze length/burn and candidate sigma choices
#plot(beta[1:400], type='l') #plot(beta, type='l') #plot(delta,
type='l') #plot(theta, type='l') #plot(tau, type='l')


#mb<-mean(beta) #md<-mean(delta) #mth<-mean(theta) #mt<-mean(tau)
#plot(seq(0,30,.01),(mb+(md/(1+exp((mth-seq(0,30,.01))/mt)))))

#plot(dose.value,dose.probs)
```

```
#plot(0:7,(mb+(md/(1+exp((mth-0:7))/mt))))
#plot(dose,(mb+(md/(1+exp((mth-dose)/mt)))))
#plot(seq(0,20),(.05+(.8/(1+exp((3.1-seq(0,20))/1.1)))),col="red")
#This code means nothing, but my roommate Steve Williams is a really
#   cool guy!


            ##############################
            ##############################
            ###                        ###
            ### ADAPTIVE TRIAL DESIGN ###
            ###                        ###
            ##############################
            ##############################



    #*************************************************************#
    ##  Allocation Scheme #1                                     ##
    ##      Calculates the ED95 dose for every iteration         ##
    #*************************************************************#


    #posterior means of lowest/highest doses in the trial
b<-burn+1 L<-length+burn model<-function(dose){
    beta[b:L]+(delta[b:L] / (1+ exp((theta[b:L] - dose)/tau[b:L])))}
mean.low<-mean(model(0)) mean.hi<-mean(model(7))


    #Overall ED95 Threshold calculated using means of high, low doses
threshold<-mean.low+(.95*(mean.hi-mean.low))


    #Iterative ED95 Dose, each column of "lauren" contains the output
    #  for each iteration from the logistic model for one dose

lauren<-matrix(0,nrow=length,ncol=length(dose))
 for (i in1:length(dose)) {
    lauren[,i]<-model(i-1)}



 #Puts '1' in column of the dsgntd ed95 dose for each iteration, if exists
ed95<-matrix(0,nrow=length,ncol=length(dose)) for (i in 1:length){
```

```
    if(max(lauren[i,]) < threshold) ed95[i,1:8]<-0
    else ed95[i,min(order(lauren[i,])[lauren[i,] >threshold])]<-1}


    #Calculates the prob that each dose is ED95
edprob.unnorm<-apply(ed95,2,sum)/length
edprob<-edprob.unnorm/sum(edprob.unnorm)



    #~~~~~~~~~~~~~~~~~~~~~~~################~~~~~~~~~~~~~~~~~~#
    #                      #Stopping Rules:#
    #                      #################
    #   -puts '1' in column indicating why trial was stopped,
    #       '0' in other two colums
    #   -does this for each iteration in outer loop
    #   -this assumes that the success of the most probably
    #       ED95 >= success of placebo
    #   -cap-20 assures no more than the cap size will be used
    #       if batches <= 20
    #   -uses normalized ED95 probs for succuss
    #       (i.e. p(dose=ED95 | there exists ED95 dose)
    #   -uses unnormalized ED95 probs for failure
    #~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~#

if(max(edprob) > success)
why.stop[j,1]<-order(edprob)[length(dose)]-1 else
    if(max(edprob.unnorm) < failure) why.stop[j,2]<-1 else
    if(sum(n)>(cap-20)) why.stop[j,3]<-1
    #if(sum(n)>160) why.stop[j,3]<-1



### Clclts var of one additl data point at each dose in curve ###

vars<-NULL for(i in 1:length(dose)){
        vars<-c(vars,var(beta + (delta / (1+ exp((theta - (i-1))/tau)))))}
vars<-(n/(n+1)) * vars


    ### Calculates probability of allocation ###
prob.allocate<- edprob*vars / (sum(edprob*vars))
```

```
    #~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~#
    # Next Batch Allocation, with 25% automatically allocated to placebo #
    #~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~#
sam.new<-sample(dose,15,replace=T,prob=prob.allocate)

n.new<-5 y.new<-NULL

for(i in 2:length(dose))
n.new<-c(n.new,length(sam.new[sam.new==i-1]))

for(i in 1:length(dose))
y.new<-c(y.new,rbinom(1,n.new[i],dose.probs[i]))

n<-n+n.new y<-y+y.new

            } #this ends the while loop
            cat("Iteration",j,"\n")
            } #this ends the outer loop

#write(t(why.stop),"g",ncolumns=3)
```

A.4 Tables

The tables in this section represent the average number of subjects allocated to each dose and the average number of sucesses over all 2,500 clinical trial simulations. The results of the 2,500 simulations are reported for each of the seven clincial situations outlined in Section 3. The first number in each cell represents the average and the second number, in parenthesis, represents the standard deviation of each cell.

	Placebo	Dose 1	Dose 2	Dose 3	Dose 4	Dose 5	Dose 6	Dose 7	Total
Avg Patient Successes	8.7(2.9)	2.8(1.9)	4.2(1.4)	8.9(3.3)	37.4(21.7)	108.9(28.9)	132.7(39.2)	38.2(12.0)	341.8(11.5)
Avg Patient Totals	128.0(4.1)	20.0(0)	20.0(0)	20.5(0.8)	56.7(28.9)	140.7(36.1)	159.6(44.4)	46.5(14.1)	592.0(14.1)

Table A.3: Situation "A": Mean number of subjects at each dose and successes at each dose for one simulated clinical trial

	Placebo	Dose 1	Dose 2	Dose 3	Dose 4	Dose 5	Dose 6	Dose 7	Total
Avg Patient Successes	5.5(2.5)	1.7(1.3)	5.9(2.9)	22.0(13.6)	70.6(46.7)	105.0(54.9)	65.0(50.3)	24.1(8.2)	299.8(86.5)
Avg Patient Totals	105.5(34.9)	20.0(0)	20.0(0)	35.4(18.6)	92.3(60.7)	125.3(65.4)	75.6(58.6)	27.9(7.9)	502.0(139.7)

Table A.4: Situation "B": Mean number of subjects at each dose and successes at each dose for one simulated clinical trial

	Placebo	Dose 1	Dose 2	Dose 3	Dose 4	Dose 5	Dose 6	Dose 7	Total
Avg Patient Successes	3.4(3.5)	1.1(0.9)	1.0(0.9)	1.5(1.4)	4.8(1.0)	41.4(24.1)	74.3(60.6)	390(42.6)	166.5(98.9)
Avg Patient Totals	64.5(38.0)	20.0(0)	20.0(0)	20.0(0)	20.1(0.3)	57.1(29.9)	89.1(72.4)	47.2(49.4)	338.0(151.9)

Table A.5: Situation "C": Mean number of subjects at each dose and successes at each dose for one simulated clinical trial

	Placebo	Dose 1	Dose 2	Dose 3	Dose 4	Dose 5	Dose 6	Dose 7	Total
Avg Patient Successes	16.4(5.3)	5.5(1.6)	23.7(15.4)	101.4(33.1)	107.6(45.5)	51.8(30.1)	32.1(24.3)	19.5(5.5)	358.0(61.3)
Avg Patient Totals	117.5(22.4)	20.0(0)	37.4(19.9)	126.0(39.2)	128.6(53.0)	61.3(35.1)	36.5(28.6)	22.7(4.8)	550.0(89.6)

Table A.6: Situation "D": Mean number of subjects at each dose and successes at each dose for one simulated clinical trial

	Placebo	Dose 1	Dose 2	Dose 3	Dose 4	Dose 5	Dose 6	Dose 7	Total
Avg Patient Successes	2.7(1.9)	1.3(0.9)	1.1(1.2)	2.0(0.9)	2.6(1.1)	7.4(2.5)	18.8(10.3)	36.9(14.1)	72.8(25.2)
Avg Patient Totals	36.0(11.9)	20.0(0)	20.0(0)	20.0(0)	20.0(0)	22.2(4.0)	36.3(15.7)	49.5(20.5)	224(47.9)

Table A.7: Situation "E": Mean number of subjects at each dose and successes at each dose for one simulated clinical trial

	Placebo	Dose 1	Dose 2	Dose 3	Dose 4	Dose 5	Dose 6	Dose 7	Total
Avg Patient Successes	14.7(8.1)	3.9(1.1)	6.4(1.9)	8.6(2.1)	10.6(3.9)	23.5(16.9)	53.1(40.1)	82.1(52.2)	202.9(107.4)
Avg Patient Totals	79.0(48.6)	20.0(0)	20.0(0)	20.0(0)	23.8(7.6)	38.6(24.0)	79.4(56.1)	115.2(72.7)	396.0(194.5)

Table A.8: Situation "F": Mean number of subjects at each dose and successes at each dose for one simulated clinical trial

	Placebo	Dose 1	Dose 2	Dose 3	Dose 4	Dose 5	Dose 6	Dose 7	Total
Avg Patient Successes	8.8(2.7)	7.2(1.8)	8.2(1.9)	8.6(1.6)	7.9(2.8)	8.7(2.8)	8.4(2.1)	8.7(2.3)	66.5(7.8)
Avg Patient Totals	20.0(0)	20.0(0)	20.0(0)	20.0(0)	20.0(0)	20.0(0)	20.0(0)	20.0(0)	160(0)

Table A.9: Situation "G": Mean number of subjects at each dose and successes at each dose for one simulated clinical trial