Spring 2013

The application of Bayesian adaptive design and Markov model in clinical trials

Xiaoyu Lu
New Jersey Institute of Technology

ABSTRACT

THE APPLICATION OF BAYESIAN ADAPTIVE DESIGN AND MARKOV MODEL IN CLINICAL TRIALS

by
Xiaoyu Lu

In this research, two new designs in clinical trials are proposed. The first problem is a new Bayesian adaptive dose-finding design and its application in an oncology clinical trial. This design is used for phase IB studies with the biomarker as the endpoint and with the fewer patients. The second problem is another new Bayesian adaptive dose-finding design with longitudinal analysis and its application in phase II depression clinical trial. This design is best fit for phase II dosing-finding clinical trials with clinical endpoints. MTD information has been obtained before the trials.

In adaptive dose-finding clinical trials, the strategy is to reduce the allocation of patients to non-informative doses and also save the trial cost. Bayesian adaptive dose finding design has the ability to utilize accumulating data obtained in real time to alter the course of the trial, thereby enabling dynamic allocation to different dosing groups and dropping of ineffective dosing group earlier. In this research, Bayesian adaptive method is used as a new and useful approach that applies to phase IB and phase II dose-finding clinical trials to evaluate safety and efficacy of the study treatment. Response model and Normal Dynamic Linear Models (NDLMs) are applied in stages 1-4. Conditional probability for each parameter in the model is derived using appropriate prior distributions. Markov Chain Monte Carlo (MCMC) method is used to do the simulation. Model parameters with meaningful prior distributions and the posterior quantities are

obtained to evaluate the trial results and they help to determine the optimal dose level which can be used in later studies. Simulations are done for different scenarios in the two designs and used to validate the model. Five-thousand simulation trials are conducted to verify the repeatability of the results. The posterior probability of success for the trial is greater than 90% based on the simulation results. The results give clearer idea if one needs to go further to test new dose levels based on the thorough evaluation of the existing partial data. Compared with the other adaptive dose finding strategy, the proposed Bayesian adaptive designs are sensitive and robust to help the investigators draw conclusion as early as possible. The designs can also reduce sample size substantially which in turn leads to savings in cost and time.

Continuous-time Markov model has the advantage over the traditional survival model and can be used to describe disease as a series of probable transitions between health states. This is an attractive feature since it provides the ability to describe the course of disease over time. It can also describe and estimate expected survival in clinical cohort. In this research, continuous-time Markov model is used in the time-to-event analysis in a phase II oncology trial. Six states are defined in the Markov chain which is used in time to progression analysis for 36 patients with neuroendocrine carcinoma. The transition probability matrix P is defined and used to iterate future transition and survival probabilities. The estimate from matrix analysis shows that the results are reliable and comparable with the Kaplan-Meier results.

THE APPLICATION OF BAYESIAN ADAPTIVE DESIGN AND MARKOV MODEL IN CLINICAL TRIALS

by
Xiaoyu Lu

A Dissertation
Submitted to the Faculty of
New Jersey Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Mathematical Sciences

Department of Mathematical Sciences, NJIT
Department of Mathematics and Computer Science, Rutgers-Newark

May 2013

APPROVAL PAGE

THE APPLICATION OF BAYESIAN ADAPTIVE DESIGN AND MARKOV MODEL IN CLINICAL TRIALS

Xiaoyu Lu

Dr. Sunil K. Dhar, Dissertation Advisor — Date
Associate Professor, Mathematical Sciences, NJIT

Dr. Steven Sun, Committee Member — Date
Director, Oncology Biostatistics, Johnson & Johnson

Dr. Sudhakar Rao, Committee Member — Date
Sr. Director, Oncology Biostatistics, Johnson & Johnson

Dr. Sundarraman Subramanian, Committee Member — Date
Associate Professor, Mathematical Sciences, NJIT

Dr. Wenge Guo, Committee Member — Date
Assistant Professor, Mathematical Sciences, NJIT

BIOGRAPHICAL SKETCH

Author: Xiaoyu Lu

Degree: Doctor of Philosophy

Date: May 2013

Undergraduate and Graduate Education:

- Doctor of Philosophy in Mathematical Sciences,
  New Jersey Institute of Technology, Newark, NJ, 2013

- Master of Science in Biomedical Engineering,
  Louisiana Tech University, Ruston, LA, 2000

- Master of Science in Biochemical Engineering,
  Nanjing University of Technology, Nanjing, P. R. China, 1995

- Bachelor of Science in Biochemical Engineering,
  Nanjing University of Technology, Nanjing, P. R. China, 1992

Major: Applied Statistics

Presentations and Publications:

Xiaoyu Lu and Sunil Dhar,
"The Application of Bayesian Adaptive Design in Clinical Trial",
American Journal of Mathematics and Statistics, Vol. 3, March 2013.

To my beloved parents,
to my dear husband,
my lovely kids
and my parents in law
非常感谢我的父母，家庭和我可爱的孩子对我的支持！

ACKNOWLEDGMENT

First, I would like to express my sincere gratitude to Dr. Sunil Dhar, who helped me a lot with the research. He patiently provided the vision, encouragement and advice necessary for me to proceed through the completion of my dissertation. Nothing would have been possible without his constant support.

I'd also like to give my heartfelt, special thanks to my committee members, Dr. Steven Sun, Dr. Sudhakar Rao for their strong support and guidance and helpful suggestions. Their extensive experience in the industry helped me a lot.

Thirdly, I am very grateful to the remaining members of my dissertation committee, Dr. Wenge Guo, Dr. Sundarraman Subramanian. Their support are greatly appreciated. Thank you.

I also wish to thank my family for all your encouragement, love and support throughout the years. Nothing would happen without the support from you.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

TABLE OF CONTENTS
(Continued)

LIST OF TABLES

LIST OF TABLES
(Continued)

LIST OF FIGURES

LIST OF FIGURES
(Continued)

Figure **Page**

CHAPTER 1

INTRODUCTION

1.1 Objective

This research reports the new design, implementation, and outcome of the Bayesian adaptive, dose-ranging trials incorporating the innovative dose finding approach to flexibly address both efficacy and safety aspect of the drug. Two designs are described and simulated. The first one is a new Bayesian adaptive dose-finding design and its application in an oncology clinical trial. This design is applied to a phase IB study with the biomarker as the endpoint and with the fewer patients (Chapter 2, 3, 4). The second one is another Bayesian adaptive dose-finding design and its application in depression clinical trial (Chapter 5). This design is applied to a phase II dosing-finding clinical trial with clinical endpoints. The last problem in this research is the application of Markov models in time to event analysis (Chapter 6).

1.2 Background Information of Bayesian Adaptive Design

Clinical trials involving new drugs are commonly classified into four phases, phase I to IV. Sometimes phase 0 trials are conducted. Phase 0 trials are first-in-human trials. Single sub-therapeutic doses of the study drug are given to a small number of subjects (10 to 15) to gather preliminary data on the agent's pharmacodynamics (PD) and pharmacokinetics (PK). PD analysis shows what the drug does to the body PK analysis shows what the body does to the drugs. Phase I studies are usually conducted in healthy volunteers. The goal here is to determine what the drug's most frequent side effects are and, often, how

the drug is metabolized and excreted. The number of subjects typically ranges from 20 to 80. Sometimes, phase I studies were conducted by two stages: phase IA and phase IB. Phase IA is mainly focused on safety analysis and phase IB is mainly focused on dose-finding based on biomarker analysis. Phase II studies begin if Phase I studies don't reveal unacceptable toxicity. While the emphasis in Phase I is on safety and biomarker analysis, the emphasis in Phase II is on effectiveness. This phase aims to obtain preliminary data on whether the drug works in people who have a certain disease or condition. For controlled trials, patients receiving the drug are compared with similar patients receiving a different treatment--usually an inactive substance (placebo), or a different drug. Safety continues to be evaluated, and short-term side effects are studied. Typically, the number of subjects in Phase II studies ranges from a few dozen to about 300. Phase III studies begin if evidence of effectiveness is shown in Phase II. These studies gather more information about safety and effectiveness, studying different populations and different dosages and using the drug in combination with other drugs. The number of subjects usually ranges from several hundred to about 3,000 people. Phase IV trials are post-marketing studies delineate additional information, including the treatment's risks, benefits, and optimal use [1].

Clinical trials are research studies to find better ways to treat patients with the selected product. Depending on the type of testing product and the stage of its clinical trial development, investigators initially enroll volunteers and/or patients into small pilot studies, and subsequently conduct larger scale studies applied to patients. The clinical trials compare the safety and efficacy of the new product with others that have already

been approved for the affliction of interest or sometimes compare the new product with the placebo, which is a simulated or otherwise medically ineffectual treatment for a disease. A full series of clinical trials may incur sizable costs.

A dose-finding study is a clinical trial where different doses of an agent (e.g., a drug) are tested against each other to establish which dose works best and/or is least harmful. A correct dose-finding study is of the utmost importance during clinical development of a new drug. It must define the no-effect dose and the mean effective and maximal effective doses. Then taking tolerability into account, the optimal therapeutic dose range can be selected. The purpose of the dose finding trials is to find a dose of treatment that is optimal with respect to simple criteria: Toxicity, efficacy and low risk of side effects. There are many possible dose optima: Minimum effective dose, maximum non-toxic dose, maximum tolerated dose, ideal therapeutic dose. To define the dosage schedule the duration of action in human being must be tested, if possible together with blood concentration measurements. An adequate dose-finding study shows the optimal doses for trials in Phase I or II, thereby saving time and effort and reducing the number of patients required.

The main goal of a dose-finding study is to estimate the response versus dose given, so as to analyze the efficacy and safety of the drug. Although such a response will nevertheless be available from phase III or phase IV trials, it is important to carry out dose-finding studies in the earlier phase I or phase II stages. The main reason for this is to avoid trials in the later phases using doses that are significantly different from those that will subsequently be recommended for clinical use and also to avoid the need for

modification of dosing schedules at later stages where a large amount of data has already been accumulated for a different dose range [2].

Typically, a dose finding study includes a placebo group of subjects (or a control group), and a few groups that receive different doses of the test drug. For instance, a typical dose-finding study may include four groups: a placebo group, low-dose group, medium-dose group and a high-dose group. The maximum tolerable dose (MTD) information is necessary to be able to design such groups and therefore dose-ranging studies are usually designed after the availability of MTD information. The tendency of clinical experts to try to demonstrate superiority of one drug over another by using doses higher than patients really need must be resisted. The price paid in poor tolerability exceeds any potential benefits [3].

Adaptive designs in clinical trials are schemes for patient assignment to treatment, the goal of which is to place more patients on the better treatment based on patient responses already accrued in the trial. In an adaptive dose-finding study, the dose assignment(s) to the next subject, or next cohort of subjects, is based on responses of previous subjects, and the dose assignment is chosen to maximize the information about the dose–response curve, according to some pre-defined objective metric (for example, minimum variability in parameter estimates). In a traditional dose-finding trial, selecting a few doses may not adequately represent the dose–response relationship and many patients will be allocated to 'non-informative' doses (wasted doses), as shown in Figure 1.1 [4]. In adaptive dose-finding, the strategy is to initially include only a few patients on some doses to explore the dose–response, then to allocate the dose range of interest to

more patients. This reduces the allocation of patients to non-informative doses. Compared with fixed randomization, this approach has the ethical advantage that fewer subjects are assigned doses that are too high or too low. It can also avoid additional separate trials that might be necessary when fixed dose-finding trials do not adequately define the dose range.

Adaptive dose-finding trials also require an infrastructure that allows the rapid communication of responses from trial sites to a central un-blinded analysis center and of adaptive dose assignments to the trial sites. Randomization software capable of rapidly computing dynamic allocation of doses to subjects is additionally mandated by adaptive trials because pre-specified randomization lists will not work. In addition, a flexible drug-supply process is required because demand for doses is not fixed in advance, but rather evolves as information on responses at various doses is gathered as the trial progresses [3].

There are several facts about the adaptive design that one needs to follow before the start of the adaptive dose-finding studies:

- Thorough upfront planning
- Decision rules for adaption are pre-specified.
- Ensure more accurate and faster decision-making on dosing selection.
- Emphasis on modeling/estimation as opposed to hypothesis testing.

The following is a list of some existing adaptive dose-finding methods [5]:

- ANOVA – Conventional method based on pairwise comparisons and multiplicity adjustment (Dunnett) of identify dose response (DR). This is a common approach used in dose finding studies

- MCP-Mod – Combination of multiple comparison procedure (MCP) to identify presence of dose response, and modeling to estimate target dose(s) and dose response profile [6]
- MTT: novel method based on Multiple Trend Tests [5]
- BMA: Bayesian Model Averaging [7]
- LOCFIT: Nonparametric local regression fitting [8]

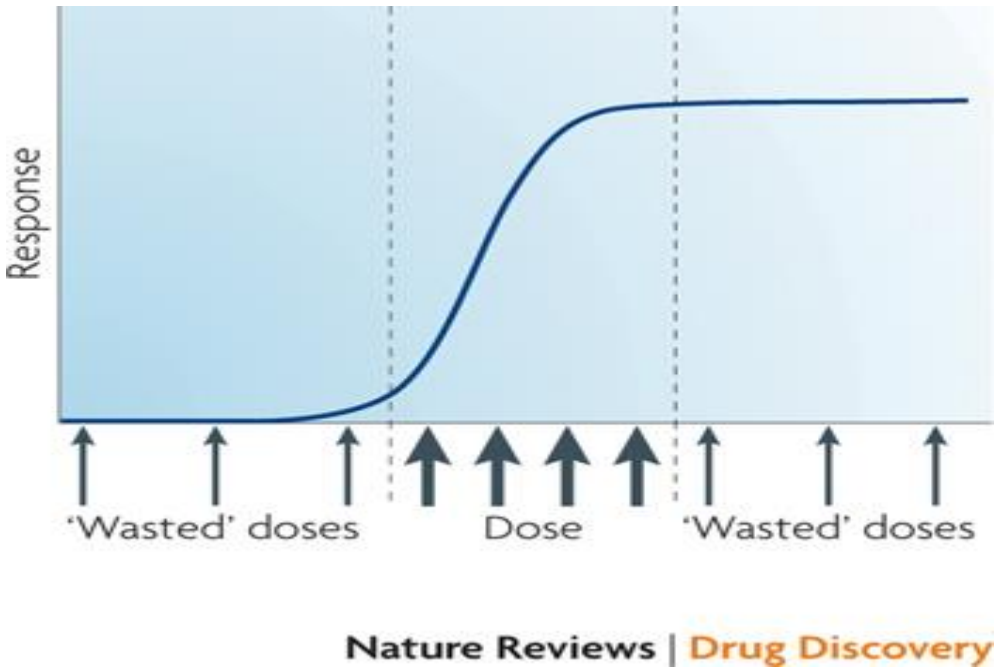


Figure 1.1 Drug responses vs. dose level. [4]

1.3 Background Information of Continuous-Time Markov Model

The term Markov chain is used to describe a process observed at discrete intervals. A Markov process describes a process observed continuously. In the application of clinical trial, Markov process is really a continuously-time process, only it is observed at discrete intervals [9].

In probability theory, a continuous-time Markov chain (CTMC) is a stochastic process $\{X(t) : t \geq 0\}$ that satisfies the Markov property and takes values from a set

called the state space [10]; it is the continuous-time version of a Markov Chain. The Markov property states that at any times $s > t > 0$, the conditional probability distribution of the process at time s given the whole history of the process up to and including time t, depends only on the state of the process at time t. In effect, the state of the process at time s is conditionally independent of the history of the process before time t, given the state of the process *at* time t. In simple terms the process can be thought of as memory-less [11].

Many clinical studies involve complex changes, for example, relapse, recurrence, remission, progress and death. The time-to-event analysis is usually done by using Kaplan-Meier methods. But there are some limitations of Kaplan-Meier method: 1) It is inadequate to describe the complexity of disease beyond two simple states: event or no event 2) it is not able to describe disease as a series of probable transition between health states. It is restricted by assumptions of non-informative censoring and limit the description of disease to permanent transition from one state (e.g., alive) to another (e.g., dead). 3) Kaplan-Meier estimator is sensitive to patient lost to follow-up, event happen due to other cause rather than disease. A CTMC is a stochastic process that has been used to describe disease treatment processes during the past. Marshall and Jones used Markov model to evaluate and describe diabetic retinopathy [12]. Schaubel used Markov model in renal disease analysis [13], Hendriks used Markov model in papilloma virus and human immunodeficiency virus [14]. In this research, continuous-time Markov chain model will be used to explore time-to-event data in oncology clinical trial. The model can be used to estimate expected median survival time and time to progression for the patients. It can

also be a potential powerful exploratory analysis method which can be applied to different disease treatments.

CHAPTER 2

A NEW DOSE-FINDING DESING USING BAYESIAN STATISTICS AND MARKOV CHAIN MONTE CARLO (MCMC)

2.1 Bayesian Statistical Method and its Application in Clinical Trials

Statistical thinking has had a central role in raising the scientific standards of clinical research over the last two centuries, especially during the past 50 years. A major reason has been the appreciation of statistical inference by drug- and medical-device-regulatory agencies. Traditional frequentist statistics has had the dominant, and often exclusive, role in this scientific renaissance. The greatest virtue of the traditional approach maybe its extreme rigor and narrowness of focus to the experiment at hand, but a side effect of this virtue is inflexibility, which in turn limits innovation in the design and analysis of clinical trials. Because of this, clinical trials tend to be overly large, which increases the cost of developing new therapeutic approaches, and some patients are unnecessarily exposed to inferior experimental therapies. Owing to such issues, there is increasing interest in Bayesian methods in clinical trials. Advances in computational techniques and power are also facilitating the application of these methods [15].

Bayesian statistics is the subset of the entire field of statistics in which the evidence about the true state of the world is expressed in terms of degrees of belief or, more specifically, Bayesian probabilities [16]. Bayesian inference is an approach in statistical inference, which is distinct from the more traditional frequentist inference. It is specifically based on the use of Bayesian probabilities to summarize evidence [17]. The formulation of statistical models for use in Bayesian statistics has the additional feature, not present with other types of statistical techniques, of requiring the formulation of a set

of prior distributions for any unknown parameters. Such prior distributions are as much part of the statistical model as the part that expresses the probability distribution of observations given the model parameters. It is usually carried out in the following steps [18]:

a) Choose a probability density function $f(\theta)$ as prior distribution that express our beliefs about a parameter θ before we see any data.
b) Choose a statistical model $f(x|\theta)$ that reflects our beliefs about X given θ.
c) After observing data $X_1, \ldots. X_n$, we update our beliefs and calculate the posterior distribution $f(\theta|X_1, \ldots X_n)$.

The usual considerations in the design of clinical trials are extended in the case of Bayesian design of clinical trials to include the influence of prior beliefs. Importantly, the application of sequential analysis techniques allows the outcome of earlier stages in the trial to influence the design of the next stage of the trial, based on the updating of beliefs as expressed by the prior and posterior distribution. Part of the problem of the design of clinical trials is that they should make good use of resources of all types: Bayesian design of clinical trials is used to aim at such efficiency [19].

In the design proposed, prior distributions of the clinical trial efficacy response mean $f(\theta)$, error term $f(\varepsilon)$ are assumed. The statistical model $f(x|\theta)$ is derived based on our beliefs about X given θ. After observing efficacy response data $X_1, \ldots. X_n$, the sample mean of the efficacy response is obtained and posterior distribution of the efficacy response is observed.

2.2 Markov Chain Monte Carlo (MCMC) and its usage in Bayesian Statistics

Markov chain Monte Carlo (MCMC) methods (which include random walk Monte Carlo methods) are a class of algorithms for sampling from probability distributions based on constructing a Markov chain that has the desired model as its equilibrium distribution. The state of the chain after a large number of steps is then used as a sample of the desired distribution. The quality of the sample improves as a function of the number of steps.

MCMC method is a general simulation method for sampling from posterior distributions and computing posterior quantities of interest. The most common application of MCMC algorithms is numerically calculating multi-dimensional integrals. In these methods, an ensemble of "walkers" moves around randomly. At each point where the walker steps, the integrand value at that point is counted towards the integral. The walker then may make a number of tentative steps around the area, looking for a place with reasonably high contribution to the integral to move into next. Random walk methods are a kind of random simulation or Monte Carlo method. A Markov chain is constructed in such a way as to have the integrand as its equilibrium distribution [19].

A major limitation towards more widespread implementation of Bayesian approaches is that obtaining the posterior distribution often requires the integration of high-dimensional functions. This can be computationally very difficult. In this paper, MCMC methods, which attempt to simulate direct draws from some complex distribution of interest, is used. MCMC approaches are used because one uses the previous sample values to randomly generate the next sample value, generating a Markov chain (as the

transition probabilities between sample values; these probabilities are only a function of the most recent sample value)[19].

2.3 Gibbs Sampling

Gibbs sampling, is a special case of the MCMC methods. It is particularly well-adapted to sampling the posterior distribution of a Bayesian network. The point of Gibbs sampling is that given a multivariate distribution it is simpler to sample from a conditional distribution than to marginalize by integrating over a joint distribution. Suppose one want to obtain k samples of $X=\{X_1, X_2, \ldots, X_n\}$ from a joint distribution $p(X_1, X_2, \ldots, X_n)$. It can be achieved as follows:

a) Begin with some initial value $X^{(0)}$ for each variable.
b) For each sample $i \in \{1,\ldots k\}$, sample each variable $X_j^{(i)}$ from the conditional distribution $p(X_j^{(i)} | X_1^{(i)},\ldots, X_{j-1}^{(i)}, X_{j+1}^{(i-1)} \ldots, X_n^{(i-1)})$. That is, sample each variable from the distribution of that variable conditioned on all other variables, making use of the most recent values and updating the variable with its new value as soon as it has been sampled.

The marginal distribution of any subset of variables can be approximated by simply examining the samples for that subset of variables, ignoring the rest. In addition, the expected value of any variable can be approximated by averaging over all the samples [20].

In this research, Bayesian statistical method is used in an adaptive dose-finding clinical trial and Gibbs sampling method are used for the simulation.

2.4 A New Proposed Bayesian Adaptive Dose-finding Design

This research reports the new design, implementation, and outcome of a Bayesian adaptive, dose-ranging trial incorporating an innovative dose finding approach to flexibly address both efficacy and safety aspect of the drug. A four-stage Bayesian adaptive design is proposed for a dose-finding study treating breast cancer patients.

The fluorodeoxyglucose positron emission tomography (FDG-PET) is a widely used biomarker which is most commonly known in cancer diagnosis and is used as the method to measure the efficacy response.

PET is a non-invasive diagnostic tool that provides tomographic images and quantitative parameters of perfusion, cell viability, proliferation and/or metabolic activity of tissues. These images result from the use of different substances of biological interest (sugars, amino acids, metabolic precursors, hormones) labeled with positron emitting radioisotopes (PET radiopharmaceuticals).

FDG is an analogue of glucose and is taken up by living cells via the first stages of normal glucose pathway. The rationale behind its use as a tracer for cancer diagnosis depends on an increased glycolytic activity in neoplastic cells. FDG is trapped into the cancer cells due to their high glycolytic activity and excreted from the body through the renal system, which is unable to reabsorb the tracer. A 50-60 minute interval between FDG administration and image scan is usually enough to obtain a good tumour/background ratio of the tracer. Figure 2.1 [21] shows the transport and metabolism of FDG in neoplastic cells. Following facilitated transport by Glut-1(k_1), FDG is phosporylated by hexokinase(k_3). FDG-6-phosphate can neither undergo further

metabolism nor diffuse out of cells. As the dephosphorylation(k_4) reaction also occurs slowly, FDG-6-phosphate is trapped intracellularly and accumulates.

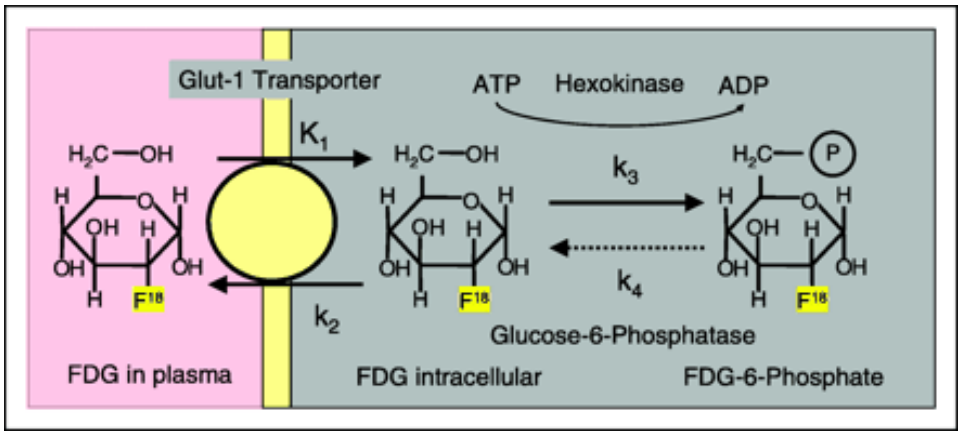


Figure 2.1 Transport and metabolism of FDG. [21]

The cell alterations related to neoplastic transformation are associated with functional impairments that are discernible before structural alterations occur. Therefore, FDG-PET can reveal the presence of a tumor when conventional morphological diagnostic modalities (i.e., X-ray, CT, MRI and ultrasound) do not yet detect any evident lesions.

FDG uptake in tumors correlates with tumor growth and viability, so the PET scan and the possible metabolic quantification may provide useful information about tumor characterization, patient prognosis, and monitoring of the response to anticancer therapy. At present there is considerable evidence that the application of FDG-PET is becoming more and more widespread for the diagnostic assessment of patients with suspected malignancies, in tumor staging, and in therapy monitoring [22]. Moreover, a reduction in the FDG-PET signal within days or weeks of initiating therapy significantly correlates

with prolonged survival and other clinical endpoints now used in drug approvals. These findings suggest that FDG-PET could facilitate drug development as an early surrogate of clinical benefit [21].

FDG accumulation was measured using the standardized uptake value (SUV) as follows [23]:

$$SUV = \frac{Radioactivity\ Concentration\ in\ ROI\ [Bq\ cm^{-3}]}{injected\ dose\ [Bq]/weight\ of\ the\ patient\ (g)} \tag{2.1}$$

Where ROI is Region of Interest, Radioactivity concentration in the ROI was determined as the maximum average radioactivity concentration in the tissue at 55 to 60 min post-injection, corrected for calibration and decay.

A clinical trial endpoint is defined as a measure that allows us to decide whether the null hypothesis of a clinical trial should be accepted or rejected. In this example, a successful efficacy endpoint/response is defined as a subject has ≥ 20% decrease on the sum of FDG-PET uptake SUVmean and SUVmax at 7 days post-dose compared to pre-dose. A successful safety response is evaluated by adverse event and laboratory values.

It is expensive to use FDG-PET scan comparing with CT/MRI in the clinical trial. The fundamental goal of the proposed adaptive design used in this trial is to reduce the sample size, to find the optimal dose efficiently, so as to save the cost and also avoid too many subjects to be exposed to wasted doses. The efficiency of this approach is increased by the use of frequent interim analysis of accumulating data. In the trial, subjects are assigned to 5 treatment groups (corresponding to 2.5 mg, 5 mg, 10 mg, 25 mg, 50 mg

doses) and three interim analyses are performed during each stage. The use of Bayesian approach produces predictive probabilities for success in Phase III. It also yields a transparent analysis that supports quantitative decision making. The design allows the range of doses to be adaptively expanded either up or down.

Figure 2.2 shows the four-stage design with the maximum possible number of subjects assigned in each treatment group by stage.

Stage 1:

In stage 1, only 10 mg dose are evaluated. Eight subjects are assigned into 10 mg group in this stage. The reason of choosing eight subjects is to make the final maximum sample size for each group to be 12. Twelve subjects per arm are calculated by sample size calculation method using historical data.

Three interim analyses are performed in each stage. For the first interim analysis, four subjects are assigned to 10 mg group. After the subjects taking the dose, the efficacy response, the sum of the SUVmean and SUVmax of FDG-PET uptake, are measured and the safety data are recorded. A successful response is defined as a subject has ≥ 20% decrease on the FDG-PET uptake at 7 days post-dose compared to pre-dose, is a clinical meaningful response in treated group.

$$\Pr(0.2 \leq \theta_d \mid \text{Data}), \tag{2.2}$$

θ_d refers to the percent of decrease on the sum of SUVmean and SUV max of FDG-PET uptake for dose d.

Dose Group	Stage 1	Stage 2	Stage 3	Stage 4	Total
50 mg			8	4	12
25 mg		8	4		12
10 mg	8	4			12
5 mg		8	4		12
2.5 mg			8	4	12
Total	8	20	24	8	60

Figure 2.2 Maximum possible numbers of subjects in each treatment group by stage.

Let θ_d be the mean response to dose d. d = 1 - 5 for each dose level in the ascending order. The probability of having $0.2 \leq \theta_d$ will be evaluated. The posterior quantities will be calculated and utilized. In the first stage, θ_1 is used for the mean dose response.

As the next step for the second interim analysis, two more subjects will be added into 10 mg group. The posterior quantities will be calculated again using expanded data. If Pr $(0.2 \leq \theta_1 \mid \text{Data}) < 0.2$ for these two consecutive analyses, this dose level will be declared as futility and the study will move onto the next stage. Otherwise, two more subjects (the 7th to 8th subject for 10 mg arm) would be added. The posterior quantities will be calculated again.

Given the results of those three interim analyses on efficacy, the efficacy of 10 mg group would be evaluated using the following criteria:

- If Pr $(0.2 \leq \theta_1 | \text{Data}) < 0.2$ for any of two consecutive analyses, this dose level will be declared as futility.
- Otherwise, the dose level will be declared as non-futility.

In any case, if the dose level has safety concern, the higher dose levels would not be tested in the next stage. Similarly, if the dose level is futile, the lower dose levels would not be tested in the next stage.

If Pr $(0.2 \leq \theta_1 \mid \text{Data}) \geq 0.8$ for any of two consecutive analyses, this dose level will be declared as effective. In this case, stage 1 will be ended early and the trial will enter stage 2.

Stage 2:

One of the following four actions would be taken in stage 2 based on the results in stage 1:

1. If the safety is good and the efficacy is non-futile, then the next higher and the next lower dose groups (i.e., 5 mg and 25 mg) will be assessed in stage 2.
2. If the safety is not good and the efficacy is non-futile, then the next lower dose group (i.e., 5 mg) will be assessed in stage 2.
3. If the safety is good and the efficacy is futile, then the next higher dose group (i.e., 25 mg) will be assessed in stage 2.
4. If the safety is not good and the efficacy is futile, then four more subjects will be added into 10 mg group for re-evaluation. The efficacy and safety will be re-evaluated using expanded data with the criteria a) - c) above. If the safety is still not good and the efficacy is still futile, then the trial will be ended. Otherwise, the study will enter stage 2 without anyone to be randomized into 10 mg group in the next stage.

Although there may still be some subjects to be assigned into 10 mg group in stage 2, 10 mg group will not be focused in this stage, since it has been tested previously.

If there were 12 subjects assigned into 10 mg group in stage 1 due to re-evaluation in case 4, there would be no subjects to be assigned into 10 mg in stage 2.

If 5 mg and/or 25 mg will be tested, the randomization ratios between the new dose levels in stage 2 (i.e., 5 mg and/or 25 mg) and the old dose levels in stage 1 (i.e., 10 mg) would be 2:1. Totally eight subjects in each new dose level and four subjects in each old level will be randomized in stage 2. Similar to the procedures in stage 1, those eight subjects in the new dose levels will be randomized in three steps by 2:1:1. Three interim analyses would be done to evaluate the efficacy at the end of each step.

Normal dynamic linear models (NDLMs) will be used in stage 2 and all later stages to borrow information across adjacent doses.

Stage 3:

Similar to stage 2, the other new dose levels (2.5 mg and 50 mg) may be tested according to the analyses in the previous stages. In stage 3, totally eight subjects in each new dose level and four subjects in each levels in the previous stage, will be assigned, see Figure 2.2.

Stage 4:

If a dose level in stage 4, either 2.5 mg group or 50 mg group, is good in safety and non-futile in efficacy, four more subjects will be assigned into that dose group to make the total number of subjects to be twelve in each of these dose groups. The final tested dose level will have twelve subjects in order to be considered adequate to evaluate both efficacy and safety assessments.

During the course of the trial, the dose response curve should be monitored. In case no significant response changes between the two doses have been observed or the posterior quantity of the higher dose is less than that of the lower dose, the trial should not go to the higher dose and the threshold of the response curve is assumed. The non-significant response changes can be defined as:

- Posterior probabilities > 0.75
- Difference of two posterior means for two adjacent doses is less than 0.05 or posterior mean of the higher dose is less than the lower dose.

Figure 2.2 shows the maximum possible number of subjects assigned in each treatment group by stage. This happens when safety is good and efficacy is acceptable in all stages. Total sixty patients are exposed to the treatment and the acceptable dose ranges are from 2.5 mg - 50 mg.

Figures 2.3 - 2.5 shows some other scenarios of the design. Figure 2.3 shows the scenario that the efficacy is not acceptable but safety is good in stage 1. So the dose goes up to 25 mg in the stage 2. The efficacy is acceptable and safety is good in stage 2 and 3. Total forty-eight patients are exposed to the treatment and the optimal dose range is 10 mg - 50 mg.

Figure 2.4 shows the scenario that the safety is not good, but the efficacy is acceptable in stage 1. Dose goes down to 5 mg in stage 2. The efficacy is acceptable and safety is good in stages 2 and 3. A total of forty-eight patients are exposed to the treatment and the optimal dose range is 2.5 mg – 10 mg.

Figure 2.5 shows the scenario that both safety and efficacy are not acceptable in stage 1. The trial ended in stage 1. Eight subjects are tested in stage 1, since Pr $(0.2 \leq \theta_1|$

Data) < 0.2, so four additional subjects in 10 mg arm and are added. If the safety is still not good and the efficacy is still futile, then the trial will be ended. Sixteen patients are exposure to the treatment. From Figures 2.3 - 2.5, one can see that less patients are put into trial when the safety issues or efficacy futility are identified earlier and sample size are adjusted due to the early detection.

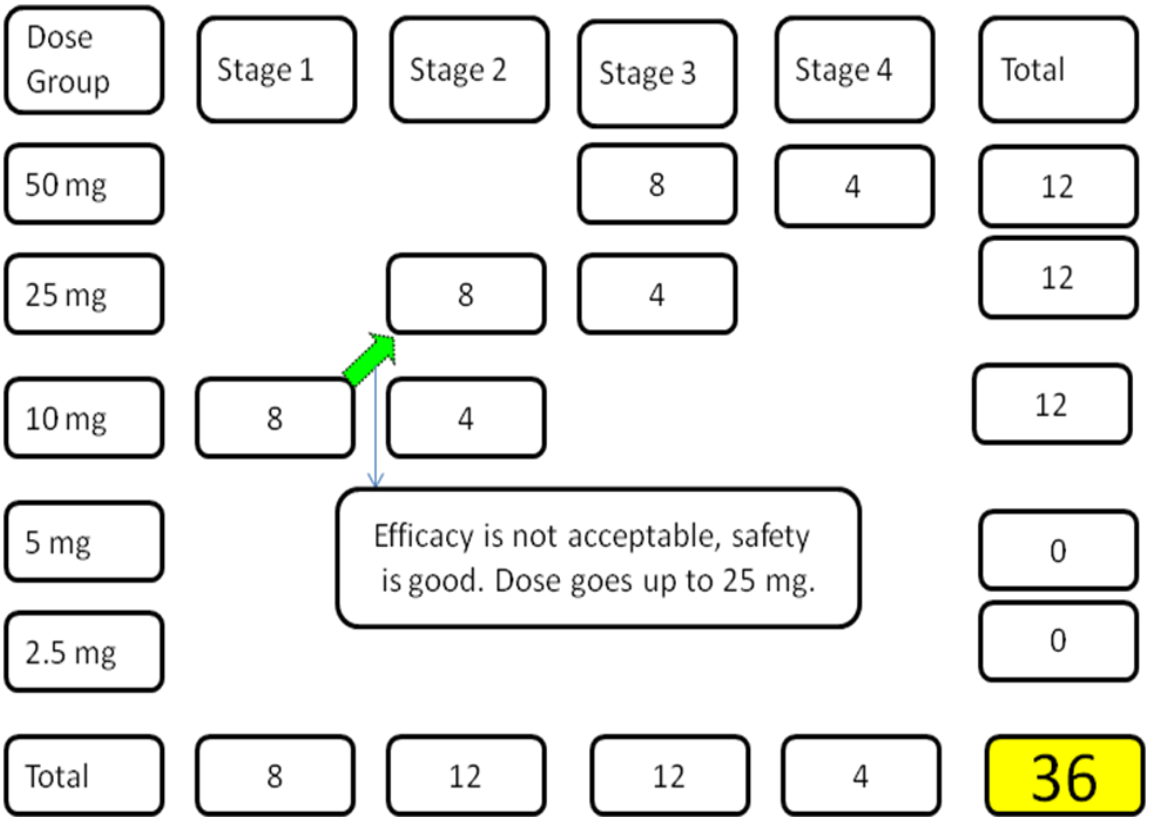


Figure 2.3 Example scenario of the design – efficacy is not acceptable and safety is good in stage 1.

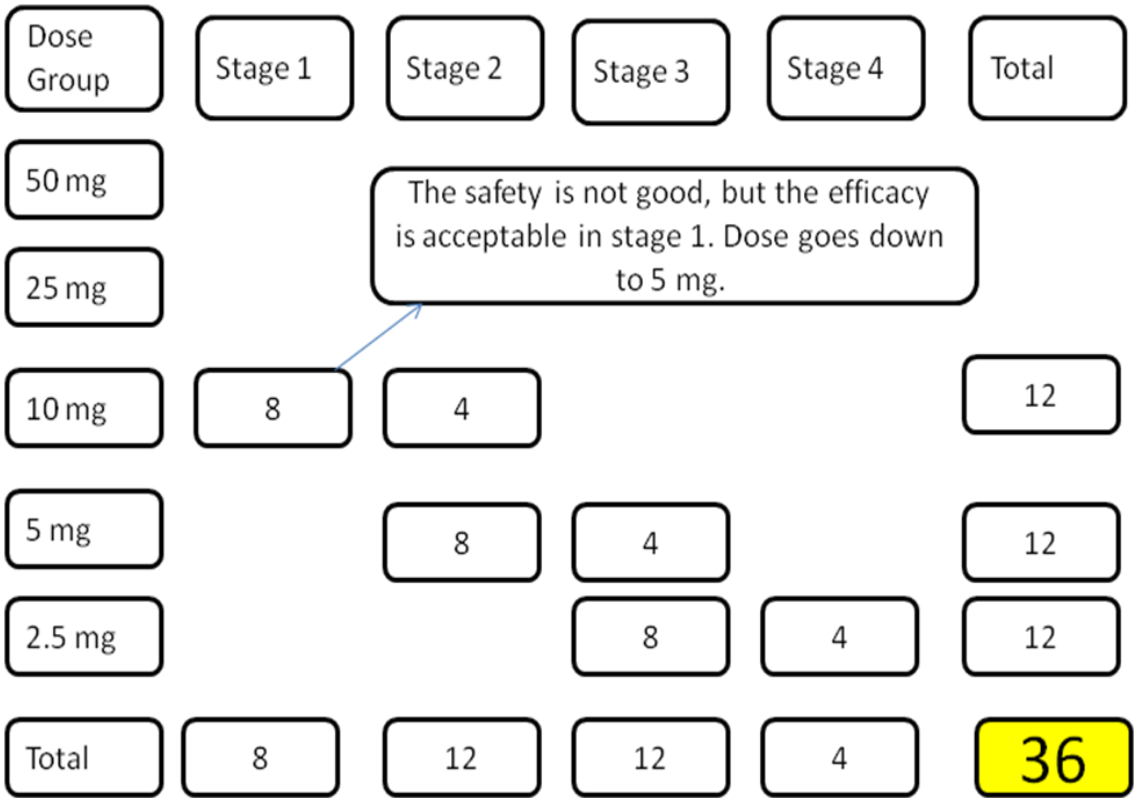


Figure 2.4 Example scenario of the design – safety is not acceptable but the efficacy is good in stage 1.

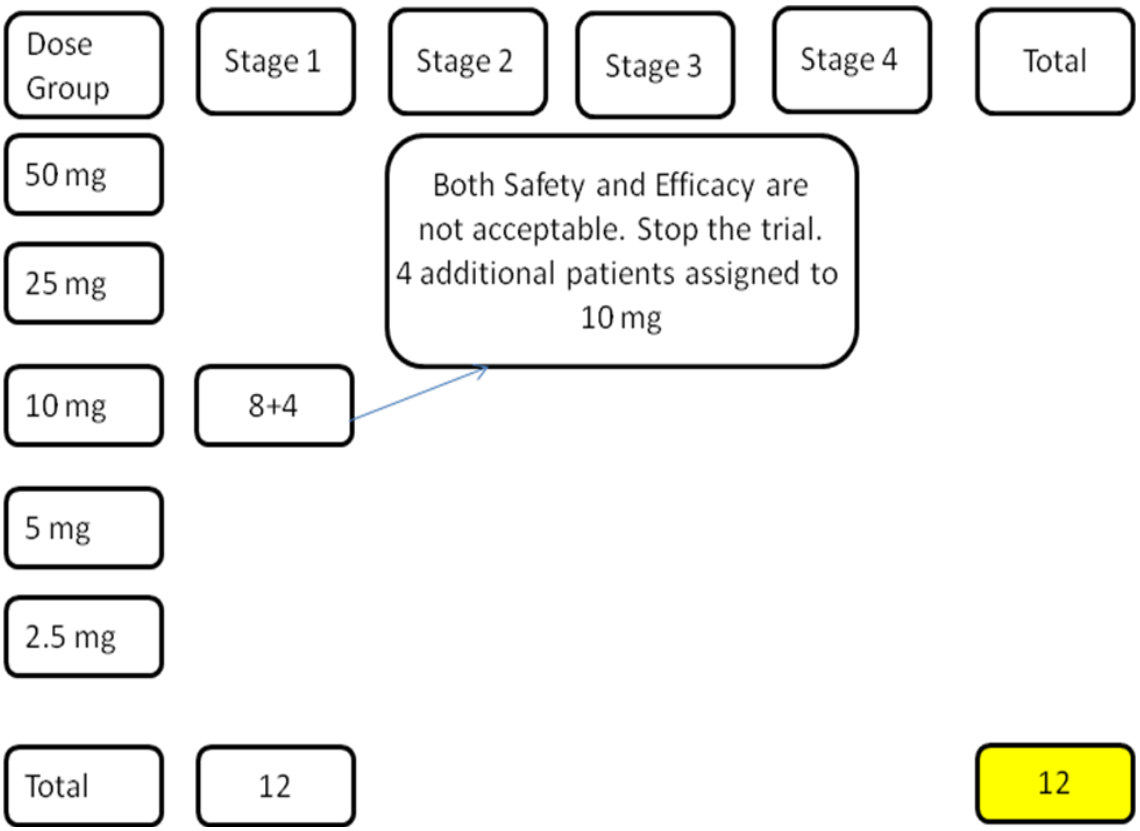


Figure 2.5 Example scenario of the design – both safety and efficacies are not acceptable in stage 1.

2.5 Statistical Models Used in the Proposed Design

2.5.1 Response Model Used in Stage 1

Predictive probability from response model is used to guide the decisions to terminate the trial for futility or move onto the next stage.

Let θ_d be the mean response to dose d for response variable Y.

Here,

$$Y \sim \theta_d + \varepsilon \tag{2.3}$$

$$\theta_d \sim N(\mu_0, \sigma_0^2), \tag{2.4}$$

$$\varepsilon \sim N(0, \sigma_\varepsilon^2) \tag{2.5}$$

where the prior distribution of θ_d follows normal distribution for the different values of d.

2.5.2 Normal Dynamic Linear Models (NDLMs) in Stage 2-4

A dose-response model based on a Normal Linear Dynamic Model (NDLMs) described by West and Harrison (1997) [24] are used in this paper. NDLM is essentially a piecewise linear model and has been used in clinical trials before. It provides the necessary flexibility to encompass both monotonic and non-monotonic dose-response relationships. It can be also easily implemented in a Bayesian updating frame work. Within this framework it provides direct probabilistic statements about many features of the dose-response. An additional advantage of NDLM is the existence of analytical results for the determination of the posterior distribution of the dose-response curve. NDLMs are also used to borrow information across adjacent doses [25].

Let Y_i be a generic outcome response variable and let $\theta_{di}=EY_i$ be the mean response for dose d. The following error structure is assumed for Yi,

$$Y_i \sim \theta_{d_i} + \varepsilon_{d_i}, \quad i = 2, 3, 4, 5, \tag{2.6}$$

where d_i is the dose given to the i-th stage. It is assumed that ε_{di} are an iid sample from N $(0, \sigma_\varepsilon^2)$ and the θ_{di} is an independent iid sample from $\theta_{di} \sim N(\theta, \sigma_\theta^2)$. An NDLM is used to defined with the following assumptions

$$\theta_{d_i} \sim N(\mu, \sigma_\theta^2), \quad i=2, 3, 4, 5, \tag{2.7}$$

$$\varepsilon \sim N(0, \sigma_\varepsilon^2). \tag{2.8}$$

The parameter σ_θ^2 represent the borrowing from one dose to the neighboring doses. The drift parameter is the variance between responses at neighboring doses. The larger the value of σ_θ^2, the less borrowing from neighboring doses. The prior distribution for the parameter σ_θ^2 in the NDLM is

$$\sigma_\theta^2 \sim IG\ (a_1, b_1) \tag{2.9}$$

The prior distribution for the error variance is

$$\sigma_\varepsilon^2 \sim IG\ (a_2, b_2) \tag{2.10}$$

Inverse Gamma was specified in Berry model and it is typical in Bayesian statistics. It serves as conjugate prior of the variance of the normal distribution. Under the prior specification $p(\sigma_\varepsilon^2, \sigma_\theta^2, \mu)=p(\sigma_\varepsilon^2)p(\sigma_\theta^2)p(\mu)$ [18,26].

Based on the above description, one can conclude that the response model in stage 1 can be represented as the following:

$$\begin{array}{ccc} & Y_{d_1} = \theta_{d_1} + \varepsilon & \\ & \swarrow \qquad \searrow & \\ N(\mu_0, \sigma_0^2) & & \varepsilon \sim N(0, \sigma_\varepsilon^2) \\ & & \downarrow \\ & & IG(a_2, b_2) \end{array} \tag{2.11}$$

θ_{d1} is the first stage mean response which follows the normal distribution $N(\mu_0, \sigma_0^2)$. The error term ε follows normal distribution $N(0, \sigma_\varepsilon^2)$ and σ_ε^2 follows inverse gamma distribution $IG(a_2, b_2)$.

In stage 2-4, NDLM dose response model is:

$$\begin{array}{ccc} & Y = \theta_{d_i} + \varepsilon, \quad i=2, 3, 4, 5, & \\ & \swarrow \qquad \searrow & \\ N(\mu_{d_{i-1}}, \sigma_\theta^2) & & \varepsilon \sim N(0, \sigma_\varepsilon^2) \\ \downarrow & & \downarrow \\ IG(a_1, b_1) & & IG(a_2, b_2) \end{array} \tag{2.12}$$

θ_{di} is the stage 2-4 response which follows the normal distribution $N(\mu_{di-1}, \sigma_\theta^2)$ and σ_θ^2 follows inverse gamma distribution IG (a_1, b_1). The error term ε follows the normal distribution of N (0, σ_ε^2) and σ_ε^2 follows inverse gamma distribution IG (a_2, b_2).

2.6 Implementation of the Proposed Design

The implementation of the proposed design can be done by two steps.

1. Derive the Formula of Conditional Probability for Each Parameter.
2. Find Posterior Distribution of θ (The efficacy response) based on simulation, which will be discussed in Chapter 3.

2.6.1 Derive the Formula of Conditional Probability for Each Parameter

As the first step, the conditional probability of each parameter can be derived using formula derivation methods in mathematical statistics, based on the three steps in the introduction of this chapter.

1) The conditional probability of μ_{di} can be derived:

$$
\begin{aligned}
& p(\mu_{d_i} \mid \theta_{di}, \sigma_\varepsilon^2, \sigma_\theta^2, Y) \\
& \propto p(\mu_{d_i}) P(\theta_{d_i} \mid \mu_{d_i}, \sigma_\varepsilon^2, \sigma_\theta^2, Y) \\
& \propto \frac{1}{\sqrt{2\pi}\sigma_{d_{i-1}}} \exp(-\frac{1}{2\sigma_{d_{i-1}}{}^2}(\mu_{d_i} - \mu_{d_{i-1}})^2) \prod_{i=1}^{k} \frac{1}{\sqrt{2\pi}\sigma_\theta} \quad \exp(-\frac{1}{2\sigma_\theta{}^2}(\theta_{d_i} - \mu_{d_i})^2) \\
& \propto \frac{1}{\sqrt{2\pi}\sigma_{d_{i-1}}} \exp(-\frac{1}{2\sigma_{d_{i-1}}{}^2}(\mu_{d_i} - \mu_{d_{i-1}})^2)(\frac{1}{2\pi\sigma_\theta})^k \quad \exp(-\frac{1}{2\sigma_\theta{}^2}\sum_{i=1}^{k}(\theta_{d_i} - \mu_{d_i})^2) \\
& \propto \frac{1}{\sigma_{d_{i-1}}} * \frac{1}{\sigma_\theta{}^k} \exp(-\frac{1}{2\sigma_{d_{i-1}}{}^2}(\mu_{d_i} - \mu_{d_{i-1}})^2 - \frac{1}{2\sigma_\theta{}^2}\sum_{i=1}^{k}(\mu_{d_i}{}^2 - 2\theta_{d_i}\mu_{d_i} + \theta_{d_i}{}^2)) \\
& \propto \exp(-\frac{1}{2\sigma_{d_{i-1}}{}^2}(\mu - \mu_{d_{i-1}})^2 - \frac{1}{2\sigma_\theta{}^2} \quad (k\mu_{d_i}{}^2 - 2\mu_{d_i}\sum_{i=1}^{k}\theta_{d_i})) \\
& \propto \exp(-\frac{1}{2\sigma_{d_{i-1}}{}^2}(\mu_{d_i}{}^2 - 2\mu_{d_i}\mu_{d_i} + \mu_{d_{i-1}}{}^2) - \frac{1}{2\sigma_\theta{}^2} \quad (k\mu_{d_i}{}^2 - 2\mu_{d_i}\sum_{i=1}^{k}\theta_i)) \\
& \propto \exp(-(\frac{1}{2\sigma_{d_{i-1}}{}^2} + \frac{k}{2\sigma_\theta})\mu_{d_i}{}^2 + (\frac{\mu_{d_{i-1}}}{\sigma_{d_{i-1}}{}^2} + \frac{\sum_{i=1}^{k}\theta_{d_i}}{\theta_{d_{i-1}}{}^2})\mu_{d_i}).
\end{aligned}
$$

Hence, the probability of μ_{di} followed a normal distribution. Based on normal distribution probability density function, one can get the following:

$$
\sigma^2(\mu_{d_i} \mid \theta_{d_i}, \sigma_\varepsilon{}^2, \sigma_\theta{}^2, Y) = \frac{1}{\frac{1}{\sigma_{d_{i-1}}{}^2} + \frac{k}{\sigma_\theta}} = \frac{\sigma_{d_{i-1}}^2 \sigma_\theta^2}{\sigma_\theta^2 + k\,\sigma_{d_{i-1}}^2} \tag{2.13}
$$

$$
\mu(\mu_{d_i} \mid \theta_{d_i}, \sigma_\varepsilon{}^2, \sigma_\theta{}^2, Y) = \frac{(\frac{\mu_{d_{i-1}}}{\sigma_{d_{i-1}}} + \frac{\sum_{i=1}^{k}\theta_{d_i}}{\sigma_\theta{}^2})}{(\frac{1}{\sigma_{d_{i-1}}{}^2} + \frac{k}{\sigma_\theta{}^2})} = \frac{\sigma_{d_{i-1}}^2 \sigma_\theta^2}{\sigma_\theta^2 + k\,\sigma_{d_{i-1}}^2}. \tag{2.14}
$$

So,

$$\mu_{d_i}|\theta_{d_i},\sigma_\varepsilon^2,\sigma_\theta^2,Y \sim N\left(\frac{\mu_{d_{i-1}}\sigma_\theta{}^2+\sigma_{d_{i-1}}{}^2\sum_{i=1}^{k}\theta_{d_i}}{\sigma_\theta{}^2+k\sigma_{d_{i-1}}{}^2},\frac{\sigma_{d_{i-1}}{}^2\sigma_\theta{}^2}{\sigma_\theta{}^2+k\sigma_{d_{i-1}}{}^2}\right)$$

i=2,3,4,5.

2) The conditional probability of θ_{di} can be derived as:

$$p(\theta_{d_i}|\mu_{d_i},\sigma_\varepsilon^2,\sigma_\theta^2,Y)$$

$$\propto p(\theta_{d_i}|\mu_{d_i},\sigma_\theta^2)\,p(Y|\theta_{d_i},\sigma_\varepsilon^2)$$

$$\propto \frac{1}{\sqrt{2\pi}\sigma_\theta}\exp\left(-\frac{1}{2\sigma_\theta^2}(\theta_{d_i}-\mu)^2\right)\prod_{j=1}^{J}\frac{1}{\sqrt{2\pi}\sigma_\varepsilon}\exp\left(-\frac{1}{2\sigma_\varepsilon{}^2}(Y_{ij}-\theta_i)^2\right)$$

$$\propto \exp\left(-\frac{1}{2\sigma_\theta^2}(\theta_{d_i}-\mu)^2\right)\exp\left(-\frac{1}{2\sigma_\varepsilon{}^2}\sum_{j=1}^{J}(Y_{ij}-\theta_i)^2\right)$$

$$\propto \exp\left(-\frac{1}{2\sigma_\theta^2}(\theta_{d_i}{}^2-2\theta_{d_i}\mu+\mu^2)-\frac{1}{2\sigma_\varepsilon{}^2}\sum_{j=1}^{J}(Y_{ij}{}^2-2Y_{ij}\theta_i+\theta_i{}^2)\right)$$

$$\propto \exp\left(-\left(\frac{1}{2\sigma_\theta{}^2}+\frac{J}{2\sigma_\varepsilon{}^2}\right)\theta_i{}^2+\left(\frac{\mu}{\sigma_\theta{}^2}+\frac{Y_i}{\sigma_\varepsilon{}^2}\right)\theta_i+\left(-\frac{\mu^2}{2\sigma_\theta{}^2}-\frac{Y_i{}^2}{2\sigma_\varepsilon{}^2}\right)\right)$$

Hence, the probability of θ_{di} followed a normal distribution. Based on normal distribution probability density function, one can get the following:

$$\sigma(\theta_{d_i}\left|\mu_{d_i},\sigma_\varepsilon^2,\sigma_\theta^2,Y\right.)=\frac{1}{\frac{1}{\sigma_\theta^2}+\frac{J}{\sigma_\varepsilon^2}}=\frac{\sigma_\varepsilon^2\sigma_\theta^2}{\sigma_\varepsilon^2+J\,\sigma_\theta^2} \tag{2.15}$$

$$\mu(\theta_{d_i}\left|\mu_{d_i},\sigma_\varepsilon^2,\sigma_\theta^2,Y\right.)=\frac{(\frac{\mu}{\sigma_\theta^2}+\frac{Y_{ij}}{\sigma_\varepsilon^2})}{(\frac{1}{\sigma_\theta{}^2}+\frac{J}{\sigma_\varepsilon^2})}=\frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2+J\sigma_\theta{}^2}\mu+\frac{J\sigma_\theta{}^2}{\sigma_\varepsilon{}^2+J\sigma_0{}^2}\overline{Y_{ij}} \tag{2.16}$$

So,

$$\theta_{d_i}\left|\mu_{d_i},\sigma_\varepsilon^2,\sigma_\theta^2,Y\right. \sim N(\frac{J\,\sigma_\theta^2}{J\,\sigma_\theta^2+\sigma_\varepsilon^2}\times\overline{Y}_i+\frac{\sigma_\varepsilon^2}{J\,\sigma_\theta^2+\sigma_\varepsilon^2}\times\mu_{d_i},\frac{\sigma_\theta^2\sigma_\varepsilon^2}{J\,\sigma_\theta^2+\sigma_\varepsilon^2})$$

i =2, 3, 4, 5.

d_i

3) The conditional probability for σ_θ^2 can be derived:

$$p(\sigma_\theta^2\left|\mu_{d_i},\theta_{d_i},\sigma_\varepsilon^2,Y\right.)$$

$$\propto p(\sigma_\theta^2)p(\theta_{d_i}\left|\sigma_\theta^2,\mu_{d_i},\sigma_\varepsilon^{2^2}\right.,Y)$$

$$\propto\frac{b_1{}^{a_1}}{\Gamma(a_1)}(\sigma_\theta{}^2)^{-(a_1+1)}\exp(-\frac{b_1}{\sigma_\theta^2})\frac{1}{\sqrt{2\pi}\sigma_\theta}\exp(-\frac{1}{2\sigma_\theta{}^2}(\theta_{d_i}-\mu_{d_i})^2)$$

$$\propto(\sigma_\theta^2)^{-(a_1+1)}\exp(-\frac{b_1}{\sigma_\theta{}^2})\sigma_\theta{}^{-1}\exp(-\frac{1}{2\sigma_\theta{}^2}(\theta_{d_i}-\mu_{d_i})^2)$$

$$\propto(\sigma_\theta^2)^{-(a_1+\frac{1}{2}+1)}\exp(-\frac{1}{\sigma_\theta{}^2}(b_1+\frac{1}{2}(\theta_{d_i}-\mu_{d_i})^2)$$

Hence, one can see that the posterior distribution for σ_θ^2 is a inverse gamma distribution.

$$\sigma_\theta^2 \left| \mu_{d_i}, \theta_{d_i}, \sigma_\varepsilon^2, Y \sim IG(a_1 + \frac{k}{2}, b_1 + \frac{1}{2}\sum_{i=1}^{k}(\theta_{d_i} - \mu_{d_i})^2) \right.$$

i=2, 3, 4, 5.

d) The conditional probability for σ_ε^2 can be derived:

$$p(\sigma_\varepsilon^2 \left| \mu_{d_i}, \theta_{d_i}, \sigma_\theta^2, Y \right.)$$

$$\propto p(\sigma_\varepsilon^2 \left| a_2, b_2 \right.) p(Y \left| \sigma_\varepsilon^2, \mu_{d_i} \theta_{d_i}, Y \right.)$$

$$\propto \frac{b_2^{a_2}}{p(a_2)} (\sigma_\varepsilon^2)^{-(a_s+1)} \exp(-\frac{b_2}{\sigma_\varepsilon^2}) \prod_{j=1}^{J} \frac{1}{\sqrt{2\pi}\,\sigma_\varepsilon^2} \exp(-\frac{1}{2\sigma_\varepsilon^2}(Y_j - \theta_{d_i})^2)$$

$$\propto (\sigma_\varepsilon^2)^{-(a_s+1)} \exp(-\frac{b_2}{\sigma_\varepsilon^2}) \sigma_\varepsilon^{-J} \exp(-\frac{1}{2\sigma_\varepsilon^2} \sum_{j=1}^{J}(Y_j - \theta_{d_i})^2)$$

$$\propto (\sigma_\varepsilon^2)^{-(a_s+\frac{J}{2}+1)} \exp(-\frac{b_2}{\sigma_\varepsilon^2}(b_2 + \frac{1}{2}\sum_{j=1}^{J}(Y_j - \theta_{d_i})^2).$$

Hence, one can see that the posterior distribution for σ_ε^2 is inverse gamma distribution.

$$\sigma_\varepsilon^2 \left| \mu_{d_i}, \theta_{d_i}, \sigma_\theta^2, Y \sim IG(a_2 + \frac{J}{2}, b_2 + \frac{1}{2}\sum_{j=1}^{J}(Y_j - \theta_{d_i})^2 \right.)$$

i=2, 3, 4, 5.

2.6.2 Find Posterior Distribution from Simulation

The Gibbs Sampler, a Markov Chain Monte Carlo (MCMC) method would be used in the simulation to find the posterior distribution of each parameter. This method starts with a set of initial values, described in the next chapter, then updating the estimator

successively from the full conditional distributions. After the estimator converges, the stationary or equilibrium distribution would be posterior distribution. In order to get the results without any bias, the burn-in sets would be discarded. Burn-in sets is defined as the first 1000 sample values of θ created from simulation, which are not stable.

CHAPTER 3

SIMULATIONS

When developing an adaptive design, a critical step is to simulate its performance across a variety of hypothesis response pattern scenarios. In order to simulate the design, assumptions have to be made to generate data representative of each response pattern. These assumptions do not affect the design or the analysis, but they are necessary to simulate the trial results. The SAS programs are used to do the simulation in this research. Appendix A shows the SAS codes for the first part of the model setup.

3.1 Estimator

The point estimator $\hat{\theta}$ is defined as the mean of sample values of θ. As defined in the previous chapter, θ refers to the percent increasing of the sum of SUVmean and SUVmax of FDG-PET uptake at Day 7 in the trial. The point estimation of $\hat{\theta}$ calculation starts with a set of initial values: μ_0, σ_0, a_1, a_2, b_1, b_2, and the sample values of θ can be obtained from full conditional distributions derived in chapter 2.

3.2 Test of the Convergence of the Markov Chain

Simulation-based Bayesian inference requires using simulated draws to summarize the posterior distribution or calculate any relevant quantities of interest. In MCMC method, there are several ways to decide whether the Markov Chain has reached its stationary or

the desired posterior, distribution. Table 3.1 [28] shows the convergence diagnostic tests available. The convergence of all the parameters, not just those of interest, should be checked.

In this research, Geweke test is used and trace plot is drawn. The assumption to use Geweke test is that MCMC process and the importance function g(θ) , jointly imply the existence of a spectrum, and the existence of a spectral density with no discontinuities at the frequency 0. The Geweke test (Geweke; 1992) compares values in the early part of the Markov chain to those in the latter part of the chain in order to detect failure of convergence. The statistic is constructed as follows. Two subsequences of the Markov chain $\{\theta^t\}$ are taken out, with $\{\theta_1^t : t = 1, \ldots, n_1\}$ and $\{\theta_2^t : t = n_a, \ldots, n\}$,

where $1 < n_1 < n_a < n$. Let $n_2 = n - n_a + 1$, and define

$$\bar{\theta}_1 = \frac{1}{n_1}\sum_{t=1}^{n_1}\theta^t \text{ and } \bar{\theta}_2 = \frac{1}{n_2}\sum_{t=n_a}^{n}\theta^t \tag{3.1}$$

Let $\hat{S}_1(0)$ and $\hat{S}_2(0)$ denote consistent spectral density estimates at zero frequency for the two MCMC chains, respectively. If the ratios n_1/n and n_2/n are fixed, $(n_1 + n_2)/n < 1$, and the chain is stationary, then the following statistic converges to a standard normal distribution as $n \to \infty$:

$$Z_n = \frac{\bar{\theta}_1 - \bar{\theta}_2}{\sqrt{\frac{\hat{S}_1(0)}{n_1} + \frac{\hat{S}_2(0)}{n_2}}} \tag{3.2}$$

This is a two-sided test, and large absolute z-scores indicate rejection.

Table 3.1 Convergence Diagnostic Tests Available in the Bayesian Procedures

Name	**Description**	**Interpretation of the Test**
Gelman-Rubin	Uses parallel chains with dispersed initial values to test whether they all converge to the same target distribution. Failure could indicate the presence of a multi-mode posterior distribution (different chains converge to different local modes) or the need to run a longer chain (burn-in is yet to be completed).	One-sided test based on a variance ratio test statistic. Large $\hat{R}_c$ values indicate rejection.
Heidelberger-Welch (stationarity test)	Tests whether the Markov chain is a covariance (or weakly) stationary process. Failure could indicate that a longer Markov chain is needed.	One-sided test based on a Cramer–von Mises statistic. Small p-values indicate rejection.
Heidelberger-Welch (half-width test)	Reports whether the sample size is adequate to meet the required accuracy for the mean estimate. Failure could indicate that a longer Markov chain is needed.	If a relative half-width statistic is greater than a predetermined accuracy measure, this indicates rejection.
Raftery-Lewis	Evaluates the accuracy of the estimated (desired) percentiles by reporting the number of samples needed to reach the desired accuracy of the percentiles. Failure could indicate that a longer Markov chain is needed.	If the total samples needed are fewer than the Markov chain sample, this indicates rejection.
autocorrelation	Measures dependency among Markov chain samples.	High correlations between long lags indicate poor mixing.
effective sample size	Relates to autocorrelation; measures mixing of the Markov chain.	Large discrepancy between the effective sample size and the simulation sample size indicates poor mixing.

3.3 Verification of Repeatability of the Results

In this research, seventy-thousand iterations are used in simulation trial to get the estimator converge. Five thousand simulation trials are conducted to verify the repeatability of the results.

3.4 Priors Selection

3.4.1 Objective Priors versus Subjective Priors

Bayesian probability measures the degree of belief that you have in a random event. By this definition, probability is highly subjective. It follows that all priors are subjective priors. Not everyone agrees with this notion of subjectivity when it comes to specifying prior distributions. There has long been a desire to obtain results that are objectively valid. Within the Bayesian paradigm, this can be achieved by using prior distributions that are "objective" (that is, that have a minimal impact on the posterior distribution).

A prior distribution is non-informative if the prior is "flat" relative to the likelihood function. Thus, a prior is non-informative if it has minimal impact on the posterior distribution of θ. Many people favor non-informative priors because they appear to be more objective.

There are several priors in this Bayesian design. The selection of the priors used in dose-response model and NDLM is based on the historical data and non-informative rule. The selection of each parameter specified in prior distribution is specified below:

a) $\mu_0 = 0.2$

A successful efficacy response is defined as a subject has ≥ 20% decrease on the sum of SUVmean and SUVmax of FDG-PET uptake at 7 days post-dose compared to pre-dose. To obtain equal probability of positive and negative efficacy responses, 0.2 is chosen as a flat prior. This will change after more data are brought in.

b) $\sigma_0 = 0.1$

The possibility of increasing on the FDG-PET uptake is very small (=0.025). If the drift effect is noticed in the data, σ_0 could be adjusted to a larger one accordingly. $0.2/1.96 \approx 0.1$

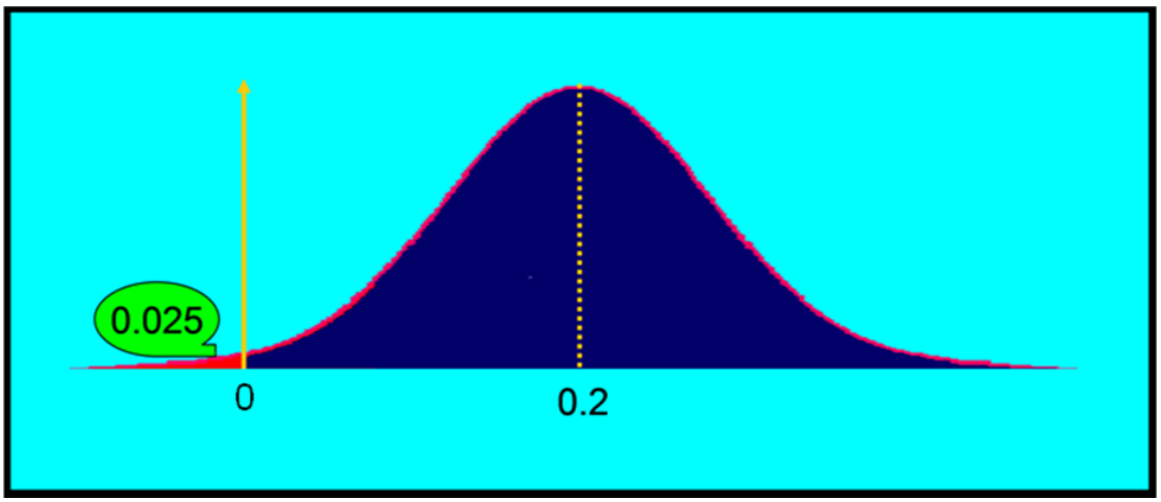


Figure 3.1 Probability distribution of θ.

c) $a_1 = a_2 = 2$

Standard deviation doesn't exist when $a_1=a_2=2$ for inverse gamma. The distribution is close to 'non-informative'. The result will be data driven which fits one's need since there is no reliable estimation.

a) $b_1 = 0.0266$

Based on the historic data [21], the standard deviation of θ is 0.163 and the variance is 0.0266. And the mean of IG(a_1, b_1) is $b_1/(a_1-1) = b_1 = 0.0266$

b) $b_2 = 0.0026$

The estimate of standard errors is based on the prior data with some assumption to fit our needs. According to the historic data, standard errors of SUV reduction are 0.02164 and 0.00776 in 10 mg and 25 mg groups, respectively [21].

Using those two observed numbers, the variance of standard error is 0.0026 with the mean of 0. To be conservative, taking 0.0026 as the mean of IG, $b_2 = \text{mean}*(a_2-1) = 0.0026$ when $a_2 = 2$.

If over-estimated, Bayesian design would not be sensitive enough. The worst case scenario is to enroll all 60 subjects without any savings on the sample.

3.5 Probability Density Functions of IG(a_1,b_1) and IG(a_2,b_2)

Based on the parameter estimation, the probability density functions are showing in the figure 3.2.

Based on the simulation results, the characteristic operation data and curves can be obtained with the different scenario of the real increasing on SUV. After that, power and probability of the FDG-PET increasing for each stage can be obtained.

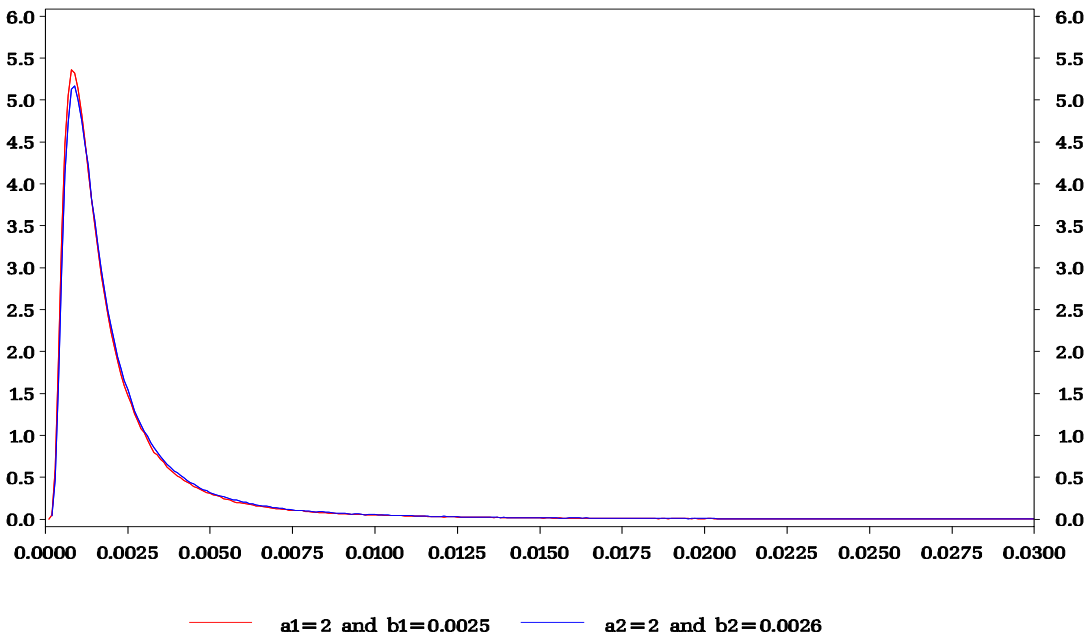


Figure 3.2 Distribution of σ_θ^2 and σ_ε^2.

3.6 The Next Step in the Simulation

After the setup of the initial values of the parameters, the posterior derived in previous chapter will be used and the following steps will be followed:

1. Starts with the defined initial values.
2. Updating successively from the full conditional distributions.
3. After converge, the stationary or equilibrium distribution is posterior distribution.
4. Discard burn-in sets (the first 1000 sample values).
5. Set the true mean of the θ and start the iteration in the simulation by using the posterior distribution derived earlier.
6. Find the samples' mean of θ from simulation and compare with the true mean.
7. Get the sample size from simulation to see how many sample size saved.
8. Find the power of the test by using: total of successful trials/trials simulated.

CHAPTER 4

RESULTS

A series of simulations have been done to evaluate the operating characteristics of the proposed Bayesian design.

4.1 Results from Different Scenarios

The model used in stage 1 of the proposed Bayesian design is as follows:

$$Y_{d1} = \theta_{d1} + \varepsilon \qquad (4.1)$$

$$\theta_{d1} \rightarrow N(\mu_0, \sigma_0^2) \qquad \varepsilon \rightarrow N(0, \sigma_\varepsilon^2) \rightarrow IG(a_2, b_2)$$

θ_{d1} is the first stage mean response which follows the normal distribution $N(\mu_0, \sigma_0^2)$ for 10 mg dose level. The error term ε follows normal distribution $N(0, \sigma_\varepsilon^2)$ and σ_ε^2 follows inverse gamma distribution $IG(a_2, b_2)$. Simulation has been done for several scenarios of the results to check the validity of the model and the designs.

4.1.1 Scenario 1

Table 4.1 shows one scenario of the assumed true mean of SUV decreasing used in the simulation. Four random values are taken from SAS random function as the observed values from normal distribution with previously assigned N (0.2, 0.026) for $\theta_{d1.}$ The prior chosen for $\boldsymbol{\sigma_\varepsilon^2}$ is IG (2, 0.0026). In stage 1, eight patients are assigned to the 10 mg dosing group. The simulation was done for the first interim analysis with four subjects'

values. Another four subjects' values were simulated for the second and third interim analyses after the first interim analysis data obtained. Five-thousand iterations are used in the program and the first 1000 burn-in results are discarded. The sample response data of eight subjects are shown in Tables 4.2. Table 4.3 shows additional four sample subjects added in the next stage and used to confirm the results in the previous stage.

Table 4.1 Scenario 1 - True Mean of SUV Decreasing Used in Simulation

Dose Group	True Mean of SUV Decreasing Used in Simulation
50 mg	0.32
25 mg	0.28
10 mg	0.21
5 mg	0.14
2.5 mg	0.10

Table 4.2 Sample Response from Eight Patients of Each Dose Level (Randomly Selected from Normal Distribution)

Sample Response from the Patients (10 mg)	Sample Response from the Patients (25 mg)	Sample Response from the Patients (5 mg)	Sample Response from the Patients (50 mg)
0.225358	0.286135	0.125904	0.345133
0.250183	0.304092	0.119045	0.318048
0.196323	0.309638	0.074631	0.279937
0.224806	0.232502	0.126400	0.317466
0.199867	0.327465	0.083362	0.364753
0.213084	0.274455	0.140660	0.348125
0.203742	0.373889	0.173447	0.289752
0.230214	0.314788	0.108202	0.287554

Table 4.3 Sample Response of Additional Four Patients to Confirm the Results in Each Dose Level

Sample Response from the Patients (10 mg)	Sample Response from the Patients (25 mg)	Sample Response from the Patients (5 mg)	Sample Response from the Patients (50 mg)
0.240127	0.276531	0.125417	0.338955
0.223251	0.309022	0.120123	0.317939
0.194789	0.302467	0.083290	0.282436
0.235901	0.233029	0.119879	0.320148

Five-thousand simulation trials were conducted to verify the repeatability of the simulation results. Table 4.4 shows the results of the first 20 simulation trials. The first column is the posterior mean of $\theta_{d1.}$ The second column is the standard deviation. If posterior mean of θ_{d1} is greater than or equal to 0.2, the tested dose level is defined to be effective. If it is less than 0.2, the tested dose level is determined to be futile. By repeating the trial for 4999 times, the rate to correctly declare the effective of 10 mg dose based on success in Table 4.4 is 91%. That means if the true SUV decreasing is 0.21 for 10 mg dose, the chance that one accepts the efficacy of 10 mg and go into the second stage to test 5 mg and 25 mg is 91% when safety assessment turns out to be acceptable.

Table 4.4 Results from the First Twenty Simulated Trials for 10 mg Dose Group

Trial	Posterior Mean	Std Dev	Success
1	0.1986	0.0110	0
2	0.2135	0.0117	1
3	0.2048	0.0082	1
4	0.2007	0.0090	1
5	0.2205	0.0108	1
6	0.2137	0.0134	1
7	0.2015	0.0123	1
8	0.1983	0.0088	0
9	0.2093	0.0097	1
10	0.2045	0.0085	1
11	0.2263	0.0089	1
12	0.2148	0.0135	1
13	0.2043	0.0201	1
14	0.2139	0.0088	1
15	0.2090	0.0101	1
16	0.2143	0.0104	1
17	0.2013	0.0088	1
18	0.2128	0.0133	1
19	0.1960	0.0103	0
20	0.2136	0.0105	1

Similar to the simulation for 10 mg dose group, 5000 simulation trials were conducted for 5 mg. Table 4.5 shows the results of the first 20 simulation trials. By repeating the trial for 4999 times, the rate to incorrectly declare the effective of 5 mg dose based on success in Table 5 is 0.14%. That means if the true SUV decreasing is 0.14 for 5 mg dose, the chance that one accepts the efficacy of 5 mg and go into the third stage to test 2.5 mg is 0.14% when safety assessment turns out to be acceptable.

Table 4.5 Results from the First Twenty Simulated Trials for 5 mg Dose Group

Trial	Posterior Mean	Std Dev	Success
1	0.1671	0.0137	0
2	0.1420	0.0141	0
3	0.1430	0.0159	0
4	0.1574	0.0117	0
5	0.1381	0.0126	0
6	0.1425	0.0164	0
7	0.1456	0.0139	0
8	0.1436	0.0171	0
9	0.1453	0.0155	0
10	0.1298	0.0148	0
11	0.1267	0.0112	0
12	0.1357	0.0115	0
13	0.1352	0.0130	0
14	0.1258	0.0140	0
15	0.1416	0.0122	0
16	0.1452	0.0109	0
17	0.1393	0.0125	0
18	0.1648	0.0115	0
19	0.1403	0.0117	0
20	0.1550	0.0107	0

Tables 4.6 and 4.7 present the first 20 simulated trials for dose groups of 25 mg and 50 mg. Since both dose levels have true responses much better than 20%, the powers to correctly detect the efficacy are as high as 100%.

Table 4.6 Results from the First Twenty Simulated Trials for 25 mg Dose Group

Trial	Posterior Mean	Std Dev	Success
1	0.2613	0.0094	1
2	0.2706	0.0094	1
3	0.2741	0.0122	1
4	0.2795	0.0110	1
5	0.2827	0.0147	1
6	0.2818	0.0105	1
7	0.2937	0.0134	1
8	0.2599	0.0125	1
9	0.2818	0.0148	1
10	0.2704	0.0137	1
11	0.2617	0.0123	1
12	0.2804	0.0137	1
13	0.2894	0.0133	1
14	0.2867	0.0105	1
15	0.2763	0.0165	1
16	0.2963	0.0140	1
17	0.2961	0.0124	1
18	0.2650	0.0140	1
19	0.2775	0.0105	1
20	0.2822	0.0133	1

Table 4.7 Results from the First Twenty Simulated Trials for 50 mg Dose Group

Trial	Posterior Mean	Std Dev	Success
1	0.3324	0.0123	1
2	0.3118	0.0101	1
3	0.3347	0.0175	1
4	0.3171	0.0130	1
5	0.3163	0.0132	1
6	0.3334	0.0135	1
7	0.3326	0.0122	1
8	0.3163	0.0107	1
9	0.3304	0.0116	1
10	0.2932	0.0106	1
11	0.2914	0.0143	1
12	0.3170	0.0118	1
13	0.3054	0.0108	1
14	0.3357	0.0085	1
15	0.3286	0.0129	1
16	0.3263	0.0099	1
17	0.3031	0.0110	1
18	0.2965	0.0155	1
19	0.3038	0.0195	1
20	0.3189	0.0139	1

Figure 4.1 shows the convergence of θ_{d1}. The data show the convergence through the 5000 iterations. It also shows the posterior distribution of $\theta_{d1.}$ The autocorrelation between the samples are 0. Figure 4.2 shows the convergence of σ_{ε}^{2} and posterior distribution of σ_{ε}^{2}. The autocorrelation is also 0.

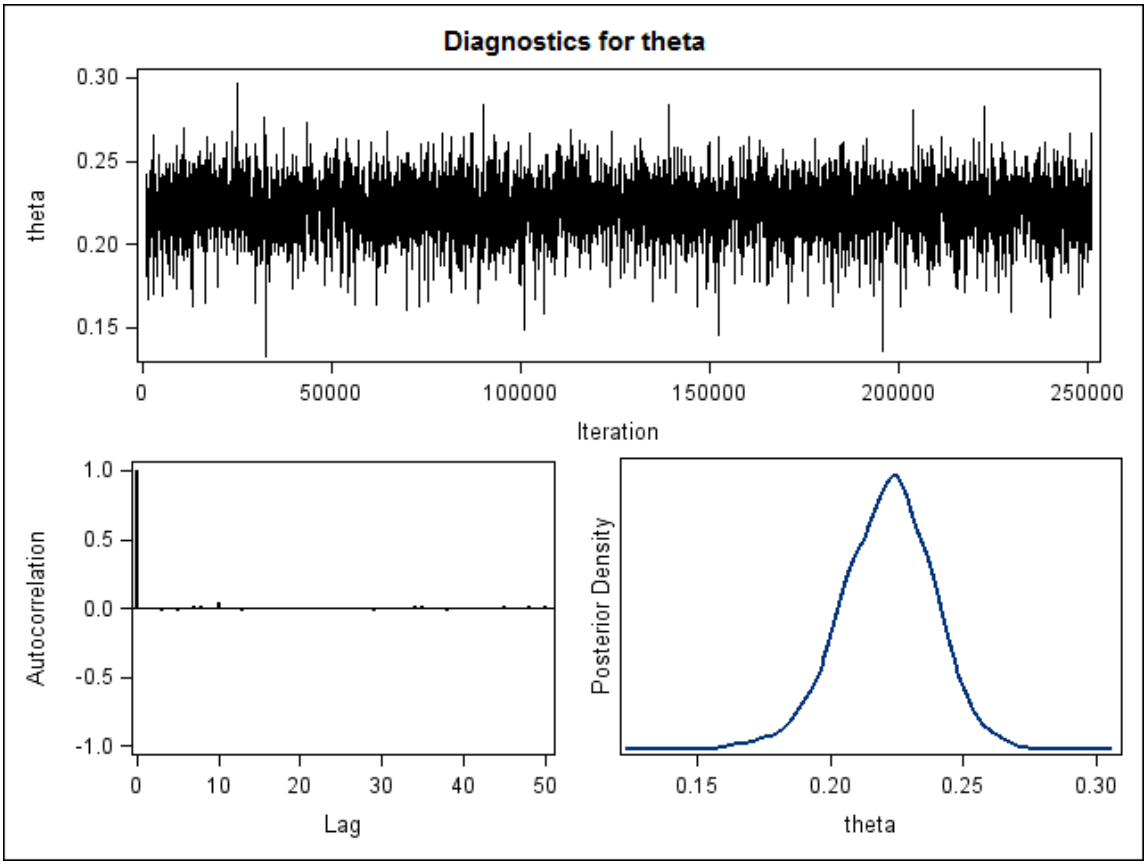


Figure 4.1 Convergence, autocorrelation and posterior distribution of θ_{d1}.

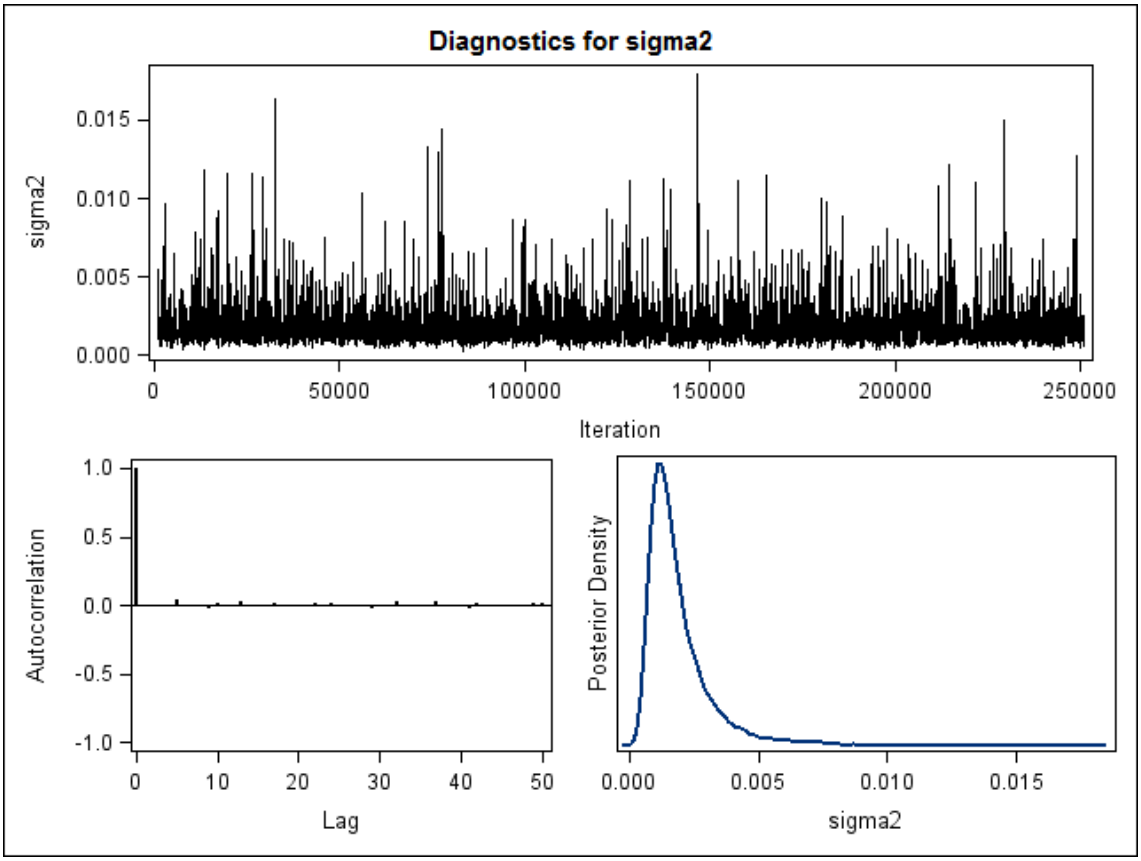


Figure 4.2 Convergence, auto-correlation, and posterior distribution of σ_ε^2**.**

The values of θ_{d1}**,** σ_ε^2 from the first stage can be used as the prior of the second stage. The similar procedure is done for stages 2, 3 and 4. The final results of a random selected trial in simulation are showing in Table 4.8. 2.5 mg was not tested since the 5 mg dose failed on efficacy. Assuming the safety results are all good, the maximum patients need to be recruited in this scenario is 60. In case any safety issues were found in higher dose, the dose level will not go up. The sample size will be saved more.

Table 4.8 Posterior Information for Each Dose Group

Dose Group	True Mean of SUV Decreasing Used in Simulation	Posterior Mean at the end of Testing Stage	Posterior Std at the end of Testing Stage	Posterior Probability (%) of SUV Decreasing ≥ 0.20	Sample Size used in the trial
50 mg	0.32	0.308	0.0185	100.0	12
25 mg	0.28	0.297	0.0205	100.0	12
10 mg	0.21	0.227	0.0136	91.2	12
5 mg	0.14	0.125	0.0412	0.14	12

4.2.2 Scenario 2

Table 4.9 shows the second scenario of the assumed true mean of SUV decreasing used in the simulation.

Table 4.9 Scenario 2 - True Mean of SUV Decreasing in Simulation

Dose Group	True Mean of SUV Increasing Used in Simulation
50 mg	0.25
25 mg	0.20
10 mg	0.15
5 mg	0.14
2.5 mg	0.10

Table 4.10 Sample Response from Eight Patients (Randomly Selected from Normal Distribution)

Real Response from the Patients (10 mg)	Real Response from the Patients (25 mg)	Real Response from the Patients (50 mg)
0.175852	0.209279	0.198659
0.128976	0.192248	0.295828
0.141329	0.150385	0.274836
0.155357	0.180104	0.233229
0.216211	0.175942	0.219439
0.145488	0.213383	0.222021
0.17906	0.192711	0.263308
0.199519	0.204148	0.241855

Table 4.11 Sample Response of Additional Four Patients to Confirm the Results in Stage 1

Real Response from the Patients (10 mg)	Real Response from the Patients (25 mg)	Real Response from the Patients (50 mg)
0.216211	0.165097	0.219439
0.145488	0.203786	0.222021
0.17906	0.205371	0.263308
0.199519	0.250228	0.241855

Figure 4.3 shows the convergence of θ_{d1}. The data show the convergence through the 50000 iterations. It also shows the posterior distribution of $\theta_{d1.}$ Figure 4.4 shows the convergence of σ_{ε}^{2}.

Table 4.12 θ_{d1} Results from the First 20 Simulated Trials

Trial	Mean	Std Dev	success
1	0.1497	0.0169	1
2	0.1641	0.0152	1
3	0.1687	0.0145	1
4	0.163	0.0156	1
5	0.1555	0.0172	1
6	0.1676	0.0146	1
7	0.1636	0.017	1
8	0.158	0.0157	1
9	0.1702	0.0148	1
10	0.1669	0.0149	1
11	0.1714	0.0143	1
12	0.1591	0.0157	1
13	0.1746	0.015	1
14	0.1533	0.0163	1
15	0.1673	0.0148	1
16	0.1571	0.0169	1
17	0.1562	0.0155	1
18	0.1639	0.015	1
19	0.1661	0.016	1
20	0.1531	0.0171	1

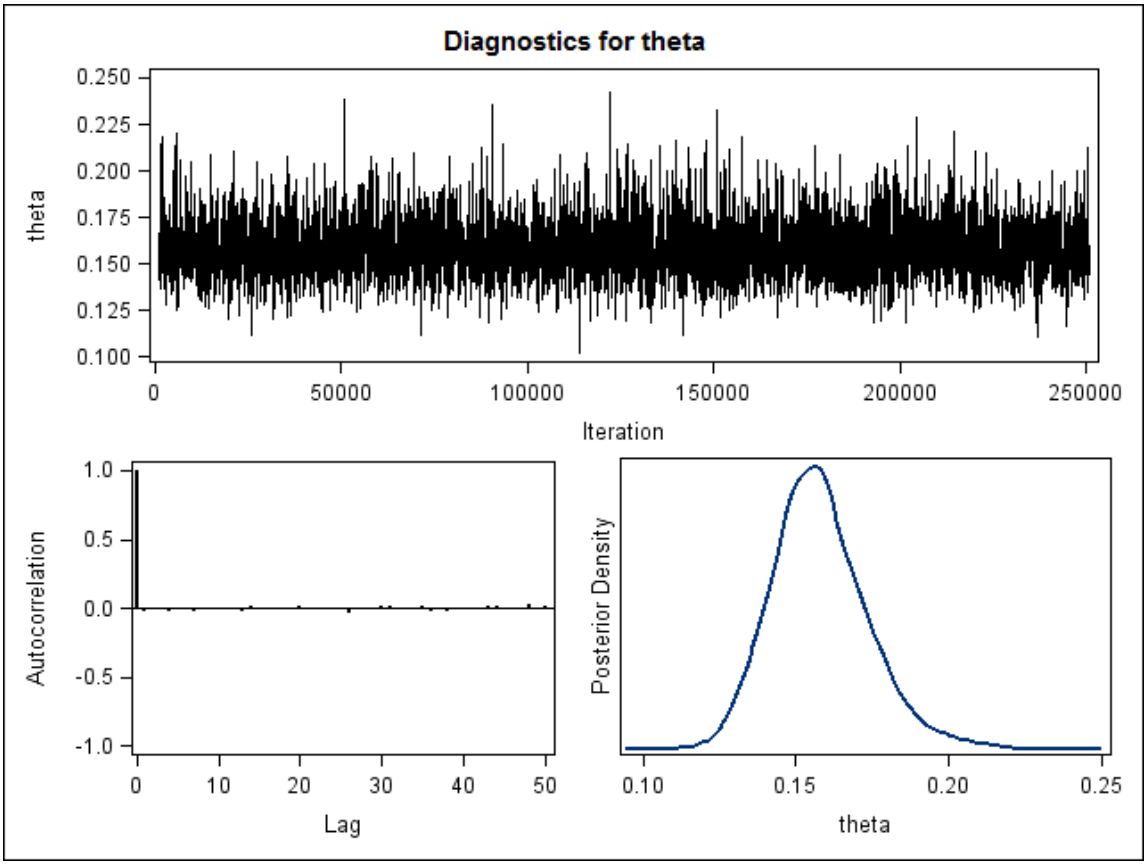


Figure 4.3 Convergence, auto-correlation and posterior distribution of θ_{d1}.

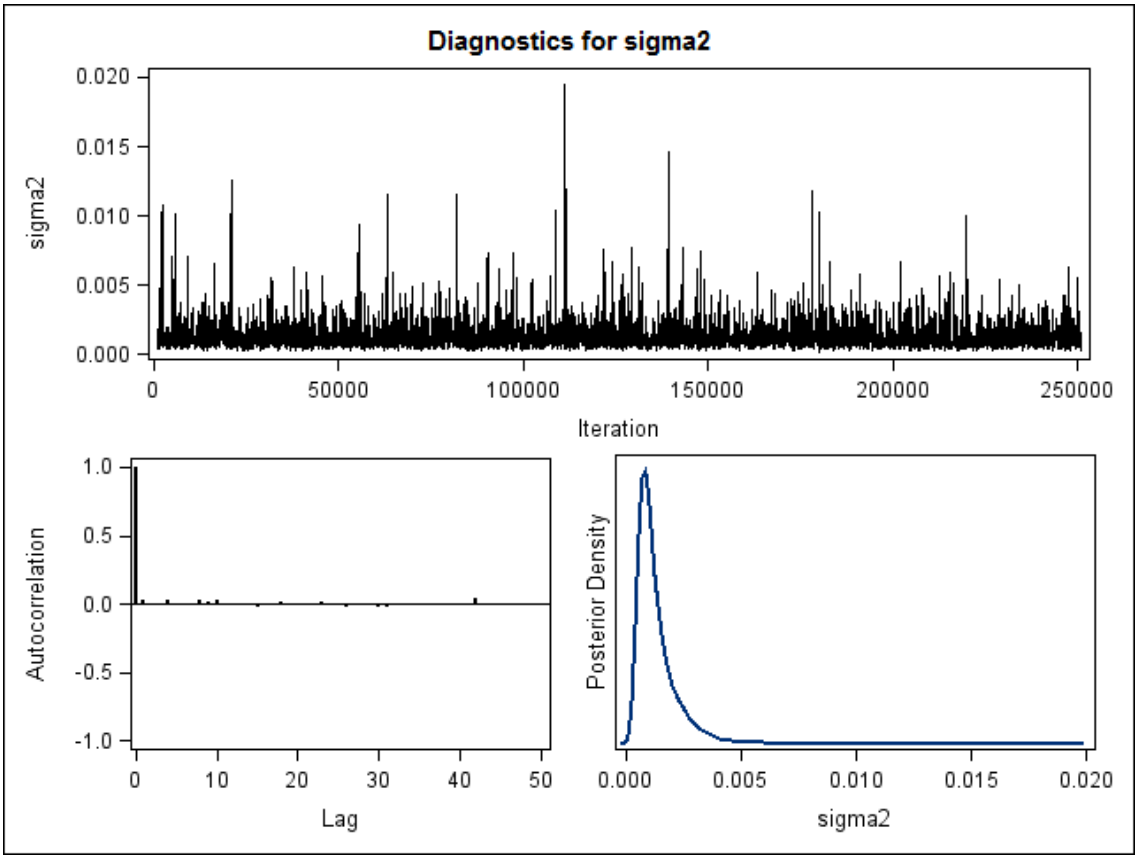


Figure 4.4 Convergence, auto-correlation, and posterior distribution of σ_ε^2.

The values of θ_{d1}, σ_ε^2 from the first stage can be used as the prior of the second stage. The similar procedure is done for stages 2, 3 and 4.

Similar simulation has been done for 10 mg, 25 mg, and 50 mg dose groups. The final results are showing in Table 4.13. 2.5 mg and 5 mg doses were not tested since the 10 mg dose failed on efficacy. Assuming the safety are all good, the maximum patients need to be recruited in this scenario is 48. Twenty four subjects in sample size are saved. In case any safety issues were found in higher dose, the dose level will not go up. The sample size will be saved more.

Table 4.13 Posterior Information for Each Dose Group

Dose Group	True Mean of SUV Decreasing Used in Simulation	Posterior Mean at the end of Testing Stage	Posterior Std at the end of Testing Stage	Posterior Probability (%) of SUV Decreasing between 0.20 and 0.90	Sample Size used in the trial
50 mg	0.25	0.261	0.0185	99.9	12
25 mg	0.20	0.199	0.0205	98.8	12
10 mg	0.15	0.148	0.0136	13.6	12

The results from scenarios 1 and 2 are showing that the model chosen and the design are repeatable and have the high power to get the correct results.

4.2 Operating Characteristics

Table 4.14 shows operating characteristics for each dose level. In each scenario, 5000 simulated trials were conducted.

Table 4.14 Operating Characteristics

Scenario	Assumed Decreasing (%) at Dose Levels					Percent of Trials Selecting the Right Doses (%)	Average of the Number of Subjects Used (Saving %)
	2.5 mg	5 mg	10 mg	25 mg	50 mg		
1	10	14	21	28	32	91.2	48(20)
2	10	14	15	20	25	73.2	36(40)
3	2.5	5	10	15	21	89.2	23 (61.7)
4	1	8	10	15	21	85	32 (46.7)
5	0	3	10	15	20	72.5	23 (61.7)
6	0	0	0	0	0	100	18 (70)
7	0	0	0	0	25	100	18 (70)
8	0	24	30	30	30	100	31 (48.3)
9	30	30	30	30	30	100	31 (48.3)
10	25	25	24	24	21	100	60 (0)

CHAPTER 5

BAYESIAN ADAPTIVE DESIGN FOR LONGITUDINAL CLINICAL TRIAL

Another Bayesian adaptive design with longitudinal model is proposed in this research. The clinical trial chosen is a phase II dose-selection study to assess the safety and efficacy of a drug in treatment of depression with major depressive disorder (MDD) patients. Patients are randomly assigned to 8 weeks of the treatment with study drug. Phase I study has been conducted to determine the maximum tolerable dose (MTD) based on safety availability. After the baseline visit, study visits will be conducted at the end of week 1, 2, 4, 6, 8. The primary endpoint defined is change from baseline in the Montgomery Asberg Depression Scale (MADRS) total score at week 8. MADRS is a ten-item diagnostic questionnaire which psychiatrists use to measure the severity of depressive episodes in patients with major disorders. It was designed in 1979 by British and Swedish researchers as an adjunct to the Hamilton Rating Scale for Depression (HAMD). Before this study conducted, MTD dose has already been defined in phase I study. Based on MTD, the dose level chosen to be tested are: 180 mg, 150 mg, 125 mg, 100 mg, 75 mg, 50 mg and 25 mg.

5.1 The New Proposed Bayesian Adaptive Design

Figure 5.1 shows the propose design with the maximum possible number of randomized subjects in each treatment group.

Dose Group	Stage 1	Stage 2	Stage 3	Total
180 mg	90			90
150 mg		90		90
125 mg				
100 mg		90		90
75 mg				
50 mg		90		90
25 mg			90	90
Placebo	45	45	45	135
Total	135	315	135	585

Figure 5.1 The maximum possible number of randomized subjects in each treatment group by stage.

Stage 1:

In stage 1, only the highest dose and placebo will be evaluated. Ninety subjects will be randomized into 180 mg group and forty-five subjects into Placebo group in a 2:1 ratio in this stage. The final maximum sample size 585 is calculated by sample size calculation method using historic data with two sided $\alpha = 0.05$ and $\beta = 0.7$ [27].

Three interim analyses will be performed in each stage. For the first interim analysis, forty-five subjects are randomized, thirty subjects in study drug treatment group and fifteen in Placebo group. In order to ensure the balance of the numbers in the study treatment and placebo during the trial, the block size of three is chosen. After the subjects take the dose, the post-baseline values of MADRS at each study visit and the safety data are recorded. The change from baseline of MADRS at week 8 is analyzed using Analysis of Covariance (ANCOVA) model with study centers as factor and baseline score of MADRS as covariate.

As the next step for the second interim analysis, thirty more subjects are added into 180 mg group and fifteen subjects are added into placebo group. The posterior quantities are calculated again using expanded data. At each interim analysis the following Bayesian posterior quantities $P(\theta_1 - \theta_8 > 2|\text{Data})$ were calculated and utilized, where θ_1 is the response from 180 mg dose group and θ_8 is the response from placebo group. If $P(\theta_1 - \theta_8 > 2|\text{Data})$ is less than 0.25 for two consecutive interim analysis, then this dose level will be declared as inefficacious due to futility and the study will be terminated. On the other hand, if $P(\theta_1 - \theta_8 > 2|\text{Data})$ is ≥ 0.75 for two consecutive interim analysis or $P(\theta_1 - \theta_8 > 2|\text{Data})$ is ≥ 0.95 for one interim analysis, the tested dose level will

be determined to be effective and the stage 1 will be stopped early. Otherwise, thirty more subjects in study treatment group and fifteen more in placebo group would be added. The posterior quantities are calculated again.

Given the results of those three interim analyses on efficacy, the efficacy of 180 mg group would be evaluated using the following criteria:

1. If the Bayesian posterior probabilities of the comparison of study drug and placebo $P(\theta_1 - \theta_8 > 2|Data)$ are less than 0.25 for two consecutive interim analysis, this dose level will be declared as inefficacious due to futility and study terminated.
2. Otherwise, the dose level will be declared as non-futile.
3. If the Bayesian posterior probabilities of the comparison of study drug and placebo $P(\theta_1 - \theta_8 > 2|Data)$ are ≥ 0.75 for two consecutive interim analysis, or $P(\theta_1 - \theta_8 > 2|Data) \geq 0.95$ for one interim analysis, this dose level will be declared as efficacious at stage 1 and the study will be terminated early. The study will continue to stage 2.

In any case, if the dose level has safety concern in any interim analysis, the next lower dose should be tested.

Stage 2:

If the efficacy response shows as non-futile, trial goes into stage 2. Otherwise, the trial is terminated. In stage 2, three different skipped doses 150 mg, 100 mg, 50 mg and placebo are tested with ninety subjects in each treatment group using the posterior mean and posterior standard deviation in the 180 mg dose group as prior. Three interim analyses will be done with thirty subjects in each study treatment group and fifteen subjects in placebo group at each interim analysis.

Given the results of those three interim analyses on efficacy, the efficacy of 150 mg, 100 mg 50 mg groups would be evaluated using the same criteria in stage 1. If the positive result showed in stage 2 for any dose levels for the first two consecutive interim

analyses, the trial goes to stage 3 directly without the third interim analysis. Or if the results show non-futile for all three interim analyses, the trial goes to stage 3. Otherwise, the trial will be terminated and the maximum dose of 180 mg in stage 1 will be the only selected dose.

Stage 3:

In stage 3, one more dose level will be tested in order to plot the dose response curve more accurately. Which dose group to be tested is determined based on the efficacy response of three dose groups in stage 2. If both 150 mg and 100 mg are effective, 125 mg would not be tested in stage 3. Otherwise, if 150 mg is effective and 100 mg is not, the dose group of 125 mg would be tested in stage 3. In some rare case, if the dose response is not following the common dose-response curve, e.g., 100 mg dose is effective and 150 mg dose is not, the dose group of 125 mg will still be tested. Whether to test the dose groups of 75 mg is to be determined in the same way. If 50 mg is effective, the dose group of 25 mg would be tested. Otherwise, 25 mg would be skipped. Three interim analyses will be done with thirty subjects in study treatment group and fifteen subjects in placebo group in each interim analysis.

The Bayesian posterior probabilities are calculated to determine the effective dosing group in stage 3. Figure 5.1 shows the maximum possible number of randomized subjects in each treatment group by stage. This happens when safety is good and efficacy is acceptable in all stages. Total five hundred and eighty-five patients are exposed to the treatment and the acceptable doses range from 25 mg-180 mg. Figure 5.2 and Figure 5.3 shows some other scenarios of the design.

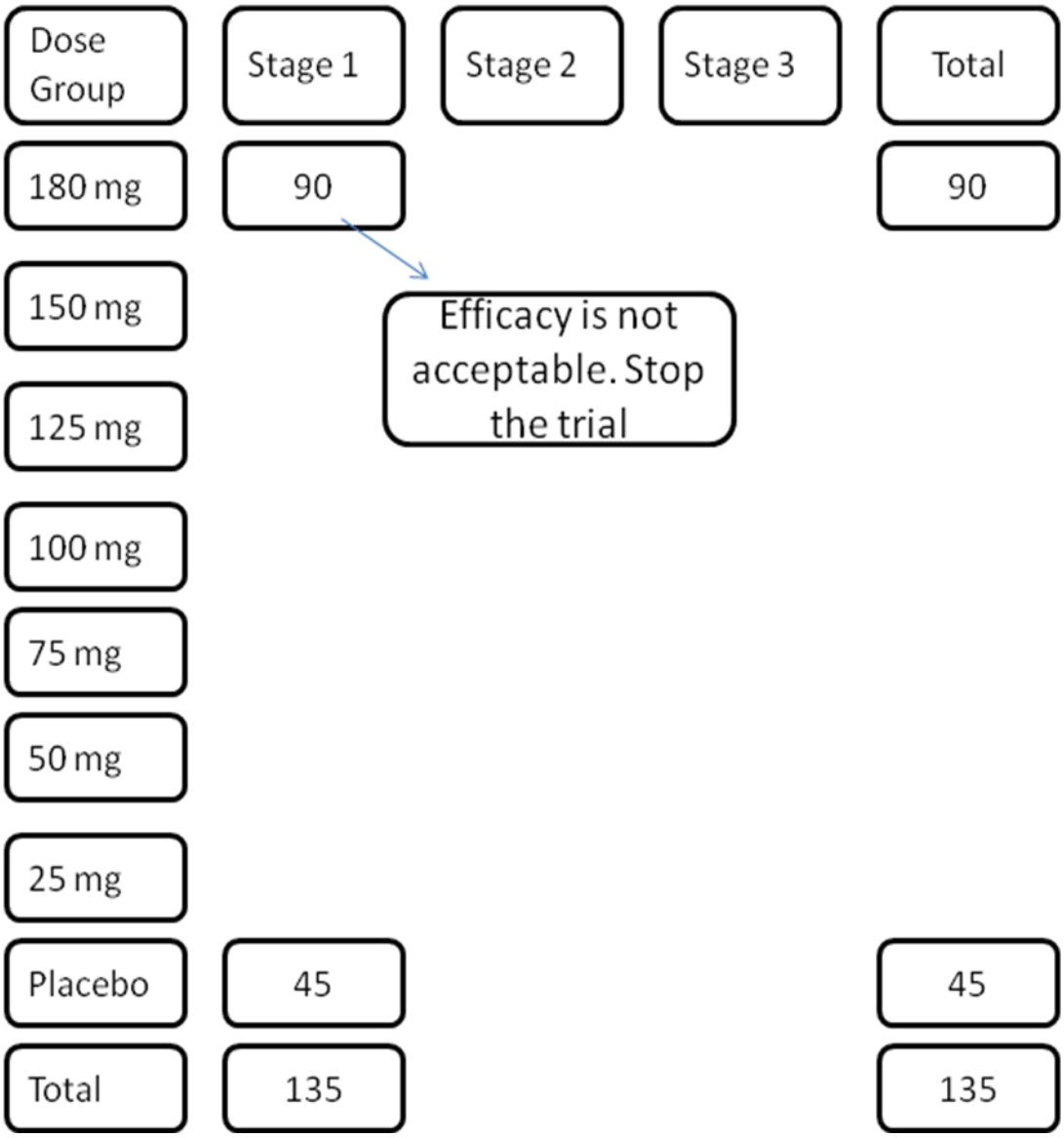


Figure 5.2 Example scenario of the design - efficacy is not acceptable at the highest dose and trial stops.

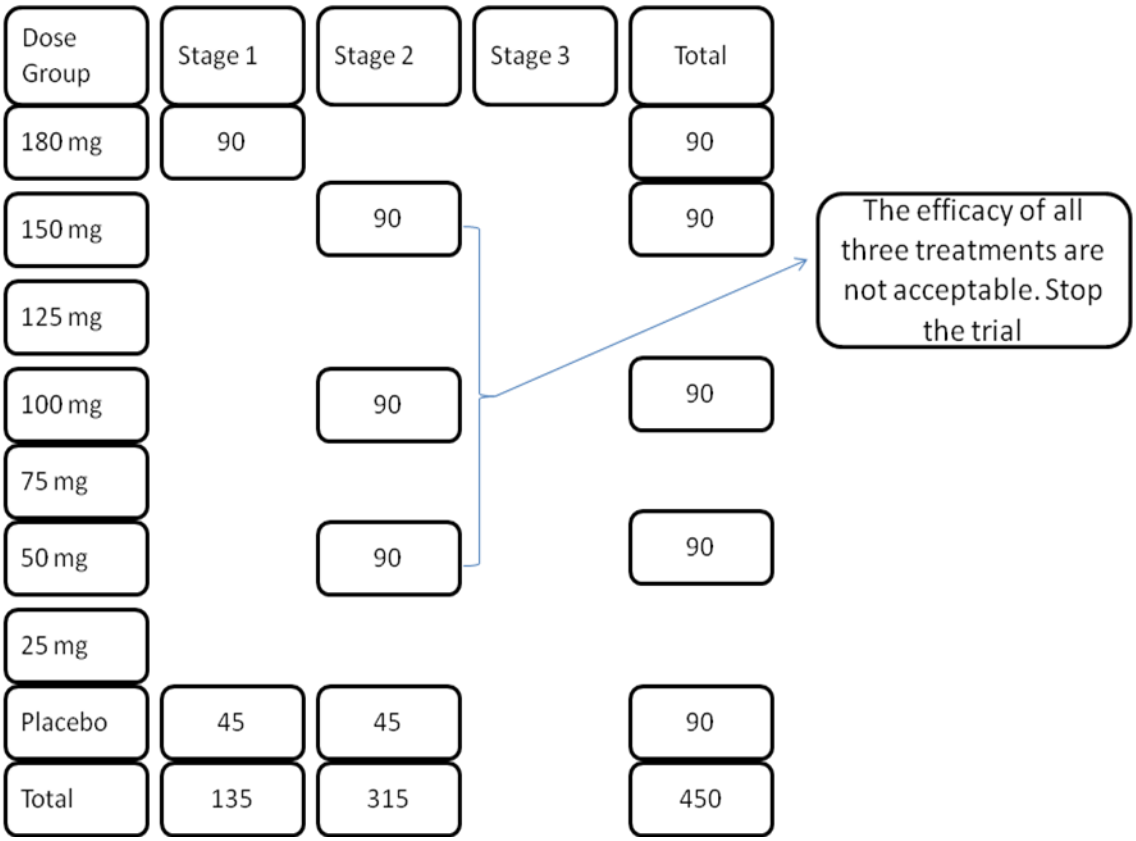


Figure 5.3 Example scenario of the design – The efficacy and safety are acceptable in the first stage. But the efficacies of all three treatments in the second stages are not acceptable. Trial stops.

5.2 Longitudinal Modeling

Since this is an eight-week study, each patient should complete 8-week assessments for each interim analysis. But some subjects may be terminated early from the study due to various reasons, including severe AE, patients withdrew consent, protocol violations, etc. Longitudinal model was developed to impute the missing week 8 data for those early drop-outs. The prediction values of the change from baseline of MADRS at week 8 from longitudinal model were used in efficacy analysis for the subjects who have not completed 8-week study. Historical data from another investigational compound were

examined to assess response patterns over time. Based on this review, a longitudinal model and prior distributions were selected for use in the current trial. The previous research found that some significant factors that affect the response of the treatment: baseline score of MADRS, sex, age [29]. The interaction between sex and time is also found. The following linear multivariate model is used to predict 8-week change from baseline of MADRS:

$$Y_{i,8} = a_{di} + b_t X_{i,t} + c_t S + d_t A + \varepsilon, \quad (5.1)$$

$$i=1, 2, 3, 4, 5, 6, 7,$$

where $X_{i,t}$ is the latest available value for the i-th subject at time t and $Y_{i,8}$ is the final response value for the i-th subject, which is unknown. S is Gender factor, A is Age factor. Error terms ε are the independent identically distributed normal random variables. Intercept a_{di} is dose dependent and is different for different doses. Slope parameters b_t, c_t, and d_t, are assumed to be the same for each dose but depend on the time of the measurement, t. So the intercept is different for each dose, one common slope for all dose for each time point. Distinct models are fit for each time period.

5.3 Test of the Model

The patient data are simulated based on the historic data for different time points: weeks 1, 2, 4, 6, and 8. Figure 5.4 shows the regression output by using SAS procedure. It is used to predict week 8 data for dose level 125 mg, based on week 2 data. That means if the last available data for the patient is at week 2, week 8 data can be predicted by the longitudinal model proposed.

Table 5.1 Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F value	Pr > F
Model	3	6726.57	20.52411	3.856	0.011
Error	143	761.06	5.322		
Corrected Total	146	7487.63			
Root MSE	5.87326		**R-Square**	0.90	
Dependent Mean	13.92612		**Adj R-Sq**	0.81	
Coeff Var	42.17442				

Table 5.2 Parameter Estimate

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	6.52653	0.90021	7.25	0.0873
Week 2	1	0.20976	0.02310	9.08	0.0698
Age	1	0.34270	0.04119	8.32	0.0761
Sex	1	0.87885	0.08394	10.47	0.0606

So the final regression model for Y_8 for dose level 125 mg is:

$$\widehat{Y}_8 = 6.527 + 0.209Y_2 + 0.342 \times Age + 0.879 \times Sex$$

In the real clinical trial, different regression output should be obtained for the different dose and time point.

5.4 Simulation

5.4.1 NDLM Model

The point estimator $\hat{\theta}$ is defined as the change from baseline of MADRS. The point estimation of $\hat{\theta}$ calculation starts with a set of initial values: μ_0, σ_0, a_1, a_2, b_1, b_2, and the sample values of θ can be obtained from full conditional distributions derived in chapter 2. Stage 1 model is showing in 5.2. NDLM model used in stage 2 – 4 is showing in 5.3.

$$Y_{d1} = \theta_{d1} + \varepsilon \tag{5.2}$$

$$\theta_{d1} \rightarrow N(\mu_0, \sigma_0^2) \qquad \varepsilon \rightarrow N(0, \sigma_\varepsilon^2) \rightarrow IG(a_2, b_2)$$

$$Y = \theta_{d_i} + \varepsilon, \quad i=2, 3, 4, 5, 6, 7 \tag{5.3}$$

$$\theta_{d_i} \rightarrow N(\mu_{d_{i-1}}, \sigma_\theta^2) \rightarrow IG(a_1, b_1) \qquad \varepsilon \rightarrow N(0, \sigma_\varepsilon^2) \rightarrow IG(a_2, b_2)$$

5.4.2 Prior Parameter Selection

1) $\mu_0 = 2$

Since a successful efficacy response is defined as a treatment group difference ≥ 2 points comparing study treatment to placebo on MASRS total score at week 8. To obtain equal probability of positive and negative efficacy responses, we choose 2 as a flat prior.

2) $\sigma_0 = 1$

The possibility of increasing on MADRS total score is very small (= 0.025). If the drift effect is noticed in the data, σ_0 could be adjusted to a larger one accordingly. Note that σ_0 is selected as $2/1.96 \approx 1$.

3) $a_1 = a_2 = 2$

Standard deviation doesn't exist when $a_1 = a_2 = 2$ for inverse gamma. The same approach is used when we choose μ_0. The distribution is close to 'non-informative'. The result will be data driven which fits one's need since there is no reliable estimation.

4) $b_1 = 187$, $b_2 = 0.41$

The estimate of standard errors is based on the prior data with some assumption to fit our needs. According to the data in Forest Laboratories' trial [30], standard errors of MADRS change are 0.85 and 0.9 in placebo and study drug groups, respectively.

Using those two observed numbers, the variance of standard error is 0.41 with the mean of 0. Taking 0.41 as the mean of IG, b_2 = mean*(a_2-1) = 0.41 when a_2 = 2.7, b_1 = 187.

Forest Laboratories conducted a phase III clinical trial for MDD indication to test the safety and efficacy of Vilazodone. The LS means and 95% confidence intervals of study drug and placebo on MADRS change at Week 8 are -13.3 (-15.1, -11.5) and -10.8 (-12.6, -9.1), respectively. The sample size is each treatment group is 231. Given this information, the standard error is around 0.9 and the standard deviation is about 13.7 for both cases. The variance is calculated to be 187. Since the value of 'a' is chosen to be 2, the value of 'b' is equal to the mean of variance which is (a-1)*187 = 187.

5.4.3 Simulation

Simulation has been done for this design. Table 5.3 shows one scenario of the assumed probability of in the simulation. Forty-five random values are taken from SAS random function as the observed values from normal distribution with previously assigned N (2, 1) for $\theta_{d1.}$ The prior chosen for σ_ε^2 is IG (2, 0.41). In stage 1, one hundred thirty-five patients are assigned to the 180 mg dosing group. The simulation was done for the first interim analysis with forty-five subjects' values. Another forty-five subjects' values were simulated for the second and third interim analyses after the first interim analysis data

obtained. Six thousand iterations are used in the program and the first 1000 burn-in results are discarded. The sample response data of ten subjects are shown in Tables 5.4. Table 5.5 shows additional ten sample subjects added in the next stage and used to confirm the results in the previous stage.

The final results for one scenario are shown in Table 5.6. The 125 mg and 25 mg doses were not tested since the 150 and 100 mg doses are successful and 50 mg failed on efficacy. The maximum patients need to be recruited in this scenario is 224. Thus, 541 subjects in sample size are saved. Figure 5.4 – Figure 5.8 show the convergence of θ_{di} for each dose level. The data show the convergence through the 50000 iterations. It also shows the posterior distribution of $\theta_{di.}$

Table 5.3 Scenario 1 - True Mean of SUV Decreasing Used in Simulation

Dose Groups	True Mean of Treatment Difference
180 mg	3.5
150 mg	3.0
125 mg	2.5
100 mg	2
75 mg	1.5
50 mg	1.0
25 mg	0.5
Placebo	0

Table 5.4 Sample Response from Ten Patients of Each Dose Level

25 mg	50 mg	75 mg	100 mg	125 mg	150 mg	180 mg
NT	1.2463	1.6133	2.1816	NT	3.5231	3.4364
	1.2233	1.7225	2.1356		2.9425	3.3537
	0.8972	1.6920	2.2626		2.9667	3.3286
	1.1578	1.7527	1.9762		3.2563	3.3696
	0.9384	1.7399	2.2038		2.9896	2.8457
	1.2893	1.6223	2.0728		2.7325	3.4341
	0.9438	1.6592	2.0695		3.0665	3.4379
	0.9102	1.8158	2.2569		3.2164	2.8973
	1.2954	1.8657	1.9857		3.0786	3.6705
	1.1879	1.9379	2.1061		2.9173	3.2780

NT: Not Tested

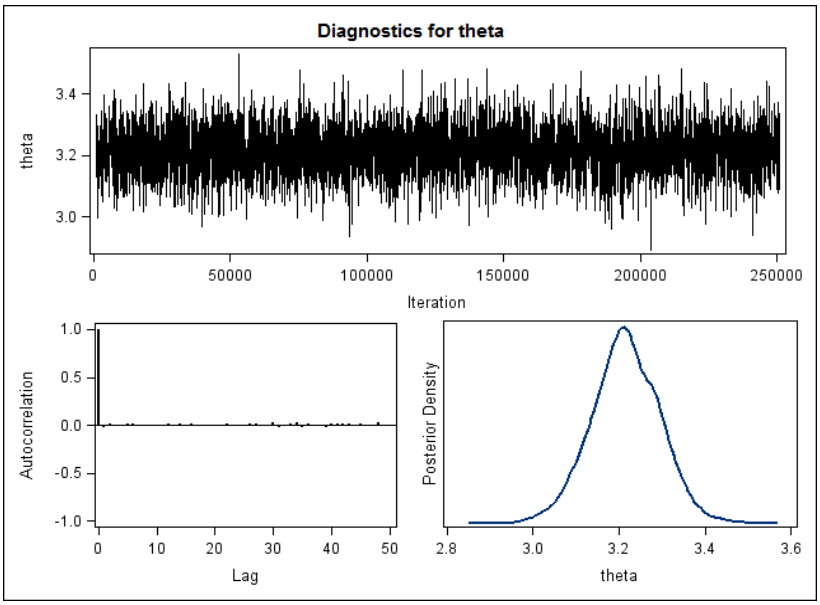


Figure 5.4 Convergence, auto-correlation and posterior distribution of θ_d for 180 mg.

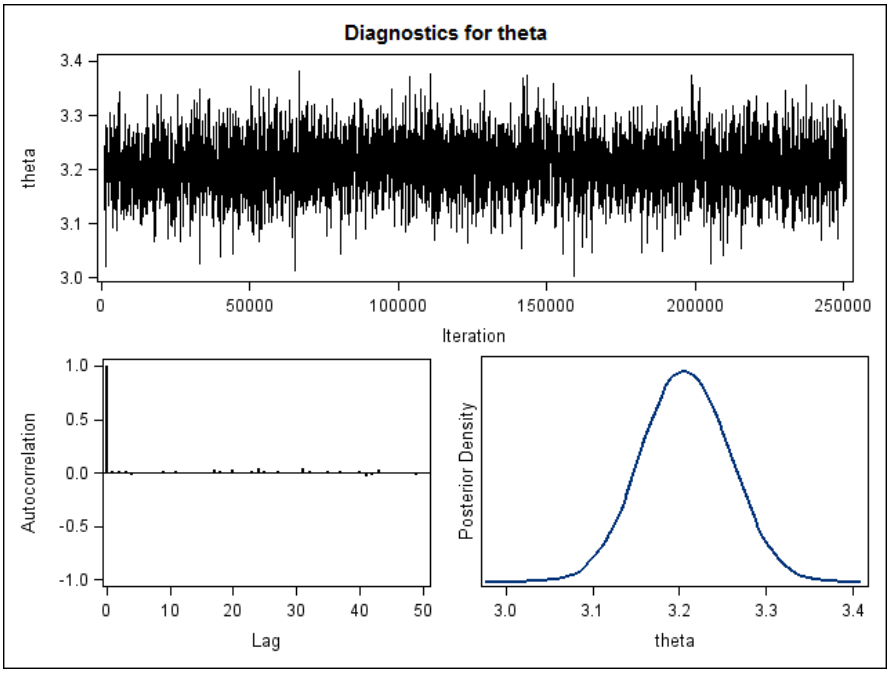


Figure 5.5 Convergence, auto-correlation and posterior distribution of θ_d for 150 mg.

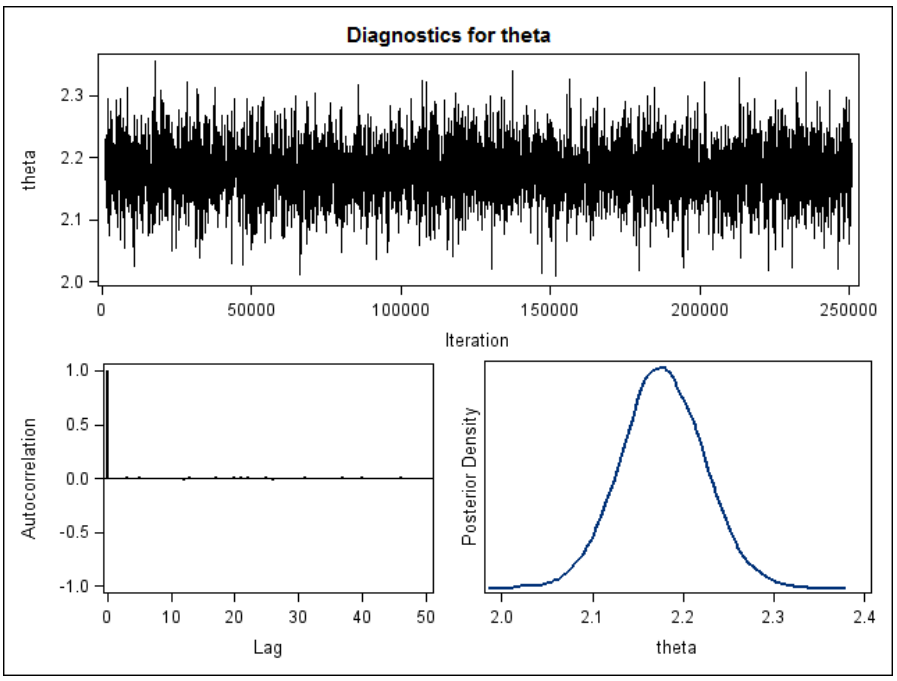


Figure 5.6 Convergence, auto-correlation and posterior distribution of θ_d for 100 mg.

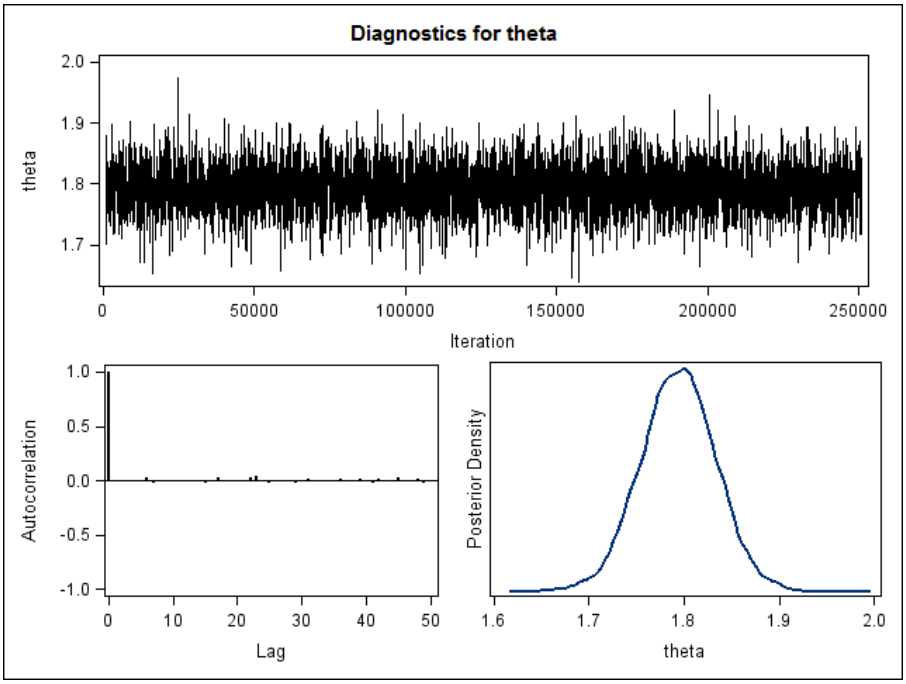


Figure 5.7 Convergence, auto-correlation and posterior distribution of θ_d for 75 mg.

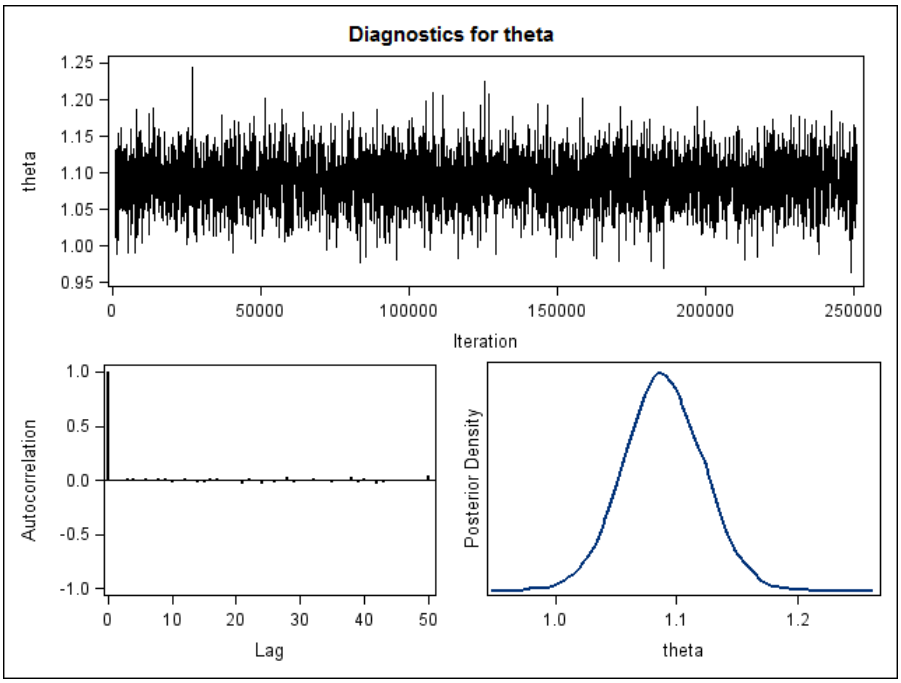


Figure 5.8 Convergence, auto-correlation and posterior distribution of θ_d for 50 mg.

Table 5.5 Posterior Information for Each Dose Group

Dose Group	True Mean of Tested Dose Groups Compared to Placebo Used in Simulation	Posterior Mean at the end of Testing Stage	Posterior Std at the end of Testing Stage	Posterior Probability (%) of Mean difference ≥ 2 points on MADRS	Average Sample Size used in the trial
180 mg	3.5	3.18	0.407	99.8	30
150 mg	3.0	3.20	0.247	100	30
125 mg	2.5	Not Tested	Not Tested	Not Tested	0
100 mg	2	2.20	0.196	86.5	41
75 mg	1.5	1.81	0.158	10.6	38
50 mg	1.0	1.08	0.099	0.00	30
25 mg	0.5	Not Tested	Not Tested	Not Tested	0
Placebo	0	Controlled	Controlled	Controlled	55
Total					224

5.5 Operating Characteristics

Table 5.6 shows operating characteristics for each dose level. In each scenario, 5000 simulated trials were conducted. The results from the table are showing that the model chosen and the design have the high power to get the correct results.

Table 5.6 Operating Characteristics

Scenario	Assumed Increasing (%) at Dose Levels							Percent of Trials Selecting the Right Doses (%)	Mean of the Number of Subjects Used (Savings %)
	180 mg	150 mg	125 mg	100 mg	75 mg	50 mg	25 mg		
1	3.5	3	2.5	2	1.5	1	0.5	94	224 (61.7)
2	3.1	3	2.7	2.5	1.5	1.2	1	94	214 (63.4)
3	3.1	3	2.1	2	1.7	1.2	1	84	250 (57.3)
4	0	0	0	0	0	0	0	100	45 (92.3)
5	3	0	0	0	0	0	0	100	151 (74.2)
6	3	3	3	0	0	0	0	100	181 (69.1)
7	3	3	3	3	3	3	3	100	195 (66.7)

CHAPTER 6

CONTINUOUS-TIME MARKOV CHAIN MODEL AND ITS APPLICATION IN TIME-TO-EVENT DATA IN CLINICAL TRIALS

6.1 Markov Chain

A Markov chain, named after Andrey Markov, is a stochastic process with the Markov property. Markov property is a mathematical model for the random evolution of a memoryless system. In other words, the sequence is a Markov chain if the probability that the system enters the state i_t at time t depends only on the immediately preceding state i_{t-1} at time t-1. Simplistically, the future is independent of the past. Often, the term Markov chain is used to mean a discrete-time Markov process. Mathematically, Markov chain is defined as follows:

Let $\Omega = \{1,2,\ldots,m\}$ $(m < \infty)$ be a finite state space, and let $\{Y_t\} = \{Y_0, Y_1, \ldots, Y_t, \ldots\}$ be a sequence of random variables defined on Ω. Then the sequence will be called a finite Markov chain if, for any sequence $\{Y_0 = i_0, Y_1 = i_1, \ldots, Y_{t-1} = i_{t-1}, Y_t = i_t\}$, t = 1,2,…, satisfies,

$$P(Y_t = i_t | Y_{t-1} = i_{t-1}, \ldots, Y_0 = i_0) = P(Y_t = i_t | Y_{t-1} = i_{t-1}) \tag{6.1}$$

The conditional probabilities $P(Y_t = j | Y_{t-1} = i) = p_{ij}(t)$, $i, j \in \Omega$

$p_{ij}(t)$ are called one-step transition probabilities for the system at time t. A Markov chain $\{Y_0, Y_1, \ldots\}$ is homogeneous if the transition probabilities are constant in time, *i.e.* $P(Y_t = j | Y_{t-1} = i) = p_{ij}$ for any $i, j \in \Omega$, and all t = 1, 2, …. It is equivalent to saying that the transition probability matrices of a homogeneous Markov chain may be represented by the single matrix $M = M_t = (p_{ij})$ for all t = 1, 2, …, where the transition probabilities p_{ij}

are free of the time index t. The transition probabilities p_{ij}(t), $1 \leq i, j \leq$ m, may be represented as p_{ij}, an m × m matrix:

$$M = (p_{ij}) = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{pmatrix}_{m \times m} \tag{6.2}$$

The matrix M is called one-step transition probability matrices.

6.2 Continuous-time Markov Chain Model

Continuous-time Markov model has the advantage over the traditional survival model and can be used to describe disease as a series of probable transitions between health states. This is an attractive feature since it provides the ability to describe the course of disease over time. It can also describe and estimate expected survival in clinical cohort. The transition probability matrix P summarizes the probabilities of events and can be used to describe the probabilistic course of the disease for a population or for an individual with a known health state.

Markov process is time-independent and time homogenous when the transition probabilities are constant during the process. One can assume that the distribution of the number of transitions into a state follows a homogenous Poisson process. The Poisson distribution is described as $\Pr\{N(t) = k\} = ((\lambda t)^k e^{-\lambda t})/k!$, where λ is the average number of transitions per unit time with λt the average in period t, and k the exact number of observed transitions. The time between transitions in a homogenous Poisson process follows an exponential distribution defined by the same parameter λ. Exponential

distribution has memory-less property and can be certainly applied with Markov chain to disease progression.

Time to progression and overall survival time are the important endpoints in many oncology trials. One can assume the distribution of the events follow Poisson distribution and the time to progression and overall survival follows an exponential distribution. In the next few sections, the applications of continuous-time Markov process applied to time to progression and overall survival analysis will be discussed.

6.2.1 Transition Count Matrix and Transition Probability Matrix

In oncology clinical trials, the transition between health states is actually a rate for a continuous-time Markov process. The transition rate does not depend on the length of the observation interval since it is the number of transitions that occur per unit time. The transition count matrix S(t) contains components f_{ij} which are the counts from state i to j in each cycle (at time t) [31]. The count matrix could be written as for a three-state Markov chain model:

$$S = \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{ccc} 1 & 2 & 3 \end{array} \\ \begin{pmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{23} & f_{33} \end{pmatrix} \end{array} \tag{6.3}$$

The probability of transition in a Markov chain can be derived based on the summation of the count matrix for all the cycles and the observation interval. The summation matrix is used to construct the probability (P) matrix using maximum likelihood estimates of $\hat{p}_{ij}$,

the probability of transition from state i to state j, given by $\hat{p}_{ij} = f_{ij}(k)/f_{i\cdot}(k)$, where f_{ij} is the frequency or count of patients making the transition from state i to state j, $f_{i.}$ is the sum of patients initially in state i and k is the cycle with total of K cycles [32].

6.2.2 Markov Models with Covariates

In clinical trials, there are many covariates that may cause the process to be non-homogenous over time. Some common covariates are age, disease characteristics, race, etc. One strategy to adjust the non-homogenous transitions is to add states to the transition matrix but it will increase model complexity and will require more computation. Another strategy is to create separate transition matrices. For example, if age dependent transition is suspected, the model can be stratified by age groups and separate transition matrices for different age groups can be created.

In case there is time-dependent transition, separate matrices can be created for different time periods and can be run simultaneously.

6.2.3 Continuous-time Markov Model in Time to Progression Analysis

Time to progression is a widely used endpoint in oncology clinical trials. Kaplan-Meier product limit is commonly used to analyze the time to progression endpoint. Only two states are considered when Kaplan-Meier product limit is used: progression and non-progression. The non-progression patient will always be right censored. The right censored patients include: 1) Patient with response. 2) Early drop-off from the study due to toxicity. 3) Lost to follow-up. 4) Died due to other cause rather than disease. For the patients who have response and dropped off from the study early, time to progression were always right censored to the last available date collected in the trial data. However,

the patients could develop progression after lost to follow-up or discontinuation and shouldn't be considered as no-event. In that case, simply censored the subjects could cause biased prediction of the study result. Also, the next patient state only depends on the current state, and independent of previous state. That is, whatever the previous states the patient was in, the next cycle state could be any states defined in Markov model. For example, patients may have standard disease (SD), response (R) in the previous cycle and they could develop into progression disease (PD), stay in SD/R, and drop off from the study or even death in the next cycle. The state they develop into is independent from the previous state. By using Markov model, those states such as lost to follow-up, death due to other caused, can be added into Markov chain and viewed as continuous-time Markov process. The use of Markov model for right-censored data has an advantage over traditional survival analysis in that each censored subject contributes more information to the model than it can contribution to the survival analysis with only two states. The prior state transitions these subjects experience add useful information [33]. Hence, Markov process can help the investigator to describe the course of the disease status over time by adding in the different states into the model.

The elements of the probability matrix $\mathbf{P} = (p_{ij})$ describe the probability of going from state i to state j in one cycle. The operation of $P \cdot P \cdot P \cdots P$ (n times), depicted as $P(n) = P^n$ yields a matrix denoted the nth matrix, whose individual elements $p_{ij}(n)$ are the probabilities of transition to state j from state i after n cycles [34]. This information is applied to a population rather than the individual. The matrix multiplication estimates the probability of reaching a certain state for an average subject after n cycles. It also estimates the proportion of a population that resides in a certain state after n cycles. The

analysis that has been done before Markov process converges is called transition analysis. Many researchers have used graphic depiction for the transition analysis. The probability curve calculated from transition matrix P^n is sometimes referred to as "Markov survival curve" [35]. Transition analysis may cause convergence of the probability distribution vector to tend to a limiting specific value as n increases. When the number of cycles increases, the probability vector approaches a limiting value and the Markov process reach a steady state.

6.2.4 Model Validation

There are two fundamental assumptions in Markov model: Markovian assumption of current state depending on the immediate past and time homogeneity. Those two assumptions need to be validated before the analysis can move forward.

Jain used likelihood ratio test to test the time homogeneity [36]. The test compares the observed transition probabilities with expected probabilities derive from the model.

Data Splitting is another way to test the time homogeneity [37]. It is separating data and using one portion of the data to fit the model. The model is used to predict the expected state distribution for the future time period that is compared to the observed state distribution of the remaining data already collected.

6.3 Data and Results

6.3.1 Data

Data from a single arm, open-label phase II study for patient with neuroendocrine carcinoma was used [38]. The treatment has been studies in several Phase I and Phase II trials in other indications. The safety profile was believed to be satisfactory. Patients were treated in an outpatient setting, receiving once weekly doses of the treatment via a 30-minute infusion. A cycle of treatment was defined as 28 days. So the patient received four treatments per cycle. Patients were to continue with treatment until disease progression, patient withdrew consent, severe adverse event, remove from study due to physician discretion, or lost to follow-up. Patient data with patient response status is listed in Table 6.1. Adverse events (AE) are any untoward medical occurrence in a patient or clinical investigation subject administered a pharmaceutical product and which does not necessarily have a causal relationship with this treatment. An adverse event (AE) can therefore be any unfavorable and unintended sign (including an abnormal laboratory finding), symptom, or disease temporally associated with the use of a medicinal investigational product, whether or not related to the medicinal investigational product [39].

Six states are defined in Table 6.2. There are three transient states, O (original state, which is the state patients entered the trial), response (R) and standard disease (SD). Three absorbing states were also defined: off study due to various reasons (C), progressive Disease (PD), and death (D). The definition of health states was constructed such that the states were mutually exclusive.

For each cycle, a count matrix was constructed based on the number of patients making the respective transitions. The count matrices were summated to give the overall summation (S) matrix [36]. The summation matrix was used to construct the transition probability (P) matrix using maximum likelihood estimates of $\hat{p}_{ij}$, the probability of transition from state i, the previous state, to state j, the future state, given by $\hat{p}_{ij}$ = f_{ij}(k)/$f_{i\cdot}$(k) where f_{ij} is the frequency or count of the patients making the transition from state i to state j at time k and $f_{i\cdot}$ is the count for all transitions from state i at time k. Thus the overall summation matrix is:

$$S = \begin{array}{c} \\ O \\ R \\ SD \\ PD \\ C \\ DTH \end{array} \begin{array}{c} \begin{array}{cccccc} O & R & SD & PD & C & DTH \end{array} \\ \begin{pmatrix} \sum_{k=1}^{K} f_{11}(k) & \sum_{k=1}^{K} f_{12}(k) & \cdot & \cdot & \cdot & \sum_{k=1}^{K} f_{16}(k) \\ \sum_{k=1}^{K} f_{21}(k) & \sum_{k=1}^{K} f_{22}(k) & \cdot & \cdot & \cdot & \sum_{k=1}^{K} f_{26}(k) \\ \sum_{k=1}^{K} f_{31}(k) & \sum_{k=1}^{K} f_{32}(k) & & & & \sum_{k=1}^{K} f_{36}(k) \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \sum_{k=1}^{K} f_{61}(k) & \sum_{k=1}^{K} f_{62}(k) & \cdot & \cdot & \cdot & \sum_{k=1}^{K} f_{66}(k) \end{pmatrix} \end{array} \quad (6.4)$$

Table 6.1 Patient Data

	Cycle								
ID	1	2	3	4	5	6	7	8	9
001	SD	SD	SD	SD	PD				
002	SD	SD	SD	WC					
003	SD	PD							
004	SD	SD	PD						
005	PD								
006	SD	SD	SD	SD	SD	SD	SD	SD	PD
007	SD	SD	SD	DTH					
008	AE								
009	PD								
010	PD								
011	PD								
012	PD								
013	PD								
014	SD	SD	SD	SD	SD	SD	SD	DTH	
015	SD	SD	R	R	R	R	R	R	ONTR(R)
016	SD	SD	SD	SD	SD	SD	SD	SD	PD
017	DTH								
018	SD	SD	SD	SD	SD	SD	SD	AE	
019	PD								
020	PD								
021	SD	SD	R	LTF					
022	SD	PD							
023	SD	WC							
024	R	R	WC						
025	SD	PD							
026	SD	PD							
027	AE								
028	SD	SD	SD	SD	R	R	R	R	ONTR(R)
029	AE								
030	SD	SD	SD	SD	SD	SD	SD	DTH	
031	SD	SD	SD	SD	SD	SD	SD	SD	ONTR(SD)
032	SD	SD	SD	PD					
033	SD	SD	SD	SD	SD	SD	SD	SD	ONTR(SD)
034	PD								
035	SD	SD	SD	PD					
036	SD	WC							

AE: Adverse events, ONTR: On-treatment, PHYS: Remove from the study due to physician discretion, WC: Withdrew consent. PD: Progression, DTH: Death, LTF: Lost to follow-up. SD: Standard disease. R: Response.

Table 6.2 Definition of Patient States

State	Definition	Type
O	Original state patient enter the trial	Transient
R	Response	Transient
SD	Standard disease	Transient
PD	Progression	Absorbing
C	Off study due to withdrew consent, lost to follow-up or Adverse Events	Absorbing
DTH	Death	Absorbing

Based on the data, the count matrix for each cycle is:

$$S_1 = \begin{array}{c} \\ O \\ R \\ SD \\ PD \\ C \\ DTH \end{array} \begin{array}{c} \begin{array}{cccccc} O & R & SD & PD & C & DTH \end{array} \\ \begin{pmatrix} 0 & 1 & 17 & 12 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array} \quad (6.5)$$

$$S_2 = \begin{array}{c} \\ O \\ R \\ SD \\ PD \\ C \\ DTH \end{array} \begin{array}{c} \begin{array}{cccccc} O & R & SD & PD & C & DTH \end{array} \\ \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 16 & 4 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array} \quad (6.6)$$

….

$$S_8 = \begin{array}{c} \\ O \\ R \\ SD \\ PD \\ C \\ DTH \end{array} \begin{array}{c} \begin{array}{cccccc} O & R & SD & PD & C & DTH \end{array} \\ \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array} \quad (6.7)$$

The summation matrix for all eight cycles is:

$$S = \begin{array}{c} \\ O \\ R \\ SD \\ PD \\ C \\ DTH \end{array} \begin{array}{c} \begin{array}{cccccc} O & R & SD & PD & C & DTH \end{array} \\ \begin{pmatrix} 0 & 1 & 17 & 12 & 5 & 1 \\ 0 & 6 & 1 & 0 & 1 & 0 \\ 0 & 2 & 61 & 9 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array} \quad (6.8)$$

So the transition matrix can be estimated from $\hat{p}_{ij} = f_{ij}(k)/f_{i\cdot}(k)$.

$$P = \begin{array}{c} \\ O \\ R \\ SD \\ PD \\ C \\ DTH \end{array} \begin{array}{c} \begin{array}{cccccc} O & R & SD & PD & C & DTH \end{array} \\ \begin{pmatrix} 0 & 0.028 & 0.472 & 0.389 & 0.139 & 0.028 \\ 0 & 0.75 & 0.125 & 0 & 0.125 & 0 \\ 0 & 0.027 & 0.824 & 0.122 & 0.014 & 0.014 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array} \quad (6.9)$$

6.3.2 Time Homogeneity Assumption

The likelihood ratio test is used to test the homogeneity assumption [36]. The criterion λ is given by $\lambda = \prod_{i,j=1}^{5} \prod_{k=1}^{9} [\frac{\hat{p}_{ij}}{\hat{p}_{ij}(k)}]^{f_{ij}(k)}$.

The Hypothesis is:

H_0: $p_{ij}(k) = p_{ij}$ for all k = 1, 2, 3, 4, … 9

H_1: the transition probabilities depend on k.

Evaluation across the entire nine cycles shows that χ^2 = 161.34, df = 160, p = 0.4554. Thus the null hypothesis of constant transition probability matrix can not be rejected. The model can be represented by a single transition matrix.

6.3.3 N Step Transition Analysis

By using n-step transition analysis, the transition matrix will be multiply by 'n' times in order to get the steady state. SAS/IML is used for the transition matrix analysis. The complete absorption occurred by 40 cycles.

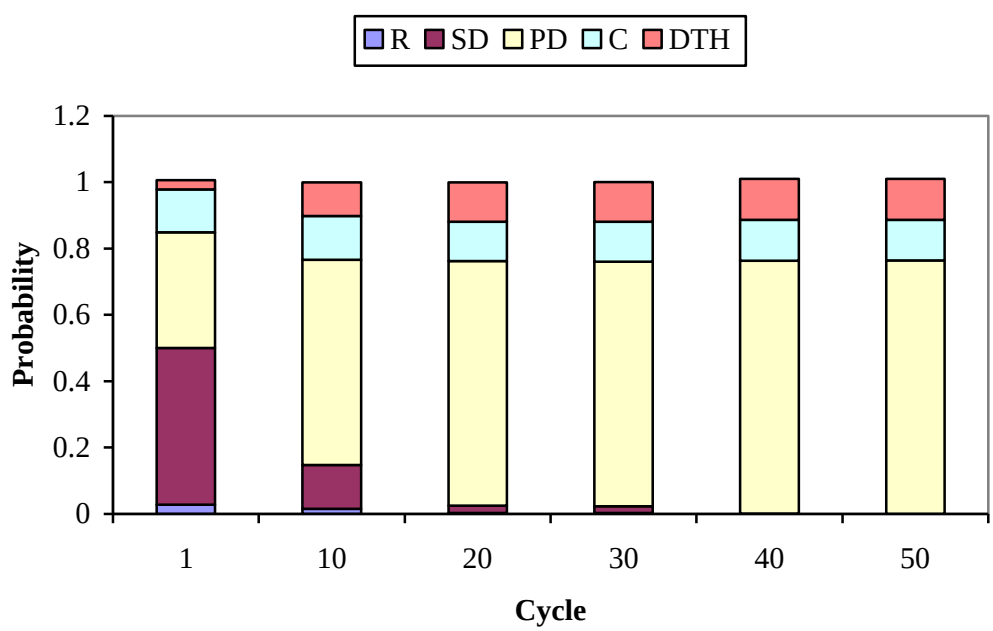


Figure 6.1 The n-step transition analysis of Markov Model. The transition state: R, SD. The absorbing state: PD, C, DTH.

The probabilities in the transient state decrease over the time and the probabilities of absorpting states increase such that all patients are eventually absorbed.

6.3.4 Estimation of Time to Progression using Matrix Algebra

By using matrix algebra, the residence time in each transition state can be calculated. The transition probability matrix of a Markov chain that contains absorbing states can be divided into four sections under appropriate arrangement of the state space [40]. As showing in Figure 6.2, the section labeled Q reflects the transition probabilities between the transient states, which are the probabilities of going from one transient state to another transient state; the section O is zero matrix, which represents zero probability of transition from absorbing state to transient state; the section R reflects the probability of being absorbed from transient state to absorbing state; and section I is an identity matrix.

$$P = \begin{array}{c} \\ \text{From: Transient States} \\ \\ \text{Absorbing States} \end{array} \begin{array}{c} \begin{array}{cc} \text{To: Transient States} & \text{Absorbing States} \end{array} \\ \left(\begin{array}{c|c} Q & R \\ \hline O & I \end{array}\right) \end{array} \qquad (6.10)$$

Figure 6.2 Four Sections in probability matrix containing absorbing states.

The fundamental matrix N is calculated by taking the inverse of (I – Q) [41]. The fundamental matrix indicates the expected number of cycles the patient will be in any of the transient states before absorption occurs [42]. Based on the fundamental matrix $N = (I - Q)^{-1}$, the average number of cycles that a patient resided in either transition state before

absorption can be estimated. The matrix N specifies the number of cycles that the patients resided in the transient state such that $N = (I - Q)^{-1}$, where I is the identity matrix and Q is the square matrix of the transient probabilities in Figure 6.2 [43].

The variance of N is $V = N(2N' - I) - N^2$, where N' has the same diagonal entries as N and zeros elsewhere and N^2 is square of N [41]. Each element of V is the variance of the corresponding element of N. The square root of each element in V was used as the standard error of the corresponding element of N. Based on the transition matrix P derived earlier, Q, N and V are derived as follows:

$$Q = \begin{array}{c} \\ O \\ R \\ SD \end{array} \begin{array}{c} \begin{array}{ccc} O & R & SD \end{array} \\ \begin{bmatrix} 0 & 0.028 & 0.472 \\ 0 & 0.75 & 0.125 \\ 0 & 0.027 & 0.824 \end{bmatrix} \end{array}$$

$$N = (I - Q)^{-1} = \begin{array}{c} \\ O \\ R \\ SD \end{array} \begin{array}{c} \begin{array}{ccc} O & R & SD \end{array} \\ \begin{bmatrix} 1 & 0.435 & 2.991 \\ 0 & 4.332 & 3.077 \\ 0 & 0.665 & 6.154 \end{bmatrix} \end{array}$$

$$V = N(2N' - I) - N^2 = \begin{array}{c} \\ O \\ R \\ SD \end{array} \begin{array}{c} \begin{array}{ccc} O & R & SD \end{array} \\ \begin{bmatrix} 0 & 0 & 3.8 \\ 0 & 3.57 & 3.41 \\ 0 & 0 & 8.41 \end{bmatrix} \end{array}$$

$$SE = \begin{array}{c} \\ O \\ R \\ SD \end{array} \begin{array}{c} \begin{array}{ccc} O & R & SD \end{array} \\ \begin{bmatrix} 0 & 0 & 1.949 \\ 0 & 1.889 & 1.847 \\ 0 & 0 & 2.431 \end{bmatrix} \end{array}$$

Thus, the duration (days) of SD and R can be found from N by multiply 28 since each cycle has 28 days:

$$N\ (\text{days}) = \begin{array}{c} \\ O \\ R \\ SD \end{array} \begin{array}{c} \begin{array}{ccc} O & R & SD \end{array} \\ \begin{pmatrix} 0 & 12.2 & 83.7 \\ 0 & 121.3 & 86.2 \\ 0 & 18.6 & 172.3 \end{pmatrix} \end{array}$$

Since patient started from the state O, the average duration of the patients stay in transient state R is 12.2 days for all the patients, and the average duration of the patients stay in transition state SD is 83.7 days for all the patients. Estimation of the sum of the total average residence time in transient states before going into the absorption states as seen in (6.9) is 12.2 + 83.7 = 95.9 days with 95% CI (36.6 – 155.2). This sum is considered to be survival time, which is time to progression in this study. The expected time to progression is 95.9 days. For the patients who transit from starting state to R, the average time staying in R is 121.3 days and average survival time staying in SD is 86.2 days. For the patients who transit from R to state SD, the average survival time to stay in R is 18.6 days and the average survival time to stay in SD is 172.3 days.

Next, the probability that the patient will eventually be absorbed given any starting state can be obtained by $N \times R$ matrix [42].

$$N = \begin{array}{c} \\ O \\ R \\ SD \end{array} \begin{array}{c} \begin{array}{ccc} O & R & SD \end{array} \\ \begin{bmatrix} 1 & 0.435 & 2.991 \\ 0 & 4.332 & 3.077 \\ 0 & 0.665 & 6.154 \end{bmatrix} \end{array}$$

$$R = \begin{array}{c} \\ PD \\ C \\ DTH \end{array} \begin{array}{c} \begin{array}{ccc} PD & C & DTH \end{array} \\ \begin{bmatrix} 0.389 & 0.139 & 0.028 \\ 0 & 0.125 & 0 \\ 0.122 & 0.014 & 0.014 \end{bmatrix} \end{array}$$

$$N \times R = \begin{array}{c} \\ O \\ R \\ SD \end{array} \begin{array}{c} \begin{array}{ccc} PD & C & DTH \end{array} \\ \begin{bmatrix} 0.754 & 0.235 & 0.080 \\ 0.375 & 0.585 & 0.043 \\ 0.751 & 0.169 & 0.086 \end{bmatrix} \end{array}$$

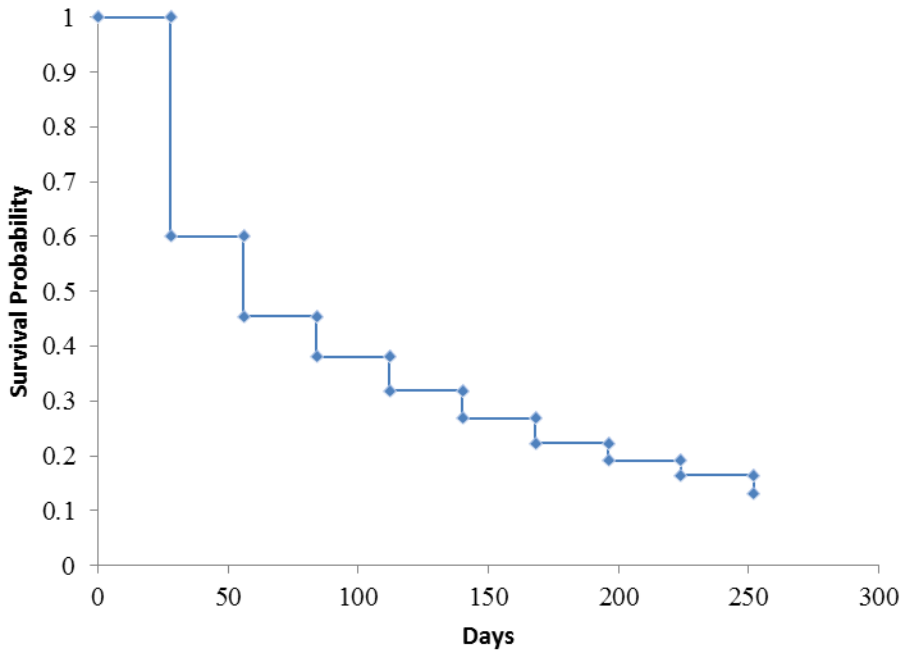
From the above calculation, one can see that the probability that patient will be absorbed into PD state is 0.754. The probability that patient will eventually drop off from the study is 0.235. The probability that patient will die is 0.080.

6.3.5 Survival Curves from Markov Model

Survival probabilities were obtained based on the transition matrix after applying matrix algebra. For the first cycle, the survival probability is taken from P. For n cycles, the survival probability is taken from transition matrix P^n. The survival function $S(t) = p_{12}(t) + p_{13}(t)$, where p_{12} and p_{13} are the transition probabilities for patients transit to the transient state (SD and R, respectively) at time t [35]. Table 6.3 shows the survival probabilities calculated from transition matrix for each cycle. Figure 6.2 shows the survival curve for the patients.

Table 6.3 Survival Probabilities

28 Days	56 Days	84 Days	112 Days	140 Days
0.600	0.454	0.381	0.319	0.268
168 Days	**196 Days**	**224 Days**	**252 Days**	
0.223	0.191	0.165	0.131	



Days

Figure 6.3 Survival curve from Markov model.

6.4 Estimation using Kaplan-Meier Estimate

Kaplan-Meier estimate is also used to find out the survival curve and expected time to progression. Lost to follow-up, discontinuation due to AE and other reasons, and on treatment patients are right censored to the last available date in the data. The median time to progression is 112 days with 95% confidence interval 56 – 224 days. The mean

time to progression is 111 days with standard error 17.34. Thirty-six patients were observed and thirteen patients were right censored (Figure 6.2). The time table is listed in Table 6.3.

By comparing mean of the time to progression calculated from Markov model and Kaplan-Meier analysis, one can find that the expected time to progression calculated from Markov model is shorter than the expected time to progression from Kaplan-Meier analysis. It is because that more states are added to the transition matrix model and this reduces the bias caused by non-informative right censoring used in Kaplan-Meier analysis. In Kaplan-Meier estimation, only two states are considered, progression and not progression. All the patients who have states SD, R, C were right censored, which has the effect of increasing the survival estimate and their confidence intervals. The Markov model estimates were stable and the variance of such estimates was less than those obtained by Kaplan-Meier estimation.

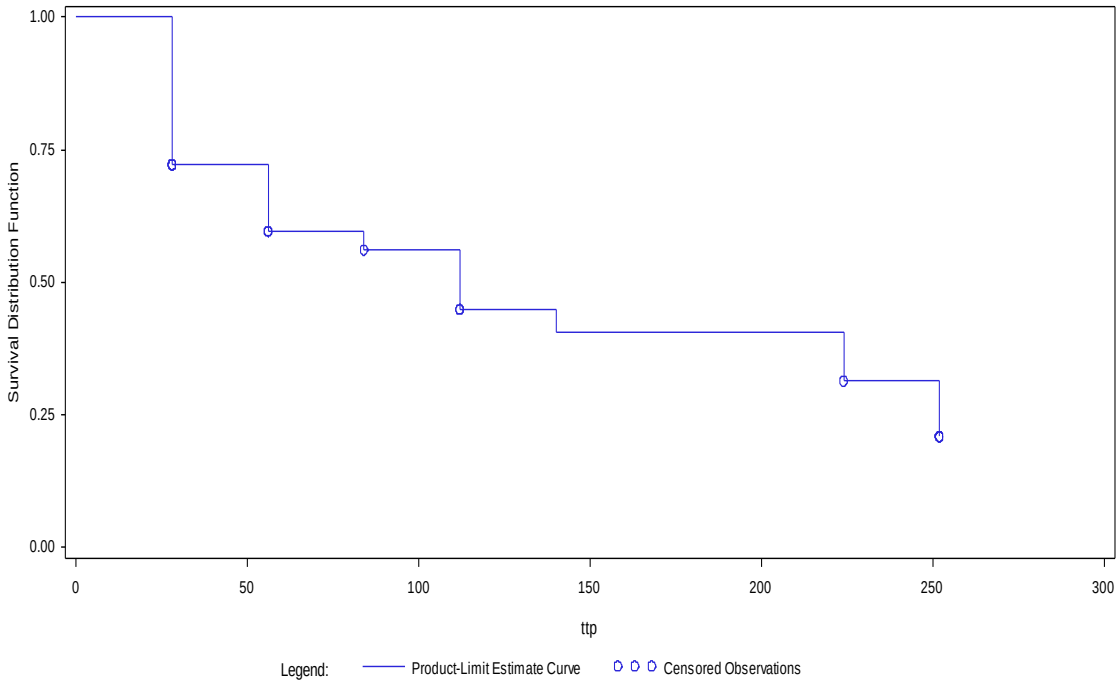


Figure 6.4 SAS output of Kaplan-Meier product limit estimation of the survival function for time to progression analysis.

Table 6.4 Kaplan-Meier Time Table for Time to Progression Analysis

Time to Progression (Days)	Survival	95% CI Lower limit	95% CI upper limit	Censoring
0	1	1	1	No
28	0.7222	0.5427	0.8398	No
28	0.7222	0.5427	0.8398	Yes
28	0.7222	0.5427	0.8398	Yes
28	0.7222	0.5427	0.8398	Yes
56	0.5966	0.4142	0.7389	No
56	0.5966	0.4142	0.7389	Yes
56	0.5966	0.4142	0.7389	Yes
84	0.5615	0.3785	0.7098	No
84	0.5615	0.3785	0.7098	Yes
112	0.4492	0.2795	0.6128	No
112	0.4492	0.2795	0.6128	Yes
112	0.4492	0.2795	0.6128	Yes
140	0.4043	0.2280	0.5742	No
224	0.3145	0.1518	0.4916	No
224	0.3145	0.1518	0.4916	Yes
252	0.2096	0.0738	0.3920	No
252	0.2096	0.0738	0.3920	Yes
252	0.2096	0.0738	0.3920	Yes
252	0.2096	0.0738	0.3920	Yes
252	0.2096	0.0738	0.3920	Yes

CHAPTER 7

CONCLUSION AND DISCUSSION

In the current age, Bayesian Adaptive Design is becoming more and more popular in clinical trials, especially for early phase trials. According to the simulation results, the proposed Bayesian Adaptive Designs are sensitive and robust to help investigators draw conclusion as early as possible. The designs have the ability to utilize accumulating data obtained in real time to alter the course of the trial, thereby enabling dynamic allocation to different dosing groups and dropping of ineffective dosing group earlier. The posterior probability of success for the trial is from 72-100% based on the simulation result. It increased the probability of success comparing with the other adaptive dose finding design. So it provides the better treatment to the patients. In this thesis, both of the Bayesian designs can reduce sample size substantially which in turn leads to savings in cost and time.

However, Bayesian Adaptive Design also has some disadvantages. One major limitation is the difficulty to control type I error after so many adaption. As the consequence, it is relatively hard to apply Bayesian Adaptive Design in phase III submission trials. How to control type I error efficiently is a topic worth more research in the future. The second Bayesian adaptive design is fit for the clinical trials which are hard to recruit the patients and the trials with slow patient enrollment. This design is also fit for the short term trials in which the interim analysis results can be obtained quickly after the patients start the trials. For the long term trials or the trials which have fast patient enrollment, the longitudinal model can be applied earlier. It is also important to find

appropriate priors. The priors should be defined based on external evidence. Also there are computational challenges associated with such a choice.

The Markov chain model is an alternative way to do the exploratory analysis in oncology trial. The model has the advantage over the traditional survival model and can be used to describe disease as a series of probable transitions between health states and predict the expected time to progression time. This is an attractive feature since it provides the ability to describe the course of disease over time. However since the complication for each oncology trial and treatment, the model needs to be validated beforehand and the time homogeneity needs to be tested.

In case the Markov process is non-homogeneous, covariate can be added into the model. The model can be stratified by covariates and separate transition matrices for different covariates can be created.

The precision of the Markov estimates obtained by matrix algebra will be also affected by sample size with small sample sizes resulting in greater variance. This is a disadvantage of the Markov methods, especially if the population is stratified. In this the situation, MCMC simulation can be useful since the precision of the estimates can be increased by increasing the number of individual simulations.

Some future work can be done by using Markov model:

1. The application can be extended to overall survival time analysis, progression free survival analysis and time-to-event analysis in oncology trial.
2. Treatment difference test can be done by using Markov models.

APPENDIX

SOURCE CODES

Some selected source codes written in SAS programming language are included in this section. The first two SAS programs are setup to calculate the posterior quantities in Bayesian Adaptive Designs. The last two programs are used for the calculation of transition matrix, fundamental matrix analysis and Kaplan-Meier analysis.

Program 1:

```
proc datasets kill; run; quit;

options nonotes nosource symbolgen mprint mlogic;

%global loop a1 a2;
%let loop=10000;
%let a1=2;
%let b1=0.01;
%let a2=2;
%let b2=0.01;

************************************************;
*** Macro Runit ***;
************************************************;

%macro runit(datain=datain50_1, seq=1, mu0=0.20, v0=0.01, a1=2, b1=0.0266, a2=2,
b2=0.0026, loop=50000, k=1, theta_mean=mean50, theta_var=var50);

%global &theta_mean &theta_var;

proc sql noprint;
        select mean(rate) into: ybar
        from &datain;

        select count(*) into: j
        from &datain;
quit;

proc transpose data=&datain out=tmp1(drop=_name_) prefix=y;
        var rate;
run;
```

```
data tmp2;
        set tmp1;
        array y[*] y1-y12;

        mu=&mu0;
        %if &dose^=10 %then %do; vt=&b1; %end;
        %else %do; vt=&v0; %end;
        theta=&ybar;

        do i = 1 to &loop;
                %if &dose^=10 %then %do;
                mu=rand('normal', (vt*&mu0+&v0*sum(theta))/(vt+&K*&v0),
                sqrt((vt*&v0)/(vt+&k*&v0)));
                vt=(&b1+0.5*sum((theta-mu)*(theta-mu)))/rand('gamma', &a1+&k/2);
                %end;
                ysum=0;
                do m = 1 to &j;
                        ytmp=(y[m]-theta)**2;
                        ysum=ysum+ytmp;
                end;
                ve=(&b2+0.5*ysum)/rand('gamma', &a2+&K*&J/2);
                thetatmp=rand('normal', (ve*mu)/(&j*vt+ve), sqrt((vt*ve)/(&j*vt+ve)));
                theta=thetatmp+((&j*vt)/(&j*vt+ve))*&ybar;
                output;
        end;
run;

data tmp3_&dose;
        set tmp2;
        if 0.2<=theta<=0.9 then resp=1;
        else resp=0;
        sigmae=sqrt(ve);
        sigmat=sqrt(vt);
        if _n_>1000;
run;

proc sql noprint;
        create table final&seq(keep=eff post_prob) as
        select sum(resp) as eff, calculated eff/(&loop-1000) as post_prob
        from tmp3_&dose;

        select mean(theta) into: &theta_mean
        from tmp3_&dose;
```

```
        select (std(theta)**2) into: &theta_var
        from tmp3_&dose;
quit;

%mend runit;


*****************************************************************;
*** Macro Dose                                                ***;
*** This macro is used to run the simulation for each dose level ***;
*****************************************************************;

%macro dose(dose=10, expmean=0.25, prmean=0.25, prvar=0.1);
%global prob&dose expstd;

data _null_;
        call symput('expstd', max(&expmean*0.138, 0.001));
        *call symput('expstd', 0.09);
run;

*** interims ***;
%macro check(i=1);

%if &i=1 %then %do;
data datain&dose._1;
        do i = 1 to 4;
                rate=max(min(rand('normal', &expmean, &expstd), 1), 0);
                output;
        end;
run;
%end;
%else %do;
data datain&dose._2;
        i=&i;
        rate=max(min(rand('normal', &expmean, &expstd), 1), 0);
run;

data datain&dose._1;
        set datain&dose._1 datain&dose._2;
run;
%end;

%runit(datain=datain&dose._1, seq=&i, mu0=&prmean, v0=&prvar, a1=&a1, b1=&b1,
a2=&a2, b2=&b2, loop=&loop, k=1, theta_mean=mean&dose, theta_var=var&dose);

data _null_;
```

```
        set final1;
        call symput("proba&dose.", post_prob);
run;

%if (&&proba&dose <0.2) | (&&proba&dose >=0.9 | &i=8) %then %do;

%global stop;
data _null_;
        set final1;
        call symput("prob&dose.", post_prob);
        call symput("stop", 1);
run;
%end;
%else %do;
data _null_;
        set final1;
        call symput("stop", 0);
run;
%end;

%mend check;

**** the interims ***;
%check(i=1);
%if &stop=0 %then %do; %check(i=5); %end;
%if &stop=0 %then %do; %check(i=6); %end;
%if &stop=0 %then %do; %check(i=7); %end;
%if &stop=0 %then %do; %check(i=8); %end;

data datain&dose;
        set datain&dose._1;
run;

%mend dose;

*********************************************************************;
*** Macro Add four                                                ***;
*** This macro is used to add more subjects for each dose level   ***;
*********************************************************************;

%macro addfour(dose=10, expmean=0.25, prmean=0.2, prvar=0.1);

data datain&dose._4;
        do i = 5 to 8;
                rate=min(rand('normal', &expmean, &expstd), 1);
                output;
```

```
        end;
run;

data datain&dose;
        set datain&dose datain&dose._4;
run;

%runit(datain=datain&dose, seq=4, mu0=&prmean, v0=&prvar, a1=&a1, b1=&b1,
a2=&a2, b2=&b2, loop=&loop, k=1, theta_mean=mean&dose, theta_var=var&dose);

data _null_;
        set final4;
        call symput("prob&dose", post_prob);
run;

%mend addfour;

*****************************************************;
*** Macro mstep                                   ***;
*** This is the general macro for all dose levels ***;
*****************************************************;

%macro mstep(seq=1);
%global prob2 prob5 prob10 prob25 prob50;

*** Stage 1 ***;
%dose(dose=10, expmean=&resp10, prmean=0.2, prvar=0.1);

*** Stage 2 ***;
%if &prob10>=0.2 %then %do;
%dose(dose=5, expmean=&resp5, prmean=&mean10, prvar=&var10);
%dose(dose=25, expmean=&resp25, prmean=&mean10, prvar=&var10);
%end;
%else %do;
%let prob5=-1;
%dose(dose=25, expmean=&resp25, prmean=&mean10, prvar=&var10);
%end;

%addfour(dose=10, expmean=&resp10, prmean=0.2, prvar=0.1);

*** Stage 3 ***;
%if &prob5>=0.2 %then %do;
%dose(dose=2, expmean=&resp2, prmean=&mean5, prvar=&var5);
%end;
%else %do;
%let prob2=-1;
```

```
%end;

%if &prob25>=0.2 %then %do;
%dose(dose=50, expmean=&resp50, prmean=&mean25, prvar=&var25);
%end;
%else %do;
%let prob50=-1;
%end;

%if &prob5>=0 %then %do; %addfour(dose=5, expmean=&resp5, prmean=&mean10,
prvar=&var10); %end;
%if &prob25>=0 %then %do; %addfour(dose=25, expmean=&resp25,
prmean=&mean10, prvar=&var10); %end;

*** Stage 4 ***;
%if &prob2>=0 %then %do; %addfour(dose=2, expmean=&resp2, prmean=&mean5,
prvar=&var5); %end;
%if &prob50>=0 %then %do; %addfour(dose=50, expmean=&resp50,
prmean=&mean25, prvar=&var25); %end;

*** power ***;
%if &prob2<0.2 and &prob5<0.2 and &prob10<0.2 and &prob25>=0.2 and
&prob50>=0.2 %then %do;
data con&seq;
        success=1;
run;
%end;
%else %do;
data con&seq;
        success=0;
run;
%end;

%macro checkempty(dsn);
%if &&prob&dsn<0 %then %do;
        data datain&dsn;
                i=1;
                rate=-1;
        run;
%end;
%mend checkempty;

%checkempty(dsn=2);
%checkempty(dsn=5);
%checkempty(dsn=10);
%checkempty(dsn=25);
```

```
%checkempty(dsn=50);

data final_data&seq;
        set datain2(in=a) datain5(in=b) datain10(in=c) datain25(in=d) datain50(in=e);
        if a then dose=2;
        if b then dose=5;
        if c then dose=10;
        if d then dose=25;
        if e then dose=50;
run;

%mend mstep;

%macro power(max=100);

%do i = 1 %to &max;
%mstep(seq=&i);

proc append base=conclusion data=con&i force;
run;
%end;

proc sql noprint;
        create table power as
        select sum(case when success=1 then 1 else 0 end) as eff, calculated eff/&max as
power
        from conclusion;
quit;

proc print data=power; run;

%mend power;

%power(max=100);

%put _user_;
```

Program 2:

```
dm "log; clear; output; clear;";

%macro power(max=100);

%do i = 1 %to &max;
%mstep(seq=&i);
```

```
%if &i=1 %then %do;
        data conclusion;
                set theta&i;
        run;
%end;
%else %do;
proc append base=conclusion data=theta&i force;
run;
%end;
%end;

proc sql noprint;
        create table power as
        select sum(success) as eff, calculated eff/&max as power
        from conclusion;
quit;

proc print data=power; run;

%mend power;

%macro mstep(dose=180, expmean=3.5, expvar=0.21,prtmean=2,
prtvar=1,prshape=2,prscale=187);
data resp;
        do i = 1 to 4;
                y=max(max(rand('normal', &expmean, &expvar), 1), 0);
                output;
        end;
run;

ods output PostSummaries=temp;
ods graphics on;
  proc mcmc data=resp outpost=respout nmc=5000 thin=50 seed=246810;
          parms theta &expmean;
```

```
            parms sigma2 &expvar;
            prior theta ~ normal(mean=&prtmean, var=&prtvar);
      prior sigma2 ~ igamma(shape = &prshape, scale = &prscale);
      model y ~ n(theta, var = sigma2);
   run;
ods graphics off;

ods trace off;

proc univariate data=respout noprint;
var theta;
output out=m mean=mean std=sd;
run;

data resp;
          set m;
          call symput("mean&dose",mean));
run;

data theta&seq;
          set temp;
          if parameter='theta';
          if 0.2<mean then resp=1;
          else resp=0;
run;

%power(max=100);

%mend;

%mstep;
%mstep(dose=150, expmean=3, expvar=2,prtmean=3.2,
prtvar=1.5,prshape=2,prscale=187,sign=<);
%mstep(dose=100, expmean=2, expvar=2,prtmean=3.5,
prtvar=1.5,prshape=2,prscale=187,sign=<);
%mstep(dose=50, expmean=1, expvar=2,prtmean=3.5,
prtvar=1.5,prshape=2,prscale=187,sign=<);

run;
```

Program 3: SAS program used for transition analysis

```
proc iml;
      S={0 0.028 0.611 0.692 0.231 0.077, 0 0.875 0 0 0.125 0,
      0 0.026 0.829 0.132 0.013 0, 0 0 0 1 0 0, 0 0 0 0 1 0, 0 0 0 0 0
1};
      T={0 0.028 0.611 0.692 0.231 0.077, 0 0.875 0 0 0.125 0,
      0 0.026 0.829 0.132 0.013 0, 0 0 0 1 0 0, 0 0 0 0 1 0, 0 0 0 0 0
1};

      do i = 1 to 48;
            T=T*S;
      end;

      print T;
run;

/*** Calculate Matrix N ***/
proc iml;
      Q={0 0.028 0.611, 0 0.6 0.1,0 0.027 0.824};
      I={1 0 0, 0 1 0 , 0 0 1 };
      T=I-Q;
      N=inv(T);

      print N;
run;

/*** Calculate V ***/
proc iml;
      N={1 0.316 3.651,0 2.600 1.477,0 0.399 5.908};
      Np={1 0 0, 0 2.600 0, 0 0 5.908};
      I={1 0 0, 0 1 0, 0 0 1};
      V=N*(2*Np-I)-N*N;

      print V;
run;
```

Program 4: SAS program for Kaplan-Meier analysis

```
data pd;
      input subjid ttp cens @@;
      DATALINES;
      1 140 0
      2 112 1
      3 56 0
      4 84 0
      5 28 0
      6 252 0
      7 112 0
      8 28 1
      9 28 0
      10 28 0
      11 28 0
      12 28 0
      13 28 0
      14 224 0
      15 252 1
      16 252 0
      17 28 0
      18 224 1
      19 28 0
      20 28 0
      21 112 1
      22 56 0
      23 56 1
      24 84 1
      25 56 0
      26 56 0
      27 28 1
      28 252 1
      29 28 1
      30 224 0
      31 252 1
      32 112 0
      33 252 1
      34 28 0
      35 112 0
      36 56 1
      ;
RUN;

proc lifetest data=PD PLOTS=(S, LS, LLS) OUTSURV=TTP;
TIME TTP*CENS(1);
RUN;

PROC PRINT DATA=TTP noobs;
RUN;
```

REFERENCES

1. The FDA's Drug Review Process: Ensuring Drugs Are Safe and Effective (2012). http://www.fda.gov/drugs/resourcesforyou/consumers/ucm/143534.htm. Accessed Date: 11/2012.

2. Lee CJ., Lee LH., Wu CL., Lee BR., and Chen ML. (2005). Clinical Trials of Drugs and Biopharmaceuticals. *CRC Press*, Boca Raton, FL.

3. Ting N. (2006). Dose Finding in Drug Development. *Springer-Verlag*., New York, NY.

4. Orloff J. and Douglas F.. (2009). The Future of Drug Development: Advancing Clinical Trial Design. *Nature Reviews: Drug Discovery*, Vol. 8, pp. 949-957.

5. Roy A. (2006). Evaluating Adaptive Dose Finding Designs and Methods: Conference on Clinical Trial Design. *Institute for International Research*, Princeton, NJ.

6. Bretz F. and Pinheiro J. (2005). Combining Multiple Comparison and Modeling Techniques in Dose Response Studies. *Biometrics*, Vol. 61, pp. 738–748.

7. Clyde M., Draper D., George EI., Hoeting JA., Madigan D., Raftery AE., and Volinsky CT. (1999). Bayesian Model Averaging. *Statistical Science*, Vol. 14, pp. 382-417.

8. Pinheiro J. and Bretz F. (2006). Adaptive Dose-Finding Designs. *PhRMA/FDA Adaptive Design Workshop*, North Bethesda, MD.

9. Domenici E., Willé DR., Tozzi F., Prokopenko I., Miller S., Mckeown A., Brittain C., Rujescu D., Giegling I., Turck CW., Holsboer F., Bullmore ET., Middleton L., Merlo-Pich E., Alexander RC., and Muglia P. (2010). Plasma Protein Biomarkers for Depression and Schizophrenia by Multi Analyte Profiling of Case-Control Collections. *Public Library of Science*. http://www.plosone.org/article/info:doi/10.1371/journal.pone.0009166. Accessed Date: 3/2013.

10. Christel Baier, Joost-Pieter Katoen, Holger Hermanns(1999). Approximate Symbolic Model Checking of Continuous-time Markov Chains. *Concur'99 Concurrency Theory*, University of Twente, Netherlands.

11. Richard Serfozo (1979). An Equivalence between Continuous and Discrete Time Markove Decision Process. *Operation Research*, Vol. 27, pp. 616-620.

12. Marshall G. and Jones RH (1995). Multi-state Models and Diabetic Retinopathy. *Stat Med*, Vol 14, pp.1975-1983.

13. Schaubel DE., Morrison HS., Desmeules M., Parsons D., and Fenton S.S.A. (1998). End-stage Renal Disease Projections for Canada to 2005 using Poisson and Markov Models. *International Journal of Epidemiol*, Vol. 27, pp. 274-281.

14. Hendriks JCM., Satten GA., Longin LM., Druten HAM., Schellekens PTA., Coutinho RA., and Griensven GJP (1996). Use of Immunological Markers and Continuous-time Markov Models to Estimate Progression of HIV Infection in Homosexual Men. *AIDS*, Vol. 10, pp. 649-656.

15. Jaynes ET. (1986). Bayesian Methods: General Background in Maximum-Entropy and Bayesian Methods in Applied Statistics, by J. H. Justice (ed.). *Cambridge Univ. Press*, Cambridge, UK.

16. Wasserman L. (2003). All of Statistics (2nd ed.). Springer Science and Business Media, Inc., New York, NY, pp. 175-182.

17. Chaloner K. and Verdinelli I. (1995). Bayesian experimental design: a review. *Statistical Science*, Vol 10: pp. 273–304.

18. Richey M. (2010). The Evolution of Markov Chain Monte Carlo Methods. *The American Mathematical Monthly*, pp. 383-413.

19. Tanner MA. (2003). Methods for the Exploration of Posterior Distributions and Likelihood Function. Tools for statistical Inference (3rd ed.). Springer Science and Business Media, Inc., New York, NY, pp. 211-232.

20. Walsh B. (2004). Markov Chain Monte Carlo and Gibbs Sampling, http://web.mit.edu/~wingated/www/introductions/mcmc-gibbs-intro.pdf. Accessed Date: 12/2012.

21. Kelloff GJ., Hoffman JM., Johnson B., Scher HI., Siegel BA., Cheng EY., Cheson BD., Oshaughnessy J., Guyton KZ., Mankoff DA., Shankar L., Larson SM., Sigman CC., Schilsky RL., and Sullivan DC. (2005). Progress and Promise of FDG-PET Imaging for Cancer Patient Management and Oncologic Drug Development. Clinical Cancer Research, Vol. 11, http://clincancerres.aacrjournals.org/content/11/8/2785.full. Accessed Date: 11/2012.

22. Lim HS., Yoon W., Chung TW., Kim JK., Park JG., Kang HK., Bom HS., and Yoon JH. (2007). FDG PET/CT for the Detection and Evaluation of Breast Diseases: Usefulness and Limitations. *Radiographics*, suppl 1, pp. 197-213.

23. Lindholm P., Minn H., Leskinen-Kallio S., Bergman J., Ruotsalainen U., and Joensuu H. (1993). Influence of the Blood Glucose Concentration on FDG uptake in Cancer – A PET study. *Journal of Nuclear Medicine*, Vol. 34, pp. 1-6.

24. West M. and Harrison J. (1997). Bayesian forecasting and dynamic models (2nd ed.). New York: Springer Science & Business Media Inc., pp. 97-111.

25. Berry SM., Spinelli W., Littman GS., Liang JZ., Fardipour P., Berry DA., Lewis RJ., and Krams M. (2010). A Bayesian dose-finding trial with adaptive dose expansion to flexibly assess efficacy and safety of an investigational drug. *The Society for Clinical Trials*, Vol. 7: pp.121-135.

26. Berry DA., Mueller P., Grieve AP., Smith M., Parke T., Blazek R., Mitchard N., Krams M. (2010). Adaptive Bayesian Designs for Dose-Ranging Drug Trials, http://workshop.stat.cmu.edu/berry1.pdf. Accessed Date: 9/2012.

27. Sidney Kennedy and Henning Anderson (2006). Efficacy of Escitalopram in the Treatment of Major Depressive Disorder Compared with Conventional Selective Serotonin Reuptake Inhibitors and Venlafaxine XR: a Meta-analysis. *Journal of Psychiatry Neuroscience*, Vol. 31, pp. 122-131.

28. SAS/STAT 9.2 User's Guide (2008). Cary, NC: SAS Institute Inc., pp. 160-161.

29. Daskalakis C., Laird NM., and Murphy JM. Laird (2002). Regression Analysis of Multiple-Source Longitudinal Outcomes: A "Stirling County" Depression Study. *American Journal of Epidemiology*, Vol. 155, pp. 1-10.

30. Gorman JM, Korotzer A., Su G. (2002). Efficacy Comparison of Escitalopram and Citalopram in the Treatment of Major Depressive Disorder: Pooled Analysis of Placebo-Controlled Trials. *CNS Spectrums*, Vol. 7, Suppl 1, pp. 40-44.

31. J. Norris (1997). Markov Chains. *Cambridge University Press*, Cambridge, UK.

32. Meeker W. and Escobar L. (1998). Location-scale-based Parametric Distributions. Statistical Methods for Reliability Data. John Wiley and Sons, Inc, New York, NY, pp.75-96.

33. Hillis A., Maguire M., Hawkins B., Newhouse M. (1986). The Markov Process as a General Method for Nonparametric Analysis of Right-censored Medical Data. *Journal of Chronic Disease*, Vol. 39, pp. 595-604.

34. Stewart W. (1994). Introduction to the Numerical Solution of Markov Chains. *Princeton University Press*, Princeton, NJ.

35. Sendi P., Bucher H., Craig B., and Pfluger D., and Battegay M. (1999). Estimating AIDS-free survival in a severely immunosuppressed asymptomatic HIV-infected population in the era of antiretroviral triple combination therapy. *Journal Acquir Immune Defic Syndr Hum Retrovirol*, Vol. 20, pp. 376-381.

36. Jain S. (1986). Markov chain model and its applications. *Comp Biomed Res*, Vol. 19, pp. 374-378.

37. Lawless J. and Yan P. (1993). Some Statistical Methods for Follow-up Studies of Disease with Intermittent Monitoring. In: Happe FM, ed. Multiple Comparisons, Selection, and Applications in Biometry, Marcel Dekker, New York, NY, pp. 427-446.

38. Pond G. (2008). Design and Analysis of Sequential Clinical Trials using a Markov Chain Transition Model with Conditional Power. *Ph. D Dissertation, University of Toronto*, Toronto, Ontario, Canada.

39. U.S. Department of Health and Human Services, Food and Drug Administration (1996). Guidance for Industry E6 Good Clinical Practice: Consolidated Guidance.

40. Fu J., Wendy WY. (2003). Distribution of Theory of Runs and Patterns and Its Applications – A Finite Markov Chain Imbedding Approach. World Scientific Publishing Co. Pte. Ltd. Singapore, pp.18-19.

41. Beck, J.R., Pauker, S.G. (1983). The Markov process in medical prognosis. *Medical Decision Making*, Vol. 3, pp. 419-458.

42. Wayne, LW. (2003). Operations Research: Applications and Algorithms. *Duxbury Press*, Pacific Grove, CA.

43. Brown, R.F., Brown, B.W., 1990a. Absorbing Markov Chains. In: “Essentials of Finite Mathematics: Matrices, Linear Programming, Probability, Markov Chains”, *Ardsley House*, New York, NY.