Bayesian Two-Stage Dose-Finding Study For Binary Endpoints In Phase II Clinical Trials

Jean Kyung Bak

The Graduate School

Yonsei University

Department of Biostatistics and Computing

Bayesian Two-Stage Dose-Finding Study For Binary Endpoints In Phase II Clinical Trials

A Dissertation

Submitted to the Department of Biostatistics and Computing

and the Graduate School of Yonsei University

in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

Jean Kyung Bak

December 2016

This certifies that the dissertation
of Jean Kyung Bak is approved.

_____

Thesis Supervisor: Dr. Chung Mo Nam

_____

Dr. Inkyung Jung

_____

Dr. Sohee Park

_____

Dr. Eun Jin Jang

_____

Dr. Min Soo Park

The Graduate School
Yonsei University
December 2016

Acknowledgement

박사학위 논문을 마치기까지 따뜻한 지도와 조언을 주신 많은 분들께 진심으로 감사드립니다. 논문의 주제를 정하고 연구계획을 세우고 완성해 나갈 수 있도록 애정 가득한 조언과 지도를 해주신 남정모 지도교수님께 진심으로 감사를 드립니다. 또한 베이지안 방법으로 연구의 틀을 마련할 수 있도록 도움을 주신 장은진교수님께 진심으로 감사 드립니다. 논문의 학문적 완성도를 높일 수 있도록 아낌없는 조언을 주신 정인경교수님께도 진심으로 감사드립니다. 늘 학과에서 학생들을 따뜻하게 챙겨주시고 관심가져주신 송기준교수님께도 감사드립니다. 논문심사과정에서 여러 가지로 부족한 제 논문의 완성도를 높이는데 조언을 주신 박소희교수님, 박민수교수님께 늘 감사함을 갖고 있습니다. 박사과정 동안 때로는 친구가 되어주고 때로는 귀여운 동생도 되어준 여러 친구들에게도 감사의 마음을 전하고 싶습니다. 특히, 늘 힘과 용기를 주고 따뜻한 조언을 준 김윤남선생님, 이명희선생님, 임현선선생님, 항상 응원하고 격려해준 은혜, 수희, 호진, 세영, 세휘, 고은, 혜수, 선규, 미경 등 많은 친구들의 사랑과 배려가 있었기에 무사히 학업을 마칠 수 있었습니다. 논문을 마무리하며 겸손함이 얼마나 소중한 가치인가를 깨닫게 되었습니다. 늘 노력하며 성장하도록 하겠습니다. 마지막으로 늘 사랑으로 지켜주며 응원해 준 가족들에게 감사의 마음을 전합니다. 제 인생의 멘토이신 하늘에 계시는 아버지, 늘 기도로 지켜주시는 어머니 늘 감사합니다. 그리고 제 인생의 뿌리이며 기둥이신 주님의 은혜 감사합니다.

Contents

List of Tables

List of Figures

Abstracts

Bayesian Two-Stage Dose-Finding Study for Binary Endpoints in Phase Ⅱ clinical trials

Bak, Jean Kyung
Dept. of Biostatistics and Computing
The Graduate School
Yonsei University

In this study, we proposed Bayesian two-stage dose-finding model to identify the MED for binary endpoint in phase Ⅱ clinical trials. Our study was motivated from the idea of the predictive probability with efficacy criterion motivating from the dose-finding study of Pozzi et al. (2013).

We made calculation of the predictive probability comparing with the pre-specified threshold in order to identify the MED in phase Ⅱ clinical trials. While we making calculation of the predictive probability, we used the Bayesian Model Averaging method to solve the problem of the model uncertainty (Raffery and Volinsky, 1999). In addition, we have experienced the challenge of the integral of the predictive probability, which has vast difficulty to solve it in Bayesian perspective. Therefore, to make the predictive probability under the efficacy criterions, we adopted the Sampling-Importance Resampling (SIR) algorithm first

proposed by Rubin (1983), which is the method of the approximation to sample from the posterior distribution even though we could not directly explain the distribution.

To make semi-parametric model under assuming monotonic constraint of the mean effect under the five dose groups, we made the stick-breaking construction with employing the jump-variable. In simulation study, the stick-breaking construction has an advantage easily to apply the proposed model while adopting the Bayesian Model Averaging. As a result, we found that the simulation result from the Bayesian Model Averaging is accurately identified the MED comparing the result of the different models.

KEY WORDS: Bayesian two-stage design, MED, binary endpoint, predictive probability, success probability, Sampling-Importance Resampling(SIR) algorithm, stick-breaking construction, jump-variable, semi-parametric model

1. Introduction

Clinical trials play a critical role in drug development. Of diverse issues in clinical trials, dose finding design has been researched for several decades in drug development. In doing dose finding study, we should well understand with characteristics of the dose response relationship of any new compound is important and ubiquitous in many areas of scientific investigation. This is especially true in the context of pharmaceutical drug development, where it is mandatory to launch safe drugs which demonstrate a clinically relevant effect. Selecting a dose too high may result in unacceptable safety problems. On the contrary, selecting a dose too low may lead to ineffective drugs. Therefore, dose finding studies perform an important role in any drug development. In fact, most researchers question about what doses relevantly different from control within acceptable safety. This question is closely connected to the estimation of a minimum effective dose (MED), that is the smallest dose with a discernible useful effect following to ICH-E4 (1994), Ruberg (1995) and Bretz et al. (2010). In Frequentist approach, if confirmatory pair-wise comparisons with a control are of main interest, multiple comparisons may be an appropriate to answer the question. However, in Bayesian approach, there is not yet decided definite answer to the question even though there are several applies to solve the question.

Phase Ⅱ studies are the basis for planning of the next stage clinical trials. Especially, the major purpose conducting phase Ⅱb is to determine the optimal dose which is going into the next stage clinical trials. An optimal dose is a dose that is high enough to demonstrate efficacy in the target population. In fact, there may be diverse strategies to determine the optimal dose, but here we focus on the Bayesian two-stage dose finding study in phase Ⅱ clinical trials. Particularly, we

use the parallel dose comparison study, which several potential doses are selected and subjects are randomized to receive one of the doses for entire study. At the end of study, we can look at how each treatment group performed as compared to the control group. Before we research the Bayesian two-stage dose finding design, we look around two-stage dose finding design about general approach in phase Ⅱ clinical trials. In clinical development of a new treatment, the conventional role of a phase Ⅱ a clinical trial is a "proof of concept" by checking the potential efficacy of a new treatment. Typically, this kind of drug is the maximum tolerated dose in cancer trials under the assumption that toxicities are positively associated with cancer killing activities (Ratain, 1993). In other words, the goal of a Proof-of Concept (PoC) study is to verify dose efficacy in patients in phase Ⅱ a clinical trials. In reality, developing new pharmacological therapies is extremely expensive and only few studies may be successful. Therefore, there is a great value in enabling earlier to check optimal dose and better making decision which dose levels continue with a drug development study. However, incomplete understanding of the dose-finding studies is recognized as a major leading to inappropriate or appropriate doses being taken into phase Ⅲ clinical trials. For example, to describe efficacy optimal dose with over-dispersed count endpoint, Pozzi et al. (2013) suggested that Bayesian adaptive two-stage dose-finding design in phase Ⅱ clinical trials.

In general, there are major several types of primary endpoints such as continuous endpoint, count endpoint, binary endpoint and so on. In this paper, we focus on binary endpoint because most patients are classified as a responder or non-responder to the treatment study at the end of a study. We mention the situation of that "Why should we use binary endpoint?". A typical primary efficacy analysis is to compare with the numbers and proportions of responders

between treatment groups. The response variable is a binary variable. Even if the endpoint is continuous, there is an increasing tendency to re-define criteria and re-classify subjects as a "responder" or a "non-responder". Ting (2006) introduced an example case that patients in anti-depressants trials are frequently referred to as responders if they experience a 50% reduction in the HAM-D score from their baseline values. In addition, there are many situations that binary response makes sense. The examples include "alive" or "dead" for patients in a salvage trial. In this manner, binary outcome in clinical trials is common type no matter which problem faced we are. Hence, in this study, we discuss the dose finding design to identify the MED for binary endpoints in phase Ⅱ clinical trials. Particularly, we discuss model based on Bayesian two-stage dose-finding process to choose the MED for binary endpoints in phase Ⅱ clinical trials.

1.1. Two-stage dose finding design in phase Ⅱ clinical trials

In general, since Simon's two-stage design (1989) has been proposed, there have been emerged diverse useful designs such as sample size determination in phase Ⅱ clinical trials based on Simon's two-stage design. In this study, we apply the concept of Simon's two-stage design to find the MED in dose finding studies. Even if Simon's design has limitation to focus on determining the sample size, many researchers have cited and utilized the method into their diverse clinical studies such as dose finding studies. In fact, Simon's two-stage design was based on four major procedures such as decision making for hypotheses, adhering strict sample size to keep the power, and enrollment procedure for patients and stopping rule with respect to Frequentist perspective.

In Frequentist approach, Polley (2008) and Cheung (2008) applied Simon's two-stage design to find the MED when they compared several dose levels with a placebo. They mentioned to manage the issue of multiple comparisons due to handling several dose levels in two-stage dose finding studies. In addition, as a way of finding solution of the multiple comparisons, Steansson (1988) and Hsu and Berger (1999) introduced two-stage design for the partitioning test of binary outcome in Frequentist approach. Based on their statements, there are three different methods such as the pre-determined step-down method, the sample-determined step-down method and the sample-determined step-up method in partitioning test to explain the multiple comparisons in Frequentist perspective. However, there is no definite method to account for multiple comparisons even though many researchers have discussed since Duncan (1965) has introduced several mixed approaches for multiple comparisons that was combined Bayesian perspective with Frequentist perspective. In addition, Muller et al. (2006) were mainly used posterior probability and decision theoretic approaches for adjust multiplicities. Besides, Meng et al. (1987) discussed the multiplicity problem relied on Bayesian p-value that is the similar approach with traditional significance testing. Although there are many efforts and approaches in Bayesian ways, it is fact that there is an argument among researchers who are interested in explaining the multiple comparisons. Therefore, we do not mention the multiple comparisons with identifying the MED among the several dose levels.

Lee et al. (2008) presented the decision method when we stop and how to stop the clinical trials for efficacy or futility based on the predictive probability in Bayesian approach. Given the interim data, the predictive probability of the posterior distribution provides the estimation ways how to reject the null hypotheses in dose finding studies. That is, they mentioned two different

assumptions. One assumption could be the decision rule that predictive probability depended on that the true response rate p_1 is greater than the null response rate, p_0. Another assumption was the decision rule how to stop the trial for success or futility. In this study, we apply the idea of decision rule using predictive probability while identifying an appropriate dose level as the MED in Bayesian perspective. In addition, we adopt the concept of decision making that was Bayesian two-stage design for phase Ⅱ clinical trials with respect to the Single Threshold Design (STD) for binary endpoint proposed by Sambucini (2008). Particularly, the design proposed by Sambucini (2008) was based on the predictive probability of the STD to select the sample size in the experimental study. Before Sambucini (2008) proposed one of the STD methods, Tan and Machin (2002) firstly introduced the idea of the Single Threshold Design (STD) in two-stage design when they decided the sample size, in which a large posterior probability of the true response rate exceeds a target value when the observed response rate is larger than the pre-specified target value. They offered to the method of the sample size calculations through extending different kinds of informative prior distributions which use informative conjugate prior distributions. Other most researchers including Herson (1979) mainly have managed the issue of the sample size determination through the predictive probability in phase Ⅱ clinical trials.

However, only few cases are known where the issue of dose finding for binary outcome relies on the Single Threshold Design (STD) with the predictive probability in phase Ⅱ clinical trials. For instance, Ivanov, Xiao and Tymofyeyev (2012) proposed the Bayesian adaptive two-stage design to find the MED for phase Ⅱ dosing finding study. They handled both the continuous and binary outcomes to make decision grounded on the posterior probability of the target dose location in making decision which dose level could be the MED. More recently, Pozzi et al.

(2013) proposed a Bayesian adaptive dose selection design with over-dispersed count data. They considered the predictive probability as a decision method since they wanted to explain the future data given the interim data while identifying the MED in phase Ⅱ clinical trials. In other words, the study design was taken into account of one interim analysis to make decision which dose level could be the MED in Bayesian perspective.

1.2. Purpose

The purpose of this study is to present the method in Bayesian two-stage dose finding design for binary endpoint in phase Ⅱ clinical trials. Especially, we discuss the dose finding method to identify the MED depended on the Bayesian predictive probability by comparing with pre-specified criterions. Besides, we also referred the approach of the inference to select the MED that was found by comparing the predictive probabilities of several dose groups with a placebo according to Dunnett (1995) and Williams (1971). Particularly, while modeling with referring to efficacy criterion for binary outcomes, we take advantage of the convenience of the approximation method to sample from the posterior distribution with Sampling-Importance Resampling (SIR) algorithm proposed by Rubin (1983) and Smith and Gelfand (1992). Under the similar studies figuring out the MED, if we let t denote the clinically relevant value, i.e., the smallest relevant value which shows a clinically relevant and statistically significant effect, we expect a dose to be better than a placebo. The approach to find the MED by comparing Bayesian predictive probability with the pre-specified criterions for binary endpoint in phase Ⅱ clinical trial could be a new challenge even though it was motivated by Pozzi et al. (2013) that was Bayesian adaptive two-stage dose finding study to the MED with

over-dispersed count data. Therefore, as mentioned above, we utilize the method of the approximation that draws samples through Sampling-Importance Resampling algorithm from the posterior distribution to calculate the predictive probability as a decision rule presented by Pozzi et al. (2013).

1.3. Outlines

In this study, we discussed the model in two-stage dose finding studies for binary endpoint with respect to the Single Threshold Design (STD) proposed by Tan and Machin (2002). In addition, there has an approximation method to sample from the integral of the equation (7) presented by Rubin (1983) and Smith and Gelfand (1992) in section 2.2.3. Furthermore, Bayesian two-stage design in dose finding studies in phase Ⅱ clinical trials was presented with the predictive probability as a decision rule. The proposed method is provided in chapter 3. To figure out the MED, simulation studies with the Bayesian Model Averaging method are presented in section 4.1. Furthermore, identifying the MED via the simulation studies of eight different models is provided in chapter4. Finally, there is the conclusion and discussion in chapter 5.

2. Background

On the average, major considerations and plans with dose finding studies should be started with the non-clinical development stage. With the clinical development plan over the entire phase, diverse clinical trial studies are designed and carried out for several decades. This clinical development study is updated over time based on newly available information. Specially, estimation of dose-response relationship might be one of the very important issues in the clinical development study. Method to find an appropriate dose finding is needed enough information such as related data and plentiful and diverse expert opinions and experiences with respect to across all phases of clinical trials.

The crucial stage for finding a proper dose level should be around phase Ⅱ. There are two parts in phase Ⅱ clinical trials. One is phase Ⅱa called Proof-of-Concept (PoC) and another is phase Ⅱb. Relied on information collected from the results of phase Ⅰ clinical studies, many clinical trials in phase Ⅱa should be planned and carried out Proof-of-Concept (PoC), dose finding studies. A commonly used Proof-of-Concept study typically has two parallel treatment groups such as a placebo group and test treatment group using high dose very close to MTD or the MTD itself. Dose finding studies usually include a placebo group, plus a few doses of test drug–e.g., low dose, several intermediate doses, and high dose. Commonly, these kinds of studies have parallel group with fixed doses in traditional approach. The main objective of the dose finding studies is to estimate the dos-response relationships for efficacy and safety.

The MED is often defined as the lowest dose with mean response significantly different from a placebo, which means that the mean effect adds the clinically

important minimum difference (ICH E4 Guideline, 1994). As mentioned in section 1, the primary objective of phase Ⅱ clinical trials is often to find the MED which is statistically significantly superior to a placebo and produces a clinically relevant effect (Ruberg, 1989 and ICH E4, 1994). As mentioned in section 1.2, we discuss two-stage dose finding design to identify the MED with respect to binary outcomes in Bayesian perspective in phase Ⅱ clinical trials. In making decision of the MED in dose finding studies, we rely on the Bayesian predictive probability of the posterior distribution following to the pre-specified efficacy and exclusion criterion.

2.1. Models in two-stage dose finding studies

2.1.1. Notations and Assumption

We assume that J dose-response relationships are increasing as follows;

$$d_1 < d_2 < \cdots < d_J,$$

where d_1 is a placebo, $d_2, \cdots, d_{J-1}$ are the intermediate doses to be the MED and d_J is the highest dose. We make notation that the structure of the data such as the current data Y and the future data Y^* with each dose level is presented by Table 1. We notify that Y_{ij} and d_{ij} (i $= 1, \cdots n_j$, $j = 1, \cdots, J$) are the response and the dose level for the ith patient at the jth dose level after finishing a treatment. The term of $\boldsymbol{Y_j}$ is the response in current data and the term of $\boldsymbol{Y_j^*}$ is the response in future data about jth dose level

Table 1. Data Structure According to Dose Level

Dose Level	Current Data (Y)	Future Data (Y^*)
1	$Y_{11} \cdots Y_{n_1 1}$	$Y^*_{11} \cdots Y^*_{m_1 1}$
2	$Y_{12} \cdots Y_{n_2 2}$	$Y^*_{12} \cdots Y^*_{m_2 2}$
⋮	⋮	⋮
k	$Y_{1\mathrm{k}} \cdots Y_{n_\mathrm{k} k}$	$Y^*_{1\mathrm{k}} \cdots Y^*_{m_\mathrm{k} k}$
⋮	⋮	⋮
J	$Y_{1\mathrm{J}} \cdots Y_{n_J J}$	$Y^*_{1\mathrm{J}} \cdots Y^*_{m_J J}$

The term of $\theta_j = E[Y_{ij}]$, $\mathrm{i} = 1, \cdots n_j$, $j = 1, \cdots, J$, is the mean effect of jth dose levels and $\boldsymbol{\theta}$ is the meaning of the parameter vector. Furthermore, according to the assumption of the mean effect in dose groups proposed by Pozzi et al.(2013), we assume that the relationship of the mean effects corresponding to each dose level is given by

$$\theta_1 \geq \theta_2 \geq \cdots \geq \theta_J. \qquad (1)$$

The equation (1) indicates that the probability of detection for disease is decreased by the dose with increasing the mean effect of the dose group after administering drug. In this study, we have taken account of two-stage dose finding studies to identify the MED in phaseⅡ clinical trials in Bayesian perspective. The decision which dose could be the MED was based on the predictive probability (PP) to compare with the pre-specified clinical threshold (t) that is clinically relevant value obtaining from diverse clinical opinions and experiences of the medical

experts. There is the brief design flow of the study in Figure 1. The terms of n, m and N are the sample sizes at the current data, the future data and the entire data, respectively.

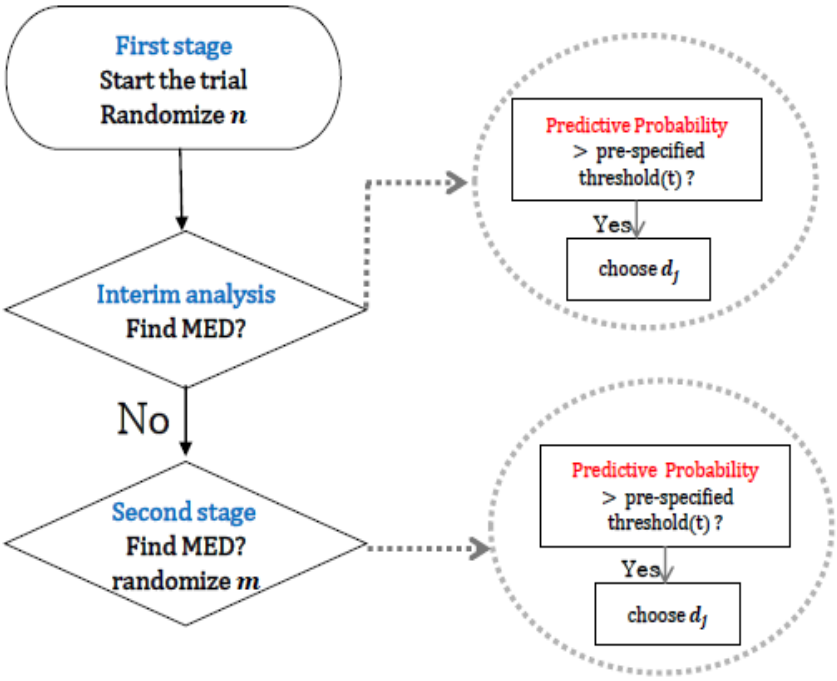

Figure 1. Flowchart of Bayesian Two-Stage Dose Finding Study in phase Ⅱ

2.2. Bayesian two-stage design in dose finding studies.

Generally, most of dose finding study is to identify the most suitable dose level for medical treatment in clinical trials. As a matter of fact, there are many uncountable approaches to find the best fitted dose for patients in Frequentist and Bayesian perspectives. Unfortunately, there is no guarantee which dose level is the best dose for patients without any toxicity. Therefore, we research Bayesian method of dose finding design for binary endpoints in phase Ⅱ clinical trials. Typically, dose finding studies in phase Ⅱa clinical trials use a single dose to assess the efficacy of new drugs, with the goal of screening out those that are ineffective. After passing the initial efficacy assessment of a new drug in phase Ⅱa clinical study, subsequently, dose finding studies in phase Ⅱ clinical trials use multi-doses to compare the efficacy of the new drugs with a placebo. Therefore, the most promising dose can be selected for large scale evaluation in late phase studies such as phase Ⅲ clinical trials. That is, dose finding studies in phase Ⅱ are often randomized, multi-doses levels with the goal of identifying the most promising dose to send to large-scale phase Ⅲ trials for definitive testing.

Kramar, Potvin and Hill (1996) mentioned multi-stage designs to achieve getting better results than single stage design. Basically, they pointed out different multi-stage designs in clinical trials focused on and developed in the sample size determination. Besides, they mentioned that Gehan (1961) firstly proposed the multi-stage design and still being widely used in spite of rarely cited. The two-stage design allows for prompt rejection of an ineffective treatment or drug at the end of the second stage. In addition, they described that Fleming (1982) developed multi-stage designs to enable early termination of a trial when treatment or drug is either clearly effective or ineffective. In particular, Simon (1989) improved

Fleming's two-stage design by minimizing either the average or the maximum number of patients required under the hypothesis of treatment ineffectiveness. Furthermore, Chen (1997) developed a three-stage design to apply Simon's designs. As mentioned above, multi-stage designs including two-stage designs have an advantage which is flexible to apply dose finding studies.

Even though multi-stage designs have better statistical properties than single-stage designs by employing information gained in the interim data, the Frequentist analysis of such designs is constrained due to rigid requirement of examining the outcome at the pre-specified sample size at each pre-determined stage. This lack of flexibility exposes a fundamental limitation of all such Frequentist-based methods. Besides, sometimes most researchers can experience such a situation that there is disparity between the proposed design and sampling plan. Therefore, some of researchers try to account for all statistical inference through adjustment method mentioned by Green and Dalberg (1992). Particularly, we are interested in getting flexibility to make decision which dose could be the MED. This kind of reason supports the need for more flexible designs.

On the other hand, Bayesian perspective provides a different approach to do designing and monitoring clinical trials by calculating of the posterior probability of diverse events given data. In particular, Bayesian perspective has a special merit in clinical conduct and impart the ability to examine interim data, update the posterior probability of interesting parameters and make relevant predictions and sensible decisions. In general, there are three kinds of major testing methods such as decision making, predictive probability and posterior probability in Bayesian perspective. But, in this study, we mainly refer to the predictive probability in section 2.2.1 as we use the predictive probability as a decision rule to make decision in Bayesian approach.

2.2.1. The Predictive Probability

Generally, based on the interim data, the predictive probability is gotten by calculating the probability of a positive calculation to reject the null hypothesis or alternative hypothesis that the trial should be conducted to the maximum planned sample size (Lee et al., 2008). The decision to continue or to stop the trial can be made according to the strength of this predictive probability. In particular, given the current data, we calculate the predictive probability for future data in dose finding studies in Bayesian perspective. For instance, according to Pozzi et al.(2013), the Bayesian predictive probability as a decision rule was calculated for over-dispersed count data to identify the MED in phase Ⅱ clinical trials. In addition, Pozzi et al.(2013) mentioned that how we can define the predictive probability of the posterior distribution in Bayesian perspective. To begin with, they assumed that the response Y_{ij} for ith patient at jth dose level, is distributed as follows;

$$Y_{ij} \sim poissson(\theta_j).$$

The likelihood function of posterior distribution of the response Y_{ij} follows a gamma distribution. That is, by imposing conjugate distribution of the gamma prior distribution and poisson likelihood function, they showed that the posterior distribution of the response is given by,

$$\theta_j|Y \sim Gamma(\alpha_j, \beta),$$

where they assumed that the scale parameter β is identical across dose groups to reach identifiability. Thus, the response in future patients, Y_{ij}^*, follows a negative-binomial distribution;

$$Y_{ij}^* \sim negative - binomial(m, r, \theta_j),$$

where r is the number of failure until the experiment is stopped. According to Pozzi et al. (2013) description, a general definition of the predictive probability of the response for "Success" as follows;

$$\text{PP} = \text{P}\{\text{"Success"}|\text{Y}\} = \text{P}\{Y^* \in Y_S|Y\} = \int_{Y_S} p(Y^*|Y)dY^* \qquad (2)$$

for some success region Y_S. To make definition about which is the meaning of the predictive probability after doing treatment in more detail, they firstly made the term of "Success" can be defined as follows:

$$\{\text{"Success"}\} = Y_S = \{Y^*: P\{\theta \in \Theta_E|Y, Y^*\} > c\},$$

for some efficacy region Θ_E and some threshold c. Then, they have the predictive quantity (2) becomes

$$\text{PP} = \text{P}\{\text{"Success"}|\text{Y}\} = \int_{Y_S} p(Y^*|Y)dY^* = \int I(\{Y^*: P\{\theta \in \Theta_E|Y, Y^*\} > c\}p(Y^*|Y)dY^*$$

The case of the intermediate dose levels, $\text{j} \in \{2, \cdots, \text{J}-1\}$, we make decision which dose will be the MED among the intermediate dose levels in phase Ⅱb clinical trials.

Using the exclusion and efficacy criterions introduced by Pozzi et al. (2013) when we assume that the dose-response relationship is increasing, we redefine the exclusions and efficacy criterions as follows;

Exclusion Criterion

$$\mathrm{P}\left\{\frac{\theta_j}{\theta_1} \geq \xi_1 \middle| Y,\ Y^*\right\} \geq 50\%, \qquad (3)$$

Efficacy Criterion

(i) $\mathrm{P}\left\{\frac{\theta_j}{\theta_1} < 1 \middle| Y,\ Y^*\right\} \geq 95\%$ (4)

(ii) $\max\{\mathrm{P}\{\frac{\theta_j}{\theta_1} \leq \xi_2 | Y,\ Y^*\},\ \mathrm{P}\{\frac{\theta_j}{\theta_J} \leq \xi_3 | Y,\ Y^*\}\} \geq 50\%$ (5)

where ξ_1, ξ_2 and ξ_3 $(0 < \xi_1, \xi_2, \xi_3 < 1)$ are determined by clinical experiences and expert opinions given data. The exclusion criterion(3) is the meaning of the efficacy at the dose level j is not better than placebo to a clinically extent. In the efficacy criterion(4), we require the dose to be superior to dose 1 with very high probability. In addition, the meaning of the equation(5) is necessary for the dose to be either at least ξ_2 better than or at most ξ_3 worse than dose J.

2.2.2. Bayesian Model Averaging to Monotone Dose Finding Model

As mentioned in section 2.1.1, the equation (1) is assumption for a single model. However, we have seen that a major goal of the model selection is to choose a single model that is considered the best one among all candidate models we expected. If we want to choose a single model as the best model, all subsequent decisions are made well under the chosen model without any problem. But, we often meet with uncertainty about selected an unsuitable single model. The model uncertainty leads to over-estimation or under-estimation about our inferences and decisions that are much risky than we expected. Therefore, in this study, we use the Bayesian Model Averaging method that is averaging different competing models. Hoeting, Madigan, Rafftery and Volinsky (1999) mentioned that the Bayesian Model Averaging gives us a coherent mechanism for accounting for this model uncertainty. Furthermore, they mentioned that the definition of the Bayesian Model Averaging to describe the uncertainty in model selection for accounting predictive performance is following to

$$\mathrm{P}(\text{event}|\text{data}) = \sum_{k=1}^{K} P(event|M_K, data)P(M_K|data), \qquad (6)$$

where models $M_K\,(\mathrm{k} = 1,2,\cdots,\mathrm{K})$ are considered. That is, the meaning of the equation (6) is an average of the posterior distributions under all different models considered which are weighted by the probability about their posterior models, where $M_1,\cdots,M_K$ are the models. According to Congdon (2014), the Bayesian Model Averaging may be based on MCMC samples from models sampled in parallel. For the models $(M_1,\cdots,M_K)$, we may obtain model weights at each

iteration, and estimates of the posterior probabilities for each model, or of model averaged parameters by averaging over samples. The posterior probability for model M_K is following to

$$\mathrm{P}(M_K|data) = \frac{P(data|M_K)p(M_K)}{\sum_{l=1}^{K} P(data|M_l)\, P(M_l)}$$

where $\mathrm{P}(\mathrm{data}|M_k) = \int P(data|\theta_k,\ M_k)\, P(\theta_k|M_k)d\theta_k$ which is the marginal likelihood of model M_K, θ_k is the vector of parameters of model M_K. $\mathrm{P}(\theta_k|M_K)$ is a prior density of θ_k under model M_K and $\mathrm{P}(\mathrm{data}|\theta_k, M_K)$ is the likelihood and $\mathrm{P}(M_K)$ is the prior probability that M_K is the true model.

Furthermore, Madigan and Raftery (1993) made an explanation that averaging all of the models provides better averaging of the predictive ability than using any single model,M_K, conditional on M. Therefore, we adopt the concept of the Bayesian Model Averaging in order to reduce model uncertainty while identifying the MED in this study. When making inference an appropriate model, they adopted the Bayesian Model Averaging which is represented by a different combination to decrease the model uncertainty. With adopting the Bayesian Model Averaging, it can be fully specified as follows, for patients, $\mathrm{i} = 1, \cdots, n_j$, dose groups, $\mathrm{j} = 1, \cdots, \mathrm{J}$ and models $\mathrm{k} = 1, \cdots, \mathrm{K}$.

Next, by the exclusion and efficacy criterions explained in section 2.2.1, we can define that the jth predictive probability given each model, M_K, is given by

$$PP_j^* = P\left\{P\left\{\frac{\theta_j}{\theta_1} < 1 \middle| Y_j, Y_j^*\ M_K\right\} > 95\%\right\} \cap$$
$$\left\{\max\{P\{\frac{\theta_j}{\theta_1} \le \xi_2 \middle| Y_j, Y_j^*, M_K\right\}, \mathrm{P}\left\{\frac{\theta_j}{\theta_J} \le \xi_3 \middle| Y_j, Y_j^*, M_K\right\}\} > 50\%\}|Y\}, \qquad (7)$$

which is an appropriate to a formal representation of the efficacy criterion. Thus, we can choose the dose as the MED if the jth $PP_j^* \geq t$(pre-specified threshold). In order to explain the predictive probability of the posterior distribution, we approximate PP_j^* given the different models which are based on the general framework of the Sampling-Importance Resampling algorithm. In particular, we provide the detail information about the application of the Sampling-Importance Resampling to dose finding studies in section 3.3.

2.2.3. Application of Predictive Probability to Dose Finding

Usually, dose finding studies have been included a placebo or control group, a few doses of the test drug in phase Ⅱb clinical trials. Naitee (2006) referred to an ideal dose finding study that should incorporate a wide range of doses from low to high doses. For instance, Pozzi et al. (2013) assumed that five dose levels such as a placebo, three intermediate doses and the highest dose. In particular, for three intermediate dose levels, $j \in \{2,3,4\}$, they made decision which dose is the MED in phase Ⅱb clinical trials. Using the exclusion and efficacy criterions presented in section 2.2.1, they determined the clinical threshold values of ξ_1, ξ_2 and ξ_3 according to the diverse opinions and experiences of clinical experts. In this study, we determine that $\xi_1 = 0.7$, $\xi_2 = 0.5$ and $\xi_3 = 1.2$ by utilizing the approach of Pozzi et al. (2013). However, those kinds of clinical thresholds may be changed by the purpose and characteristics of each clinical study. In particular, in order to make decision which dose could be the MED, we need to obtain samples from the posterior distribution. Unfortunately, as Dani (2006) mentioned in his study, it is

very complicated to directly sample from the posterior distribution for the vast majority of problems of practical relevance.

3. Proposed methods

3.1. Model Setting and Assumptions

We propose the Bayesian two-stage model under the monotone constraint of the mean effects between dose groups to identify the MED in dose finding studies for binary outcome in phase Ⅱ clinical trials. In particular, we use the semi-parametric model due to the monotonicity between then mean effects of dose groups presented by Pozzi et al. (2013). We start with what is the model we used and which distribution follows we used. To reduce the model uncertainty, we adopt the Bayesian Model Averaging method using ten different model combinations presented by Pozzi et al. (2013). In addition, we present the distribution and process of each stage in more detail as follows. As mentioned above, we assume five dose levels including a placebo, three intermediate doses and the highest dose for binary endpoint. First, we let $\boldsymbol{\theta}$ be the unknown probability parameters and Y_{ij} $(i = 1, \cdots, n_j,\ j = 1, \cdots, 5)$ be the binary response variables that could be success or failure.

In this study, our proposed method has been based on the general framework of the distribution that we use a conjugate prior distribution $\theta_j \sim beta(\alpha_j, \beta_j)$ for binary endpoint in Bayesian perspective. When we have $Y_{1j}, \cdots, Y_{n_j j} | \theta_j \sim \mathrm{Ber}(\theta_j)$, we can get

$$Y_{\cdot j} | \theta_j \sim \mathrm{bin}(n_j, \theta_j)$$

for $Y_{\cdot j} = \sum_{i=1}^{n_j} Y_{ij}$ and the likelihood function of posterior distribution of the success rate follows a beta distribution. That is, by imposing conjugate distribution

of the beta prior and multiplying the conjugate beta prior $\theta_j \sim beta(\alpha_j, \beta_j)$ and binomial likelihood, $Y_{.j} \sim \text{bin}(n_j, \theta_j)$, we obtain the posterior distribution of the success rate in first stage as follows;

$$\theta_j | Y_j \sim beta(\alpha_j + Y_j, \beta_j + n_j - Y_j).$$

Thus, in second stage, the number of the success in the potential future patients $m = N - n$, $Y_{.j}^*$ is distributed as a beta-binomial distribution;

$$Y_{.j}^* \sim beta - binomial(m, \alpha_j + Y_{.j}, \beta_j + n_j - Y_{.j}).$$

In particular, in order to explain the posterior distribution, we should consider what the shape of the prior distribution is in more detail. As mentioned above, we use the conjugate beta prior distribution $\theta_j \sim beta(\alpha_j, \beta_j)$ for binary outcome in Bayesian perspective. For convenience, we want to fix another parameter β_j on β. Thus, we explain re-parameterization method about the parameters in Appendix.

3.2. Monotone Dose Finding Model in Bayesian Perspective

To obtain the predictive probability of the equation (7) in section 2.2.2, we explain the method presented by to Pozzi et al. (2013). In particular, in this study, we adopt the ten different models presented by Pozzi et al. (2013) as follows;

$$M_1 : \theta_1 \geq \theta_2 \geq \theta_3 \geq \theta_4 \geq \theta_5 \qquad M_2 : \theta_1 = \theta_2 \geq \theta_3 \geq \theta_4 \geq \theta_5$$
$$M_3 : \theta_1 \geq \theta_2 \geq \theta_3 \geq \theta_4 = \theta_5 \qquad M_4 : \theta_1 \geq \theta_2 \geq \theta_3 = \theta_4 = \theta_5$$
$$M_5 : \theta_1 = \theta_2 = \theta_3 \geq \theta_4 \geq \theta_5 \qquad M_6 : \theta_1 = \theta_2 \geq \theta_3 \geq \theta_4 = \theta_5 \qquad (8)$$
$$M_7 : \theta_1 = \theta_2 = \theta_3 = \theta_4 \geq \theta_5 \qquad M_8 : \theta_1 = \theta_2 \geq \theta_3 = \theta_4 = \theta_5$$
$$M_9 : \theta_1 = \theta_2 = \theta_3 \geq \theta_4 = \theta_5 \qquad M_{10} : \theta_1 \geq \theta_2 = \theta_3 = \theta_4 = \theta_5.$$

Ten different models of the equation (8) are averaged with the Bayesian Model Averaging in Bayesian perspective.

To model the monotonic constraint under mth model and to make an explanation of the distribution of the interesting parameter vector $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5)$ about five dose levels, we adopt the concept of the jump variable such as $\delta_{j,k} = \log(\alpha_j) - \log(\alpha_{j+1})$ and we put a truncated normal[1] prior on

$$\delta_{sum} = \sum_{j-1} \delta_{j,k} = \log(\alpha_1) - \log(\alpha_5) \sim TN(\mu_{sum}, \delta_{sum}),$$

where K is ten different model presented in the equation (8), $k = 1, \cdots, 10$. Besides, if we assume the monotonic dose response relationship about five dose levels, there is the relationship between five dose groups after taking logarithm of the interesting parameter about five dose levels in ten different models of the equation (8).

1) We use a truncated normal distribution for a particular parameter, so that we can restrict the range of parameter regions but have some belief about its mean and variance (Ando, 2010). The prior is chosen to maximize the average information in the data density relative to that in the prior.

where $\alpha_j > \alpha_{j+1}$ to make definition $\delta_{j,k} = \log(\alpha_j) - \log(\alpha_{j+1}) > 0$. After completing the jump variables about each model in equation (5) of section 2.2.2, we make a stick prior about each model. Then we made stick-breaking construction about the stick prior. The stick-breaking construction brings about semi-parametric model due to monotonic constraint of dose response relationship. The stick priors are made by corresponding to the relationship of the model and their order statistics that $s_{j,k} \sim U(0,1)$ $\mathrm{j} = 1,2,3$ and $\mathrm{k} = 1, \cdots, 10$. Finally, we summarize the different configurations of the $\delta's$ in jump × model matrix as follows;

$$\Delta = \begin{array}{c} \begin{matrix} M_1 & M_2 & M_3 & M_4 & M_5 & M_6 & M_7 & M_8 & M_9 & M_{10} \end{matrix} \\ \begin{pmatrix} \delta_{1,1} & 0 & \delta_{1,3} & \delta_{1,4} & 0 & 0 & 0 & 0 & 0 & \delta_{1,10} \\ \delta_{2,1} & \delta_{2,2} & \delta_{2,3} & \delta_{2,4} & 0 & \delta_{2,6} & 0 & \delta_{2,8} & 0 & 0 \\ \delta_{3,1} & \delta_{3,2} & \delta_{3,3} & 0 & \delta_{3,5} & \delta_{3,6} & 0 & 0 & \delta_{3,9} & 0 \\ \delta_{4,1} & \delta_{4,2} & 0 & 0 & \delta_{4,5} & 0 & \delta_{4,7} & 0 & 0 & 0 \end{pmatrix} \end{array}$$

Ishwaran and James (2001) remarked that the overall treatment effect δ_{sum} is partitioned along each column using a "stick-breaking" construction involving independent uniform variables. Eventually, after putting a prior on the collection of models, the stick-breaking construction allows us to average over all of ten different models without any parametric assumption on the dose response curve. That is, it is fact that we adopt the Bayesian Model Averaging in the equation (1) as a Bayesian semi-parametric model. This model made with the Bayesian Model Averaging is used to draw samples from the posterior distribution.

3.3. The Sampling-Importance Resampling(SIR) to Dose Finding Model

We adopt the Sampling-Importance Resampling algorithm to approximate the posterior distribution $\pi(\theta|Y)$ of the parameters θ given the interim data to compute the predictive probabilities to be used for making an interim decision. First, we let the parameters, θ, be obtained from the prior density $p(\theta)$ in the interim posterior. That is, we can tell $p(\theta) = \pi(\theta|Y)$. However, our primary concern is to approximate the integrals of equation (8) in section 2.2.3 by averaging over many possible future outcomes Y^* in stage 2.

To approximate the integrals of the equation (7), we select a posterior sample, $(\boldsymbol{\theta})^{(1)}, \cdots, (\boldsymbol{\theta})^{(r)}, \cdots, (\boldsymbol{\theta})^{(N)}$given the model with the Bayesian Model Averaging. To simulate future outcomes Y^* in stage 2, we draw a pair $(\boldsymbol{\theta})^{(l)}$ from this sample, and utilize this parameter to simulate one post-interim posterior data set $Y_j^{*(l)}$, where $j \in \{2,3,4\}$ is the dose to be explored; we also include future placebo and the responses of the five doses in this data set. The probability of data set $Y_j^{*(l)}$ given the pair, $(\boldsymbol{\theta})^{(r)}$, $p\left(Y_j^{*(l)} \middle| (\boldsymbol{\theta})^{(r)}\right)$ is computed $r = 1, \cdots, R$. We draw the samples with the importance ratios or weights as follows;

$$
\begin{aligned}
w_r &= \frac{l(\theta_r; Y)}{\sum_j \ l(\theta_j; Y)} \\
&= \frac{p\left(Y_j^{*(l)} \middle| (\theta)^{(r)}\right) \times p((\theta)^{(r)})}{\sum_j \ p\left(Y_j^{*(l)} \middle| (\theta)^{(j)}\right) \times p((\theta)^{(j)})} \\
&= \frac{p\left(Y_j^{*(l)} \middle| (\theta)^{(r)}\right)}{\sum_j \ p(Y_j^{*(l)} \mid (\theta)^{(j)})}
\end{aligned}
$$

which is the interim posterior sample uniformly distributed with probability $\mathrm{p}\left((\alpha,\beta)^{(l)}\right) = 1/N,\ \ \mathrm{r} = 1,\cdots,\mathrm{R}$.

Next, we should check a sample $(\boldsymbol{\theta})^{(r)}$ to be satisfied with the decision criterions such as efficacy criterion presented in section 2.2.1 directly. With satisfied the efficacy criterions in sample $(\boldsymbol{\theta})^{(r)}$ s, we can approximate the equation(7) is following to

$$\begin{aligned}
&\mathrm{P}\{\text{Success in dose j}|Y^*, M_k\} \\
&\quad = \int P\{Success\ in\ dose\ j|(\boldsymbol{\theta}), Y^*, M_K\}\, P\{(\boldsymbol{\theta}), Y^*, M_K\} d(\boldsymbol{\theta}) \\
&\quad \approx \frac{1}{N}\sum_{i=1}^{N} I\left\{(\boldsymbol{\theta})^{(r)}\ Success\ in\ dose\ j\right\} \times w_k \\
&\quad = PP_j^{*(l)}[Success\ in\ dose\ j].
\end{aligned}$$

We repeat this procedure for l until we can recognize very large number, say R. Finally, we can identify the lowest dose d satisfying the efficacy criterions.

$$PP_j^{*(l)} = \frac{1}{L}\sum_{l=1}^{L} I\left\{ \cap_{\{success\}}\{PP_j^{\{l\}}[Success] > \mathrm{c}\}\right\}.$$

For convenience, we summarize the algorithm as follows:

Step 1: Select dose j;

Step 2: Sample $(\boldsymbol{\theta})^{(1)}, (\boldsymbol{\theta})^{(2)}, \cdots, (\boldsymbol{\theta})^{(r)}, \cdots, (\boldsymbol{\theta})^{(N)}$ from the post-interim posterior distribution;

Step 3: Draw $(\boldsymbol{\theta})^{(l)}$ from the sample of the posterior distribution at interim stage of the entire sample size N;

Step 4: Simulate one data set $Y_j^{*(l)} | (\boldsymbol{\theta})^{(l)}$;

Step 5: Apply Sampling-Importance Resampling to compute $\mathrm{p}\left(Y_j^{*(l)} \middle| (\boldsymbol{\theta})^{(r)}\right)$, $\mathrm{r} = 1, \cdots, \mathrm{R}$ w_r and $PP_j^{*(l)}[criterion]$ for each criterion success is defined upon;

Step 6: Repeat steps from 3 to 5

3.4. Procedure at Second Stage

As described above, we provide the Bayesian two-stage dose finding method to find the MED based on the predictive probability comparing the pre-specified threshold (t) for binary outcome in phase Ⅱ clinical trials. As we can see figure 1, which is the entire design flow, we randomize a placebo, three intermediate doses, $\mathrm{j} = \{2, 3, 4\}$ and the highest dose at the second stage. In this stage, we make a test which dose is the MED under the predictive probability to compare to the pre-specified threshold. However, if we do not make decision in the end of the study, we should make decision whether to stop for efficacy, or stop for futility relied on the opinion of the clinical experts.

4. Simulation Study

In this study, fundamental approach of the simulation studies was conducted to certify which dose level is the MED in phase Ⅱ clinical trials. Therefore, in this section, we performed the simulations to evaluate the performance of the proposed method under two different scenarios that we can make decision which dose level could be the MED using the Bayesian Model Averaging method with ten presented models about five dose levels presented by Pozzi et al. (2013).

In addition, we conducted the simulations to evaluate the performance of the presented eight different models, respectively, to check which model is useful to find the MED or not useful without using the Bayesian Model Averaging. In simulation setting with the Bayesian Model Averaging and eight different models, we identify the true dose level before conducting the simulation to find the MED about each scenario. After conducting the simulation, we compare the true dose with the identified MED we found. As a result, we could check that the method we used at each scenario is well to detect the MED of the Bayesian Model Averaging and eight different models.

Before we identify the MED, we find the true dose about each scenario in order to check which case of the scenario of the model is well find the MED. For example, there is an example case of the scenario how we can find the true dose before conducting the simulation under the efficacy criterion presented in section 2.2.1. There is the Table 2 in the method to find the true dose.

In order to identify the MED, we make ten scenarios in Table 3. Each scenario is distributed in binomial distribution with the probability of the success p_j called the mean effect, θ_j. The scenarios 1 and 2 are used for the Bayesian Model

Averaging to find the MED. The scenario 3, 4, 5, 6, 7, 8, 9 and 10 are used for eight different models to identify the MED.

Table 2. Method to find the True Dose

	θ_1	θ_2	θ_3	θ_4	θ_5
Binomial Probability	0.5	0.45	0.4	0.28	0.24
$\frac{\theta_j}{\theta_1}$ for success probability	1	0.9	0.8	0.56	0.48
$\frac{\theta_j}{\theta_5}$ for success probability	2.08	1.87	1.67	1.17	1

Note: The term of θ_1 is the mean effect of a placebo, θ_j is the mean effect of the j dose, θ_5 is the mean effect of the highest dose. As we mentioned the efficacy criterions in section 2.2.2, we can decide the true dose satisfying in which $\frac{\theta_j}{\theta_1}$ is at least greater than 0.5 and $\frac{\theta_j}{\theta_5}$ is less than 1.2.

Table 3. Scenarios of the Mean Effect

	θ_1	θ_2	θ_3	θ_4	θ_5
Scenario1	0.5	0.45	0.4	0.28	0.24
Scenario2	0.45	0.38	0.25	0.22	0.21
Scenario3	0.5	0.35	0.3	0.26	0.22
Scenario4	0.51	0.51	0.4	0.29	0.245
Scenario5	0.55	0.45	0.3	0.255	0.255
Scenario6	0.45	0.24	0.21	0.21	0.21
Scenario7	0.5	0.5	0.5	0.26	0.22
Scenario8	0.58	0.58	0.3	0.26	0.26
Scenario9	0.45	0.45	0.26	0.19	0.19
Scenario10	0.45	0.45	0.45	0.24	0.205

Note: The term of θ_1 is the mean effect of a placebo, θ_2, θ_3, and θ_4 are the mean effects of the intermediate doses d, θ_5 is the mean effect of the highest dose.

We assume that five dose levels such as a placebo, three intermediate doses and the highest dose was used to identify the MED with model. To begin with, we generate observations from random variable, $Y_{ij} \sim bin(n_j, \theta_j)$, for patients $i = 1, \cdots, n_j$ at dose levels, $j = 1, \cdots, 5$ through the checking the sufficient condition as ten different models in equation (7) presented in section 2.2.2. The observable data set is determined by following steps;

Step 1: Generate data set following to binomial random variable in 5 dose levels, which is $Y_{ij} \sim bin(n_j, \theta_j)$ for patients $i = 1, \cdots, n_j$ at dose levels, $j = 1, \cdots, 5$. Each value of the probability of the success θ_j is composed in accordance with efficacy criterions in section 2.2.1.

Step 2: Using the observable data set in step 1, based on the clinical thresholds, $t = 0.3$, $t = 0.4$ and $t = 0.5$, we run MCMC algorithm with the Bayesian Model Averaging method that we iterate 10,000 times, burn-in is 1,000 and thin is 60. We get that total sample size is 150 in first stage about both in the Bayesian Model Averaging of ten models and eight different models. Besides, we do 100 iterations about Bayesian Model Averaging of ten models and eight different models to identify the MED about each sample.

4.1. Simulation Results using the Bayesian Model Averaging

In order to identify the MED with Bayesian Model Averaging, we conducted the simulation study of two different scenarios relying on the relationship of the mean effect of the dose levels. To check the effect of corresponding to the difference of the clinical thresholds, we did simulation when the clinically thresholds were t= 0.3, t = 0.4 and t= 0.5 in accordance with each scenario.

In the table 4, based on the definition of efficacy criterions in section 2.2.2, we recognize that the true dose is dose 4 because it is satisfied with that $\frac{\theta_j}{\theta_1}$ is at least greater than 0.5 and $\frac{\theta_j}{\theta_5}$ is less than 1.2, simultaneously. Therefore, we regard the dose 4 as the true dose before finding the MED. After conducting the simulation study, we can certify that the dose 4 as the MED is fitted of the true dose when we have t = 0.4 and t = 0.5 except for t = 0.3 in scenario 1. In addition, we can find that there was an under-estimated tendency in identifying the MED

In the table 5, based on the definition of efficacy criterions in section 2.2.1, we recognize that the true dose is dose 3 because it is satisfied with that $\frac{\theta_j}{\theta_1}$ is at least greater than 0.5 and $\frac{\theta_j}{\theta_5}$ is less than 1.2, simultaneously. Therefore, we regard the dose 3 as the true dose before finding the MED. After conducting the simulation study, by the simulation result of scenario2, we certify that the dose 3 is the MED that this result is fitted of the true dose when we have t = 0.3, t= 0.4 and t = 0.5. In addition, we can find that there was an under-estimated tendency in identifying the MED

Table 4. Assumption: the true dose is dose 4 based on Scenario 1 ($\theta_1 = 0.5,\ \theta_2 = 0.45,\ \theta_3 = 0.4,\ \theta_4 = 0.28,\ \theta_5 = 0.24$)

Threshold (t)	Dose	Mean Effect	Frequency (%)	MED
0.3	1	0.5	4(0.4)	
	2	0.45	10(1)	
	3	0.4	45(45)	◙
	4*	0.28	36(36)	
	5	0.24	5(0.5)	
0.4	1	0.5	1(01)	
	2	0.45	8(0.8)	
	3	0.4	38(38)	
	4*	0.28	48(48)	◙
	5	0.24	5(0.5)	
0.5	1	0.5	0(0)	
	2	0.45	3(0.3)	
	3	0.4	36(36)	
	4*	0.28	53(53)	◙
	5	0.24	8(0.8)	

* : true dose; ◙ : MED

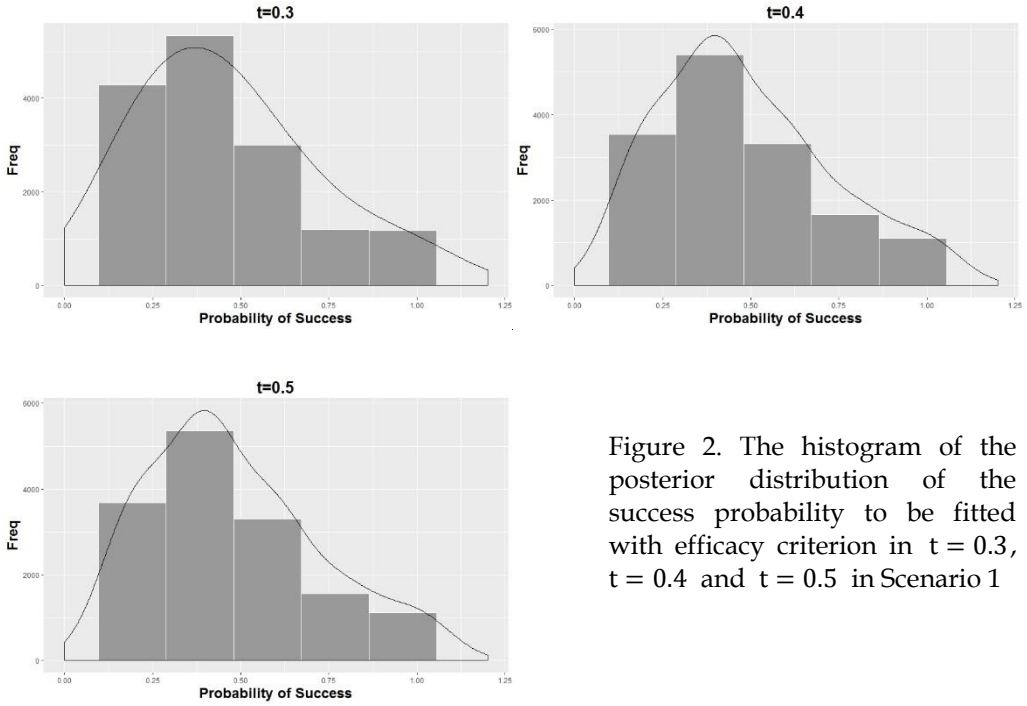

Figure 2. The histogram of the posterior distribution of the success probability to be fitted with efficacy criterion in t = 0.3, t = 0.4 and t = 0.5 in Scenario 1

Table 5. Assumption: the true dose is dose 3 based on scenario 2 ($\theta_1 = 0.45,\ \theta_2 = 0.38,\ \theta_3 = 0.25,\ \theta_4 = 0.22,\ \theta_5 = 0.21$)

Threshold (t)	Dose	Mean Effect	Frequency (%)	MED
0.3	1	0.45	4(0.4)	
	2	0.38	17(17)	
	3*	0.25	68(68)	◙
	4	0.22	11(11)	
	5	0.21	5(0.5)	
0.4	1	0.45	0(0)	
	2	0.38	11(11)	
	3*	0.25	64(64)	◙
	4	0.22	23(23)	
	5	0.21	2(0.2)	
0.5	1	0.45	0(0)	
	2	0.38	11(11)	
	3*	0.25	61(61)	◙
	4	0.22	28(28)	
	5	0.21	0(0)	

* : true dose; ◙: MED

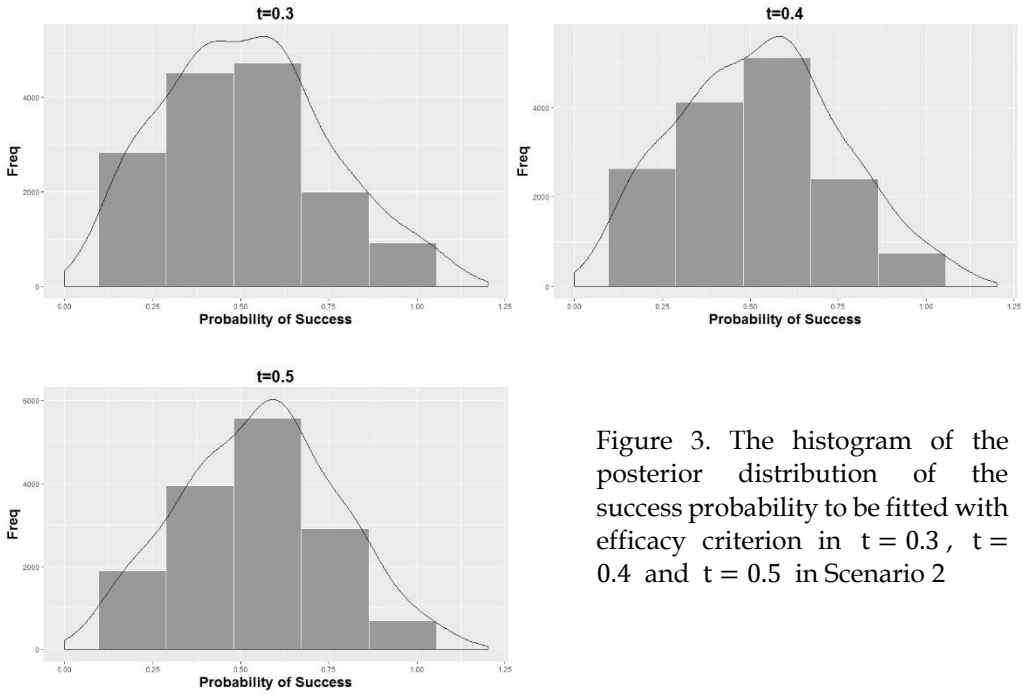

Figure 3. The histogram of the posterior distribution of the success probability to be fitted with efficacy criterion in $t = 0.3$, $t = 0.4$ and $t = 0.5$ in Scenario 2

4.2. Simulation Results in eight different models

As mentioned in section 4.1, we certify that the simulation of the Bayesian Model Averaging is to show accurately identifying to the MED. If so, we are interested in which model among ten different models may well fitted to figure out the MED. Therefore, in order to check and compare with the simulations results which method is to test stably the MED between different models and the Bayesian Model Averaging, we conduct the simulation study relying on the eight different models presented by Pozzi et al.(2013) and we did simulation when the clinically thresholds are $t = 0.3$, $t = 0.4$ and $t = 0.5$ about each scenario.

In the table 6, according to the definition of efficacy criterion in section 2.2.2, we recognize that the true dose is dose 4 because it is satisfied with that $\frac{\theta_j}{\theta_1}$ is at least greater than 0.5 and $\frac{\theta_j}{\theta_5}$ is less than 1.2, simultaneously. Therefore, we regard the dose 4 as the true dose before finding the MED. After conducting the simulation study, by the simulation result of scenario 3, we certify that the dose 4 as the MED is fitted of the true dose when we have $t = 0.3$, $t = 0.4$ and $t = 0.5$.

In the table 7, according to efficacy criterions, we recognize that dose 4 is satisfied with that $\frac{\theta_j}{\theta_1}$ is at least greater than 0.5 and $\frac{\theta_j}{\theta_5}$ is less than 1.2, simultaneously. Therefore, we regard the dose 4 as the true dose before finding the MED. After conducting the simulation study, by the simulation result of scenario 4, we certify that the dose4 as the MED is fitted of the true dose when we have $t = 0.3$, $t = 0.4$ and $t = 0.5$.

Table 6. Assumption: the true dose is dose 4 based on scenario 3 ($\theta_1 = 0.5,\ \theta_2 = 0.35,\ \theta_3 = 0.3,\ \theta_4 = 0.26,\ \theta_5 = 0.22$)

Threshold (t)	Dose	Mean Effect	Frequency (%)	MED
0.3	1	0.5	0(0)	
	2	0.35	2(0.2)	
	3	0.3	25(25)	
	4*	0.26	70(70)	◙
	5	0.22	3(0.3)	
0.4	1	0.5	1(0.1)	
	2	0.35	1(0.1)	
	3	0.3	17(17)	
	4*	0.26	80(80)	◙
	5	0.22	1(0.1)	
0.5	1	0.5	0(0)	
	2	0.35	0(0)	
	3	0.3	32(32)	
	4*	0.26	67(67)	◙
	5	0.22	1(0.1)	

* : true dose; ◙: MED

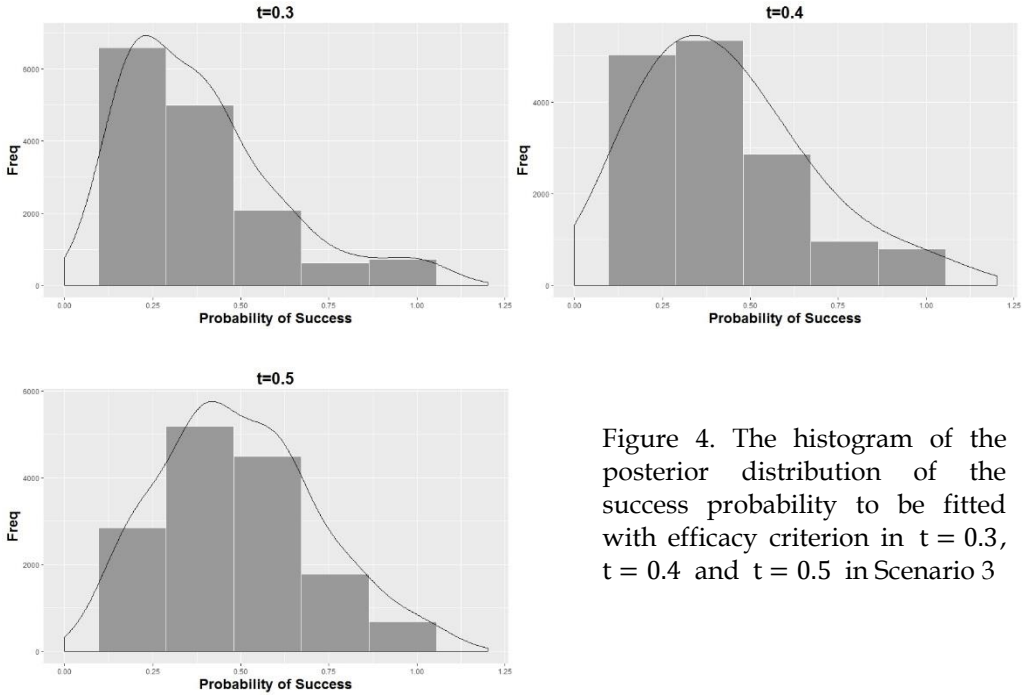

Figure 4. The histogram of the posterior distribution of the success probability to be fitted with efficacy criterion in t = 0.3, t = 0.4 and t = 0.5 in Scenario 3

Table 7. Assumption: the true dose is dose 4 based on scenario 4 ($\theta_1 = 0.51,\ \theta_2 = 0.51,\ \theta_3 = 0.4,\ \theta_4 = 0.29,\ \theta_5 = 0.245$)

Threshold (t)	Dose	Mean Effect	Frequency (%)	MED
0.3	1	0.51	2(0.2)	
	2	0.51	0(0)	
	3	0.4	11(11)	
	4*	0.29	76(76)	◙
	5	0.245	11(11)	
0.4	1	0.51	1(0.1))	
	2	0.51	0(0)	
	3	0.4	7(0.7)	
	4*	0.29	81(81)	◙
	5	0.245	11(11)	
0.5	1	0.51	0(0)	
	2	0.51	0(0)	
	3	0.4	2(0.2)	
	4*	0.29	87(87)	◙
	5	0.245	11(11)	

* : true dose; ◙: MED

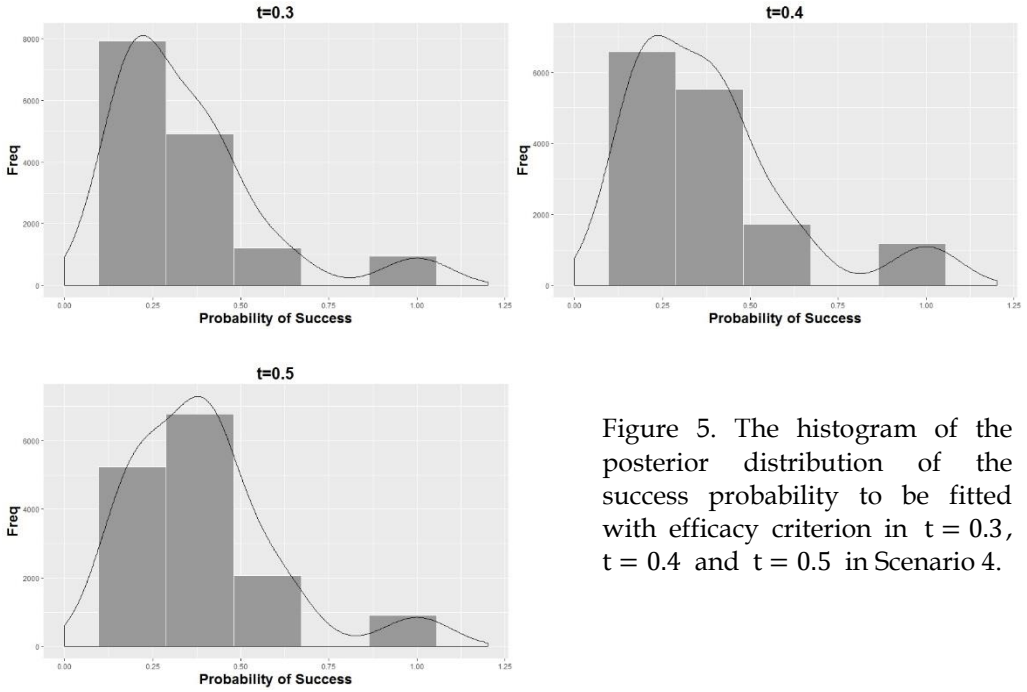

Figure 5. The histogram of the posterior distribution of the success probability to be fitted with efficacy criterion in t = 0.3, t = 0.4 and t = 0.5 in Scenario 4.

In the table 8, according to efficacy criterions, we recognize that dose3 is satisfied with that $\frac{\theta_j}{\theta_1}$ is at least greater than 0.5 and $\frac{\theta_j}{\theta_5}$ is less than 1.2. Therefore, we regard the dose 3 as the true dose before finding the MED. After conducting the simulation study, by the simulation result of scenario 5, we certify that the dose3 as the MED is fitted of the true dose when we have $t = 0.3$, $t = 0.4$ and $t = 0.5$.

In the table 9, according to efficacy criterions, we recognize that dose2 is satisfied with that $\frac{\theta_j}{\theta_1}$ is at least greater than 0.5 and $\frac{\theta_j}{\theta_5}$ is less than 1.2. Therefore, we regard the dose2 as the true dose before finding the MED. After conducting the simulation study, by the simulation result of scenario 6, we certify that the dose2 as the MED is fitted of the true dose when we have $t = 0.3$, $t = 0.4$ and $t = 0.5$.

Table 8. Assumption: the true dose is dose 3 based on scenario 5 ($\theta_1 = 0.55,\ \theta_2 = 0.45,\ \theta_3 = 0.3,\ \theta_4 = 0.255,\ \theta_5 = 0.255$)

Threshold (t)	Dose	Mean Effect	Frequency (%)	MED
0.3	1	0.55	2(0.2)	
	2	0.45	4(0.4)	
	3*	0.3	76(76)	◙
	4	0.255	18(18)	
	5	0.255	0(0)	
0.4	1	0.55	1(01)	
	2	0.45	0(0)	
	3*	0.3	83(83)	◙
	4	0.255	16(16)	
	5	0.255	0(0)	
0.5	1	0.55	0(0)	
	2	0.45	0(0)	
	3*	0.3	90(90)	◙
	4	0.255	10(10)	
	5	0.255	0(0)	

* : true dose; ◙: MED

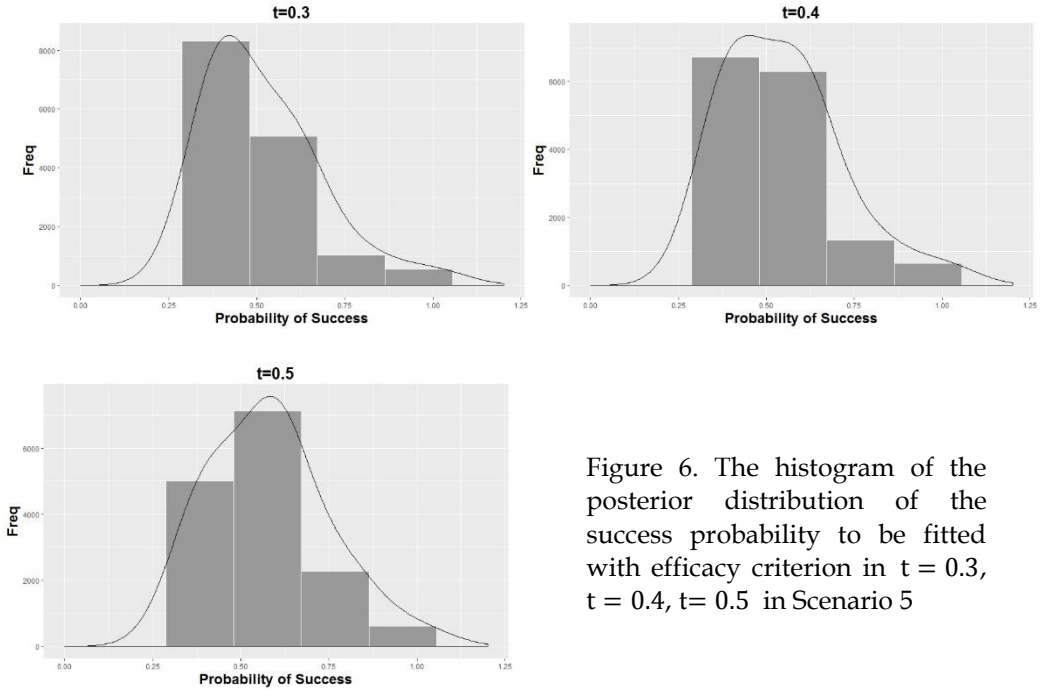

Figure 6. The histogram of the posterior distribution of the success probability to be fitted with efficacy criterion in t = 0.3, t = 0.4, t= 0.5 in Scenario 5

Table 9. Assumption: the true dose is dose 2 based on scenario 6 ($\theta_1 = 0.45,\ \theta_2 = 0.24,\ \theta_3 = 0.21\ \theta_4 = 0.21,\ \theta_5 = 0.21$)

Threshold (t)	Dose	Mean Effect	Frequency (%)	MED
0.3	1	0.45	1(0.1)	
	2*	0.24	65(65)	◙
	3	0.21	34(34)	
	4	0.21	0(0)	
	5	0.21	0(0)	
0.4	1	0.45	0(0)	
	2*	0.24	58(58)	◙
	3	0.21	42(42)	
	4	0.21	0(0)	
	5	0.21	0(0)	
0.5	1	0.45	0(0)	
	2*	0.24	71(71)	◙
	3	0.21	29(29)	
	4	0.21	0(0)	
	5	0.21	0(0)	

* : true dose; ◙: MED

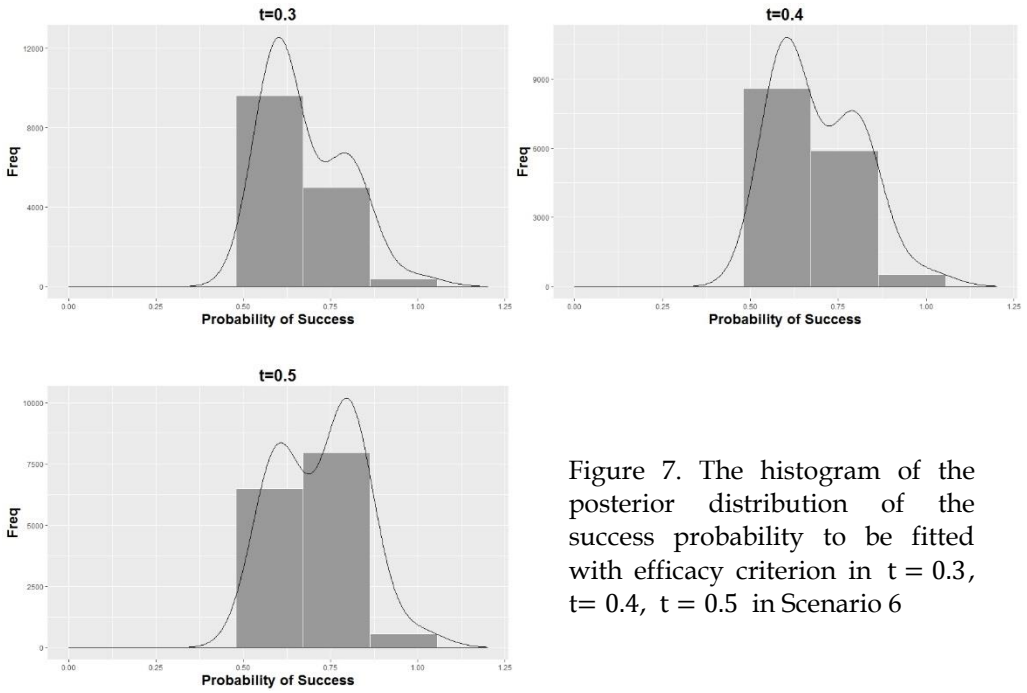

Figure 7. The histogram of the posterior distribution of the success probability to be fitted with efficacy criterion in $t = 0.3$, $t = 0.4$, $t = 0.5$ in Scenario 6

In the table 10, according to efficacy criterions, we recognize that dose4 is satisfied with that $\frac{\theta_j}{\theta_1}$ is at least greater than 0.5 and $\frac{\theta_j}{\theta_5}$ is less than 1.2. Therefore, we regard the dose 4 as the true dose before finding the MED. After conducting the simulation study, , by the simulation result of scenario 7, we certify that the dose4 as the MED is fitted of the true dose when we have $t = 0.3$, $t = 0.4$ and $t = 0.5$.

In the table 11, according to efficacy criterions, we recognize that dose3 is satisfied with that $\frac{\theta_j}{\theta_1}$ is at least greater than 0.5 and $\frac{\theta_j}{\theta_5}$ is less than 1.2. Therefore, we regard the dose 3 as the true dose before finding the MED. However, after conducting the simulation study, by the simulation result of scenario 8, we certify that the dose4 as the MED is fitted of the true dose in $t = 0.3$ except for $t = 0.4$ and $t = 0.5$. In addition, we can find that there was an over-estimated tendency in identifying the MED

Table 10. Assumption: the true dose is dose 4 based on scenario 7 ($\theta_1 = 0.5,\ \theta_2 = 0.5,\ \theta_3 = 0.5\ \ \theta_4 = 0.26,\ \theta_5 = 0.22$)

Threshold (t)	Dose	Mean Effect	Frequency (%)	MED
0.3	1	0.5	3(03)	
	2	0.5	0(0)	
	3	0.5	0(0)	
	4*	0.26	55(55)	◙
	5	0.22	42(42)	
0.4	1	0.5	2(0.2)	
	2	0.5	0(0)	
	3	0.5	0(0)	
	4*	0.26	50(50)	◙
	5	0.22	48(48)	
0.5	1	0.5	0(0)	
	2	0.5	0(0)	
	3	0.5	0(0)	
	4*	0.26	58(58)	◙
	5	0.22	42(42)	

* : true dose; ◙: MED

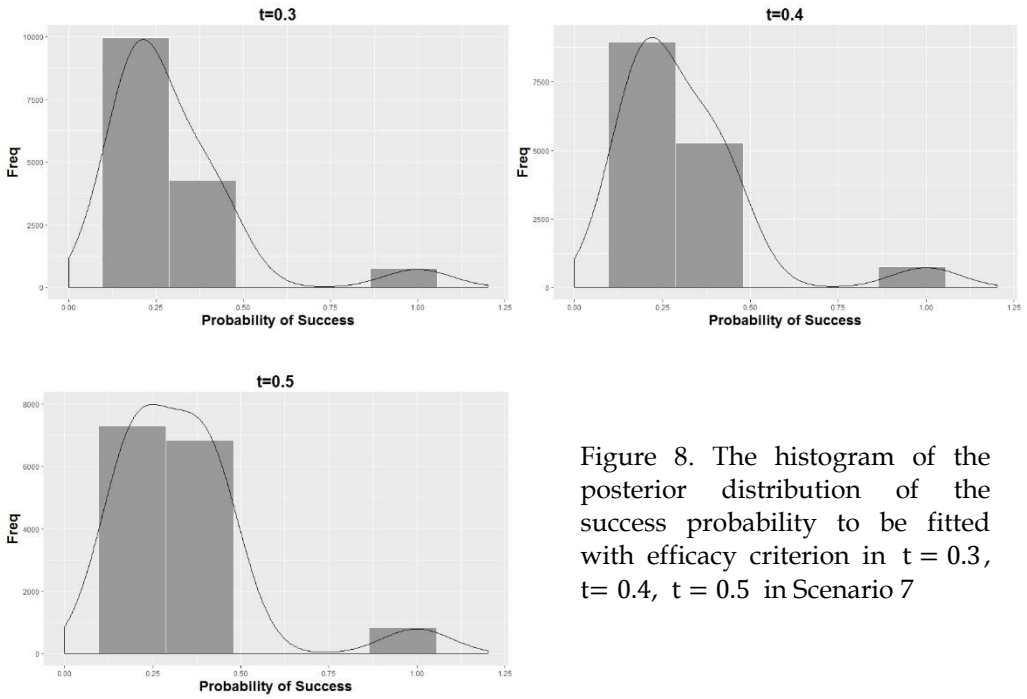

Figure 8. The histogram of the posterior distribution of the success probability to be fitted with efficacy criterion in $t = 0.3$, $t = 0.4$, $t = 0.5$ in Scenario 7

Table 11. Assumption: the true dose is dose 3 based on scenario 8 ($\theta_1 = 0.58,\ \theta_2 = 0.58,\ \theta_3 = 0.3\ \ \theta_4 = 0.26,\ \theta_5 = 0.26$)

Threshold (t)	Dose	Mean Effect	Frequency (%)	MED
0.3	1	0.58	1(0.1)	
	2	0.58	0(0)	
	3*	0.3	59(59)	◙
	4	0.26	40(40)	
	5	0.26	0(0)	
0.4	1	0.58	0(0)	
	2	0.58	0(0)	
	3*	0.3	36(36)	
	4	0.26	64(64)	◙
	5	0.26	0(0)	
0.5	1	0.58	0(0)	
	2	0.58	0(0)	
	3*	0.3	49(49)	
	4	0.26	51(51)	◙
	5	0.26	0(0)	

* : true dose; ◙: MED

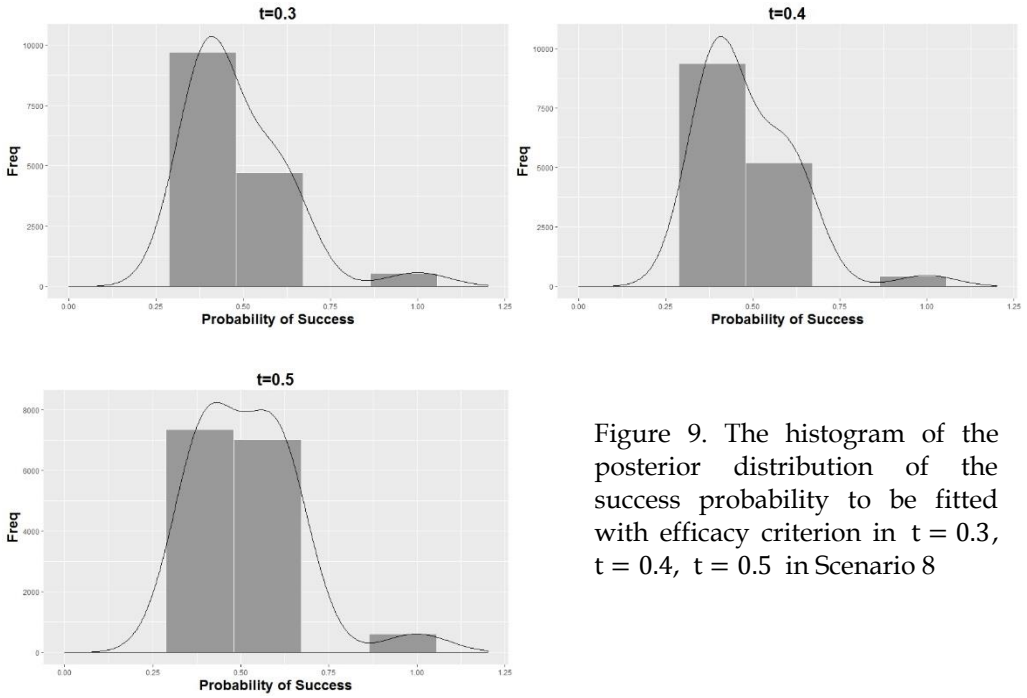

Figure 9. The histogram of the posterior distribution of the success probability to be fitted with efficacy criterion in t = 0.3, t = 0.4, t = 0.5 in Scenario 8

In the table 12, according to efficacy criterions, we recognize that dose3 is satisfied with that $\frac{\theta_j}{\theta_1}$ is at least greater than 0.5 and $\frac{\theta_j}{\theta_5}$ is less than 1.2. Therefore, we regard the dose 3 as the true dose before finding the MED. After conducting the simulation study, by the simulation result of scenario 9, we can certify that the dose5 is the MED but there is no fitted dose comparing with true dose. As a result, we can't find the MED in scenario9 because dose5 was not considered the MED in advance.

In the table 13, according to efficacy criterions, we recognize that dose 4 is satisfied with that $\frac{\theta_j}{\theta_1}$ is at least greater than 0.5 and $\frac{\theta_j}{\theta_5}$ is less than 1.2. Therefore, we regard the dose 4 as the true dose before finding the MED. After conducting the simulation study, by the simulation result of scenario 10, we can certify that the dose3 is the MED but there is no fitted dose comparing with true dose. In addition, we can find that there was an under-estimated tendency in identifying the MED

Table 12. Assumption: the true dose is dose 3 based on scenario 9($\theta_1 = 0.45$, $\theta_2 = 0.45$, $\theta_3 = 0.26$ $\theta_4 = 0.19$, $\theta_5 = 0.19$)

Threshold (t)	Dose	Mean Effect	Frequency (%)	MED
0.3	1	0.45	16(16)	
	2	0.45	0(0)	
	3*	0.26	0(0)	
	4	0.19	0(0)	
	5	0.19	84(84)	◙
0.4	1	0.45	5(0.5)	
	2	0.45	0(0)	
	3*	0.26	0(0)	
	4	0.19	1(0.1)	
	5	0.19	94(94)	◙
0.5	1	0.45	1(0.1)	
	2	0.45	0(0)	
	3*	0.26	0(0)	
	4	0.19	0(0)	
	5	0.19	99(99)	◙

* : true dose; ◙: MED

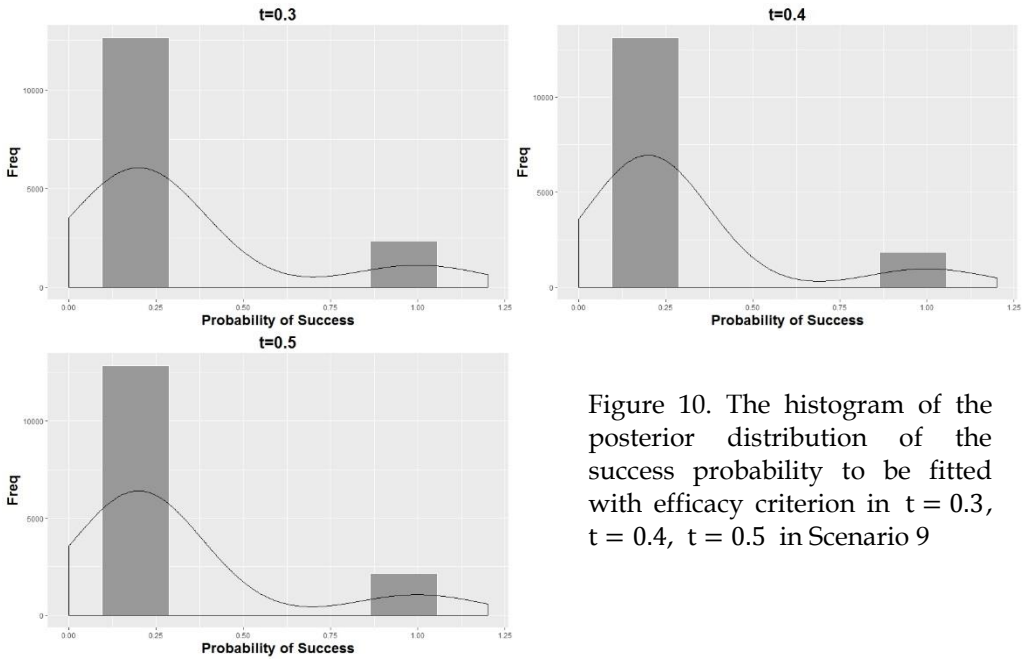

Figure 10. The histogram of the posterior distribution of the success probability to be fitted with efficacy criterion in t = 0.3, t = 0.4, t = 0.5 in Scenario 9

Table 13. Assumption: the true dose is dose 4 based on scenario 10 ($\theta_1 = 0.45,\ \theta_2 = 0.45,\ \theta_3 = 0.45\ \theta_4 = 0.24,\ \theta_5 = 0.205$)

Threshold (t)	Dose	Mean Effect	Frequency (%)	MED
0.3	1	0.45	17(17)	
	2	0.45	0(0)	
	3	0.45	83(83)	◙
	4*	0.24	0(0)	
	5	0.205	0(0)	
0.4	1	0.45	5(0.5)	
	2	0.45	0(0)	
	3	0.45	95(95)	◙
	4*	0.24	0(0)	
	5	0.205	0(0)	
0.5	1	0.45	1(0.1)	
	2	0.45	0(0)	
	3	0.45	99(99)	◙
	4*	0.24	0(0)	
	5	0.205	0(0)	

* : true dose; ◙: MED

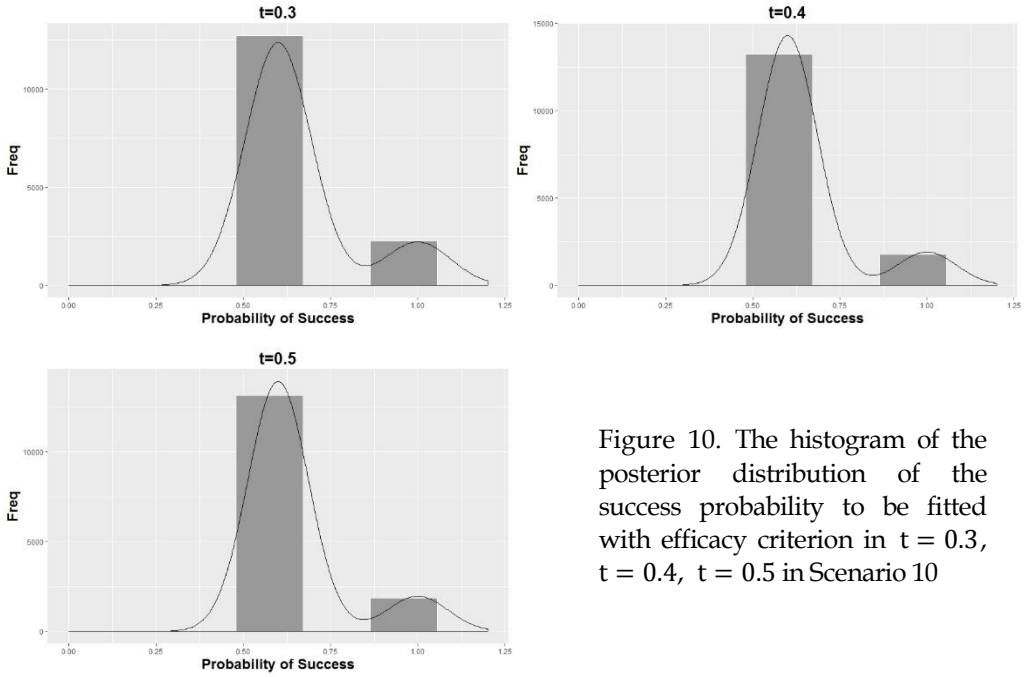

Figure 10. The histogram of the posterior distribution of the success probability to be fitted with efficacy criterion in $t = 0.3,\ t = 0.4,\ t = 0.5$ in Scenario 10

5. Conclusion and Discussion

Pozzi et al. (2013) suggested that the Bayesian adaptive two-stage dose finding design to identify the MED in phase Ⅱ clinical trials based on the predictive probability when we have over-dispersed count endpoints. Our proposed method was the dose finding method to find the MED for binary endpoints. Unfortunately, there were a few cases of adoption into the clinical trials even though we are interested in binary outcomes such as response or no response to treatment intervention. In particular, our proposed method was the semi-parametric model with monotonic constraint of the mean effect of the five dose groups. In addition, our proposed method adopted the Bayesian Model Averaging which is averaging presented ten different models to decrease the model uncertainty. Furthermore, pre-specified efficacy criterions were included into our proposed method to make decision which dose level is the MED.

To accomplish our goal in this study, the prior distributions for satisfied with efficacy criterions and their predictive distributions were defined with the efficacy criterions mentioned by Pozzi et al. (2013). We met a challenge to overcome draw samples from the posterior distribution. However, this problem had a variety of difficulties to obtain sample directly from the posterior distribution. Therefore, we approximate the integral of the posterior distribution through the Sampling-Importance Resampling algorithm proposed by Rubin (1983) and presented by Smith and Gelfand (1992).

Two different simulation scenarios were used to test our proposed model in adopting the Bayesian Model Averaging. For two simulation scenarios, the clinical threshold was set to $t = 0.3$, $t = 0.4$ and $t = 0.5$, respectively. As well, eight different simulation scenarios were used to test our proposed model in each

different model of the mean effect of the dose groups, respectively. The likelihood data sets were sampled from the binomial random distribution for each case.

After checking the model relevancy to identify the MED, the cases of the adoption of the Bayesian Model Averaging were well accurately suited to identify the MED coincided with the true dose. On the contrary, the case of each different model was not stable to identify the MED coincided with the true dose.

Unfortunately, there was limitation to show the appropriateness of our proposed model. One was that we did not adopt in diverse real data to test our proposed model. Another was to find the most suitable scenario in generating the binomial random variables in order to be satisfied with the efficacy criterions.

In the future study, it will be more useful approach if we extend the sample size determination to identify the MED to apply in phase Ⅱ clinical trials.

Appendix

For convenience, in order to handling only one parameter $\boldsymbol{\alpha_j}$ with fixing the parameter, β_j, we use the re-parameterization method. For binary response variable, Y_{ij} given the probability of the success p_{ij}, is binomial distributed: $Y_{ij} \sim bin(n_j, p_j)$. In addition, we regard that hyper-parameter of the binomial distribution, p_j, is beta distributed since we employ a conjugate beta prior distribution of the binomial distribution, $p_j \sim beta(\alpha_j, \beta_j)$. Thus, because of the mean of beta distribution, $\theta_j = \frac{\alpha_j}{\alpha_j + \beta_j}$, we use $\boldsymbol{\alpha_j} \propto \frac{\theta_j}{1-\theta_j}$ with increasing θ_j. Finally, we can find that $\frac{\theta_j}{1-\theta_j}$ is monotonically increasing. That is, with increasing θ_j, $\boldsymbol{\alpha_j}$ is monotonically increasing. We have the $\boldsymbol{\alpha_j}'s$ are proportional to the mean effect θ_j. For the beta distribution with two parameters α_j and β_j, we should make simple expression which can account for shape parameter α_j with fixing β_j on β. Without logical description, we could not explain the model with ease. Therefore, we use the re-parameterization concept to describe only one parameter among $\boldsymbol{\alpha_j} = (\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5)$ of the beta prior distribution. We know that the beta distribution is given by

$$\pi(p_j;\ \alpha_j, \beta_j) = \frac{\Gamma(\alpha_j + \beta_j)}{\Gamma(\alpha_j)\Gamma(\beta_j)} p_j^{\alpha_j - 1} (1 - p_j)^{\beta_j - 1}, \quad 0 < p_j < 1,$$

where $\alpha_j > 0$, $\beta_j > 0$ and $\Gamma(\cdot)$ is the gamma function. The mean and variance of p_j are, respectively,

$$\mathrm{E}(p_j) = \frac{\alpha_j}{\alpha_j + \beta_j}$$
$$\mathrm{var}(p_j) = \frac{\alpha_j \beta_j}{(\alpha_j+\beta_j)^2(\alpha_j+\beta_j+1)}.$$

To make re-parameterization of beta density, we let the mean effect $\theta_j = \frac{\alpha_j}{\alpha_j+\beta_j}$ and precision parameter be $\phi_j = \alpha_j + \beta_j$. Because $\alpha_j = \mu_j\phi_j$ and $\beta_j = (1-\theta_j)\phi_j$, we can re-write as follows;

$$\mathrm{E}(p_j) = \theta_j$$
$$\mathrm{var}(p_j) = \frac{V(\theta_j)}{1+\phi_j},$$

where $\mathrm{V}(\theta_j) = \theta_j(1-\theta_j)$. Therefore, we can get a new beta density as follows;

$$\mathrm{f}(p_j;\, \theta, \phi) = \frac{\Gamma(\phi_j)}{\Gamma(\theta_j\phi_j)\Gamma((1-\theta_j)\phi_j)} p_j^{\theta_j\phi_j-1}(1-p_j)^{(1-\theta_j)\phi_j-1},\ 0 < p_j < 1.$$

If we let $\theta_j = \frac{\alpha_j}{\alpha_j+\beta_j}$ and $\phi_j = \alpha_j + \beta$, we find that $p_j \sim beta(\theta_j\phi, (1-\theta)\phi)$ and $\theta_j\alpha_j + \theta_j\beta = \alpha_j$. Thus, we have $\alpha_j = \frac{\theta_j}{1-\theta_j}\beta$. Using $\alpha_j \propto \frac{\theta_j}{1-\theta_j}$ with increasing θ_j, we find that $\frac{\theta_j}{1-\theta_j}$ is monotonically increasing. That is, when θ_j increases, $\boldsymbol{\alpha_j}$ is monotonically increasing. The parameter β is assumed identical across dose groups to reach an identifiable condition; $\log(\beta) \sim N(0, \sigma_\beta^2)$, which correspond to β around one on average. Next, we should consider the method that the parameter α_j of the beta prior distribution is to be fitted well in the semi-parametric model through the Bayesian Model Averaging under the monotonic constraint.

Bibliography

1. Ivanova A, Xiao C and Tymofyeyev Y. (2012). Two-stage designs for Phase 2 dose-finding trials. *Statistics in Medicine* 31, 2872-2881

2. Duncan DB. (1965). A Bayesian Approach to Multiple Comparisons. *Technometrics* 7, 171-222

3. Rubin DB. (1983). The Bayesian Bootstrap. *The Annals of Statistics* 9, 130-134

4. Bretz F, Dette H and Pinheiro JC. (2010). Practical considerations for optimal designs in clinical dose finding studies. *Statistics in Medicine* 29, 731-742

5. Gehan EA. (1961). The determination of the number of patients required in a preliminary and a follow-up trial of a new chemotherapy agent. *Journal of chronic disease* 13, 346-353

6. Green SJ and Dahlberg S. (1992). Planned versus attained design in Phase Ⅱ Clinical Trials. *Statistics in Medicine* 11, 853-862

7. ICH Harmonized Tripartite Guideline. (1994). Topic E4: Dose-response information to support drug registration

8. Hsu JC. and Berger RL. (1999). Stepwise confidence intervals without multiplicity adjustment for dose response and toxicity studies. *Journal of the American Statistical Association* 94, 468-482

9. Lee JJ. and Liu DD. (2008). A predictive probability design for phase Ⅱ cancer clinical trials. *Clinical Trials* 5, 93-106

10. Hoeting JA, Madigan D, Raftery AE and Volinsky CT. (1999). Bayesian Model Averaging: A Tutorial. *Statistical Science 14*, 382-417

11. Koch KR. (2007). Gibbs Sampling by Sampling-Importance Resampling. *Journal of Geodesy* 81, 581-591

12. Kramar A, Povtin D. and Hill C. (1996). Multistage designs for phase Ⅱ clinical trials: statistical issues in cancer research. *Brithsh Journal of Cancer 74*, 1317-1320

13. Pozzi L, Schmidli H. M, Gasparini M, and Racine-Poon A. (2013). A Bayesian adaptive dose selection procedure with an over-dispersed count endpoint. *Statistics in Medicine* 32, 5008-5027

14. Naitee T. (2006). Dose finding in drug development. New York, Springer

15. Polley M-Y. and Cheung YK. (2008). Two-stage designs for dose-finding trials with a biologic endpoint using stepwise tests. *Biometrics* 64, 232-241

16. Muller P, Parmigiani G, and Rice K. (2006). FDR and Bayesian Multiple Comparisons Rules. *ISBA 8th World Meeting on Bayesian Statistics*, 1-18

17. Simon R. (1989). Optimal two-stage designs for phase Ⅱ clinical trials. *Controlled clinical trials* 10, 1-10

18. Smith A and Gelfand A. (1992). Bayesian Statistics without Tears: A Sampling −Resampling Perspective. *The American Statistician* 46, 84-88

19. Ruberg SJ. (1989). Contrasts for identifying the minimum effective dose. *Journal of the American Statistical Association* 84, 816-822

20. Chen TT. (1997). Optimal Three-Stage Designs for Phase Ⅱ Cancer Clinical Trials. *Statistics in Medicine* 16, 2701-2711

21. Tan SB, Machin D. (2002). Bayesian two-stage designs for phase Ⅱ clinical trials. *Ststistics in Medicine* 21, 1991-2012

22. Fleming TR. (1982). One-Sample Multiple Testing Procedure for Phase Ⅱ Clinical Trials. *Biometrics* 38, 143-151

23. Sambucini V. (2008). A Bayesian predictive two-stage design for phase Ⅱ clinical trials. *Statistics in Medicine 27*, 1199-1224

국 문 요 약

이항자료를 갖는 2상 임상시험에서 적정 투여량 결정에 대한 베이지안 모형

본 논문은 2 상 임상시험에서 이항자료를 갖는 경우, 베이지안 접근방법을 이용한 이 단계 설계방법으로서 최소 투여량으로 최대 효과를 보이는 투여량을 결정하는 방법에 관한 연구이다.

제안한 연구방법은, 첫 번째 단계의 자료를 가지고 있다는 전제하에 중간분석 과정에서 미래의 자료에 대한 예측확률을 구하여 사전에 정의한 임상적 한계치와 비교를 통해서, 최소의 투여량이면서 최대의 효과를 보이는 투여량을 결정하는 방법에 대하여 다루었다. 본 논문은 이항자료를 갖는 경우에 대한 연구로서 사전분포와 사후분포 모두 베타 분포족을 따른다고 가정한다.연구에서 사용된 모형은, 투여량에 따른 평균 효과가 단조 감소하는 패턴을 가정하고 또한 베이지안 준-모수 모형을 가정하였다. 연구 정보를 설명하기 위한 모형을 선택하는데 있어서, 본 연구는 하나의 모델을 사용하지 않고 사용 가능한 모든 모델을 평균화시키는 방법인 베이지안 모형 평균화 방법을 사용하였다. 이 베이지안 모형 평균화 방법의 사용은 하나의 모델만 사용함으로 인한 불확실성을 제거하기 위한 방법으로서 Raffery 와 Volinsky(1999)에 의해서 사용되었다.

특히, 본 연구에서는 다섯 개 투여량의 증가에 따른 평균 효과가 단조 감소의 가정하에 베이지안 준-모수 모형을 설명하기 위해서, stick-breaking construction 방법을 적용하여 제안한 모형을 손쉽게 설명하는데 활용하였다.

이 stick-breaking construction 방법은 베이지안 접근방법에서 준-모수 모형에 대한 문제를 해결하는 한 가지 방법으로 소개되고 있다.

그러나 베이지안 접근 방법으로 예측확률에 대한 분포의 설명은 많은 어려움이 따르므로, 가중부여에 의한 재표집방법을 이용한 시뮬레이션 연구를 통해서, 베이지안 모형 평균화 방법에 의해 최소의 투여량으로 최대 효과를 보이는 투여량이 어떤 것인지를 찾는 방법을 사용하였다. 뿐만 아니라 전체 모델을 평균화시키지 않고 각 모델별로 결과를 산출하여 베이지안 모델 평균화 방법에 의해 얻은 결과와 비교를 통해서, 베이지안 모델 평균화 방법의 결과가 그렇지 않은 결과보다 더 정확한 결과를 제공해 주고 있음을 확인하였다. 특히, 베이지안 모형 평균화 방법에 의해 최소 투여량으로 최대 효과를 보이는 투여량을 찾는 결과에서 과소 추정하는 경향을 보임을 확인하였다. 반면에 베이지안 모형 평균화 방법을 사용하지 않은 경우에는, 어떤 모델이 사용되는지에 따라 과소 추정의 경향성이 다르게 나타남을 확인할 수 있었다.

이항자료에 대한 실제 2 상 임상시험 결과를 확보하지 못하여 실제 사례로는 제안한 방법을 적용하지 못했다는 점이 제한점이 된다. 하지만, 제안한 방법이 베이지안 모형 평균화 방법에서 매우 안정적으로 최소 투여량에 따른 최대 효과를 보이는 투여량에 대한 정보를 제공해 준다는 점은 향 후 본 연구를 표본 수 산출에 관한 연구로 확장을 고려해 볼 만하다 할 수 있겠다.

핵심 되는 말 : 베이지안 이 단계 설계, 최대 효과 투여량, 베이지안 모형 평균화, 예측 확률