Published in final edited form as:

Stat Med. 2015 January 30; 34(2): 249–264. doi:10.1002/sim.6339.

Bayesian probability of success for clinical trials using historical data

Joseph G. Ibrahim^{a,*†}, Ming-Hui Chen^b, Mani Lakshminarayanan^c, Guanghan F. Liu^d, and Joseph F. Heyse^d

^aDepartment of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, U.S.A.

^bDepartment of Statistics, University of Connecticut, Storrs, CT 06269, U.S.A.

^cBiotechnology Clinical Development, Pfizer, Inc., Collegeville, PA 19426, U.S.A.

^dMerck & Co., Inc., Whitehouse Station, NJ 08889, U.S.A.

Abstract

Developing sophisticated statistical methods for go/no-go decisions is crucial for clinical trials, as planning phase III or phase IV trials is costly and time consuming. In this paper, we develop a novel Bayesian methodology for determining the probability of success of a treatment regimen on the basis of the current data of a given trial. We introduce a new criterion for calculating the probability of success that allows for inclusion of covariates as well as allowing for historical data based on the treatment regimen, and patient characteristics. A new class of prior distributions and covariate distributions is developed to achieve this goal. The methodology is quite general and can be used with univariate or multivariate continuous or discrete data, and it generalizes Chuang-Stein's work. This methodology will be invaluable for informing the scientist on the likelihood of success of the compound, while including the information of covariates for patient characteristics in the trial population for planning future pre-market or post-market trials.

Keywords

average probability of success; fitting prior; linear regression; marginal likelihood; power prior; simulation; validation prior

1. Introduction

Conduct and completion of a successful clinical trial is an expensive and time-consuming initiative. It requires thoughtful planning and execution on the basis of study conduct details that need to be pre-planned and stated in the protocol. Owing to the potential lack of efficacy and/or safety concerns as well as concerns in cost, it has become quite common in clinical trials practice to formulate go/no-go decisions on the basis of the data of a phase II trial to

*Correspondence to: Joseph G. Ibrahim, Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, U.S.A..

†ibrahim@bios.unc.edu

try and quantify the benefits of moving to the phase III setting. In an environment where more than 40% of phase III clinical trials are failing, researchers are placing a greater level of emphasis on developing statistical methods to ensure a go/no-go decision is efficient, informative, and reliable. Traditionally, the steps involved in planning and making decisions in clinical trials are based on classical hypothesis testing that controls type I and type II errors through evaluating the statistical significance using pre-specified criteria for concluding efficacy or futility of the treatment. In case of traditional group sequential designs, the control is accomplished by the use of α -spending functions. Under the classical frequentist setting, success of a trial is determined once a significant p -value (say, p -value < 0.05) is achieved. In such cases, it is common to determine the sample size for the trial by assuming a fixed significance level (say 5%) and an effect size that will provide a required power for the trial. In this traditional sample size calculation, the effect size and its variance are estimated from historical data. But the traditional approach does not lend itself to incorporating variations in those estimates, and as a result, it is a common practice to consider a conservative sample size on the basis of several combinations of these estimates. On the basis of this standard calculation for a sample size, it is worthy to note that the power is a conditional value that depends on a desired effect size assumed to be clinically meaningful, and it is not the true probability of success (POS) for that particular trial. If the research team is interested in determining the POS for the trial, it needs to be determined as an unconditional value that is true across a range of possible values of effect size.

O'Hagan *et al.* [1] and Chuang-Stein [2] proposed that a POS can be thought of as a weighted average of power across the range of possible values of effect sizes. In other words, a POS can be calculated as an *expected value of power*. O'Hagan *et al.* [1] defined this as an *assurance* and calculates it the same as an unconditional probability that the trial will end up with a positive outcome. When translated to a simple situation, it actually becomes an expected power with respect to a given prior for the parameter and has no relationship to the frequentist power. One of the earliest references on expected power is that of Spiedelhalter *et al.* [3]. Chuang-Stein [2] defined the average POS as

$$\int P(\text{Trial produces a significant } p\text{-value} | \Delta) p(\Delta | d) d\Delta, \quad (1.1)$$

where Δ is the treatment effect, d is the observed treatment effect in the prior trial, and $p(\Delta | d)$ is the posterior distribution of Δ given d . In (1.1), the 'average probability of success' is obtained by weighting the success probability at various Δ by the likelihood of Δ under the posterior distribution of Δ given d . Stallard *et al.* [4] proposed a similar approach by calculating the predictive probability of a significant result in a phase III trial given the phase II data. The approach proposed by Stallard *et al.* [4] can also be used to calculate the posterior probability that a frequentist analysis of a subsequent phase III trial will yield a significant result at the interim and final analyses of the phase II trial. Within the Bayesian framework, Dmitrienko and Wang [5] proposed a Bayesian predictive approach for monitoring clinical trial data, whereas Eaton *et al.* [6] examined the limiting behavior of the 'conditional average power'.

For binary data, Chen *et al.* [7] used a combination of point estimates, frequentist p -values, and a posterior probability of the effect exceeding a given value to be the ‘go/no go’ criteria to move from a single-arm phase II oncology trial to the phase III setting. In a recent paper, Wang *et al.* [8] used the average conditional power concept to calculate the probability of study success (PrSS) as a two-step process: (i) synthesize historical data using Bayesian modeling to derive the prior distribution of the treatment effect and (ii) evaluate the PrSS analytically or via trial simulation. They used the same expression as in (1.1) that was originally provided by O’Hagan *et al.* [1]. They also illustrated the calculation of PrSS in three different examples involving survival data. One of the limitations in all the approaches that have been suggested so far for computing a POS is the absence of covariates and their contributions in determining the effect size and eventually the POS. Also, most of these aforementioned approaches promote a hybrid approach, that is, a combination of Bayesian and frequentist methods.

In this paper, building on the previous work of Dmitrienko and Wang [5] and Eaton *et al.* [6], we propose a fully Bayesian approach to POS in superiority trials that effectively yield the methods of O’Hagan *et al.* [1] and Chuang-Stein [2] as special cases. We first formally introduce the notion of ‘fitting prior’ and ‘validation prior’ in the POS calculation. Unlike in [5, 6], we allow for the inclusion of covariates as well as historical data in the POS calculation in our development. With our new approach, it is shown in Section 2 that the hybrid approach that uses both frequentist and Bayesian calculations is a special case of our proposed procedure when an improper fitting prior [9] is specified for the treatment effect. Our proposed method is quite general and computationally efficient. Our approach yields Bayesian quantities that are easy to interpret and thus does not rely on p -values or hybrid approaches. Moreover, our proposed approach naturally involves the use of historical data to carry out the POS calculation. The historical data are incorporated through a power prior [10] to allow for several historical datasets. Closed-form expressions are obtained for POS in the normal linear regression model, and we provide a methodology for modeling the covariate distributions as well as develop an efficient Gibbs algorithm for carrying out the calculations. Our development is completely general and yields all of the previous work in this area as a special case.

The rest of this paper is organized as follows. Section 2 defines the general notions of Bayesian POS, obtains closed-form expressions for normal data, makes connections to the work of Chuang-Stein [2], and obtains formulas for POS when incorporating several historical datasets via the power prior. Section 3 examines general Bayesian POS methodology for the normal linear regression model using the power prior along with several historical datasets. Closed-form expressions for POS are derived, and covariate modeling is also discussed in this section. Section 3 further presents a simulation algorithm for carrying out the general POS calculation for the linear regression model with unknown variances. In Section 4, we present two real data examples to demonstrate our proposed methodology. We close the article with some discussion in Section 5.

2. Bayesian probability of success for the normal model

Consider a simple one-arm trial with treatment effect Δ . Let TV denote the target value. Consider the hypotheses for a superiority trial: $H_0: \Delta \leq TV$ versus $H_1: \Delta > TV$, where the trial is considered successful if H_1 is accepted. According to (1.1) [2], the POS can be formulated more generally as

$$POS = \int P(\text{success}|\Delta) p(\Delta|\text{prior data}) d\Delta. \quad (2.1)$$

If $P(\text{success}|\Delta) = P(\text{Trial produces a significant } p\text{-value}|\Delta)$, (2.1) reduces to (1.1). Therefore, the POS defined in (1.1) is a hybrid approach by combining Bayesian and frequentist concepts. A similar hybrid Bayesian and frequentist approach was also considered by Stallard *et al.* [4]. In this paper, we propose a new and very general Bayesian approach for computing POS without using the frequentist concept.

Let $\mathbf{y} = (y_1, \dots, y_n)$ denote the future data of size n . Assume that $y_i \stackrel{i.i.d.}{\sim} N(\Delta, \sigma^2)$, where σ^2 is assumed to be known. Let $\pi^{(f)}(\Delta)$ denote a fitting prior [9, 11], which combines the likelihood function based on the future data $\mathbf{y}$ to construct the posterior distribution for computing the probability of a successful trial. Also, let $\pi^{(v)}(\Delta)$ denote a validation prior, which is a more general version of $p(\Delta|\text{prior data})$ in (2.1) for weighting the success

probability at various Δ . We assume $\pi^{(f)}(\Delta) = \frac{1}{\sqrt{2\pi}\tau^{(f)}} \exp\left\{-\frac{1}{2(\tau^{(f)})^2}(\Delta - \Delta^{(f)})^2\right\}$ and $\pi^{(v)}(\Delta) = \frac{1}{\sqrt{2\pi}\tau^{(v)}} \exp\left\{-\frac{1}{2(\tau^{(v)})^2}(\Delta - \Delta^{(v)})^2\right\}$, where $\Delta^{(f)}$, $\tau^{(f)}$, $\Delta^{(v)}$, and $\tau^{(v)}$ are hyperparameters.

In the Bayesian framework, given $\pi^{(f)}(\Delta)$ and $\mathbf{y}$, we declare that the trial is successful if

$$P(\Delta > TV | \mathbf{y}, \pi^{(f)}) \geq \gamma, \quad (2.2)$$

where the probability is taken with respect to the posterior distribution of $\pi(\Delta | \mathbf{y}, \pi^{(f)})$, and $\gamma > 0$ is a pre-specified credible level. A common choice of γ is 0.95. Let

$w = (\tau^{(f)})^2 / \left\{ \frac{\sigma^2}{n} + (\tau^{(f)})^2 \right\}$. Then, we have $\Delta | \mathbf{y}, \pi^{(f)} \sim N(w\bar{y} + (1-w)\Delta^{(f)}, w\frac{\sigma^2}{n})$, where $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$. It is easy to show that (2.2) is equivalent to

$$w\bar{y} + (1-w)\Delta^{(f)} \geq TV - \left(\frac{w\sigma^2}{n} \right)^{1/2} z_{1-\gamma}, \quad (2.3)$$

where $z_{1-\gamma}$ is the $(100(1-\gamma))^{th}$ percentile of the $N(0, 1)$ distribution. Using (2.3), we have

$P(\text{Success}|\Delta) = 1 - \Phi\left\{\left([TV - \left(\frac{w\sigma^2}{n}\right)^{1/2} z_{1-\gamma} - (1-w)\Delta^{(f)}]/w - \Delta\right) / \sqrt{\frac{\sigma^2}{n}}\right\}$, where $\Phi(\cdot)$ is the $N(0, 1)$ cumulative distribution function.

Consider the Z-test statistic $Z = \frac{\bar{Y} - TV}{\sigma / \sqrt{n}}$, and $1 - \gamma$ is assumed to be a significance level. Then, following [2], given the future sample mean $\bar{Y}$, the ‘trial produces a significant p -value’ under the one-sided alternative hypothesis can be formulated as

$$p\text{-value} = P\left(Z > \frac{\bar{Y} - TV}{\sigma / \sqrt{n}} \mid \Delta = TV\right) \leq 1 - \gamma, \quad (2.4)$$

where $Z \sim N(0, 1)$. Similar to (2.2) and (2.3), (2.4) is equivalent to

$$\bar{y} \geq TV - \left(\frac{\sigma^2}{n}\right)^{1/2} z_{1-\gamma}. \quad (2.5)$$

It is easy to see that (2.5) is identical to (2.3) when $\pi^{(f)}(\cdot) \propto 1$ (i.e., $(\tau^{(f)})^2 \rightarrow \infty$, which gives $w = 1$). Thus, our proposed Bayesian approach includes the hybrid Bayesian and frequentist approach as a special case when an improper uniform fitting prior $\pi^{(f)}(\cdot)$ is specified for $\cdot$. We also note that Dmitrienko and Wang [5] and Eaton *et al.* [6] considered a similar Bayesian formulation as (2.2) without including the information of covariates. Dmitrienko and Wang [5] computed the predictive probability of a successful outcome at the planned end of the trial using the interim data to construct the fitting prior. Eaton *et al.* [6] calculated the probability of a successful trial as $E[1\{P(\bar{y} > TV \mid \mathbf{y}, \pi^{(f)}) > \gamma\}]$, where the indicator function $1\{A\} = 1$ if A is true and 0 otherwise, and the expectation is taken with respect to the marginal distribution of $\mathbf{y}$.

Single historical dataset

Let $\mathbf{y}_0 = (y_{01}, \dots, y_{0n_0})$ denote the historical data and assume that $y_{0i} \stackrel{i.i.d.}{\sim} N(\Delta, \sigma^2)$. We further assume an improper uniform initial validation prior for $\cdot$. Using $\mathbf{y}_0$, we have

$\Delta^{(v)} = \bar{y}_0 = \frac{1}{n_0} \sum_{i=1}^{n_0} y_{0i}$ and $(\tau^{(v)})^2 = \frac{\sigma^2}{n_0}$. Then, the predictive distribution of $\bar{y}$ given $\cdot \sim N(\Delta^{(v)}, (\tau^{(v)})^2)$ can be written as

$$\bar{y} \sim N\left(\Delta^{(v)}, \frac{\sigma^2}{n} + (\tau^{(v)})^2\right). \quad (2.6)$$

Using (2.3) and (2.6), the POS in (2.1) can be calculated as

$$\text{POS} = \Phi\left\{\frac{(1-w)\Delta^{(f)} + w\Delta^{(v)} + \left(w\frac{\sigma^2}{n}\right)^{1/2} z_{1-\gamma} - TV}{w\sqrt{\frac{\sigma^2}{n} + (\tau^{(v)})^2}}\right\}. \quad (2.7)$$

Two historical datasets—Suppose that we have two historical datasets $\mathbf{y}_{01} = (y_{011}, \dots, y_{01n_{01}})'$ and $\mathbf{y}_{02} = (y_{021}, \dots, y_{02n_{02}})'$. Again, we assume that $y_{0ki} \stackrel{i.i.d.}{\sim} N(\Delta, \sigma^2)$ for $i = 1, \dots, n_{0k}$ and $k = 1, 2$. We use $\mathbf{y}_{01}$ to construct the fitting prior, whereas $\mathbf{y}_{02}$ is used to construct the validation prior. We assume improper uniform initial fitting and validation priors for $\cdot$.

Then, using y_{01} , we obtain $\Delta^{(f)} = \bar{y}_{0f} = \frac{1}{n_{01}} \sum_{i=1}^{n_{01}} y_{01i}$ and $(\tau^{(f)})^2 = \frac{\sigma^2}{n_{01}}$, and using y_{02} , we have $\Delta^{(v)} = \bar{y}_{0v} = \frac{1}{n_{02}} \sum_{i=1}^{n_{02}} y_{02i}$ and $(\tau^{(v)})^2 = \frac{\sigma^2}{n_{02}}$. Then, using the above $\Delta^{(f)}$ and $\tau^{(f)}$, POS in (2.1) can be calculated using (2.7). When two historical datasets are available, from a practical point of view, it is reasonable to use the one obtained more recently (y_{02}) for constructing the validation prior, whereas the other dataset that is obtained earlier is used as y_{01} for constructing the fitting prior.

Remark 1: We first note that our proposed POS procedure is a generalization of the hybrid procedure of Chuang-Stein [2]. In our setting, the fitting prior is used to construct the ‘posterior distribution’ based on the future data $\mathbf{y}$ for computing $P(\cdot > TV|\mathbf{y}, \pi^{(f)})$ in (2.2), and this prior is intended to represent the prior distribution of the parameters based on the historical data. On the other hand, the validation prior represents the prior distribution based on the ‘current study’, that is, the prior distribution based on the most recently completed study for which we wish to compute the likelihoods of the treatment effect at various values in calculating the POS. In the context of the hybrid Bayesian and frequentist approach of Chuang-Stein [2], the validation prior is precisely the posterior distribution of the treatment effect from the most recently completed study. Thus, the fitting prior and validation prior in our context play very different roles and have different interpretation, one prior is based on historical data, and the other prior is based on the current study, for which the POS is heavily based. The procedure of Chuang-Stein [2] is thus a very special case of our framework. The initial prior is a noninformative prior that represents the prior distribution of the parameters before observing the data. Thus, the initial fitting prior is the prior distribution of the parameters before the historical data are observed, and the initial validation prior is the prior distribution of the parameters before the ‘current data’ are observed. For more on fitting priors constructed from historical data and initial priors, we refer the reader to [9].

3. Bayesian probability of success for the linear regression models

3.1. The general methodology

Let z_i denote the treatment indicator such that $z_i = 1$ if the i^{th} patient receives treatment and $z_i = 0$ if the i^{th} patient receives control, and also, let $\mathbf{x}_i$ denote a p -dimensional vector of covariates for $i = 1, \dots, n$. We assume a randomized trial and let $D = \{(y_i, z_i, \mathbf{x}_i), i = 1, \dots, n\}$ denote the future data, where z_i is determined by randomization and $\mathbf{x}_i$ is unknown but independent of z_i . We assume a linear regression model for y_i , which is given by $y_i = \beta_0 + \beta_1 z_i + \mathbf{x}_i' \beta_2 + \varepsilon_i$, where ε_i 's are independent, $\varepsilon_i \sim N(0, \sigma^2)$, and β_2 is a p -dimensional vector of regression coefficients corresponding to $\mathbf{x}_i$. We further assume that the $\mathbf{x}_i$'s are independent and follow the distribution $\mathbf{x}_i \sim f(\mathbf{x}_i | \boldsymbol{\alpha})$, where $\boldsymbol{\alpha}$ is a vector of unknown parameters, for $i = 1, \dots, n$. Write $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2')$, $\boldsymbol{\theta} = (\boldsymbol{\beta}, \sigma^2)$, and $\mathbf{x} = (\mathbf{x}_1', \dots, \mathbf{x}_n')'$. Then, the distribution of $\mathbf{x}$, called the covariate distribution, is given by $f(\mathbf{x} | \boldsymbol{\alpha}) = \prod_{i=1}^n f(\mathbf{x}_i | \boldsymbol{\alpha})$.

Let TV denote the target value. Consider the hypotheses $H_0: \beta_1 \leq TV$ versus $H_1: \beta_1 > TV$, for which the trial is successful if H_1 is accepted. Then, the POS is defined as

$$POS = \int \int P(\text{success} | z, \mathbf{x}, \boldsymbol{\theta}) f(z) f(\mathbf{x} | \boldsymbol{\alpha}) \pi^{(v)}(\boldsymbol{\theta}, \boldsymbol{\alpha}) d\mathbf{x} dz d\boldsymbol{\theta} d\boldsymbol{\alpha}, \quad (3.1)$$

where $f(z)$ is a known distribution determined by the randomization scheme, $\pi^{(v)}(\boldsymbol{\theta}, \boldsymbol{\alpha})$ is a validation prior of $\boldsymbol{\theta}$ and $\boldsymbol{\alpha}$,

$$P(\text{success} | z, \mathbf{x}, \boldsymbol{\theta}) = E[1\{P(\beta_1 > TV | D, \pi^{(f)}) \geq \gamma\} | z, \mathbf{x}, \boldsymbol{\theta}], \quad (3.2)$$

$P(\beta_1 > TV | D, \pi^{(f)})$ is computed with respect to the posterior distribution of $\boldsymbol{\theta}$ given the data D and the fitting prior $\pi^{(f)}(\boldsymbol{\theta})$, and the expectation is taken with respect to the sampling distribution of D given $\mathbf{x}$ and $\boldsymbol{\theta}$. In (3.1), we assume independent validation priors for $\boldsymbol{\theta}$ and $\boldsymbol{\alpha}$, namely,

$$\pi^{(v)}(\boldsymbol{\theta}, \boldsymbol{\alpha}) = \pi_1^{(v)}(\boldsymbol{\theta}) \pi_2^{(v)}(\boldsymbol{\alpha}). \quad (3.3)$$

3.2. Single historical dataset

Suppose that we have historical data $D_0 = \{(y_{0i}, z_{0i}, \mathbf{x}_{0i}), i = 1, \dots, n_0\}$, where $z_{0i} = 1$ if treatment and $z_{0i} = 0$ if control and $\mathbf{x}_{0i}$ is a p -dimensional vector of covariates for $i = 1, \dots, n_0$. We assume that $y_{0i} = \beta_0 + \beta_1 z_{0i} + \mathbf{x}_{0i}' \boldsymbol{\beta}_2 + \varepsilon_{0i}$, where the ε_{0i} 's are independent and $\varepsilon_{0i} \sim N(0, \sigma^2)$ and $\mathbf{x}_{0i} \sim f(\mathbf{x}_{0i} | \boldsymbol{\alpha})$ for $i = 1, \dots, n_0$. We first use the historical data D_0 to specify the validation prior $\pi_1^{(v)}(\boldsymbol{\theta})$ in (3.3) as

$$\pi_1^{(v)}(\boldsymbol{\theta}) \propto \left[\prod_{i=1}^{n_0} \frac{1}{\sigma} \exp \left\{ -\frac{1}{2\sigma^2} (y_{0i} - \beta_0 - \beta_1 z_{0i} - \mathbf{x}_{0i}' \boldsymbol{\beta}_2)^2 \right\} \right] \pi_{10}^{(v)}(\boldsymbol{\theta}), \quad (3.4)$$

where $\pi_{10}^{(v)}(\boldsymbol{\theta})$ is an initial validation prior. We may specify an improper prior for $\pi_{10}^{(v)}(\boldsymbol{\theta})$ as

$$\pi_{10}^{(v)}(\boldsymbol{\theta}) \propto \frac{1}{\sigma^2}. \quad (3.5)$$

We then specify a validation prior $\pi_2^{(v)}(\boldsymbol{\alpha})$ in (3.3) based on the characteristics of the subjects in the population of the future trial. Specifically, we assume

$$\pi_2^{(v)}(\boldsymbol{\alpha}) = \pi_2^{(v)}(\boldsymbol{\alpha} | D_0, b_0) \propto \left[\prod_{i=1}^{n_0} f(\mathbf{x}_{0i} | \boldsymbol{\alpha}) \right]^{b_0} \pi_{20}^{(v)}(\boldsymbol{\alpha}), \quad (3.6)$$

where $0 < b_0 < 1$ and $\pi_{20}^{(v)}(\boldsymbol{\alpha})$ is an initial validation prior for $\boldsymbol{\alpha}$. A large value of b_0 with a noninformative initial prior $\pi_{20}^{(v)}(\boldsymbol{\alpha})$ (e.g., $\pi_{20}^{(v)}(\boldsymbol{\alpha}) \propto 1$) implies that the subjects in the future trial have similar characteristics as those in the historical data. On the other hand, when the subjects in the future trial are different than those in the historical data, we may take a small

value of b_0 and specify a more informative initial prior to reflect the characteristics of the subjects in the future trial. Thus, (3.6) provides us with a class of flexible validation priors.

As $P(\text{success}|\mathbf{x}, \boldsymbol{\theta})$ in (3.2) is independent of $\boldsymbol{\alpha}$, the fitting prior needs to be specified only for $\boldsymbol{\theta}$. Letting $\pi^{(f)}(\boldsymbol{\theta})$ be a fitting prior, the posterior distribution of $\boldsymbol{\theta}$ given D can be written as

$$\pi(\boldsymbol{\theta}|D, \pi^{(f)}) \propto \left[\prod_{i=1}^n \frac{1}{\sigma} \exp \left\{ -\frac{1}{2\sigma^2} (y_i - \beta_0 - \beta_1 z_i - \mathbf{x}'_i \boldsymbol{\beta}_2)^2 \right\} \right] \pi^{(f)}(\boldsymbol{\theta}). \quad (3.7)$$

We take a conjugate prior for $\boldsymbol{\theta} = (\boldsymbol{\beta}, \sigma^2)$ as

$$\pi^{(f)}(\boldsymbol{\theta}) \propto \frac{1}{(\sigma^2)^{\frac{p+2}{2}}} \exp \left\{ -\frac{1}{2\sigma^2} (\boldsymbol{\beta} - \boldsymbol{\beta}_0)' \Sigma_0^{-1} (\boldsymbol{\beta} - \boldsymbol{\beta}_0) \right\} \frac{1}{(\sigma^2)^{\frac{b_{0f1}}{2}+1}} \exp \left(-\frac{b_{0f2}}{2\sigma^2} \right), \quad (3.8)$$

where $\boldsymbol{\beta}_0$ is a $(p+2)$ -dimensional vector, Σ_0 is a $(p+2) \times (p+2)$ positive definite covariance matrix, $b_{0f1} \geq 0$, and $b_{0f2} \geq 0$, and all are pre-specified hyperparameters. Then, we have

$$P(\beta_1 > TV | D, \pi^{(f)}) = \int_{(\boldsymbol{\beta}, \sigma^2): \beta_1 > TV, \sigma^2 > 0} \pi(\boldsymbol{\beta}, \sigma^2 | D, \pi^{(f)}) d\boldsymbol{\beta} d\sigma^2. \quad (3.9)$$

Let $\mathbf{y} = (y_1, \dots, y_n)'$ and also let X be an $n \times (p+2)$ design matrix with i^{th} row equal to $(1, z_i, \mathbf{x}'_i)$. Assuming that X is of full rank, the least squares estimate of $\boldsymbol{\beta}$ based on D is given by $\hat{\boldsymbol{\beta}} = (X'X)^{-1}X'\mathbf{y}$. Using (3.8), (3.7) can be rewritten as

$$\pi(\boldsymbol{\theta}|D, \pi^{(f)}) \propto \frac{1}{(\sigma^2)^{\frac{p+2}{2}}} \exp \left\{ -\frac{1}{2\sigma^2} (\boldsymbol{\beta} - \tilde{\boldsymbol{\beta}})' \tilde{\Sigma}^{-1} (\boldsymbol{\beta} - \tilde{\boldsymbol{\beta}}) \right\} \frac{1}{(\sigma^2)^{\frac{b_{f1}}{2}+1}} \exp \left(-\frac{b_{f2}}{2\sigma^2} \right), \quad (3.10)$$

where $\tilde{\boldsymbol{\beta}} = (X'X + \Sigma_0^{-1})^{-1} (X'X\hat{\boldsymbol{\beta}} + \Sigma_0^{-1}\boldsymbol{\beta}_0)$, $\tilde{\Sigma} = (X'X + \Sigma_0^{-1})^{-1}$, $b_{f1} = n + b_{0f1}$, and $b_{f2} = (\mathbf{y} - X\hat{\boldsymbol{\beta}})'(\mathbf{y} - X\hat{\boldsymbol{\beta}}) + (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)' [X'X - X'X(X'X + \Sigma_0^{-1})^{-1}X'X] (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0) + b_{0f2}$. In (3.10), we write $\tilde{\boldsymbol{\beta}} = (\tilde{\beta}_0, \tilde{\beta}_1, \dots, \tilde{\beta}_{p+1})'$ and $\tilde{\Sigma} = (\Sigma_{jk}, j = 0, 1, \dots, p+1, k = 0, 1, \dots, p+1)$. Then, we are led to the following theorem.

Theorem 1—Given D and $\pi^{(f)}$, β_1 follows a t -distribution, that is, $\beta_1 \sim t_v(\mu(\mathbf{y}, X), \tau(\mathbf{y}, X))$, where the degrees of freedom $v = b_{f1}$, the location, and scale parameters are given by $\mu(\mathbf{y}, X) = \tilde{\beta}_1$ and $\tau(\mathbf{y}, X) = [\Sigma_{11}b_{f2}/v]^{1/2}$, and $\tilde{\beta}_1$, Σ_{11} , b_{f1} , and b_{f2} are defined in (3.10).

The proof of Theorem 1 follows from straightforward algebra, and thus, the details are omitted for brevity. Using the result in Theorem 1, (3.2) can be rewritten as

$$\begin{aligned}
P(\text{success}|D, \boldsymbol{\theta}) &= E[1\{P(\beta_1 > TV|D, \pi^{(f)}) \geq \gamma\}|X, \boldsymbol{\theta}] \\
&= E\left[1\left\{P\left(\frac{\beta_1 - \mu(\mathbf{y}, X)}{\tau(\mathbf{y}, X)} > \frac{TV - \mu(\mathbf{y}, X)}{\tau(\mathbf{y}, X)} \middle| D, \pi^{(f)}\right) \geq \gamma\right\} \middle| X, \boldsymbol{\theta}\right] \\
&= E[1\{TV \\
&\quad - \mu(\mathbf{y}, X) \leq t_{\nu, 1-\gamma}\tau(\mathbf{y}, X)\}|X, \boldsymbol{\theta}],
\end{aligned} \tag{3.11}$$

where $t_{\nu, 1-\gamma}$ is the $(100(1 - \gamma))^{th}$ percentile of the t_ν distribution. We can use the results established in Theorem 1 to compute (3.11) via a Monte Carlo (MC) method.

3.3. Two historical datasets

Suppose that we have two datasets $D_{0k} = \{(y_{0ki}, z_{0ki}, \mathbf{x}_{0ki}), i = 1, \dots, n_{0k}\}$, where $z_{0ki} = 1$ if treatment and $z_{0ki} = 0$ if control, and $\mathbf{x}_{0ki}$ is a p -dimensional vector of covariates for $i = 1, \dots, n_{0k}$, for $k = 1, 2$. We assume that

$$y_{0ki} = \beta_0 + \beta_1 z_{0ki} + \mathbf{x}_{0ki}' \boldsymbol{\alpha} + \varepsilon_{0ki}, \quad (3.12)$$

where the ε_{0ki} 's are independent and $\varepsilon_{0ki} \sim N(0, \sigma^2)$, and $\mathbf{x}_{0ki} \sim f(\mathbf{x}_{0ki}|\boldsymbol{\alpha})$ for $i = 1, \dots, n_{0k}$ and $k = 1, 2$. The POS is given by (3.1).

We use the historical data D_{01} via the power prior [10] to elicit the fitting prior. Specifically, we take $\pi^{(f)}(\boldsymbol{\theta})$ as

$$\pi_1^{(f)}(\boldsymbol{\beta}, \sigma^2 | D_{01}, a_0) \propto \left[\prod_{i=1}^{n_{01}} \frac{1}{\sigma} \exp \left\{ -\frac{1}{2\sigma^2} (y_{01i} - \beta_0 - \beta_1 z_{01i} - \mathbf{x}_{01i}' \boldsymbol{\beta}_2)^2 \right\} \right]^{a_0} \pi_{10}^{(f)}(\boldsymbol{\theta}), \quad (3.13)$$

where $0 < a_0 < 1$ is a pre-specified precision parameter and $\pi_{10}^{(f)}(\boldsymbol{\theta})$ is an initial fitting prior, which may be specified by (3.5) or (3.8). Let X_{01} denote the design matrix with i^{th} row equal to $(1, z_{01i}, \mathbf{x}_{01i}')$, and also let $\mathbf{y}_{01} = (y_{011}, \dots, y_{01n_{01}})'$ for the historical data D_{01} . Assuming that X_{01} is of full rank, the least square estimate of $\boldsymbol{\beta}$ based on D_{01} is given by

$\hat{\boldsymbol{\beta}}_{01} = (X_{01}' X_{01})^{-1} X_{01}' \mathbf{y}_{01}$. Suppose that $\pi_{10}^{(f)}(\boldsymbol{\theta})$ is specified by (3.8). Then, (3.13) can be rewritten as

$$\pi_1^{(f)}(\boldsymbol{\beta}, \sigma^2 | D_{01}, a_0) \propto \frac{1}{(\sigma^2)^{\frac{p+2}{2}}} \exp \left\{ -\frac{1}{2\sigma^2} (\boldsymbol{\beta} - \tilde{\boldsymbol{\beta}}_0)' \tilde{\Sigma}_0^{-1} (\boldsymbol{\beta} - \tilde{\boldsymbol{\beta}}_0) \right\} \frac{1}{(\sigma^2)^{\frac{\tilde{b}_{0f1}}{2} + 1}} \exp \left(-\frac{\tilde{b}_{0f2}}{2\sigma^2} \right), \quad (3.14)$$

where $\tilde{\boldsymbol{\beta}}_0 = (a_0 X_{01}' X_{01} + \Sigma_0^{-1})^{-1} (a_0 X_{01}' X_{01} \hat{\boldsymbol{\beta}}_{01} + \Sigma_0^{-1} \boldsymbol{\beta}_0)$, $\tilde{\Sigma}_0 = (a_0 X_{01}' X_{01} + \Sigma_0^{-1})^{-1}$, $\tilde{b}_{0f1} = a_0 n_{01} + b_{0f1}$, and

$\tilde{b}_{0f2} = a_0 (\mathbf{y}_{01} - X_{01} \hat{\boldsymbol{\beta}}_{01})' (\mathbf{y}_{01} - X_{01} \hat{\boldsymbol{\beta}}_{01}) + (\hat{\boldsymbol{\beta}}_{01} - \boldsymbol{\beta}_0)' [a_0 X_{01}' X_{01} - a_0^2 X_{01}' X_{01} (a_0 X_{01}' X_{01} + \Sigma_0^{-1})^{-1} X_{01}' X_{01}] (\hat{\boldsymbol{\beta}}_{01} - \boldsymbol{\beta}_0) + b_{0f2}$. Then, the result established in Theorem 1 can be extended to this case with $\beta_1 \sim t_\nu(\mu(\mathbf{y}, X), \tau(\mathbf{y}, X))$, where the degrees of freedom $\nu = n + b_{0f1}$, and $\mu(\mathbf{y}, X)$ and $\tau(\mathbf{y}, X)$.

X) are calculated via (3.10) with $(\beta_0, \Sigma_0, b_{0f1}, b_{0f2})$ replaced by $(\tilde{\beta}_0, \tilde{\Sigma}_0, \tilde{b}_{0f1}, \tilde{b}_{0f2})$. In (3.3), the validation prior $\pi_1^{(v)}(\theta)$ is specified as

$$\pi_1^{(v)}(\theta, D_{02}) \propto \left[\prod_{i=1}^{n_{02}} \frac{1}{\sigma} \exp \left\{ -\frac{1}{2\sigma^2} (y_{02i} - \beta_0 - \beta_1 z_{02i} - x'_{02i} \beta_2)^2 \right\} \right] \pi_{10}^{(v)}(\beta, \sigma^2), \quad (3.15)$$

where the initial validation prior $\pi_{10}^{(v)}(\beta, \sigma^2)$ is given by (3.5). As discussed in Remark 1, the validation prior represents the posterior distribution of the data from the most recently completed study in the context of the hybrid Bayesian and frequentist approach of Chuang-Stein [2], and thus, discounting a ‘posterior distribution’ is not appealing for POS purposes. To obtain an accurate and informative assessment of POS, it makes more sense to use the full data D_{02} in the POS calculation, that is, not to discount the validation prior. On the other hand, it does make sense to discount the fitting prior because this prior is really based on historical data D_{01} , and we may want to downweight the impact of D_{01} on the POS calculation. Thus, downweighting the fitting prior makes more sense in this case.

Again, we specify the validation prior $\pi_2^{(v)}(\alpha)$ in (3.3) based on the characteristics of the subjects in the population of the future trial. When the subjects in the future trial are similar to those in the two historical datasets, the validation prior $\pi_2^{(v)}(\alpha)$ may be specified as

$$\pi_2^{(v)}(\alpha | D_{01}, b_{01}, D_{02}, b_{02}) \propto \left[\prod_{i=1}^{n_{01}} f(x_{01i} | \alpha) \right]^{b_{01}} \left[\prod_{i=1}^{n_{02}} f(x_{02i} | \alpha) \right]^{b_{02}} \pi_{20}^{(v)}(\alpha), \quad (3.16)$$

where $\pi_{20}^{(v)}(\alpha)$ is an initial validation prior, $b_{01} + b_{02} > 0$, and $0 \leq b_{0j} \leq 1$ for $j = 1, 2$. In (3.16), when the subjects in the future trial are more similar to those in the historical D_{01} , we choose $b_{01} \gg b_{02}$, or if they are more similar to those in D_{01} , we choose $b_{01} \ll b_{02}$; otherwise, we may pick similar values for b_{01} and b_{02} . If the subjects in the future trial are different than those in either D_{01} or D_{02} , then we choose both b_{01} and b_{02} to be small and specify a more informative initial validation prior for α to reflect the characteristics of the subjects in the future trial.

To determine a guide value of a_0 in (3.14), we fit the linear model (3.12) for $k = 2$ using $\pi_1^{(f)}(\beta, \sigma^2 | D_{01}, a_0)$ as the prior for (β, σ^2) . Let X_{02} denote the design matrix with i^{th} row equal to $(1, z_{02i}, x'_{02i})$, and also let $y_{02} = (y_{021}, \dots, y_{02n_{02}})'$ for D_{02} . Assuming that X_{02} is of full rank, the least squares estimate of β based on D_{02} is given by $\hat{\beta}_{02} = (X'_{02} X_{02})^{-1} X'_{02} y_{02}$. Then, the marginal likelihood of y_{02} is given by

$$m(a_0) = |\tilde{\Sigma}_0|^{-\frac{1}{2}} (\tilde{b}_{0f2}/2)^{-\frac{\tilde{b}_{0f1}}{2}} [\Gamma(\tilde{b}_{0f1}/2)]^{-1} (2\pi)^{-\frac{n_{02}}{2}} |\tilde{\Sigma}_{0v}|^{\frac{1}{2}} (\tilde{b}_{0v2}/2)^{-\frac{\tilde{b}_{0v1}}{2}} \Gamma(\tilde{b}_{0v1}/2), \quad (3.17)$$

$$\begin{aligned}\tilde{b}_{0v1} &= n_{02} + \tilde{b}_{0f1}, \tilde{b}_{0v2} \\ &= (\mathbf{y}_{02} - X_{02}\hat{\boldsymbol{\beta}}_{02})'(\mathbf{y}_{02} \\ &\quad - X_{02}\hat{\boldsymbol{\beta}}_{02}) + (\hat{\boldsymbol{\beta}}_{02} - \tilde{\boldsymbol{\beta}}_0)' \left[X_{02}'X_{02} - X_{02}'X_{02}(X_{02}'X_{02} + \tilde{\Sigma}_0^{-1})^{-1}X_{02}'X_{02} \right] (\hat{\boldsymbol{\beta}}_{02} \\ &\quad - \tilde{\boldsymbol{\beta}}_0) + \tilde{b}_{0f2}, \tilde{\Sigma}_{0v}\end{aligned}$$

where $\tilde{\Sigma}_0 = (X_{02}'X_{02} + \tilde{\Sigma}_0^{-1})^{-1}$, and $\tilde{\boldsymbol{\beta}}_0$, $\tilde{\Sigma}_0$, b_{0f1} , and b_{0f2} are defined in (3.14). The guide value for choosing a_0 according to the marginal likelihood is given by

$$a_0^{\text{opt}} = \arg \min_{0 \leq a_0 \leq 1} [-2\log\{m(a_0)\}]. \quad (3.18)$$

Remark 2—The a_0 that maximizes this marginal likelihood corresponds to the a_0 value that yields the greatest compatibility between these two datasets, so it has a nice interpretation. As a_0 can be viewed as a compatibility parameter, the maximum of the marginal likelihood yields the a_0 value, which results in greatest compatibility.

3.4. The covariate distributions

One crucial issue in the proposed method is the specification of the covariate distribution $f(\mathbf{x}_i|\boldsymbol{\alpha})$ for $\mathbf{x}_i$. Lipsitz and Ibrahim [12], Ibrahim *et al.* [13], and Chen and Ibrahim [14] proposed a strategy for specifying the covariate distribution in order to reduce the number of nuisance parameters in the context of missing covariates. Specifically, the joint distribution of the covariates is modeled as a product of one-dimensional conditional distributions. This modeling strategy has the potential of greatly reducing the number of nuisance parameters that need to be specified. For example, with dichotomous covariates, an obvious choice for $f(\mathbf{x}_i|\boldsymbol{\alpha})$ is a joint multinomial log-linear model as described in [15]. Following Lipsitz and Ibrahim [12], one can reduce the number of nuisance parameters by deleting higher order interaction terms from the log-linear model. More generally, we can write the distribution of the p -dimensional covariate vector $\mathbf{x}_i = (x_{i1}, \dots, x_{ip})'$ through a series of one-dimensional conditional distributions, given by

$$f(x_{i1}, \dots, x_{ip}|\boldsymbol{\alpha}) = f(x_{i1}|\boldsymbol{\alpha}_1)f(x_{i2}|x_{i1}, \boldsymbol{\alpha}_2) \times \dots \times f(x_{ip}|x_{i1}, \dots, x_{i,p-1}, \boldsymbol{\alpha}_p), \quad (3.19)$$

where $\boldsymbol{\alpha}_j$ is a vector of indexing parameters for the j^{th} conditional distribution,

$\boldsymbol{\alpha} = (\boldsymbol{\alpha}'_1, \dots, \boldsymbol{\alpha}'_p)'$, and the $\boldsymbol{\alpha}_j$'s are distinct.

There are many possibilities in modeling the covariate distribution as a sequence of one-dimensional conditional distributions, especially when p is large. Thus, in these cases, we follow the guidance proposed in [14] to specify $f(\mathbf{x}_i|\boldsymbol{\alpha})$. If the covariates are all categorical, say dichotomous, then we specify a sequence of logistic regressions for each $f(x_{ij}|x_{i1}, \dots, x_{i,j-1}, \boldsymbol{\alpha}_j)$. We may also use the probit or complementary log-log links to model the categorical covariates. For categorical covariates with more than two levels, we consider a multinomial logistic regression model [15], and if the categorical covariates are ordinal with more than two levels, ordinal regression models such as the continuation odds model [16]

and the cumulative logit model [17] can be used for $f(x_{ij}|x_{i1}, \dots, x_{i,j-1}, \alpha_j)$. If the covariates consist of counts, then we can model them using the Poisson distribution.

For continuous covariates that take values on the real line, we model them via a sequence of univariate normal distributions. If x_{ij} is continuous and strictly positive, we can transform to logs, $x_{ij}^* = \log(x_{ij})$, and then specify a normal distribution for $f(x_{ij}^*|x_{i1}, \dots, x_{i,j-1}, \alpha_j)$, or if $\log(x_{ij})$ is not approximately normal, we specify an exponential (or gamma) model for $f(x_{ij}|x_{i1}, \dots, x_{i,j-1}, \alpha_j)$. With categorical and continuous covariates (i.e., mixed covariates), our usual convention is to specify the joint covariate distribution by specifying the distribution for the categorical covariates by conditioning on the continuous covariates first. See [14] for more detailed discussion.

3.5. Computational development

To compute POS in (3.1), we use the Gibbs sampling algorithm to generate (θ, α) from the validation prior $\pi^{(v)}(\theta, \alpha)$. We then use the results established in Theorem 1 to develop an easy-to-implement MC algorithm for computing the POS. The development of the Gibbs sampling algorithm to generate from $\pi^{(v)}(\theta, \alpha)$ is given in the Appendix. The detailed steps of the computational algorithm for POS in (3.1) are given as follows:

- Step 0: specify n , TV , γ , q (the proportion of subjects randomized to the test drug group), the hyper-parameters in the initial validation priors $\pi_{10}^{(v)}(\theta)$ and $\pi_2^{(v)}(\alpha)$ such as (b_{01}, b_{02}) in (3.16), M (the Gibbs sample size), and B (the number of simulations);
- Step 1: generate θ and α from the validation prior $\pi^{(v)}(\theta, \alpha)$;
- Step 2: using the value of α from step 1, generate x_i from $f(x_i|\alpha)$, also randomly assign the i^{th} subject into two treatment arms, that is, generate $z_i \sim \text{Bernoulli}(q)$ independently, for $i = 1, \dots, n$, and set X to be the resulting $n \times (p + 2)$ design matrix with i^{th} row $(1, z_i, x_i')$;
- Step 3: using the value of θ from step 1 and the value of X from step 2, generate a sample $\{y^{(m)}, m = 1, \dots, M\}$ from a $N(X\beta, \sigma^2 I_n)$ distribution;
- Step 4: compute $\mu(y^{(m)}, X)$ and $\tau(y^{(m)}, X)$ given in Theorem 1 for $m = 1, \dots, M$;
- Step 5: compute $\hat{P} = \frac{1}{M} \sum_{m=1}^M 1\{TV - \mu(y^{(m)}, X) \leq t_{\nu, 1-\gamma} \tau(y^{(m)}, X)\}$;
- Step 6: repeat steps 1–5 B times; and
- Step 7: compute the average of $\hat{P}$ in these B runs, which gives an MC estimate of POS.

4. Applications

4.1. Example 1: POS evaluation for acceptability test for a single group

In this example, we consider a POS evaluation for an acceptability test in a vaccine trial for the prevention of shingles (herpes zoster), a viral disease characterized by a painful skin rash with blisters and often in a stripe. The zoster vaccine is considered the most effective way to reduce the incidence of herpes zoster. The vaccine works through stimulating the vaccine

recipient to generate antibody. The higher the antibody titer, the better the vaccine's immune response is for protection. The considered trial is a randomized, double-blinded, multicenter, placebo-controlled clinical trial to evaluate the safety, tolerability, and immunogenicity of ZOSTAVAX™, a vaccine for the prevention of herpes zoster. Approximately 300 adults who received chronic/maintenance systemic corticosteroid therapy (2 weeks) were enrolled and randomized in a 2:1 ratio to receive vaccine or placebo. Enrollment was further stratified by age, with approximately 180, 90, and 30 patients in 60 to 69, 70 to 79, and 80 years of age, respectively. Antibody titers were evaluated from serum samples collected at baseline (pre-vaccination) and at about 42 days post-vaccination. The geometric mean fold rise (GMFR) from baseline to post-vaccination was one of the primary endpoints used to evaluate immunogenicity response. The statistical analysis was based on the natural log-transformed antibody titer. Therefore, the mean change from baseline in the natural log-transformed antibody titer will be the analysis endpoint for the following discussion. For the purpose of the POS evaluation, we will focus on the subgroup of patients who were 70 to 79 years of age. A total of 83 patients (57 in vaccine group and 26 in placebo group) had change from baseline values in this subgroup. The mean change from baseline (in nature-log scale) values and its corresponding standard errors were 0.7782 and 0.1085 for vaccine and 0.1303 and 0.1321 for placebo. The GMFRs and 95% confidence intervals (CIs) based on a t -distribution were 2.18 and (1.75, 2.71) for vaccine and 1.14 and (0.87, 1.50) for placebo.

One of the criteria for success is that the GMFR from baseline of the antibody titer after vaccination would be above a certain cutoff value ϕ . This is often called the acceptability criterion. The null hypothesis for acceptability is $H_0: \text{GMFR} < \phi$. The boundary ϕ is a pre-specified constant, which may be based on historical data and previous trials. For illustrative purposes in this example, we set ϕ between 1.4 (a lower bound of a 95% CI observed from a previous trial) and 2.0 (a high point estimate observed from another trial). In the subgroup of patients aged 70 to 79 years old, the lower bound of the calculated 95% CI is 1.75. So the acceptability test would be significant if ϕ is between 1.4 and 1.7. Under that margin, we can claim that the vaccine generated an acceptable immune response. In contrast, the placebo group had an GMFR estimate of 1.14, and the 95% CI contains 1, indicating that the mean antibody titers for subjects in the placebo group do not have a significant increase from baseline.

Under the Bayesian framework, the posterior distribution for the mean change from baseline is normal with mean = 0.7782 and standard deviation $0.1085 \sqrt{57}$ (with a noninformative prior). The posterior predictive probability of the average GMFR $> \phi$ for a future sample of size $n = 57$, that is, the predictive mean $(\bar{y}) > \log(\phi)$ of the future sample from a $N(0.7782, \sigma^2)$ distribution can be calculated. Suppose that we specify $\sigma = 0.1085 \sqrt{57}$, $\theta = 0$, and $(\tau^{(f)})^2 = 100,000$. Using the historical data, we have $n_0 = 57$, $\mu^{(v)} = 0.7782$, and $(\tau^{(v)})^2 = \sigma^2/n_0$. We set $n = 57$ and $\gamma = 0.95$. Using (2.7), we calculate the POS with $TV = \log(\phi)$. Table I presents the posterior probabilities and POSs calculated with different acceptability bounds. In this one-sample acceptability test problem, there are no covariates involved in the analysis. The results show that for each given acceptable margin ϕ , the posterior predictive probabilities equal the full Bayesian POS values when $\gamma = 0.5$. This implies that the average probability for $P(\text{GMFR} > \phi)$ above 50% using the completed trial data to construct a

validation prior is equal to the posterior predictive probability of $P(\text{GMFR} > \phi)$ in a future trial when the completed trial is used as a prior. The fully Bayesian POS value decreases as we increase the required credible level γ from 50% upwards. For example, if we want to have a credible level of 95%, a future trial with the same sample size will have a POS of 95.7% for an acceptability test margin of 1.4. This high POS demonstrated good evidence that trials will likely succeed with the given acceptability criterion. In fact, several published studies all met the similar acceptability criterion in different settings [18, 19].

4.2. Example 2: POS evaluation for a superiority test

In this example, we look at a superiority test comparing a test anti-depressant compound with placebo in two subsequently conducted trials MK0869-P008 and MK0869-P059. MK0869-P008 was a proof of concept (POC) phase II trial in which the test compound was studied with placebo and an active control (paroxetine 20 mg) in outpatients with major depressive disorder [20]. Patients were randomized equally to receive a test compound, active control, or placebo for 6 weeks. Efficacy measurements were made at baseline and at the end of weeks 1, 2, 4, and 6, or on termination. The test compound was given in a tablet form of 300 mg. Patients enrolled in the trial had baseline HAMD-17 ≥ 22 . The observed mean baseline HAMD-17 score from the trial was about 25.3. A total of four investigative sites in the USA were involved in the trial. MK-0869-P059 was one of the phase III trials conducted in the program [21]. The trial used a nano-formulation of 160 mg for the test compound, which was considered to be bioequivalent to the 300-mg tablet used in the phase II trial. In order to minimize the placebo response and therefore reduce the potential failed trials, the phase III trial also enrolled patients with relatively high baseline HAMD-17. The enrollment criterion required a baseline HAMD-17 ≥ 25 , corresponding to moderate to severe depression. The observed mean baseline HAMD-17 score from the trial was about 27.5, which was balanced among the treatment groups. A total of 27 investigative sites were involved in this phase III trial.

Table II summarizes the baseline demographics for the patients in these trials. Compared to the patients in the phase II trial (P008), there were more female patients enrolled in the phase III trial (P059). There were also slightly more White patients in P059 than in P008. The HAMD-17 change from baseline scores in the placebo group in P059 was larger than that of P008 (−11.9 vs. −8.3).

In this example, we consider four covariates, namely, age, gender (female versus male), race (White or Black versus other), and baseline HAMD-17. Let x_{i1} denote age, $x_{i2} = 1$ if male and 0 if female, $x_{i3} = 1$ if White and 0 otherwise, and $x_{i4} = 1$ if Black and 0 otherwise; and finally, let x_{i5} denote the baseline HAMD-17. We note that in our notation, (x_{i3}, x_{i4}) takes only three possible values, namely, (1, 0) (White), (0, 1) (Black), and (0, 0) (others).

Following (3.19), we assume (i) $f(x_{i1}|\alpha_1) \propto \alpha_{12}^{-1/2} \exp \left\{ -\frac{1}{2\alpha_{12}}(x_{i1} - \alpha_{11})^2 \right\}$, where $\alpha_{12} > 0$ and $\alpha_1 = (\alpha_{11}, \alpha_{12})'$ (a normal model); (ii) $f(x_{i2}|\alpha_2, x_{i1}) = \frac{\exp\{x_{i2}(\alpha_{21} + \alpha_{22}x_{i1})\}}{1 + \exp(\alpha_{21} + \alpha_{22}x_{i1})}$, where $\alpha_2 = (\alpha_{21}, \alpha_{22})'$ (a logistic regression model); (iii) $f(x_{i3}, x_{i4}|\alpha_3, x_{i1}, x_{i2}) = \exp \{x_{i3}(\alpha_{31} + \alpha_{32}x_{i1} + \alpha_{33}x_{i2}) + x_{i4}(\alpha_{34} + \alpha_{35}x_{i1} + \alpha_{36}x_{i2})\} [1 + \exp(\alpha_{31} + \alpha_{32}x_{i1} + \alpha_{33}x_{i2}) + \exp(\alpha_{34} +$

$a_{35}x_{i1} + a_{36}x_{i2})]^{-1}$, where $\alpha_3 = (\alpha_{31}, \dots, \alpha_{36})'$ (a multivariate logistic regression model); and (iv)

$$f(x_{i5}|\alpha_4, x_{i1}, x_{i2}, x_{i3}, x_{i4}) \propto \alpha_{46}^{-1/2} \exp \left[-\frac{1}{2\alpha_{46}} \{x_{i5} - (\alpha_{41} + \alpha_{42}x_{i1} + \alpha_{43}x_{i2} + \alpha_{44}x_{i3} + \alpha_{45}x_{i4})\}^2 \right]$$

, where $\alpha_4 = (\alpha_{41}, \dots, \alpha_{46})'$ (a normal regression model). In all calculations, we specified noninformative initial priors for all model parameters. Specifically, we set $\beta_0 = \mathbf{0}$, $\Sigma_0 = 1000I_{p+2}$, where I_{p+2} is the $(p+2) \times (p+2)$ identity matrix, and $b_{0f1} = b_{0f2} = 0.001$ in (3.8). In addition, a $N(0, 1000)$ (or inverse gamma $IG(0.001, 0.001)$) initial prior was specified for each location (scale) parameter in α independently. The proportion of subjects randomized to the test drug group (q) was chosen to be 0.5. The response variable is minus the change from baseline in HAMD-17 score. The datasets from the MK0869-P008 and MK0869-P059 trials are denoted by D_{01} and D_{02} , respectively. Figure 1 shows the plots of $-2 \log m(a_0)$ in (3.17) versus a_0 for the models with or without covariates for paroxetine and MK0869, respectively. It is interesting to see from this figure that $-2 \log m(a_0)$ is convex in a_0 for each of these four plots. In Figure 1, the values of a_0^{opt} in (3.18) and the corresponding values of $-2 \log m(a_0)$ are 0.979 and 2119.45 for plot (a); 0.142 and 2130.41 for plot (b); 0.443 and 2147.05 for plot (c); and 0.208 and 2143.62 for plot (d), respectively. These values indicate that for the active control, paroxetine, we could borrow about 97.9% and 14.2% of the historical data for the models with and without covariates, respectively; and for MK0869, the amounts of the historical data that could be borrowed are 44.3% and 20.8% for the models with and without covariates, respectively, in the POS calculation.

The plots in Figure 1 indicate that the guide value a_0^{opt} may depend on the covariates of the patients in the trials. When the covariate information is incorporated in the analysis, a larger value of a_0 is more likely to maximize the marginal likelihood of the outcome. Clinically, this may imply that the similarity of the patient characteristics between the historical trial and the current trial can be strengthened through the covariate information. Incorporating these covariates in the analysis allows us to borrow more from the historical data to construct the power prior. The estimated guide values of a_0^{opt} are also different for the MK0869 and paroxetine groups. This may reflect the dependence of the power prior parameter on the similarity of the outcomes between the historical trial (P008) and the current trial (P059). As indicated in Table II, the treatment effects are about 3.4 and 2.0, respectively, for paroxetine, and about 4.2 and 0.4, respectively, for MK0869 in the two trials in terms of change from baseline HAMD-17 total scores. Tables III and IV show the POS values for a future trial with $n = 300$ for paroxetine and MK0869, respectively, under different values of TV , power prior parameter a_0 , credible level γ , and whether or not the covariates are incorporated. The POS values versus a_0 for $\gamma = 0.90$, $TV = 0.5$, and $TV = 1.0$ are also plotted in Figure 2. The highlighted values in Tables III and IV and the two vertical lines shown in each plot of Figure 2 correspond to the values of a_0^{opt} , which maximize the corresponding marginal likelihoods. As expected, the POS values decrease as TV (or γ) increases. Comparing the highlighted values across the columns, we can see that without incorporating the covariates, the POS values are lower than those with incorporating the covariates, which is mainly due to using a smaller power prior parameter value for a_0 . The covariates between two historical trials (P08 and P059) may have a similar impact on the

POS values. When the future trial population has a similar covariate distribution as in P008 (i.e., $b_{01} = 1$ and $b_{02} = 0$), or as in P059 (i.e., $b_{01} = 0$ and $b_{02} = 1$), or as in both trials, the POS values are generally similar.

Depending on the values of TV and the credible level, the estimated POS values for paroxetine ranged from 35% to 88%, which represents a reasonably good chance of success after incorporating the uncertainty of all other covariates and nuisance parameters. For the MK0869 group, however, the estimated POS values were around 4% to 33%, which clearly shows a fairly low chance of success in a future trial. In fact, the results from the other four phase III trials [21] in the clinical development program showed that the treatment effects for paroxetine against placebo were significantly different in all three trials in which the paroxetine was included. None of the four phase III trials showed a significant treatment effect for MK0869. The retrospective POS calculations provide a very good indication for the POS for the future trials in this clinical development program.

5. Discussion

In this paper, we have developed a fully Bayesian and general framework for POS that (i) allows for the inclusion of historical data via the power prior; (ii) allows for the inclusion of covariates through a covariate distribution; and (iii) generalizes the hybrid frequentist Bayesian-frequentist development of Chuang-Stein. For the linear model, one obtains closed-form elegant expressions for POS. These calculations are very useful in setting for which the scientist needs to make a crucial decision on the likelihood of moving to a phase III trial based on the results of a phase II along with historical data. In the two real data examples of Section 4, we illustrated several useful features of the POS calculation and showed the benefits of the inclusion of covariates in the POS calculation along with estimated values of a_0 obtained from the marginal likelihood for several ranges of target values and priors and treatments. Our POS results were consistent with published phase III trials [21], which showed no significant treatment. These calculations are useful in clinical settings where a trial team or a research review committee needs to make a crucial decision on the likelihood of moving to the second stage within a trial, or to a phase III trial based on the results of a phase II along with historical data or moving within an entire portfolio. Decisions needed to be made within trial may include moving to next phase after interim data, moving from primary to secondary endpoints, or even moving to safety after confirming efficacy. POS calculations presented in this paper can easily be extended to any of these scenarios. In Example 2 of Section 4, four covariates were considered based on an *a priori* assumption regarding their significance as stated in the protocol. The POS calculations could also be used as a basis for sensitivity analysis to study the impact of these covariates in facilitating decision making in moving forward. The general framework and related ease in incorporating covariates as part of the calculations should make the POS calculations to be routine and unavoidable in any submissions to internal review committees within a pharmaceutical industry.

More often in the current environment, POS calculations are carried out in POC studies, that is, early phase IIA trials. As there might not be vast amount of clinical trial data (except for historical data from literature) before a POC trial, it is common in that setting to assume a

noninformative prior for the validation prior for the POS calculation. In addition, it is also common to use the data from this POC study to update future validation priors for further POS calculations. In this paper, we have primarily focused on POS calculations when there already exists real data from the most recently completed trial that can then be used in the construction of the validation prior. Such a setting is common when one wishes to move to a phase III setting for which phase II or even phase III data may be available from completed trials. It is clear that our methodology for POS can be applied to a variety of setting for which we may have noninformative or informative validation and/or fitting priors.

Although our focus on this paper has been on POS calculations for superiority trials, the framework developed here can also be easily applied to other settings such as non-inferiority or equivalence trials. In non-inferiority or equivalence trials, the hypotheses are set up differently, but the POS calculations would proceed in a very similar fashion. In addition, the basic ideas of POS developed in this paper can certainly be extended to other settings such as generalized linear models and semi-parametric survival models, along with multivariate models that can entertain a multivariate endpoint. These developments are currently under investigation.

Appendix A

The Gibbs sampling algorithm and full conditional distributions

We use the Gibbs sampling algorithm to sample from the validation prior $\pi^{(v)}(\boldsymbol{\theta}, \boldsymbol{\alpha})$ in (3.3). We present the development of the Gibbs sampling algorithm only for a single historical dataset $D_0 = \{(y_{0i}, z_{0i}, \mathbf{x}_{0i}), i = 1, 2, \dots, n_0\}$ because the computational development for other cases is similar. Using (3.4), (3.6), and (3.5), the validation prior for $(\boldsymbol{\theta}, \boldsymbol{\alpha})$ takes the form

$$\pi_1^{(\nu)}(\boldsymbol{\theta}, \boldsymbol{\alpha}) \propto \frac{1}{\sigma^2} \left[\prod_{i=1}^{n_0} \frac{1}{\sigma} \exp \left\{ -\frac{1}{2\sigma^2} (y_{0i} - \beta_0 - \beta_1 z_{0i} - \mathbf{x}_{0i}' \boldsymbol{\beta}_2)^2 \right\} \right] \times \left[\prod_{i=1}^{n_0} f(x_{0i} | \boldsymbol{\alpha}) \right]^{b_0} \pi_{20}^{(\nu)}(\boldsymbol{\alpha}), \quad (\text{A.1})$$

where $f(\mathbf{x}_{0i} | \boldsymbol{\alpha})$ is specified by (3.19). The full conditional distributions required in the Gibbs sampling algorithm are (i) $[\boldsymbol{\beta} | \sigma^2, D_0]$; (ii) $[\sigma^2 | \boldsymbol{\beta}, D_0]$; and (iii) $[\boldsymbol{\alpha} | D_0]$. We sample $\boldsymbol{\beta}$, σ^2 , and $\boldsymbol{\alpha}$ from the aforementioned full conditional distributions in turn. For (i), the full conditional distribution of $\boldsymbol{\beta}$ is given by

$$\boldsymbol{\beta} | \sigma^2, D_0 \sim N \left((X_0' X_0)^{-1} X_0' \mathbf{y}_0, \sigma^2 (X_0' X_0)^{-1} \right), \quad (\text{A.2})$$

where $\mathbf{y}_0 = (y_{01}, \dots, y_{0n_0})'$ and X_0 is a $n_0 \times (p+2)$ matrix with i^{th} row $(1, z_{0i}, \mathbf{x}_{0i}')$ for $i = 1, \dots, n_0$. For (ii), the conditional distribution of σ^2 is given by

$$\sigma^2 | \boldsymbol{\beta}, D_0 \sim \text{IG} \left(\frac{n_0}{2}, \frac{1}{2} (\mathbf{y}_0 - X_0 \boldsymbol{\beta})' (\mathbf{y}_0 - X_0 \boldsymbol{\beta}) \right), \quad (\text{A.3})$$

where $IG(a, b)$ denotes the inverse gamma distribution with density proportional to

$\frac{1}{(\sigma^2)^{a+1}} \times \exp(-b/\sigma^2)$. Thus, sampling β and σ^2 from (A.2) and (A.3) is straightforward. For (iii), the full conditional distribution is of the form

$$\pi^{(\nu)}(\alpha|D_0) \propto \left[\prod_{i=1}^{n_0} f(x_{0i}|\alpha) \right]^{b_0} \pi_{20}^{(\nu)}(\alpha). \quad (\text{A.4})$$

Sampling α from (A.4) depends on the specific form of (3.19) and the specification of $\pi_{20}^{(\nu)}(\alpha)$. For the covariate distributions given in Section 4.2, the full conditional distributions for the components of α are either log concave or inverse gamma. Note that the adaptive-rejection algorithm of Gilks and Wild [22] can be readily used to sample from these full conditional distributions.

References

1. O'Hagan A, Stevens JW, Campbell MJ. Assurance in clinical trial design. *Pharmaceutical Statistics*. 2005; 4:187–201.
2. Chuang-Stein C. Sample size and the probability of a successful trial. *Pharmaceutical Statistics*. 2006; 5:305–309. [PubMed: 17128428]
3. Spiedelhalter DJ, Freedman LS, Parmar MKB. Bayesian approaches to randomized trials. *Journal of the Royal Statistical Society, Series A*. 1994; 157:357–416.
4. Stallard N, Whitehead J, Cleall S. Decision-making in a phase II clinical trial: a new approach combining Bayesian and frequentist concepts. *Pharmaceutical Statistics*. 2005; 4:119–128.
5. Dmitrienko A, Wang MD. Bayesian predictive approach to interim monitoring in clinical trials. *Statistics in Medicine*. 2006; 25:2178–2195. [PubMed: 16007570]
6. Eaton ML, Muirhead RJ, Soaita A. On the limiting behavior of the “probability of claiming superiority” in a Bayesian context. *Bayesian Analysis*. 2012; 7:771–782.
7. Chen Y, Chen Z, Mori M. A new statistical decision rule for single-arm phase II oncology trials. *Statistical Methods in Medical Research*. 2012 (to appear in print).
8. Wang Y, Fu H, Kulkarni P, Kaiser C. Evaluating and utilizing probability of study success in clinical development. *Clinical Trials*. 2013; 10:407–413. [PubMed: 23471634]
9. Chen MH, Ibrahim JG, Lam P, Yu A, Zhang Y. Bayesian design of non-inferiority trials for medical devices using historical data. *Biometrics*. 2011; 67:1163–1170. [PubMed: 21361889]
10. Ibrahim JG, Chen MH. Power prior distributions for regression models. *Statistical Science*. 2000; 15:46–60.
11. Wang F, Gelfand AE. A simulation-based approach to Bayesian sample size determination for performance under a given model and for separating models. *Statistical Science*. 2002; 17:193–208.
12. Lipsitz SR, Ibrahim JG. A conditional model for incomplete covariates in parametric regression models. *Biometrika*. 1996; 83:916–922.
13. Ibrahim JG, Lipsitz SR, Chen MH. Missing covariates in generalized linear models when the missing data mechanism is nonignorable. *Journal of the Royal Statistical Society, Series B*. 1999; 61:173–190.
14. Chen MH, Ibrahim JG. Maximum likelihood methods for cure rate models with missing covariates. *Biometrics*. 2001; 57:43–52. [PubMed: 11252617]
15. Agresti, A. *Categorical Data Analysis*. 3rd ed. New York: Wiley; 2013.
16. Fienberg SE, Mason WM. Identification and estimation of age-period-cohort models in the analysis of discrete archival data. *Sociological Methodology*. 1979:1–67.

17. McCullagh P. Regression models for ordinal data. *Journal of the Royal Statistical Society, Series B*. 1980; 42:109–142.
18. Kerzner B, Murray AV, Cheng E, Ifle R, Harvey PR, Tomlinson M, Barben JL, Rarrick K, Stek JE, Chung MO, Schödel FP, Wang WW, Xu J, Chan IS, Silber JL, Schlienger K. Safety and immunogenicity profile of the concomitant administration of ZOSTAVAX and inactivated influenza vaccine in adults aged 50 and older. *Journal of the American Geriatrics Society*. 2007; 55:1499–1507. [PubMed: 17908055]
19. Gilderman LI, Lawless JF, Nolen TM, Sterling T, Rutledge RZ, Fernsler DA, Azrolan N, Sutradhar SC, Wang WW, Chan IS, Schlienger K, Schödel F, Silber JL. Zostavax Protocol 010 Study Group. A double-blind, randomized, controlled, multi-center safety and immunogenicity study of a refrigerator-stable formulation of Zostavax. *Clinical and Vaccine Immunology*. 2008; 15:314–319. [PubMed: 18077611]
20. Kramer M, Cutler N, Feighner J, Shrivastava R, Carman J, Sramek JJ, Reines SA, Liu G, Snavely D, Wyatt-Knowles E, Hale JJ, Mills SG, MacCoss M, Swain CJ, Harrison T, Hill RG, Hefti F, Scolnick EM, Cascieri MA, Chicchi GG, Sadowski S, Williams AR, Hewson L, Smith D, Carlson EJ, Hargreaves RJ, Rupniak NM. Distinct mechanism for antidepressant activity by blockade of central substance P receptors. *Science*. 1998; 281:1640–1645. [PubMed: 9733503]
21. Keller M, Montgomery S, Ball W, Morrison M, Snavely D, Liu G, Hargreaves R, Hietala J, Lines C, Beebe K, Reines S. Lack of efficacy of the substance P (neurokinin receptor) antagonist aprepitant in the treatment of major depressive disorder. *Biological Psychiatry*. 2006; 59:216–223. [PubMed: 16248986]
22. Gilks WR, Wild P. Adaptive rejection sampling for Gibbs sampling. *Applied Statistics*. 1992; 41:337–348.

Figure 1.

Plots of $-2 \log m(a_0)$ in (3.17) versus a_0 for the models with or without covariates for paroxetine and MK0869.

(a) Paroxetine with TV = 0.5

(b) Paroxetine with TV = 1.0

(c) MK0869 with TV = 0.5

(d) MK0869 with TV = 1.0

Figure 2.

Plots of POS versus a_0 when $\gamma = 0.90$ for paroxetine and MK0869.

Table I

Posterior predictive probability of $\text{GMFR} > \phi$ and POS.

	ϕ						
	1.4	1.5	1.6	1.7	1.8	1.9	2.0
$\log(\phi)$	0.3365	0.4055	0.4700	0.5306	0.5878	0.6419	0.6931
Empirical posterior predictive probability	0.9980	0.9924	0.9777	0.9467	0.8927	0.8129	0.7103
Full Bayes POS							
$\gamma = 0.50$	0.9980	0.9924	0.9777	0.9467	0.8927	0.8129	0.7103
$\gamma = 0.60$	0.9965	0.9878	0.9663	0.9243	0.8558	0.7610	0.6462
$\gamma = 0.70$	0.9939	0.9802	0.9493	0.8930	0.8079	0.6977	0.5728
$\gamma = 0.80$	0.9888	0.9667	0.9212	0.8457	0.7408	0.6154	0.4837
$\gamma = 0.85$	0.9841	0.9551	0.8990	0.8107	0.6943	0.5619	0.4291
$\gamma = 0.90$	0.9757	0.9361	0.8648	0.7603	0.6311	0.4930	0.3625
$\gamma = 0.95$	0.9569	0.8973	0.8011	0.6738	0.5310	0.3918	0.2713

Summary of patient baseline demographics and HAMD-17 total scores for MK0869-P008 and MK0869-P059.

Table II

Study	Treatment group	n	Gender		Mean age (years)	Race			HAMD-17 total	
			Female (%)	Male (%)		White (%)	Black (%)	Other (%)	Baseline	CFB
P008	MK0869	67	49	51	38.6	64	15	21	25.8	-12.5
	Paroxetine	70	44	56	40.4	66	19	16	24.9	-11.7
	Placebo	69	61	39	39.0	59	22	19	25.2	-8.3
P059	MK0869	150	65	35	40.7	79	11	10	27.6	-12.3
	Paroxetine	148	66	34	39.4	74	10	16	27.5	-13.9
	Placebo	150	59	41	41.5	75	13	12	27.7	-11.9

CFB, change from baseline.

Table III

POS values for paroxetine.

TV	w	POS											
		No covariates				$(b_{01}, b_{02}) = (0, 1)$				$(b_{01}, b_{02}) = (1, 0)$			
		$\gamma = 0.85$	0.90	0.95	0.85	0.90	0.95	0.85	0.90	0.95	0.85	0.90	0.95
0.5	0	0.638	0.571	0.468	0.603	0.536	0.435	0.599	0.531	0.430	0.605	0.540	0.436
	0.142	0.678	0.612	0.508	0.654	0.586	0.482	0.648	0.581	0.477	0.655	0.588	0.484
	0.2	0.695	0.629	0.524	0.673	0.606	0.502	0.668	0.601	0.496	0.674	0.607	0.503
	0.4	0.748	0.685	0.581	0.737	0.674	0.570	0.731	0.667	0.563	0.738	0.675	0.571
	0.6	0.796	0.737	0.636	0.795	0.736	0.636	0.788	0.729	0.629	0.795	0.737	0.637
	0.8	0.838	0.785	0.690	0.844	0.792	0.700	0.838	0.785	0.691	0.844	0.792	0.700
	0.979	0.871	0.823	0.734	0.881	0.836	0.751	0.875	0.828	0.742	0.881	0.836	0.751
1	0	0.874	0.827	0.740	0.885	0.840	0.757	0.879	0.833	0.748	0.885	0.840	0.757
	0.2	0.497	0.428	0.329	0.460	0.393	0.300	0.455	0.388	0.295	0.462	0.394	0.301
	0.142	0.532	0.460	0.357	0.502	0.432	0.333	0.497	0.427	0.328	0.504	0.434	0.334
	0.4	0.546	0.474	0.369	0.520	0.449	0.347	0.515	0.444	0.342	0.521	0.450	0.348
	0.6	0.595	0.522	0.411	0.581	0.507	0.399	0.574	0.501	0.393	0.581	0.508	0.400
	0.8	0.644	0.570	0.456	0.640	0.566	0.454	0.632	0.558	0.446	0.640	0.566	0.454
	0.979	0.690	0.617	0.501	0.696	0.624	0.510	0.687	0.615	0.501	0.696	0.624	0.510
	0	0.730	0.658	0.542	0.742	0.674	0.561	0.734	0.664	0.550	0.742	0.673	0.560
1	0	0.734	0.663	0.547	0.748	0.679	0.567	0.739	0.670	0.556	0.747	0.679	0.566
	0.2	0.356	0.293	0.211	0.321	0.262	0.187	0.317	0.259	0.184	0.323	0.264	0.188
	0.142	0.380	0.314	0.226	0.351	0.288	0.206	0.347	0.284	0.203	0.353	0.290	0.207
	0.4	0.390	0.323	0.233	0.364	0.299	0.215	0.359	0.295	0.211	0.365	0.301	0.216
	0.6	0.426	0.354	0.257	0.410	0.340	0.246	0.404	0.334	0.242	0.410	0.341	0.247
	0.8	0.463	0.388	0.283	0.457	0.383	0.281	0.450	0.376	0.275	0.457	0.383	0.281
	0.979	0.502	0.423	0.312	0.505	0.428	0.318	0.496	0.419	0.311	0.505	0.427	0.318
1	0	0.536	0.455	0.338	0.549	0.469	0.354	0.539	0.459	0.345	0.548	0.468	0.353
	0.2	0.540	0.458	0.342	0.554	0.474	0.360	0.544	0.464	0.349	0.553	0.473	0.358

Table IV

POS values for MK0869.

TV		POS											
		No covariates				$(b_{01}, b_{02}) = (0, 1)$				$(b_{01}, b_{02}) = (1, 0)$			
		$\gamma = 0.85$	0.90	0.95	0.95	0.85	0.90	0.95	0.95	0.85	0.90	0.95	0.95
0.5	0	0.205	0.159	0.104	0.185	0.143	0.093	0.180	0.138	0.089	0.184	0.142	0.092
	0.2	0.272	0.215	0.145	0.244	0.192	0.128	0.239	0.187	0.124	0.243	0.191	0.127
	0.208	0.274	0.217	0.147	0.246	0.194	0.130	0.241	0.189	0.125	0.245	0.193	0.129
	0.4	0.348	0.282	0.196	0.312	0.250	0.171	0.307	0.245	0.167	0.311	0.249	0.170
	0.443	0.365	0.297	0.208	0.327	0.263	0.182	0.323	0.259	0.178	0.327	0.262	0.181
	0.6	0.431	0.357	0.257	0.386	0.315	0.223	0.382	0.312	0.219	0.386	0.315	0.222
	0.8	0.518	0.439	0.326	0.464	0.387	0.282	0.462	0.385	0.279	0.465	0.388	0.282
	1	0.604	0.523	0.403	0.544	0.463	0.348	0.543	0.462	0.346	0.546	0.465	0.349
1.0	0	0.117	0.086	0.052	0.105	0.077	0.047	0.101	0.074	0.044	0.104	0.076	0.046
	0.2	0.136	0.101	0.073	0.139	0.104	0.064	0.135	0.100	0.061	0.138	0.103	0.063
	0.208	0.159	0.119	0.074	0.141	0.105	0.065	0.137	0.101	0.061	0.140	0.104	0.064
	0.4	0.207	0.157	0.099	0.181	0.136	0.085	0.176	0.132	0.082	0.180	0.135	0.084
	0.443	0.218	0.166	0.106	0.190	0.144	0.091	0.186	0.140	0.087	0.189	0.143	0.090
	0.6	0.263	0.204	0.133	0.228	0.175	0.112	0.225	0.172	0.109	0.228	0.174	0.111
	0.8	0.327	0.259	0.173	0.282	0.220	0.144	0.280	0.217	0.141	0.282	0.220	0.144
	1	0.397	0.320	0.220	0.342	0.272	0.183	0.340	0.269	0.180	0.343	0.272	0.182
1.5	0	0.060	0.042	0.023	0.054	0.038	0.021	0.051	0.035	0.019	0.053	0.037	0.021
	0.2	0.081	0.049	0.032	0.071	0.050	0.028	0.067	0.047	0.026	0.070	0.049	0.027
	0.208	0.082	0.058	0.032	0.071	0.050	0.028	0.068	0.047	0.026	0.070	0.049	0.028
	0.4	0.106	0.076	0.043	0.091	0.065	0.037	0.088	0.062	0.035	0.090	0.064	0.036
	0.443	0.112	0.080	0.046	0.096	0.068	0.039	0.092	0.065	0.037	0.095	0.067	0.038
	0.6	0.137	0.099	0.058	0.116	0.083	0.048	0.112	0.080	0.046	0.115	0.082	0.047
	0.8	0.173	0.127	0.075	0.145	0.105	0.061	0.142	0.102	0.059	0.144	0.104	0.061
	1	0.216	0.160	0.097	0.178	0.131	0.078	0.175	0.128	0.075	0.177	0.130	0.077