

## Research Article

# Bayesian Equivalence Testing and Meta-Analysis in Two-Arm Trials with Binary Data

Cynthia Kpekpena and Saman Muthukumarana 

Department of Statistics, University of Manitoba, Winnipeg, MB, Canada R3T 2N2

Correspondence should be addressed to Saman Muthukumarana; saman.muthukumarana@umanitoba.ca

Received 15 January 2018; Revised 9 June 2018; Accepted 24 June 2018; Published 8 August 2018

Academic Editor: Kazuhisa Nishizawa

Copyright © 2018 Cynthia Kpekpena and Saman Muthukumarana. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We consider a Bayesian approach for assessing hypotheses of equivalence in two-arm trials with binary Data. We discuss the development of likelihood, the prior, and the posterior distributions of parameters of interest. We then examine the suitability of a normal approximation to the posterior distribution obtained via a Taylor series expansion. The Bayesian inference is carried out using Markov Chain Monte Carlo (MCMC) methods. We illustrate the methods using actual data arising from two-arm clinical trials on preventing mortality after myocardial infarction.

## 1. Introduction

Consider a clinical trial where a pharmaceutical company wants to test a new drug against a currently existing drug. Sometimes in these studies, the clinical trial end point may be the success or failure of the treatment. A binary outcome is an outcome whose unit can take on only two possible states “0” and “1.” This success/failure response variable could be heart disease (Yes/No), patient condition (Good/Critical), how often patient feel depressed (Never/Often), and so on. The natural distribution for modeling these types of binary data is the binomial distribution given by

$$f(x; p) = \binom{n}{x} p^x (1-p)^{n-x} \quad \text{for } x = 0, 1, \dots, n, \quad (1)$$
$$p \in (0, 1).$$

The mean and variance for the binomial random variable are  $E(X) = np$  and  $\text{var}(X) = np(1-p)$ , respectively. In (1), it is assumed that there are only two outcomes (denoted “success” or “failure”) and a fixed number of trials ( $n$ ). The trials are independent with a constant probability of success.

The main objective of this type of clinical trial is to determine whether there is a significant difference between active treatment (new drug) and reference treatment

(current drug). Tests of significance have generally been argued not to be enough. That is, if the  $p$  value for a test of significance leads to the nonrejection of the null hypothesis, it is not a proof that the null hypothesis holds. The clinician may want to test a null hypothesis of equivalence against an alternative hypothesis that states that there is a sufficient difference between the two drugs.

Equivalence testing is widely used when a choice is to be made between a drug (or a treatment) and an alternative. The term equivalence in the statistical sense is used to mean a weak pattern displayed by the data under study regarding the underlying population distribution. Equivalence tests are designed to show the nonexistence of a relevant difference between two treatments. It is known that Fisher’s one-sided exact test is the same as the test for equivalence in the frequentist approach [1]. This testing procedure is similar to the classical two-sided test procedure but involves an equivalence zone determined by a margin known as equivalence margin ( $\delta$ ).

The equivalence margin ( $\delta$ ), which represents a margin of clinical indifference, is usually estimated from previous studies and as such is also based primarily on clinical criteria as well as statistical principle. This margin is influenced by statistical principle but largely dependent on the interest of the experimenter and research questions clinicians wish to

answer. As such, the statistical method employed together with the design of the study must be in such a manner that the margin of difference is not too restrictive to capture the bounds of the research question. For a test of equivalence of two binomial proportions, the equivalence margin is discussed in [2].

The frequentist approach to equivalence testing is the two one-sided test (TOST) procedure. By the TOST, equivalence is established at the  $\alpha$  significance level if a  $(1 - 2\alpha) \times 100\%$  confidence interval for the difference in treatment means  $\mu_i - \mu_j$  is contained within the interval  $(-\delta, \delta)$  where  $\delta$  is the equivalence margin.

The motivation for this paper is based on the fact that for a given disease, there is likely to be many other substitute drugs or new drugs that can be used to treat the patients. But these drugs may not all be at the same cost; some may possibly have adverse side effects, and the method of application could be complex for others. On grounds of these information, we do equivalence testing to see if two different drugs can be regarded as equivalent in terms of their treatment effect. There are a variety of different approaches to this problem as indicated by some recent literature. See Wellek [1], Albert [2], Gamalo et al. [3], Rahardja and Zhao [4] and Zaslavsky [5] for comprehensive details on recent developments. We remark that Gamalo et al. [3] consider a Bayesian approach to proportions along with non-inferiority trials. In this paper, we consider a Bayesian approach focusing on equivalence tests. We also construct a simple normal approximation and provide a mechanism for missing data analysis as well.

The remaining sections of this article are organized as follows: In Section 2, Bayesian inferential procedure for binary data is discussed. Section 3 presents a normal approximation to the posterior distribution obtained via a Taylor series expansion. We then examine the suitability of this normal approximation. We discuss a Gibbs sampling mechanism for estimating missing data in Section 4. In Section 5, we analyze a published dataset by Carlin [6] and Yusuf et al. [7]. This dataset consists with 22 treatment-control trials to prevent mortality after myocardial infarction. We conclude with a discussion of the approach in Section 5.

## 2. Bayesian Inferential Procedure

Let  $X_t$  be the number of individuals with positive exposure out of a total of  $n_t$  patients in treatment group with proportion  $P_t$ . Accordingly, let  $X_c$  denote the number of individuals with positive exposure out of a total  $n_c$  in the control group with proportion  $P_c$ . Then,

$$\begin{aligned} X_t &\sim \text{Bin}(n_t, P_t), \\ X_c &\sim \text{Bin}(n_c, P_c). \end{aligned} \quad (2)$$

The priors on the parameters,  $P_t$  and  $P_c$  are given by

$$\begin{aligned} P_t &\sim \text{Beta}(\alpha, \beta), \\ P_c &\sim \text{Beta}(\varepsilon, \eta). \end{aligned} \quad (3)$$

Then the posterior distributions of  $P_t$  and  $P_c$  are given by

TABLE 1: Normal approximation to the Beta distribution.

| Exact distribution | Approximation        |
|--------------------|----------------------|
| Beta(2, 1)         | $N(1, \infty)$       |
| Beta(1, 2)         | $N(0, \infty)$       |
| Beta(10, 10)       | $N(0.5000, 8.4853)$  |
| Beta(5, 1)         | $N(1, \infty)$       |
| Beta(1, 5)         | $N(0, \infty)$       |
| Beta(2, 2)         | $N(0.5000, 2.8284)$  |
| Beta(3, 3)         | $N(0.5000, 4.0)$     |
| Beta(2, 4)         | $N(0.2500, 4.6188)$  |
| Beta(4, 4)         | $N(0.5000, 4.8990)$  |
| Beta(5, 5)         | $N(1, 5.6569)$       |
| Beta(30, 20)       | $N(0.6042, 14.1673)$ |
| Beta(20, 30)       | $N(0.3958, 14.1673)$ |
| Beta(50, 20)       | $N(0.7206, 18.3776)$ |
| Beta(20, 50)       | $N(0.2794, 18.3776)$ |

$$\begin{aligned} \pi(P_t | X_t) &\propto \text{Beta}(x_t + \alpha, n_t + \beta - x_t), \\ \pi(P_c | X_c) &\propto \text{Beta}(x_c + \varepsilon, n_c + \eta - x_c). \end{aligned} \quad (4)$$

For Bayesian inference about treatment effect, a test is required to determine whether the posterior probability of treatment proportions  $P_t$  and  $P_c$  lies within the bounds of the equivalence margin or not. There is therefore, the need to sample from the posterior distribution of  $P_t - P_c$ . The marginal posteriors of  $P_t$  and  $P_c$  are Beta distributions, and therefore  $\pi(P_t - P_c | X_t, X_c)$  is not in an analytically tractable form. So,  $P_{1t}, P_{2t}, \dots, P_{mt}$  are generated from  $\pi(P_t | X_t)$  and independently  $P_{1c}, P_{2c}, \dots, P_{mc}$  generated from  $\pi(P_c | X_c)$  because  $P_t$  and  $P_c$  are independent. Then,  $P_{1t} - P_{1c}, P_{2t} - P_{2c}, \dots, P_{mt} - P_{mc}$  can be treated as a random sample from  $\pi(P_t - P_c | X_t, X_c)$ .

## 3. Normal Approximation to the Beta Posterior Distribution

Note that the posterior distributions of  $P_t$  and  $P_c$  are Beta distributions. By following Kpekpena [8], a normal approximation to posteriors can be obtained using a Taylor series expansion of the Beta distribution. By applying a Taylor series expansion with first three terms, it can be shown that  $\pi(P_t | X_t) \approx \mathcal{N}(\mu, \sigma)$ , where

$$\begin{aligned} \mu &= \frac{1 - \alpha - x_t}{2 - \alpha - n_t - \beta}, \\ \sigma &= \frac{1}{\left[ -(-2 - \alpha - n_t - \beta)^3 / ((1 - x_t - \alpha)(1 - n_t - \beta + x_t)) \right]^{1/2}}. \end{aligned} \quad (5)$$

Similarly, the approximation of  $\pi(P_c | X_c)$  can also be obtained. The details of this construction are given in the Appendix. We provide some approximations based on this development in Table 1 and Figures 1 and 2. It is clear that the approximation starts to work well for the values of the posterior parameters from  $x + \alpha = 10$  and  $n + \beta - x = 10$ . However, the approximation is not suitable when Beta posterior parameters are less than 10.

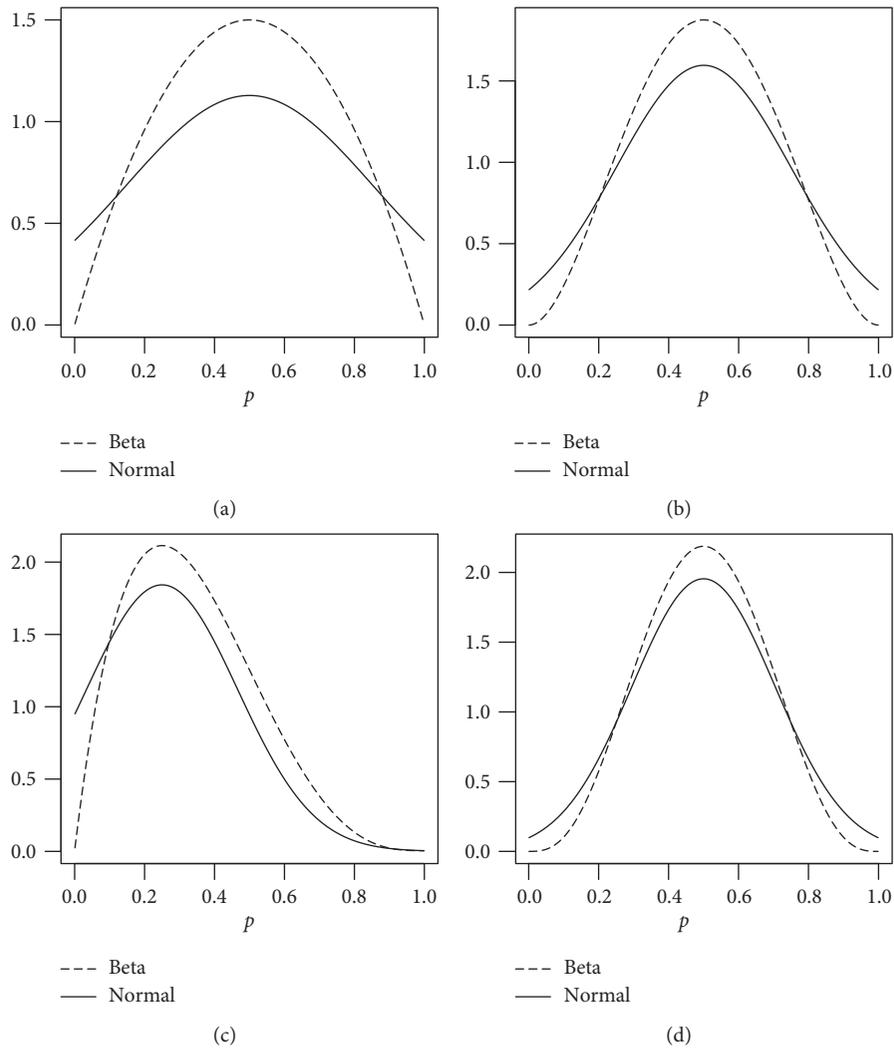


FIGURE 1: The normal approximations of (a) Beta (2,2), (b) Beta (3,3), (c) Beta (2,4), and (d) Beta (4,4).

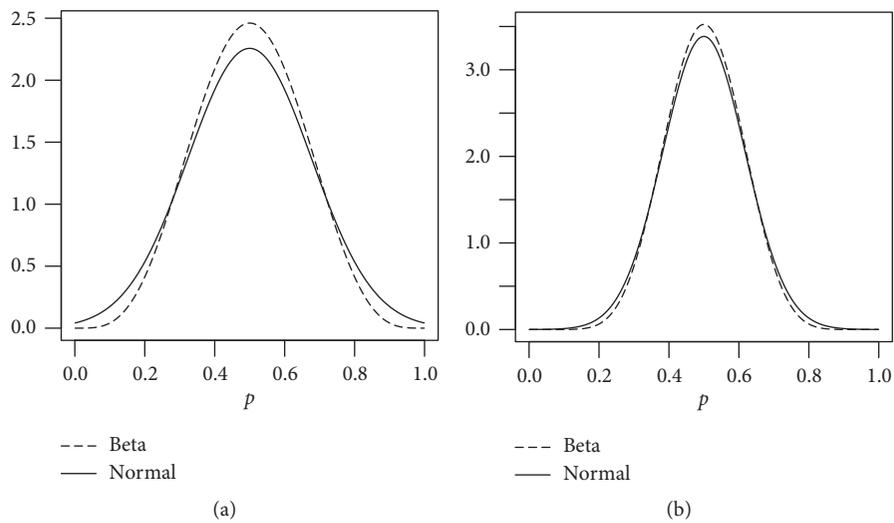


FIGURE 2: Continued.

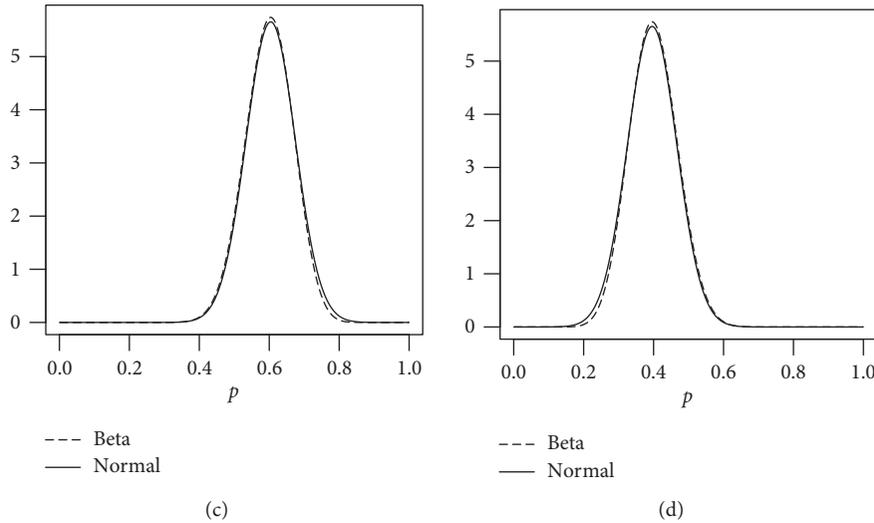


FIGURE 2: The Normal approximations (a) Beta (5, 5), (b) Beta (10, 10), (c) Beta (30, 20), and (d) Beta (20, 30).

#### 4. Estimating Missing Data in Arms

Missing data are easily handled in Bayesian inference by treating them as another set of parameters. We estimate the missing values conditioning on the observed data. For example, let  $X_1, \dots, X_n$  be a binary random sample from  $\text{Ber}(P)$  in an arm and suppose that  $X_m$  is missing. Note that  $P$  represents  $P_t$  in treatment arm and  $P_c$  in control group. Let  $P \sim \text{Beta}(\alpha, \beta)$  and  $Y = \sum_{i \neq m}^n X_i$ . Then, the likelihood of the observed data is

$$L(X_{\text{obs}} | P) = \binom{n-1}{y} P^y (1-P)^{n-1-y}. \quad (6)$$

The posterior of  $P$  based on the complete data  $X = (Y, X_m)$  is

$$\pi(P | X) \propto P^{y+x_m} (1-P)^{n-y-x_m} \frac{1}{B(\alpha, \beta)} P^{\alpha-1} (1-P)^{\beta-1}. \quad (7)$$

The full conditionals of  $P$  and  $X_m$  are

$$\begin{aligned} \pi(P | y, x_m) &\sim \text{Beta}(y + x_m + \alpha, n - y - x_m + \beta), \\ \pi(x_m | y, P) &\sim \text{Ber}(P). \end{aligned} \quad (8)$$

It is easy to generate from these full conditionals in  $R$ , so  $P$  and  $x_m$  can be estimated using Gibbs sampling.

#### 5. Data Analysis

We apply our approach on data analyzed in [7, 9]. The data includes 22 trials of beta-blockers to prevent mortality after myocardial infarction. For each of the 22 trials, a test of equivalence is done to ascertain whether the treatment proportion is equivalent to the control proportion. This example is based on the Statistical inferential procedure for binary data discussed in Sections 2 and 3. For each arm, the number of patients who had myocardial infarction out of

a total  $n_t$  is considered to be the number of successes in  $n_t$  binomial trials. Similarly, the number of cases in the control group is treated as a binomial outcome independent of the treatment group. The equivalence margin  $\delta$  is chosen to be as small as possible such that if the absolute value of the difference in the control and treatment proportions is less than  $\delta$ , and we can say that the two proportions are equivalent. For demonstration purpose, we assume a practically meaningful equivalence margin  $\delta = 0.01$ . We use non-informative priors  $\text{Beta}(2, 1)$  for the parameters,  $P_t$  and  $P_c$ . The hypothesis for a test of equivalence of study number  $i$  and its control group is as follows:

$$\begin{aligned} H_0 &: |P_{ti} - P_{ci}| \geq \delta, \\ H_1 &: |P_{ti} - P_{ci}| < \delta. \end{aligned} \quad (9)$$

We perform the equivalence test in (9) using the Bayes factor [9]. Table 2 gives the results of the equivalence tests. The first column  $D_i$  is the  $i$ th study label. Columns 2 and 3 are the treatment proportion ( $x_t/n_t$ ) and control proportion ( $x_c/n_c$ ), respectively. Columns 4 ( $P(H_1 | X)$ ) and 6 ( $P_A(H_1 | X)$ ) are the posterior probabilities that  $H_1 : |P_{ti} - P_{ci}| < \delta$  is true under the Beta posterior distributions and under the normal approximation to the Beta posterior, respectively. Column 5 ( $B$ ) is the Bayes factor for exact posterior, and  $B_A$  in column 7 is the Bayes Factor based on the normal approximation. For study 1, the Bayes factor for the exact posterior is 7.3822 whereas that of the normal approximation is 7.5466. Both Bayes factors are above 1 which implies that  $H_0$  is more likely to be true, and  $H_0$  is the hypothesis that the treatment proportion is not equivalent to the control proportion. We remark that classical hypothesis tests give one hypothesis a preferred status and only consider evidence against it which is not the case in Bayesian tests. Results also indicate that the approximation and exact computation lead to the same conclusion in each study indicating the suitability of the approximation.

We now consider a missing data analysis in an arm. As an example, suppose an observation was missing in the

TABLE 2: Posterior probabilities and Bayes factor.

| $D_i$    | $P_{\bar{u}_i}$ | $P_{c_i}$ | $P(H_1   X)$ | $B$      | $P_A(H_1   X)$ | $B_A$    |
|----------|-----------------|-----------|--------------|----------|----------------|----------|
| $D_1$    | 0.0789          | 0.0769    | 0.1193       | 7.3822   | 0.1170         | 7.5466   |
| $D_2$    | 0.0614          | 0.1207    | 0.0725       | 12.7931  | 0.0669         | 13.9506  |
| $D_3$    | 0.0725          | 0.1183    | 0.1184       | 7.4459   | 0.1132         | 7.8283   |
| $D_4$    | 0.0665          | 0.0836    | 0.2343       | 3.2680   | 0.2294         | 3.3583   |
| $D_5$    | 0.0789          | 0.0740    | 0.3718       | 1.6896   | 0.3710         | 1.6957   |
| $D_6$    | 0.0678          | 0.1154    | 0.0998       | 9.02004  | 0.0962         | 9.3964   |
| $D_7$    | 0.1037          | 0.1618    | 0.0015       | 665.6667 | 0.0010         | 971.6452 |
| $D_8$    | 0.0949          | 0.1019    | 0.3861       | 1.5900   | 0.3854         | 1.5949   |
| $D_9$    | 0.0899          | 0.1312    | 0.0787       | 11.7064  | 0.0784         | 11.7538  |
| $D_{10}$ | 0.0720          | 0.0979    | 0.0432       | 22.1482  | 0.0395         | 24.3175  |
| $D_{11}$ | 0.0733          | 0.0892    | 0.2950       | 2.3898   | 0.2958         | 2.3802   |
| $D_{12}$ | 0.1711          | 0.1767    | 0.2317       | 3.3159   | 0.2341         | 3.2722   |
| $D_{13}$ | 0.0309          | 0.0546    | 0.1951       | 4.1256   | 0.1900         | 4.2626   |
| $D_{14}$ | 0.0664          | 0.0510    | 0.3030       | 2.3003   | 0.3022         | 2.3089   |
| $D_{15}$ | 0.1623          | 0.2109    | 0.0989       | 9.1112   | 0.0994         | 9.0554   |
| $D_{16}$ | 0.1594          | 0.1784    | 0.1871       | 4.3447   | 0.1886         | 4.3010   |
| $D_{17}$ | 0.1116          | 0.0984    | 0.2193       | 3.5560   | 0.2214         | 3.5147   |
| $D_{18}$ | 0.0530          | 0.0390    | 0.2673       | 2.7411   | 0.2649         | 2.7754   |
| $D_{19}$ | 0.0345          | 0.0224    | 0.3175       | 2.1496   | 0.3293         | 2.0372   |
| $D_{20}$ | 0.1531          | 0.1835    | 0.1584       | 5.3131   | 0.1552         | 5.4439   |
| $D_{21}$ | 0.0691          | 0.1181    | 0.0331       | 29.2110  | 0.0311         | 31.1380  |
| $D_{22}$ | 0.0324          | 0.0579    | 0.0932       | 9.7296   | 0.0867         | 10.536   |

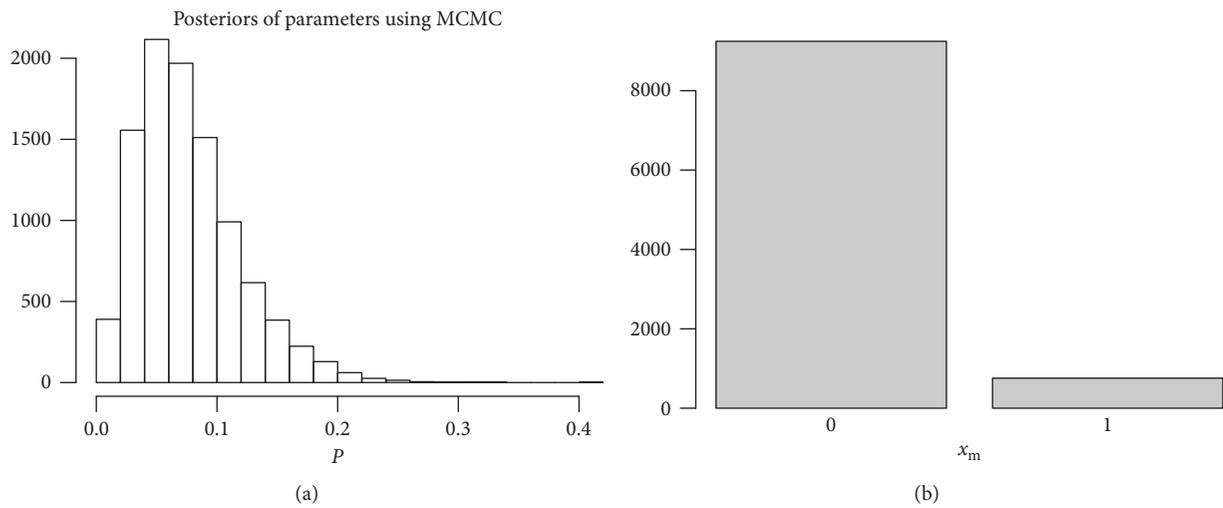


FIGURE 3: The distribution of  $x_m$  shows it is more likely to be 0.

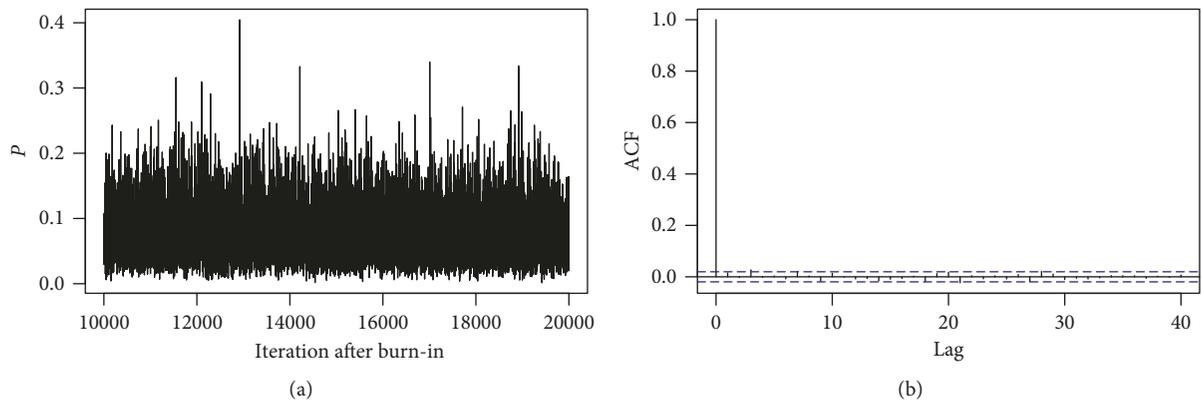


FIGURE 4: There is no discernible pattern in the trace plot and no large spikes in the autocorrelation plot.

treatment group under study 1. We estimate this missing value using Gibbs sampling derived in Section 4. The posterior distributions of parameters  $P$  and  $x_m$  are given in Figure 3 based on 20000 MCMC simulations. According to Figure 3, it is likely that  $x_m$  is 0. The trace plot in Figure 4 shows that mixing is good enough, and there are no large spikes in the autocorrelation plot after lag 0. This is an indication of convergence of the Markov Chain.

We also consider a meta-analysis of the binary data in two-arm trials in order to assess the between-study variations. Let  $y_i$  be the estimate of the true effect size  $\mu_i$  corresponding to the  $i$ th study. Then, the random effects model is given as

$$y_i = \mu_i + e_i \quad \text{where } e_i \sim N(0, \sigma_i^2). \quad (10)$$

As developed in Muthukumarana and Tiwari [10], we consider a hierarchical Dirichlet process formulation for  $\mu_i$  as follows:

$$\begin{aligned} \mu_i | F &\sim F, \\ F &\sim \text{DP}(M, H), \\ H &\sim N(\mu, \tau^2), \\ \mu &\sim N(\mu_0, d\tau^2), \\ \frac{1}{\tau^2} &\sim G(a, b), \end{aligned} \quad (11)$$

where  $M$ ,  $\mu_0$ , and  $d$  are known.

We analyze the dataset published in Nissen and Wolski [11] to assess the between-study heterogeneity. In this dataset, there are 42 trials including 15565 diabetes patients who were put on rosiglitazone (treatment group) and 12282 diabetes patients assigned to medication that does not contain rosiglitazone (control group). Note that the interest is on myocardial infarction and death from rosiglitazone as a treatment for diabetes. We use the odds ratio as the treatment effect. The parameters of the model are estimated by Gibbs sampling algorithm implemented in R. The estimates of the model parameters ( $\mu_i$ ) are given in Table 3.

We also conducted a simulation study to assess the validity of the approach. Each study was simulated by means of a binomial random variable in which the number of cases in the treatment group and the control group are generated as independent binomial random variables. We generate twenty binomial successes using the rbinom random generator. We assume  $n = 200$  in each case and fix the  $p$  at 0.7. This setting is similar to administering a treatment in twenty hospitals with 200 patients in each hospital. Fixing  $p$  at 0.7 generates number of cases that do not vary so much from each other. This is confirmed in the non-significance of the chi-square test for heterogeneity. Another set of twenty ‘‘number of cases’’ is generated from the binomial distribution, but this time we induce heterogeneity. This is done by varying the success probability of each trial. For instance rbinom(1, 200, 0.86), rbinom(1, 200, 0.10), rbinom(1, 200, 0.55), . . .

TABLE 3: The estimates of posterior treatments and standard deviations.

| Parameter  | Estimate | Standard deviation |
|------------|----------|--------------------|
| $\tau^2$   | 0.74     | 0.2794073          |
| $\mu$      | 0.71     | 0.4142608          |
| $\mu_1$    | -0.73    | 0.8151172          |
| $\mu_2$    | -0.63    | 0.7358434          |
| $\mu_3$    | 1.2      | 0.4914719          |
| $\mu_4$    | -1.1     | 0.6541732          |
| $\mu_5$    | -0.74    | 0.8379231          |
| $\mu_6$    | -0.77    | 0.7702218          |
| $\mu_7$    | -0.70    | 0.8191522          |
| $\mu_8$    | -0.61    | 0.7714286          |
| $\mu_9$    | -0.74    | 0.7989194          |
| $\mu_{10}$ | -0.75    | 0.8194161          |
| $\mu_{11}$ | -0.77    | 0.7766973          |
| $\mu_{12}$ | -0.74    | 0.8008516          |
| $\mu_{13}$ | -0.67    | 0.7848413          |
| $\mu_{14}$ | -0.73    | 0.8148389          |
| $\mu_{15}$ | -0.65    | 0.7427869          |
| $\mu_{16}$ | -0.72    | 0.6874473          |
| $\mu_{17}$ | -0.67    | 0.6670385          |
| $\mu_{18}$ | -0.74    | 0.8121317          |
| $\mu_{19}$ | -0.69    | 0.7916802          |
| $\mu_{20}$ | -0.74    | 0.8151257          |
| $\mu_{21}$ | -0.78    | 0.7748052          |
| $\mu_{22}$ | -0.78    | 0.7449046          |
| $\mu_{23}$ | -0.71    | 0.8175963          |
| $\mu_{24}$ | -0.75    | 0.8185433          |
| $\mu_{25}$ | -1.6     | 0.4702446          |
| $\mu_{26}$ | -0.57    | 0.5950003          |
| $\mu_{27}$ | -0.74    | 0.8104195          |
| $\mu_{28}$ | -0.74    | 0.8192542          |
| $\mu_{29}$ | -0.74    | 0.8091605          |
| $\mu_{30}$ | -0.73    | 0.8372302          |
| $\mu_{31}$ | -0.58    | 0.6008941          |
| $\mu_{32}$ | -0.72    | 0.8096048          |
| $\mu_{33}$ | -0.76    | 0.8134446          |
| $\mu_{34}$ | -0.69    | 0.6464147          |
| $\mu_{35}$ | -0.72    | 0.8196591          |
| $\mu_{36}$ | -0.75    | 0.8109735          |
| $\mu_{37}$ | -0.71    | 0.8172117          |
| $\mu_{38}$ | -0.72    | 0.8127885          |
| $\mu_{39}$ | -0.74    | 0.8024024          |
| $\mu_{40}$ | -0.73    | 0.8114458          |
| $\mu_{41}$ | -0.189   | 0.5258747          |
| $\mu_{42}$ | 0.01     | 0.3224310          |

Note that our interest is in comparing the posterior treatment means of the heterogeneous studies with the studies that are not heterogeneous. Table 4 compares the posterior treatment means of 20 studies with heterogeneity to the treatment means of 20 other studies in which there is no heterogeneity. Column 1 is the posterior treatment means of the nonheterogeneous ( $\mu_i$ ) studies whereas  $\mu_i^*$  in column 2 is the posterior treatments of the heterogeneous studies. Treatment means in column 1 ( $\mu_i$ ) are mostly 0.68 or just slightly below or above it. If the responses are similar, the treatment effects are supposed to be an estimate of a common treatment mean. On the other hand, all the treatment means in column 2 ( $\mu_i^*$ ) differ from each other

TABLE 4: Estimates of treatment means for twenty studies with 200 observations within each study.

| Study | $\mu_i$ | $\mu_i^*$ |
|-------|---------|-----------|
| 1     | 0.68    | 0.092     |
| 2     | 0.67    | 0.39      |
| 3     | 0.67    | 0.80      |
| 4     | 0.68    | 0.69      |
| 5     | 0.68    | 0.76      |
| 6     | 0.68    | -2.8      |
| 7     | 0.68    | -1.4      |
| 8     | 0.68    | 0.35      |
| 9     | 0.68    | -0.11     |
| 10    | 0.70    | 0.53      |
| 11    | 0.67    | 0.69      |
| 12    | 0.70    | -0.054    |
| 13    | 0.68    | 0.72      |
| 14    | 0.67    | 0.39      |
| 15    | 0.68    | 0.41      |
| 16    | 0.69    | 0.79      |
| 17    | 0.67    | 0.81      |
| 18    | 0.69    | -0.94     |
| 19    | 0.68    | -1.6      |
| 20    | 0.69    | 0.81      |

significantly indicating that the model can assess the between-study variations.

## 6. Discussion

We have considered a Bayesian analysis of binary data in testing hypotheses of equivalence. The tests of hypotheses of equivalence are popular in clinical trials, and our approach is relatively simple and easy to perform. Bayesian formulation was considered for testing hypothesis of equivalence, and we observed that the normal approximation to the Beta posterior can be used for moderately large sample sizes. We also presented a mechanism for estimating missing data in arms. This is useful in situations where data are partially missing in some arms. We also considered a meta-analytic approach for assessing between-study variations.

There are two directions we would like to pursue along the methods discussed in this article. We are interested in enhancing the method to accommodate extra covariates into the model in the presence of multiple outcomes in an arm. The incorporation of covariates makes the Bayes factor inappropriate, and we would like to examine the other model selection criterion in place of the Bayes factor.

## Appendix

Let the best estimate of  $P$ ,  $P_0$  be the value of  $P$  for which the posterior is at its maximum. That is,

$$\begin{aligned} \frac{d\pi(P | x)}{dP} \Big|_{P_0} &= 0, \\ \frac{d^2\pi(P | x)}{dP^2} \Big|_{P_0} &< 0. \end{aligned} \quad (\text{A.1})$$

The Taylor series expansion of a function  $f(x)$  at  $X = x_0$  is

$$f(x) = \sum_{m=0}^{\infty} \frac{f^{(m)}(x_0)}{m!} (x - x_0)^m. \quad (\text{A.2})$$

Let the log of the posterior distribution be

$$L(P) = \log(\pi(P | X)). \quad (\text{A.3})$$

By applying a Taylor series expansion to  $L(P)$  at  $P_0$  with first three terms,

$$\begin{aligned} L(P) &= L(P_0) + \frac{dL(P)}{dP} \Big|_{P_0} (P - P_0) + \frac{1}{2} \frac{d^2L(P)}{dP^2} \Big|_{P_0} \\ &\quad \cdot (P - P_0)^2 + \dots \\ &= \text{constant} + \frac{1}{2} \frac{d^2L(P)}{dP^2} \Big|_{P_0} (P - P_0)^2 + \dots \end{aligned} \quad (\text{A.4})$$

By taking the exponential of  $L(P)$ ,

$$\pi(P | X) \propto K \exp^{(1/2)(d^2L(P)/dP^2)} \Big|_{P_0} (P - P_0)^2, \quad (\text{A.5})$$

where  $K$  is a normalizing constant.

Let  $\mu = P_0$  and  $\sigma = 1/[-(d^2L(P))/dP^2 \Big|_{P_0}]^{1/2}$ . This gives

$$\pi(P | X) \approx \mathcal{N}(\mu, \sigma),$$

$$\pi(P_t | X_t) \sim \text{Beta}(x_t + \alpha, n_t + \beta - x_t)$$

$$\sim P_t^{x_t + \alpha - 1} (1 - P_t)^{n_t + \beta - x_t - 1}$$

$$\Rightarrow L(P) = k + (x_t + \alpha - 1) \log P_t$$

$$+ (n_t + \beta - x_t - 1) \log(1 - P_t),$$

$$\frac{dL(P_t)}{dP_t} = \frac{(x_t + \alpha - 1)}{P_t} - \frac{(n_t + \beta - x_t - 1)}{1 - P_t} = 0$$

$$\Rightarrow (1 - P_t)(x_t + \alpha - 1)$$

$$- P_t(n_t + \beta - x_t - 1) = 0,$$

$$\alpha - 1 + x_t + 2P_t - \alpha P_t - n_t P_t - \beta P_t = 0,$$

$$2P_t + x_t + \alpha - 1 - \alpha P_t - n_t P_t - \beta P_t = 0,$$

$$P_0 = \frac{1 - \alpha - x_t}{2 - \alpha - n_t - \beta}$$

$$\frac{dL(P_t)}{dP_t} = (x_t + \alpha - 1)P_t^{-1} - (n_t + \beta - x_t - 1)(1 - P_t)^{-1},$$

$$\frac{d^2}{dP_t^2} (\pi(P_t | X_t)) = -(x_t + \alpha - 1)P_t^{-2}$$

$$- [ -(-1)(1 - P_t)^{-2} (n_t + \beta - x_t - 1) ]$$

$$= \frac{-(x_t + \alpha - 1)}{P_t^2} - \frac{-(n_t + \beta - x_t - 1)}{(1 - P_t)^2},$$

$$\begin{aligned}
1 - P_0 &= 1 - \frac{1 - \alpha - x_t}{2 - \alpha - n_t - \beta} \\
&= \frac{2 - \alpha - n_t - \beta - 1 + x_t + \alpha}{2 - \alpha - n_t - \beta} \\
&= \frac{1 - n_t - \beta + x_t}{2 - \alpha - n_t - \beta}, \\
\frac{d^2}{dP_t^2}(\pi(P_t | X_t)) |_{P_0} &= \frac{(1 - x_t - \alpha)}{\left[ \frac{(1 - x_t - \alpha)/(2 - \alpha - n_t - \beta)}{(1 - n_t - \beta + x_t)} \right]^2} \\
&\quad + \frac{1}{\left[ \frac{(1 - n_t - \beta + x_t)/(2 - \alpha - n_t - \beta)}{(1 - x_t - \alpha)} \right]^2} \\
&= (1 - x_t - \alpha) \times \frac{(2 - \alpha - n_t - \beta)^2}{(1 - x_t - \alpha)^2} + (1 - n_t - \beta + x_t) \\
&\quad \times \frac{(2 - \alpha - n_t - \beta)^2}{(1 - n_t - \beta + x_t)^2} \\
&= \frac{(2 - \alpha - n_t - \beta)^2}{1 - x_t - \alpha} + \frac{(2 - \alpha - n_t - \beta)^2}{1 - n_t - \beta + x_t} \\
&= (2 - \alpha - n_t - \beta)^2 \left[ \frac{1 - n_t - \beta + x_t + 1 - x_t - \alpha}{(1 - x_t - \alpha)(1 - n_t - \beta + x_t)} \right] \\
&= \frac{(2 - \alpha - n_t - \beta)^3}{(1 - x_t - \alpha)(1 - n_t - \beta + x_t)}, \\
\sigma &= \frac{1}{\left[ -(d^2/dP_t^2)(\pi(P_t | X_t)) |_{P_0} \right]^{1/2}} \\
&= \frac{1}{\left[ \frac{(2 - \alpha - n_t - \beta)^3}{(1 - x_t - \alpha)(1 - n_t - \beta + x_t)} \right]^{1/2}}.
\end{aligned} \tag{A.6}$$

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

Saman Muthukumarana has been partially supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

## References

- [1] S. Wellek, *Testing Statistical Hypothesis of Equivalence and Noninferiority*, CRC Press, Boca Raton, FL, USA, 2010.
- [2] J. Albert, "Teaching inference about proportions using bayes and discrete models," *Journal of Statistics Education*, vol. 3, no. 3, 1995.
- [3] M. Gamalo, R. Wu, and R. Tiwari, "Bayesian approach to noninferiority trials for proportions," *Journal of Biopharmaceutical Statistics*, vol. 21, no. 5, pp. 902–919, 2011.
- [4] D. Rahardja and Y. D. Zhao, "Bayesian inference of a binomial proportion using one-sample misclassified binary data," *Model Assisted Statistics and Applications*, vol. 7, no. 1, pp. 17–22, 2012.
- [5] B. G. Zaslavsky, "Bayesian hypothesis testing in two-arm trials with dichotomous outcomes," *Biometrics*, vol. 69, no. 1, pp. 157–163, 2013.
- [6] J. B. Carlin, "Meta-analysis for 2x2 tables: a bayesian approach," *Statistics in Medicine*, vol. 11, no. 2, pp. 141–158, 1992.
- [7] S. Yusuf, R. Peto, J. Lewis, R. Collins, and P. Sleight, "Beta blockade during and after myocardial infarction: an overview of the randomized trials," *Progress in Cardiovascular Diseases*, vol. 27, no. 5, pp. 335–371, 1985.
- [8] C. Kpekpena, "Bayesian analysis of binary and count data in two-arm trials," M.S. thesis, University of Manitoba, Winnipeg, MB, Canada, 2014, <http://mspace.lib.umanitoba.ca/handle/1993/23588>.
- [9] R. E. Kass and A. E. Raftery, "Bayes factors," *Journal of the American Statistical Association*, vol. 90, no. 430, pp. 773–795, 1995.
- [10] S. Muthukumarana and R. C. Tiwari, "Meta-analysis using Dirichlet process," *Statistical Methods in Medical Research*, vol. 25, no. 1, pp. 352–365, 2016.
- [11] S. Nissen and K. Wolski, "Effect of rosiglitazone on the risk of myocardial infarction and death from cardiovascular causes," *New England Journal of Medicine*, vol. 356, no. 24, pp. 2457–2471, 2007.



Hindawi

Submit your manuscripts at  
[www.hindawi.com](http://www.hindawi.com)

