MARCEL DEKKER, INC. • 270 MADISON AVENUE • NEW YORK, NY 10016

JOURNAL OF BIOPHARMACEUTICAL STATISTICS
Vol. 12, No. 4, pp. 471–483, 2002

ON SAMPLE SIZE CALCULATION BASED ON ODDS RATIO IN CLINICAL TRIALS

Hansheng Wang,[1,*] Shein-Chung Chow,[1] and Gang Li[2]

[1]StatPlus, Inc., Heston Hall, Suite 206, 1790 Yardley-Langhorne Road, Yardley, PA 19067
[2]Department of Mathematics, Statistics and Computer Science, University of Illinois at Chicago, 851 S. Morgan St, Chicago, IL 60608

ABSTRACT

Sample size calculation formulas for testing equality, noninferiority, superiority, and equivalence based on odds ratio were derived under both parallel and one-arm crossover designs. An example concerning the study of odds ratio between a test compound (treatment) and a standard therapy (control) for prevention of relapse in subjects with schizophrenia and schizoaffective disorder is presented to illustrate the derived formulas for sample size calculation for various hypotheses under both a parallel design and a crossover design. Simulations were performed to assess the adequacy of the sample size calculation formulas. Simulation results were given at the end of the paper.

Key Words: Odds ratio; Therapeutic equivalence; Noninferiority; Superiority

INTRODUCTION

In clinical trials, it is often of interest to investigate the relative effect (e.g., risk or benefit) of the treatments for the disease under study. The odds ratio

*Corresponding author. E-mail: hwang@statplusinc.com

DOI: 10.1081/BIP-120016231
1054-3406 (Print); 1520-5711 (Online)

has been frequently used to assess the association between a binary exposure variable and a binary disease outcome (see Refs. [2,4]). Let p_T be the probability of observing an outcome of interest for a patient treated by a test compound (treatment) and p_C for a patient treated by a standard therapy (control). For patients receiving the treatment, the odds that a patient will have an outcome of interest over that he/she will not have an outcome of interest is defined as

$$O_T = \frac{p_T}{1 - p_T}.$$

Similarly, the odds for a patient receiving the control is defined as

$$O_C = \frac{p_C}{1 - p_C}.$$

As a result, the odds ratio between the treatment and the control is defined as

$$\text{OR} = \frac{O_T}{O_C} = \frac{p_T(1 - p_C)}{p_C(1 - p_T)}.$$

The odds ratio is always positive and has a range from 0 to ∞. When $\text{OR} = 1$, i.e., $p_T = p_C$, it implies that there is no difference between the treatment and control in terms of the outcome of interest. When $\text{OR} > 1$, treatment is more likely to produce the outcome of interest than control. Note that $(1 - \text{OR})$ is usually referred to as relative odds reduction in the literature. Intuitively, OR can be estimated by

$$\widehat{\text{OR}} = \frac{\hat{p}_T(1 - \hat{p}_C)}{\hat{p}_C(1 - \hat{p}_T)}, \tag{1}$$

where $\hat{p}_T$ and $\hat{p}_C$ are the maximum likelihood estimators of p_T and p_C, which are respectively given by

$$\hat{p}_T = \frac{x_T}{n_T} \text{ and } \hat{p}_C = \frac{x_C}{n_C},$$

where x_T and x_C are the respective numbers of outcomes of interest observed in the treatment and control groups, n_T and n_C are the numbers of patients receiving the treatment and the control, respectively. For a two-arm parallel trial, the asymptotic variance of $\log(\widehat{\text{OR}})$ can be obtained as

$$\text{var}[\log(\widehat{\text{OR}})] = \frac{1}{n_T p_T(1 - p_T)} + \frac{1}{n_C p_C(1 - p_C)}.$$

Note that the above asymptotic variance can be estimated by simply replacing p_T and p_C with their maximum likelihood estimators $\hat{p}_T$ and $\hat{p}_C$ (see Refs. [3,5]). However, the asymptotic variance of $\log(\widehat{\text{OR}})$ under a crossover design is unknown.

In clinical trials, commonly considered hypotheses include point hypotheses for testing equality and interval hypotheses for testing equivalence/noninferiority

MARCEL DEKKER, INC. • 270 MADISON AVENUE • NEW YORK, NY 10016

and superiority. In what follows, sample size calculations based on odds ratio between test and control will be derived under these hypotheses.

PARALLEL DESIGN

Formulas for sample size calculation for testing various hypotheses based on odds ratio will be derived under a parallel design.

Test for Equality

For testing equality of p_T and p_C between the treatment and control groups, the hypotheses of interest are given by

$$H_0 : \mathrm{OR} = 1 \ \text{ vs. } \ H_a : \mathrm{OR} \neq 1.$$

Testing the above hypotheses is equivalent to testing the following hypotheses

$$H_0 : \log(\mathrm{OR}) = 0 \ \text{ vs. } \ H_a : \log(\mathrm{OR}) \neq 0.$$

Under the null hypothesis, the test statistic

$$T = \log(\widehat{\mathrm{OR}}) \left[\frac{1}{n_T \hat{p}_T (1 - \hat{p}_T)} + \frac{1}{n_C \hat{p}_C (1 - \hat{p}_C)} \right]^{-1/2} \tag{2}$$

follows the standard normal distribution when n_T and n_C are sufficiently large. Thus, we reject the null hypothesis $H_0 : \log(\mathrm{OR}) = 0$ if $|T| > z_{\alpha/2}$, where $z_{\alpha/2}$ is the upper $(\alpha/2)$th percentile of the standard normal distribution. Under the alternative hypothesis $H_a : \log(\mathrm{OR}) \neq 0$, the power of the above test can be approximated by

$$\Phi \left(|\log(\mathrm{OR})| \left[\frac{1}{n_T p_T (1 - p_T)} + \frac{1}{n_C p_C (1 - p_C)} \right]^{-1/2} - z_{\alpha/2} \right),$$

where Φ is the cumulative distribution of the standard normal random variable. As a result, the sample size needed for achieving a desired power of $(1 - \beta)$ can be obtained by solving

$$|\log(\mathrm{OR})| \left[\frac{1}{n_T p_T (1 - p_T)} + \frac{1}{n_C p_C (1 - p_C)} \right]^{-1/2} - z_{\alpha/2} = z_\beta.$$

Under the assumption that $n_T / n_C = k$, we have

$$n_C = \frac{(z_{\alpha/2} + z_\beta)^2}{\log^2(\mathrm{OR})} \left(\frac{1}{k p_T (1 - p_T)} + \frac{1}{p_C (1 - p_C)} \right). \tag{3}$$

MARCEL DEKKER, INC. • 270 MADISON AVENUE • NEW YORK, NY 10016

Test for Noninferiority/Superiority

The problem of testing noninferiority and superiority can be unified by the following hypotheses:

$$H_0 : \text{OR} \leq \delta' \text{ vs. } H_a : \text{OR} > \delta',$$

where δ' is the noninferiority or superiority margin on the raw scale. If we let $\delta = \log(\delta')$, testing the above hypotheses is equivalent to testing the following hypotheses

$$H_0 : \log(\text{OR}) \leq \delta \text{ vs. } H_a : \log(\text{OR}) > \delta,$$

where δ is the noninferiority or superiority margin on the log-scale. When $\delta > 0$, the rejection of the null hypothesis indicates superiority over the reference value. When $\delta < 0$, the rejection of the null hypothesis implies noninferiority against the reference value.

When $\log(\text{OR}) = \delta$, the test statistic

$$T = (\log(\widehat{\text{OR}}) - \delta)\left[\frac{1}{n_T\hat{p}_T(1-\hat{p}_T)} + \frac{1}{n_C\hat{p}_C(1-\hat{p}_C)}\right]^{-1/2}$$

follows the standard normal distribution when n_T and n_C are sufficiently large. Thus, we reject the null hypothesis at the α level of significance if $T > z_\alpha$. Under the alternative hypothesis $H_a : \log(\text{OR}) > \delta$, the power of the above test is given by

$$\Phi\left((\log(\text{OR}) - \delta)\left[\frac{1}{n_T p_T(1-p_T)} + \frac{1}{n_C p_C(1-p_C)}\right]^{-1/2} - z_\alpha\right).$$

As a result, the sample size needed for achieving a desired power of $(1 - \beta)$ can be obtained by solving

$$(\log(\text{OR}) - \delta)\left[\frac{1}{n_T p_T(1-p_T)} + \frac{1}{n_C p_C(1-p_C)}\right]^{-1/2} - z_\alpha = z_\beta.$$

Under the assumption that $n_T/n_C = k$, we have

$$n_C = \frac{(z_\alpha + z_\beta)^2}{(\log(\text{OR}) - \delta)^2}\left(\frac{1}{kp_T(1-p_T)} + \frac{1}{p_C(1-p_C)}\right). \quad (4)$$

Test for Equivalence

To establish equivalence, the following hypotheses are usually considered

$$H_0 : |\text{OR}| \geq \delta' \text{ vs. } H_a : |\text{OR}| < \delta'.$$

Similarly, testing the above hypotheses is equivalent to testing the following hypotheses

$$H_0 : |\log(\mathrm{OR})| \geq \delta \text{ vs. } H_a : |\log(\mathrm{OR})| < \delta,$$

where $\delta = \log(\delta')$. The above hypotheses can be tested using the two one-sided tests (TOST) procedure (see Ref. [1]). We then reject the null hypothesis at the α level of significance if

$$(\log(\widehat{\mathrm{OR}}) - \delta)\left[\frac{1}{n_T\hat{p}_T(1-\hat{p}_T)} + \frac{1}{n_C\hat{p}_C(1-\hat{p}_C)}\right]^{-1/2} < -z_\alpha$$

and

$$(\log(\widehat{\mathrm{OR}}) + \delta)\left[\frac{1}{n_T\hat{p}_T(1-\hat{p}_T)} + \frac{1}{n_C\hat{p}_C(1-\hat{p}_C)}\right]^{-1/2} > z_\alpha.$$

Then the power of the above test can be approximated by

$$\begin{cases} 2\Phi\left(\delta\left[\frac{1}{n_Tp_T(1-p_T)} + \frac{1}{n_Cp_C(1-p_C)}\right]^{-1/2} - z_\alpha\right) - 1 & \text{if } \log(\mathrm{OR}) = 0 \\ \Phi\left((\delta - |\log(\mathrm{OR})|)\left[\frac{1}{n_Tp_T(1-p_T)} + \frac{1}{n_Cp_C(1-p_C)}\right]^{-1/2} - z_\alpha\right) & \text{if } \log(\mathrm{OR}) \neq 0. \end{cases}$$

Under the assumption that $n_T/n_C = k$, the sample size is given by

$$\begin{cases} n_C = \frac{(z_\alpha + z_{\beta/2})^2}{\delta^2}\left[\frac{1}{kp_T(1-p_T)} + \frac{1}{p_C(1-p_C)}\right] & \text{if } \log(\mathrm{OR}) = 0 \\ n_C = \frac{(z_\alpha + z_\beta)^2}{(\delta - |\log(\mathrm{OR})|)^2}\left[\frac{1}{kp_T(1-p_T)} + \frac{1}{p_C(1-p_C)}\right] & \text{if } \log(\mathrm{OR}) \neq 0. \end{cases} \quad (5)$$

Note that very similar formulas were also obtained by Tu.[5]

CROSSOVER DESIGN

Consider a one-arm crossover design. For simplicity, we assume that there are no period effects and carryover effects. Without loss of generality, we further assume that every subject will first receive the control and then the treatment. Let $x_{ij}, j = 1, \ldots, n$ be 1 if an outcome of interest is observed from the jth subject in the ith period and 0 otherwise, then the number of outcomes of interest observed in the control group is given by $x_C = \sum_{j=1}^{n} x_{1j}$. And x_T can be similarly defined. Then the odds ratio between the treatment and control can be estimated according to

MARCEL DEKKER, INC. • 270 MADISON AVENUE • NEW YORK, NY 10016

Eq. (1). By Taylor's expansion, we can obtain

$$\sqrt{n}(\log(\widehat{\mathrm{OR}}) - \log(\mathrm{OR})) = \sqrt{n}\left(\frac{1}{p_{\mathrm{C}}(1-p_{\mathrm{C}})}(\hat{p}_{\mathrm{C}} - p_{\mathrm{C}}) - \frac{1}{p_{\mathrm{T}}(1-p_{\mathrm{T}})}(\hat{p}_{\mathrm{T}} - p_{\mathrm{T}})\right)$$
$$+ o_p(1) = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}\left(\frac{x_{1j} - p_{\mathrm{C}}}{p_{\mathrm{C}}(1-p_{\mathrm{C}})} - \frac{x_{2j} - p_{\mathrm{T}}}{p_{\mathrm{T}}(1-p_{\mathrm{T}})}\right)$$
$$+ o_p(1) \rightarrow_d N(0, \sigma_d^2),$$

where

$$\sigma_d^2 = \mathrm{var}\left(\frac{x_{1j} - p_{\mathrm{C}}}{p_{\mathrm{C}}(1-p_{\mathrm{C}})} - \frac{x_{2j} - p_{\mathrm{T}}}{p_{\mathrm{T}}(1-p_{\mathrm{T}})}\right). \tag{6}$$

Let

$$d_j = \left(\frac{x_{1j}}{\hat{p}_{\mathrm{C}}(1-\hat{p}_{\mathrm{C}})} - \frac{x_{2j}}{\hat{p}_{\mathrm{T}}(1-\hat{p}_{\mathrm{T}})}\right).$$

Then, σ_d^2 can be estimated by sample variance of d_j, $j = 1,\ldots,n$, which is denoted by $\hat{\sigma}_d^2$.

Test for Equality

Similarly, the hypotheses of interest are given by

$$H_0 : \mathrm{OR} = 1 \ \text{ vs. } \ H_{\mathrm{a}} : \mathrm{OR} \neq 1.$$

Testing the above hypotheses is equivalent to testing the following hypotheses

$$H_0 : \log(\mathrm{OR}) = 0 \ \text{ vs. } \ H_{\mathrm{a}} : \log(\mathrm{OR}) \neq 0.$$

Under the null hypothesis, the test statistic

$$T = \frac{\sqrt{n}\log(\widehat{\mathrm{OR}})}{\hat{\sigma}_d}$$

follows the standard normal distribution when n is sufficiently large. Thus, we reject the null hypothesis $H_0 : \mathrm{OR} = 1$ if $|T| > z_{\alpha/2}$. Under the alternative hypothesis $H_{\mathrm{a}} : \mathrm{OR} \neq 1$, the power of the above test can be approximated by

$$\Phi\left(\frac{\sqrt{n}|\log(\mathrm{OR})|}{\sigma_d} - z_{\alpha/2}\right).$$

As a result, the sample size needed for achieving a desired power of $(1 - \beta)$ can be obtained by solving

$$\frac{\sqrt{n}|\log(\mathrm{OR})|}{\sigma_d} - z_{\alpha/2} = z_{\beta}.$$

This leads to

$$n = \frac{(z_{\alpha/2} + z_\beta)^2 \sigma_d^2}{\log^2(\mathrm{OR})}. \tag{7}$$

Test for Noninferiority/Superiority

As indicated earlier, the problem of testing noninferiority and superiority can be unified by the following hypotheses:

$$H_0 : \log(\mathrm{OR}) \leq \delta \text{ vs. } H_a : \log(\mathrm{OR}) > \delta,$$

where $\delta = \log(\delta')$ is the noninferiority or superiority margin on log-scale. When $\log(\mathrm{OR}) = \delta$, the test statistic

$$T = \frac{\sqrt{n}(\log(\widehat{\mathrm{OR}}) - \delta)}{\hat{\sigma}_d}$$

follows the standard normal distribution when n is sufficiently large. Thus, we reject the null hypothesis at the α level of significance if $T > z_\alpha$. Under the alternative hypothesis $H_a : \log(\mathrm{OR}) > \delta$, the power of the above test is approximated by

$$\Phi\left(\frac{\sqrt{n}(\log(\mathrm{OR}) - \delta)}{\sigma_d} - z_\alpha\right).$$

As a result, the sample size needed for achieving a desired power of $1 - \beta$ can be obtained by solving

$$\frac{\sqrt{n}(\log(\mathrm{OR}) - \delta)}{\sigma_d} - z_\alpha = z_\beta.$$

Thus, we have

$$n_C = \frac{(z_\alpha + z_\beta)^2 \sigma_d^2}{(\log(\mathrm{OR}) - \delta)^2}. \tag{8}$$

Test for Equivalence

To establish equivalence, the following hypotheses are considered

$$H_0 : |\log(\mathrm{OR})| \geq \delta \text{ vs. } H_a : |\log(\mathrm{OR})| < \delta.$$

We then reject the null hypothesis at the α level of significance if

$$\frac{\sqrt{n}(\log(\widehat{\mathrm{OR}}) - \delta)}{\hat{\sigma}_d} < -z_\alpha$$

and

$$\frac{\sqrt{n}(\log(\widehat{\mathrm{OR}}) + \delta)}{\hat{\sigma}_d} > z_\alpha.$$

The power of the above test is approximated by

$$\begin{cases} 2\Phi\left(\frac{\sqrt{n}\delta}{\sigma_d} - z_\alpha\right) - 1 & \text{if } \log(\mathrm{OR}) = 0 \\ \Phi\left(\frac{\sqrt{n}(\delta - |\log(\mathrm{OR})|)}{\sigma_d} - z_\alpha\right) & \text{if } \log(\mathrm{OR}) \neq 0. \end{cases}$$

Then the sample size is given by

$$\begin{cases} n = \frac{(z_\alpha + z_{\beta/2})^2 \sigma_d^2}{\delta^2} & \text{if } \log(\mathrm{OR}) = 0 \\ n = \frac{(z_\alpha + z_\beta)^2 \sigma_d^2}{(\delta - |\log(\mathrm{OR})|)^2} & \text{if } \log(\mathrm{OR}) \neq 0. \end{cases} \tag{9}$$

AN EXAMPLE

Parallel Design

Suppose a sponsor is interested in conducting a clinical trial to study the relative risk between a test compound (treatment) and a standard therapy (control) for prevention of relapse in subjects with schizophrenia and schizoaffective disorders. Based on the results from a previous study with 365 subjects (i.e., 177 subjects received the test compound and 188 received the standard therapy), about 25% (45/177) and 40% (75/188) of subjects receiving the test compound and the standard therapy experienced relapse after the treatment, respectively. Subjects who experienced the first relapse may withdraw from the study or stay in the trial. The sponsor is interested in studying the odds ratio of the test compound as compared to the standard therapy for prevention of experiencing the first relapse. In addition, it is also of interest to examine the odds ratio for prevention of experiencing the second relapse.

Test for Equality

Assuming the relapse rates in the test group and the control group are 25% and 40%, respectively. This yields a relative risk of

$$\mathrm{OR} = \frac{0.40(1 - 0.25)}{(1 - 0.40)0.25} = 2.$$

According to Eq. (3) and $n = n_\mathrm{T} = n_\mathrm{C}$ $(k = 1)$, the sample size per treatment

group needed in order to achieve an 80% ($\beta = 0.2$) power at 5% ($\alpha = 0.05$) level of significance is given by

$$n = \frac{(z_{0.05/2} + z_{0.2})^2}{\log^2(2)} \left(\frac{1}{0.4(1-0.4)} + \frac{1}{0.25(1-0.25)} \right)$$

$$= \frac{(1.96 + 0.84)^2}{0.69^2}(4.17 + 5.33) = 156.4 \approx 157.$$

Test for Superiority

Suppose that a 20% difference in odds ratio (in log scale) is considered of clinical importance. Hence, the superiority margin is chosen to be $\delta = 0.2$. According to Eq. (4), the sample size per treatment group required for achieving an 80% power ($\beta = 0.2$) is given by

$$n = \frac{(z_{0.05} + z_{0.2})^2}{(\log(2) - 0.2)^2} \left(\frac{1}{0.4(1-0.4)} + \frac{1}{0.25(1-0.25)} \right)$$

$$= \frac{(1.64 + 0.84)^2}{0.49^2}(4.17 + 5.33) = 243.4 \approx 244.$$

Test for Equivalence

Assuming that (i) the true relapse rate of the test compound is approximately the same as that of the standard therapy (i.e., $\log(\text{OR}) = 0$) and (ii) the equivalence limit of the odds ratio (in log scale) is 50% ($\delta = 0.50$). According to Eq. (5) the sample size needed to achieve an 80% ($\beta = 0.2$) power for establishment of equivalence is given by

$$n = \frac{(z_{0.05} + z_{0.2/2})^2}{0.5^2} \left(\frac{1}{0.25(1-0.25)} + \frac{1}{0.25(1-0.25)} \right)$$

$$= \frac{(1.64 + 1.28)^2}{0.25}(5.33 + 5.33) = 363.6 \approx 364.$$

Crossover Design

Now suppose that the sponsor consider to adopt a one-arm crossover design for this trial. In other words, each subject will first receive the standard therapy (control) then receive the test compound (treatment). Using the same example, it is

assumed that the relapse rate for patients in the standard therapy is 25%. At the next dosing period, the relapse rate increases to 40% after receiving the test compound. Also, assuming that σ_d in Eq. (6) is 2.5. The sample size calculation for various hypotheses can be carried out as follows.

Test for Equality

The relative risk between the standard therapy and the test compound is given by

$$\text{OR} = \frac{0.40(1 - 0.25)}{(1 - 0.40)0.25} = 2.$$

According to Eq. (7), the sample size needed in order to achieve an 80% ($\beta = 0.2$) power at the 5% ($\alpha = 0.05$) level of significance is given by

$$n = \frac{(z_{0.05/2} + z_{0.2})^2 \sigma_d^2}{\log^2(2)} = \frac{(1.96 + 0.84)^2 2.5^2}{0.69^2} = 102.9 \approx 103.$$

Test for Superiority

Suppose that a 20% difference in odds ratio (in log scale) is considered of clinical importance. Hence, the superiority margin is chosen to be $\delta = 0.2$. According to Eq. (8), the sample size needed in order to achieve a power of 80% ($\beta = 0.2$) is given by

$$n = \frac{(z_{0.05} + z_{0.2})^2 2.5^2}{(\log(2) - 0.2)^2} = \frac{(1.64 + 0.84)^2 2.5^2}{0.49^2} = 160.1 \approx 161.$$

Test for Equivalence

Assume that (i) there is no difference in the true relapse rate between the standard therapy and the test compound ($\log(\text{OR}) = 0$) and (ii) the equivalence limit of the odds ratio (in log scale) is 50%. Accordig to Eq. (9), the sample size needed in order to achieve a power of 80% ($\beta = 0.2$) is given by

$$n = \frac{(z_{0.05} + z_{0.2/2})^2 2.5^2}{0.5^2} = \frac{(1.64 + 1.28)^2 2.5^2}{0.25} = 213.2 \approx 214.$$

MARCEL DEKKER, INC. • 270 MADISON AVENUE • NEW YORK, NY 10016

REMARKS

For testing the following hypotheses under a parallel design

$$H_0 : \log(\mathrm{OR}) = 0 \text{ vs. } H_a : \log(\mathrm{OR}) \neq 0,$$

an alternative test can be constructed as follows

$$T^* = \log(\widehat{\mathrm{OR}})\left[\left(\frac{1}{n_T}+\frac{1}{n_C}\right)\left(\frac{1}{\hat{p}(1-\hat{p})}\right)\right]^{-1/2}, \tag{10}$$

where

$$\hat{p} = \frac{n_T\hat{p}_T + n_C\hat{p}_C}{n_T + n_C}.$$

Under the null hypothesis $H_0 : \log(\mathrm{OR}) = 0$, T^* is asymptotically distributed as the standard normal variable. Hence, we reject the null hypothesis at the α level of significance if $|T^*| > z_{\alpha/2}$. It should be noted that formula (10) is very similar to formula (2) except that the estimate of the variance of $\log(\widehat{\mathrm{OR}})$ is different. In Eq. (2), $\mathrm{var}(\log(\widehat{\mathrm{OR}}))$ is estimated by the maximum likelihood estimate (MLE) without any constraints, while in Eq. (10) the same quantity is estimated by the MLE under the null hypothesis that $\log(\mathrm{OR}) = 0$. We will refer to Eq. (2) as the unconditional method and Eq. (10) as the conditional method. Under the null hypothesis, both methods have an asymptotic size of α. In practice, it is a dilemma regarding which method (i.e., unconditional or conditional) should be used because one is not necessarily more powerful than the other. However, the conditional approach for testing noninferiority/superiority and equivalence is

Table 1. The Simulation Result with Nominal Significance Level 0.05 of the Test for Equality Under the Parallel Design

p_T	p_C	n_T	n_C	Power	p_T	p_C	n_T	n_C	Power
0.20	0.30	426	213	0.793	0.40	0.55	262	131	0.808
0.20	0.35	200	100	0.799	0.40	0.60	150	75	0.826
0.20	0.40	118	59	0.796	0.45	0.55	590	295	0.800
0.25	0.35	482	241	0.794	0.45	0.60	264	132	0.810
0.25	0.40	224	112	0.800	0.45	0.65	150	75	0.819
0.25	0.45	130	65	0.800	0.50	0.60	588	294	0.798
0.30	0.40	526	263	0.789	0.50	0.65	262	131	0.814
0.30	0.45	242	121	0.797	0.50	0.70	148	74	0.821
0.30	0.50	140	70	0.810	0.55	0.65	576	288	0.804
0.35	0.45	560	280	0.793	0.55	0.70	254	127	0.818
0.35	0.50	254	127	0.801	0.55	0.75	144	72	0.829
0.35	0.55	146	73	0.810	0.60	0.70	550	275	0.809
0.40	0.50	580	290	0.800	0.60	0.75	242	121	0.818

Table 2. The Simulation Result with Nominal Significance Level 0.05 of the Test for Equivalence Under the Parallel Design

p_T	p_C	n_T	n_C	Power	p_T	p_C	n_T	n_C	Power
0.30	0.30	122	61	0.785	0.45	0.60	494	247	0.798
0.30	0.35	140	70	0.780	0.50	0.50	102	51	0.769
0.30	0.40	260	130	0.800	0.50	0.55	116	58	0.768
0.30	0.45	636	318	0.798	0.50	0.60	216	108	0.784
0.35	0.35	112	56	0.784	0.50	0.65	544	272	0.785
0.35	0.40	128	64	0.785	0.55	0.55	104	52	0.790
0.35	0.45	228	114	0.802	0.55	0.60	120	60	0.766
0.35	0.50	528	264	0.796	0.55	0.65	234	117	0.787
0.40	0.40	108	54	0.783	0.55	0.70	672	336	0.784
0.40	0.45	120	60	0.773	0.60	0.60	108	54	0.785
0.40	0.50	212	106	0.800	0.60	0.65	130	65	0.772
0.40	0.55	488	244	0.786	0.60	0.70	272	136	0.780
0.45	0.45	104	52	0.794	0.65	0.65	112	56	0.779
0.45	0.50	116	58	0.774	0.65	0.70	144	72	0.773
0.45	0.55	210	105	0.800	0.70	0.70	122	61	0.781

difficult to carry out due to the fact that the MLE under the null hypothesis is almost impossible to track.

SIMULATION

A total of four simulation studies were carried out to evaluate the finite sample performance of the derived sample formulae. Since both tests for equality

Table 3. The Simulation Result with Nominal Significance Level 0.05 of the Test for Equality Under the Crossover Design

p_T	p_C	σ_d	n	Power	p_T	p_C	σ_d	n	Power
0.20	0.30	3.004753	244	0.804	0.40	0.55	2.599919	145	0.802
0.20	0.35	2.971249	118	0.829	0.40	0.60	2.648516	84	0.818
0.20	0.40	2.983305	73	0.830	0.45	0.55	2.533354	313	0.792
0.25	0.35	2.724457	254	0.786	0.45	0.60	2.575073	142	0.813
0.25	0.40	2.817950	130	0.815	0.45	0.65	2.636148	82	0.806
0.25	0.45	2.799687	77	0.820	0.50	0.60	2.571822	316	0.825
0.30	0.40	2.658232	285	0.792	0.50	0.65	2.628982	142	0.817
0.30	0.45	2.675998	135	0.813	0.50	0.70	2.706164	81	0.810
0.30	0.50	2.714261	81	0.800	0.55	0.65	2.588022	301	0.815
0.35	0.45	2.614447	307	0.822	0.55	0.70	2.679054	135	0.804
0.35	0.50	2.619666	141	0.776	0.55	0.75	2.827535	78	0.844
0.35	0.55	2.648626	82	0.810	0.60	0.70	2.692324	292	0.808
0.40	0.50	2.541418	309	0.817	0.60	0.75	2.801336	129	0.794

Table 4. The Simulation Result with Nominal Significance Level 0.05 of the Test for Equivalence Under the Crossover Design

p_T	p_C	σ_d	n	Power	p_T	p_C	σ_d	n	Power
0.30	0.20	2.959132	255	0.779	0.55	0.40	2.599655	270	0.797
0.35	0.20	2.963281	1003	0.814	0.55	0.45	2.556995	113	0.782
0.35	0.25	2.820795	182	0.784	0.60	0.45	2.596150	269	0.782
0.40	0.25	2.835270	528	0.796	0.60	0.50	2.529833	112	0.796
0.40	0.30	2.666669	142	0.797	0.65	0.50	2.618966	293	0.781
0.45	0.30	2.709451	364	0.802	0.65	0.55	2.615441	126	0.787
0.45	0.35	2.579485	122	0.773	0.70	0.55	2.659312	351	0.778
0.50	0.35	2.603712	289	0.786	0.70	0.60	2.680022	143	0.779
0.50	0.40	2.544874	114	0.796	0.75	0.60	2.794186	513	0.778

and superiority/noninferiority are similar to each other. The simulation studies were only performed for testing equality and equivalence under both parallel and crossover design. All simulations were performed based on 1000 iterations. For crossover design, correlated binary random variables [e.g., (Z_T, Z_R)] were generated according to the following procedure:

1. Generate independent uniform (0,1) random variables denoted by X_0, X_T, X_R.
2. Let Y_i, $i = T, R$ be a random variable with probability 0.5 be X_0 and 0.5 be X_i.
3. For a given p_i, $i = T, R$, Z_i is defined to be 1 if $Y_i < p_i$, and 0 otherwise.

It can be noted that Z_i is a binary response with success probability p_i. Since, there is 0.25 probability $Y_T = Y_R = X_0$, Z_T and Z_R is correlated in someway.

The simulation results are summarized in Tables 1–4. From those tables it can be seen clearly that our sample size formula works satisfactorily with usually no more than 3% deviation from the target value.

REFERENCES

1. Chow, S.C.; Liu, J.P. *Design and Analysis of Clinical Trials*; John Wiley & Sons: New York, 1998.
2. Deeks, J. Swots Corner: What Is an Odds Ratio? Bandolllier **1996**, *25*, 6–9.
3. Haldane, J.B.S. The Estimation and Significance of Logarithm of a Ratio of Frequencies. Ann. Hum. Genet. **1955**, *20*, 309–311.
4. Lachin, J.M. *Biostatistical Methods: The Assessment of Relative Risk*; John Wiley & Sons: New York, 2000.
5. Tu, D. On the Use of the Ratio or Odds Ratio of Cure Rates in Therapeutic Equivalence Clinical Trials with Binary Endpoints. J. Biopharm. Stat. **1998**, *8*, 263–282.