A STATISTICAL INVESTIGATION INTO NONINFERIORITY TESTING FOR TWO BINOMIAL PROPORTIONS

by

NICHOLAS BLOEDOW

B.S. University of Wisconsin-Oshkosh, 2011

A REPORT

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Statistics
College of Arts and Science

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2015

Approved by:

Major Professor
Dr. Christopher Vahl

Copyright

Abstract

In clinical research, noninferiority trials are becoming an important tool for investigating whether a new treatment is useful. The outcome measured can be either continuous (e.g. blood pressure level), time-to-event (e.g. days until heart attack), or binary (e.g. death). Rather than showing that the new treatment is superior to an active control, i.e. standard drug or treatment already available, one tests whether the new treatment is not meaningfully worse than the active control.

Here we consider a binary outcome such as success or failure following an intervention. Evaluation of the treatment relative to control becomes a comparison of two binomial proportions; without loss of generality it will be assumed the larger the probability of success for an intervention the better. Simulation studies under these assumptions were programmed over a variety of different sample sizes and true population proportions to determine the performance between asymptotic noninferiority methods based on calculations of risk differences (with and without a continuity correction), relative risks, and odds ratio from two independent samples. Investigation was done to compare type I error rates, power when true proportions were exactly the same, and power when the true proportion for treatment group was less than the control, but not meaningfully inferior. Simulation results indicate most analysis methods have comparable type I error rates; however, the method based on relative risk has higher power under most circumstances. Due to the ease of interpretation with the relative risk, its use is recommended for establishing noninferiority of a binomial proportion between 0.2 and 0.8.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I want to dedicate this work to everyone (family, friends, classmates, and faculty) who has supported me through this process. Special recognition goes out to my major professor, Dr. Christopher Vahl, who presented me with this topic and gave continual guidance/encouragement. I would also like to thank Dr. Leigh Murray and Dr. Abigail Jager for their willingness to serve on my committee and offer up insightful comments/suggestions. Furthermore, I want to express my sincerest gratitude to the Department of Statistics at Kansas State University for providing me with the opportunity to complete my Master's degree. Over these past three years in Manhattan (KS), it has been a pleasure to study within this department and become member of the Kansas State Family. GO STATE!

Chapter 1 - Introduction

The implementation of noninferiority testing has become quite popular with clinical and pharmaceutical research. Noninferiority testing is an important tool for determining whether a new treatment is useful in comparison to standard drug or treatment already out in the market. These (new) therapies offer advantages such as fewer side effects, lower cost, easier application, or fewer drug interactions (Walker and Nowacki, 2010). *Ceteris paribus*, the new treatment would not necessarily need to be superior, nor even equivalent, to the standard in order to be considered beneficial. A small loss of efficacy could be tolerated in order to gain advantages similar to those described above. The tolerable trade-off of efficacy can be considered a "zone of indifference" between treatments and is commonly referred to as the noninferiority margin.

The outcome measured can be either continuous (e.g. blood pressure level), time-to-event (e.g. days until heart attack), or binary (e.g. death). Here we consider a binary outcome such as success or failure following an intervention. Evaluation of the treatment relative to control becomes a comparison of two binomial proportions. Without loss of generality it will be assumed the larger the probability of success for an intervention the better.

Test of Statistical Superiority

When learning about hypothesis testing, the idea of one-sided hypotheses are introduced first and commonly referred to as superiority tests. Let p_T be the proportion of success for the treatment group (i.e. those individuals who are receiving the new experimental intervention) and p_C be the proportion of success for the control group (i.e. those receiving the standard treatment). To establish the superiority of treatment over control, one would test the following hypotheses:

$$H_O\colon\ p_T \leq p_C$$
$$H_A\colon\ p_T > p_C\ .$$

If H_O is rejected in favor of H_A, then the new treatment is considered to be statistically superior to the control. However, it does not demonstrate that the treatment is superior by a meaningful amount. Let $\delta > 0$ and be a clinically meaningful difference in the proportions. This is summarized by the following hypotheses:

$$H_O:\ p_T \le p_C + \delta$$
$$H_A:\ p_T > p_C + \delta\ .$$

If H_O is rejected is in favor of H_A, then the new treatment is considered to be clinically superior to the control.

Difference Testing

Difference testing extends the concept of statistical superiority and converts it into a two-sided test that shows the proportion of success for the treatment group is either inferior or superior to the control. To establish difference between treatment and control, one would test the following hypotheses:

$$H_O:\ p_T = p_C$$
$$H_A:\ p_T \ne p_C\ .$$

If H_O is rejected is in favor of H_A, then the new treatment is considered to be different from the control.

One must be cautious to avoid misinterpretation when failing to reject the null hypothesis. It’s tempting to consider the two groups are equal to one another once the trial has been deemed “negative.” However, as pointed out by Altman and Bland (1995), “absence of evidence is not evidence of absence.” The logic of hypothesis testing disallows you to support significance for H_O due to the “burden of proof” being placed on the wrong (research) hypothesis (Walker and Nowacki, 2010). Doing so could increase the potential for making a type II error since power is not being properly controlled. Therefore if the goal is to demonstrate that the two treatments are the same, you would have to perform a test of equivalence.

Equivalence Testing

Equivalence testing, in its essence, reverses the null and alternative hypotheses in a difference test. The null hypothesis assumes that the absolute difference between the two proportions is greater than or equal to some clinically meaningful amount, i.e. the equivalence margin. Any difference within that equivalence margin, i.e. zone of indifference, would show equivalency between the two proportions. To establish equivalence between treatment and control, one would test the following hypotheses:

$$H_O: |\, p_T - p_C| \geq \delta$$
$$H_A: \ |p_T - p_C| < \delta \ .$$

Here the value of δ is the specified equivalence margin for the difference between binomial proportions. If H_O is rejected is in favor of H_A, then the new treatment group is considered to be equivalent to the control.

Noninferiority Testing

Note that the alternative hypothesis for an equivalence test defines a lower and an upper bound for the difference, i.e. $H_A: -\delta < p_T - p_C < \delta$. Because the upper bound is irrelevant for establishing noninferiority, the alternative hypothesis for a test of noninferiority is one-sided because you are interested in showing that the new treatment is not meaningfully worse than the active control. Being not meaningfully worse incorporates the cases where the proportions are equivalent or the treatment is superior to the control. The lower equivalence bound becomes the noninferiority margin. Careful consideration must be made to make sure you have an appropriate margin. Having too large of a margin will cause you to show noninferiority more frequently than is warranted, and vice versa.

Comparison of Differences

There are three different measures (risk difference, relative risk, and odds ratio) used to compare the differences between the binomial proportions of the treatment and control group.

Risk Difference

The easiest comparison to calculate and interpret is simply the difference between binomial proportions ("risk") of success. The null hypothesis for the noninferiority test is

$$H_O: \ d = p_T - p_C \leq -\delta$$

versus the alternative

$$H_A: \ d = p_T - p_C > -\delta$$

where d is the risk difference. Here the value of δ is the specified noninferiority margin for tests involving the risk difference between binomial proportions.

Relative Risk

Another comparison of difference can be measured by the ratio of the two binomial proportions. The null hypothesis for this noninferiority test is

$$H_O\colon\ RR = \frac{p_T}{p_C} \leq 1 - \phi$$

versus the alternative

$$H_A\colon RR = \frac{p_T}{p_C} > 1 - \phi$$

where RR is the relative risk. The value $1 - \phi$ is the specified noninferiority margin for tests involving the relative risk between binomial proportions, such that $\phi \geq 0$. Let ϕ be the smallest unacceptable percentage reduction of the proportion of success for the treatment group relative to the control.

Odds Ratio

The last measure is the ratio of odds of success between the two groups. The null hypothesis for the noninferiority test is

$$H_O\colon OR = \frac{p_T/(1-p_T)}{p_C/(1-p_C)} = \frac{p_T(1-p_C)}{p_C(1-p_T)} \leq 1 - \psi$$

versus the alternative

$$H_A\colon OR = \frac{p_T/(1-p_T)}{p_C/(1-p_C)} = \frac{p_T(1-p_C)}{p_C(1-p_T)} > 1 - \psi$$

where OR is the odds ratio. The value $1 - \psi$ is the specified noninferiority margin for tests involving the odds ratio between binomial proportions, such that $\psi \geq 0$. Let ψ be the smallest unacceptable percentage reduction of the odds for the treatment group relative to the control.

Comparison of Scales

To get a better understanding the range of the risk difference, relative risk, and odds ratio, a contour plot (Figure 1.1) and 3D plot (Figure 1.2) were constructed with the proportions of success ranging from 0.05 to 0.95 for the treatment and control groups. R code is provided in Appendix D (Contour & 3D Plots of Noninferiority Parameters). The values of the risk difference can range anywhere from -1 to 1. Zero represents the case when the proportion of success for the treatment group is equal to the control. The contours of risk difference are lines equidistant from one another and form the surface of a tilted plane over the unit square. The values of the relative risk and odds ratio can range anywhere from 0 to ∞. One represents the case when the proportion of success for the treatment group is equal to the control. The contours of the relative risk are lines that radiate from the origin. The contours of odds ratio exhibit a nonlinear pattern except for the case where the two proportions are equal. Since it is assumed without loss of generality that the larger the probability of success for an intervention is better, our primary focus will be in the lower right hand corner of the plots. Therefore, we assume the proportion of success for the treatment group to be less than or equal to the control.

Figure 1.1 Contour plots of risk difference, relative risk, and odds ratio relative to probabilities of success for treatment and control ranging from 0.05 to 0.95

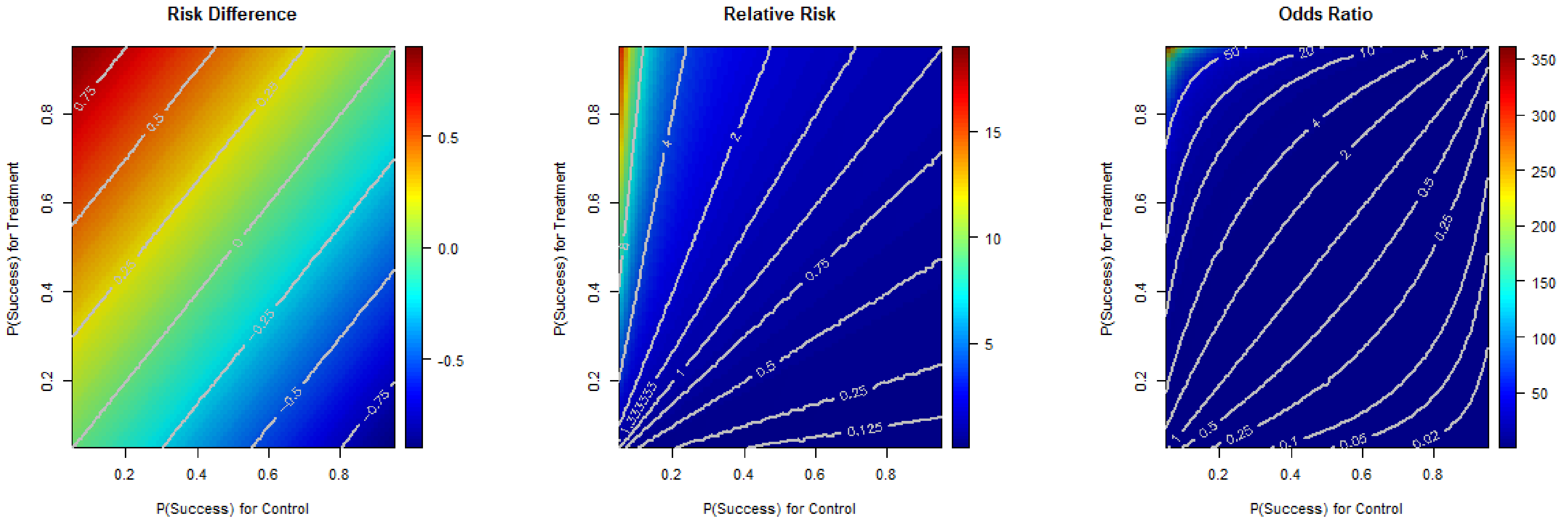
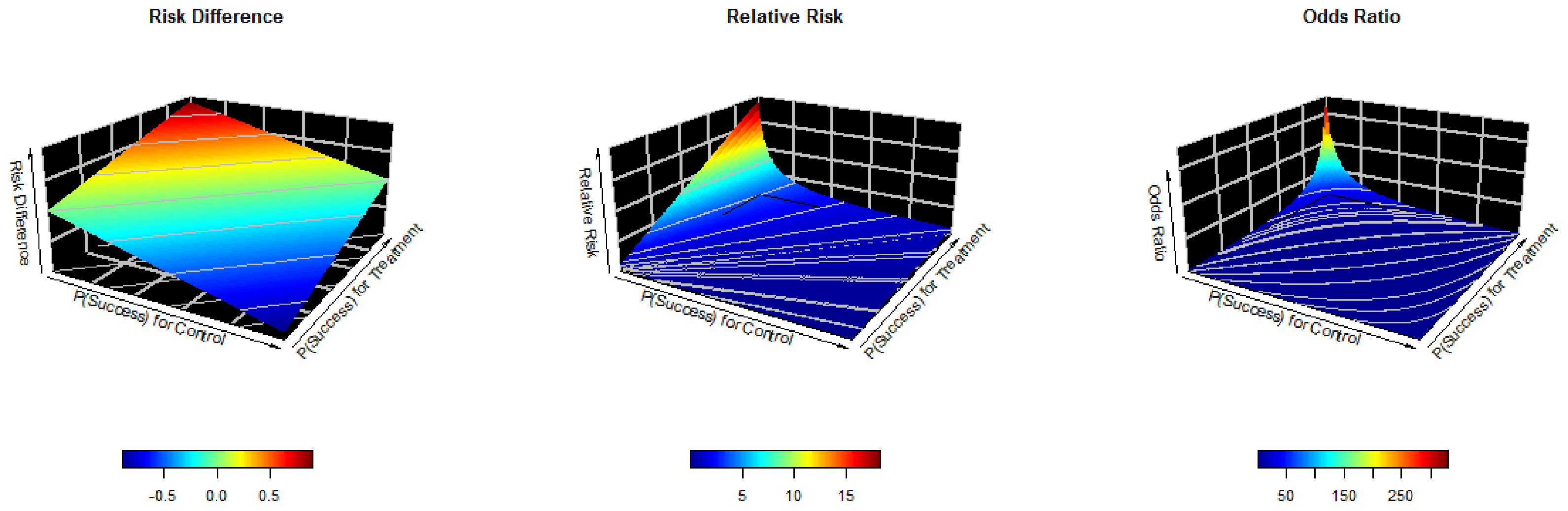


Figure 1.2 3D plots of risk difference, relative risk, and odds ratio relative to probabilities of success for treatment and control ranging from 0.05 to 0.95



Risk Difference

Relative Risk

Odds Ratio

Chapter 2 - Statistical Methods

Binomial Proportions

Here we consider a binary outcome such as success or failure following an intervention. Evaluation of the treatment relative to control becomes a comparison of two binomial proportions; without loss of generality it will be assumed the larger the probability of success for an intervention the better. The counts for the outcomes from each group can be summarized into a $2\ x\ 2$ contingency table (Table 2.1).

Table 2.1 Contingency table ($2\ x\ 2$) of the outcomes (success or failure) from each group

Group	Success	Failure	Sample Size
Treatment	x	$n_T - x$	n_T
Control	y	$n_C - y$	n_C
Total	s	$n_T + n_C - s$	$n_T + n_C$

The proportion of success for the treatment group (p_T) is estimated by $\hat{p}_T = x/n_T$, where x is the number of successes and n_T is the sample size for the treatment group. The proportion of success for the control group (p_C) is estimated by $\hat{p}_C = y/n_C$, where y is the number of successes and n_C is the sample size for the control group. The statistics $\hat{p}_T$ and $\hat{p}_C$ are the respective unrestrained maximum likelihood estimates (MLEs) of p_T and p_C. The risk difference between the proportion of success for the treatment group relative to the control group $d = p_T - p_C$ is estimated by $\hat{d} = \hat{p}_T - \hat{p}_C$. The relative risk between the proportion of success for the treatment group relative to the control group $RR = p_T/p_C$ is estimated by $\widehat{RR} = \hat{p}_T/\hat{p}_C$. The odds ratio between the odds of success for the treatment group relative to the control group $OR = p_T(1 - p_C)/p_C(1 - p_T)$ is estimated by $\widehat{OR} = \hat{p}_T(1 - \hat{p}_C)/\hat{p}_C(1 - \hat{p}_T)$.

Noninferiority Testing

There are two approaches, strict tests or test-based confidence limits, which can be used to perform noninferiority testing. For this report, we will focus on the latter. The confidence coefficient for the test-based confidence limits is $100(1 - 2\alpha)\%$ (Schuirmann, 1987). The purpose of using 2α as the level of significance is to correct for a one-sided test being performed. Once the confidence limits are calculated, the lower confidence limit is compared to the corresponding noninferiority margin $(-\delta,\ 1-\phi, or\ 1-\psi)$. If the lower confidence limit is greater than the noninferiority margin, then you would reject H_O in favor of H_A, thereby establishing at a fixed statistical significance level that the treatment group is noninferior to the control.

Note that there are many different asymptotic methods available to compare two binomial proportions. In this report, we will focus on the default methods provided within SAS® (SAS Institute Inc., 2013). The investigation here is on the performance of various statistical methods related to noninferiority testing, rather than the statistical methods themselves. Therefore, discussion will be limited to their implementation by SAS PROC FREQ. For a more detailed discussion on the statistical methodology, see Newcombe (1998) and Chapter 11 of Rothmann et al. (2012).

Risk Difference

There are four different asymptotic confidence limits (Wald, Farrington-Manning, Hauck-Anderson, and Newcombe) preprogrammed into SAS for the risk difference. Confidence limits are distinguished from one another depending upon whether a continuity correction is incorporated. The following information has been taken from the SAS/STAT® 13.1 User's Guide (SAS Institute Inc., 2013). The confidence limits are assumed to have a confidence coefficient of $100(1-\alpha)\%$.

Asymptotic (Wald) Confidence Limits

The Wald confidence limits for the risk difference are computed as $\hat{d} \pm \left(z_{\alpha/2} \times se(\hat{d})\right)$ where $\hat{d} = \hat{p}_T - \hat{p}_C$ estimates the risk difference, $z_{\alpha/2}$ is the $100(1-\alpha/2)$ percentile of the standard normal distribution, and the standard error is computed from the sample proportions as $e(\hat{d}) = \sqrt{\hat{p}_T(1-\hat{p}_T)/n_T + \hat{p}_C(1-\hat{p}_C)/n_C}$.

With Continuity Correction

A continuity correction can incorporated with the Wald confidence limits. This causes the confidence limits to become $\hat{d} \pm \left(cc + z_{\alpha/2} \times se(\hat{d})\right)$ where $cc = (1/n_T + 1/n_C)/2$.

Rothmann et al. (2012, p. 260) denote that this confidence interval can have suboptimal coverage probabilities when sample sizes are small and/or probabilities of success are near 0 or 1; the unrestricted MLE of the variance is inconsistent with the null hypothesis.

Farrington-Manning (Score) Confidence Limits

The Farrington-Manning confidence limits for the risk difference are computed as $\hat{d} \pm \left(z_{\alpha/2} \times se\left(\hat{d}\right)\right)$ where $\hat{d} = \hat{p}_T - \hat{p}_C$ estimates the risk difference, $z_{\alpha/2}$ is the $100(1-\alpha/2)$ percentile of the standard normal distribution, the standard error is computed from the sample proportions as $se\left(\hat{d}\right) = \sqrt{\tilde{p}_T(1-\tilde{p}_T)/n_T + \tilde{p}_C(1-\tilde{p}_C)/n_C}$, and $\tilde{p}_T$ and $\tilde{p}_C$ are the constrained maximum likelihood estimates of p_T and p_C. The maximum likelihood estimates are subjected to a constraint of $-\delta$ and computed as follows:

$$\tilde{p}_T = 2u\cos(w) - b/3a \qquad \tilde{p}_C = \tilde{p}_T + \delta$$

where

$$w = (\pi + \cos^{-1}(v/u^3))/3 \qquad a = 1 + \theta$$

$$v = b^3/(3a)^3 - bc/6a^2 + d/2a \qquad b = -\left(1 + \theta + \hat{p}_T + \theta\hat{p}_C - \delta(\theta + 2)\right)$$

$$u = sign(v)\sqrt{b^2/(3a)^2 - c/3a} \qquad c = \delta^2 - \delta(2\hat{p}_T + \theta + 1) + \hat{p}_T + \theta\hat{p}_C$$

$$\theta = n_C/n_T \qquad d = \hat{p}_T\delta(1-\delta)$$

The major difference between the Wald (without continuity correction) and Farrington-Manning method is the constraint on the MLE for the standard error.

Hauck-Anderson Confidence Limits

The Hauck-Anderson confidence limits for the risk difference are computed as $\hat{d} \pm \left(cc + z_{\alpha/2} \times se\left(\hat{d}\right)\right)$ where $\hat{d} = \hat{p}_T - \hat{p}_C$ estimates the risk difference, $z_{\alpha/2}$ is the $100(1-\alpha/2)$ percentile of the standard normal distribution, and the standard error is computed from the sample proportions as $se\left(\hat{d}\right) = \sqrt{\hat{p}_T(1-\hat{p}_T)/(n_T-1) + \hat{p}_C(1-\hat{p}_C)/(n_C-1)}$. The Hauck-Anderson continuity correction cc is computed as $cc = 1/\left(2\,min(n_T, n_C)\right)$.

Hauck and Anderson recommended the use of unbiased estimates with the standard error, i.e. using $n-1$ in the denominators instead of n as in MLE, and different continuity correction to the Wald method (Rothmann, et al. 2012, p. 260).

Newcombe Confidence Limits

Newcombe (hybrid-score) confidence limits for the risk difference are constructed from the individual Wilson score confidence limits for each of the two proportions. These confidence limits for the individual proportions are used in the standard error terms of the Wald confidence limits for the proportion difference. Wilson score confidence limits for p_T and p_C are the roots of $|p_i - \hat{p}_i| = z_{\alpha/2}\sqrt{p_i(1-p_i)/n_i}$ for $i = T, C$. The confidence limits are computed as

$$\left(\hat{p}_i + z_{\alpha/2}^2/2n_i \pm z_{\alpha/2}\sqrt{\left(\hat{p}_i(1-\hat{p}_i) + z_{\alpha}^2/4n_i\right)/n_i}\right)\Big/\left(1 + z_{\alpha/2}^2/2n_i\right) .$$

Denote the lower and upper Wilson score confidence limits for p_T as L_T and U_T, and denote the lower and upper Wilson score confidence limits for p_C as L_C and U_C. The Newcombe confidence limits for the proportion difference $(d = p_T - p_C)$ are computed as follows:

$$d_L = (\hat{p}_T - \hat{p}_C) - z_{\alpha/2}\sqrt{(\hat{p}_T - L_T)^2 + (U_C - \hat{p}_C)^2}$$

$$d_U = (\hat{p}_T - \hat{p}_C) - z_{\alpha/2}\sqrt{(U_T - \hat{p}_T)^2 + (\hat{p}_C - L_C)^2} .$$

With Continuity Correction

A continuity correction can incorporated with the Newcombe confidence limits. This occurs through the inclusion of a continuity correction $cc = 1/2n_i$ with the initial calculations of Wilson confidence limits for the individual proportions, now calculated as the root of $|p_i - \hat{p}_i| - 1/2n_i = z_{\alpha/2}\sqrt{p_i(1-p_i)/n_i}$. The continuity-corrected confidence limits for the individual proportions are then used to compute the proportion difference confidence limits d_L and d_U.

Relative Risk

There is one asymptotic (Wald) confidence limit preprogrammed into SAS for the relative risk. The following information has been taken from the SAS/STAT® 13.1 User’s Guide (SAS Institute Inc., 2013). The confidence limits are assumed to have a confidence coefficient of $100(1-\alpha)\%$.

Asymptotic (Wald) Confidence Limits

The asymptotic confidence limits for the relative risk is computed as $\left(\widehat{RR} \times exp\left(-z_{\alpha/2}\sqrt{v}\right), \widehat{RR} \times exp\left(z_{\alpha/2}\sqrt{v}\right)\right)$ where $\widehat{RR} = \hat{p}_T/\hat{p}_C$ estimates the relative risk,

$z_{\alpha/2}$ is the $100(1-\alpha/2)$ percentile of the standard normal distribution, and v is computed as $v = Var(\ln \widehat{RR}) = ((1-\hat{p}_T)/x) + ((1-\hat{p}_C)/y)$. If either of the number of successes (x or y) for the treatment group or control is zero, then the estimates are not computed.

The asymptotic (Wald) confidence limits for the relative risk were found on the logarithmic scale $[\ln(\widehat{RR}) \pm z_{\alpha/2}\sqrt{v}]$ and then exponentiated back to their original scale. The natural log transformation aids in assuring that convergence will be reached more rapidly, making the normal approximation better (Rothmann, et al. 2012, p. 278).

Odds Ratio

There are two different asymptotic confidence limits (Wald and Score) preprogrammed into SAS for the odds ratio. The following information has been taken from the SAS/STAT® 13.1 User's Guide (SAS Institute Inc., 2013). The confidence limits are assumed to have a confidence coefficient of $100(1-\alpha)\%$.

Asymptotic (Wald) Confidence Limits

The asymptotic confidence limits for the odds ratio is computed as $\left(\widehat{OR} \times exp(-z_{\alpha/2}\sqrt{v}), \widehat{OR} \times exp(z_{\alpha/2}\sqrt{v})\right)$ where $\widehat{OR} = \hat{p}_T(1-\hat{p}_C)/\hat{p}_C(1-\hat{p}_T)$ estimates the odds ratio, $z_{\alpha/2}$ is the $100(1-\alpha/2)$ percentile of the standard normal distribution, and v is computed as $v = Var(\ln \widehat{OR}) = 1/x + 1/(n_T - x) + 1/y + 1/(n_C - y)$. If any of the four cell frequencies $(x, n_T - x, y, or\ n_C - y\ )$ are zero, then the estimates are not computed.

Similar to the relative risk, the asymptotic (Wald) confidence limits for the odds ratio were found on the logarithmic scale $[\ln(\widehat{OR}) \pm z_{\alpha/2}\sqrt{v}]$ and then exponentiated back to their original scale. The natural log transformation aides in assuring that convergence will be reached more rapidly, making the normal approximation better (Rothmann, et al. 2012, p. 291).

Score Confidence Limits

The score confidence limits for the odds ratio are computed by inverting the score test. A score-based chi-square test statistic for the null hypothesis can be expressed as

$$Q(\theta) = \frac{\{n_T(\hat{p}_T - \tilde{p}_T)\}^2/\{(n_T + n_C)/(n_T + n_C - 1)\}}{1/\left(n_T\tilde{p}_T(1-\tilde{p}_T)\right) + 1/\left(n_C\tilde{p}_C(1-\tilde{p}_C)\right)}$$

where $\tilde{p}_T$ and $\tilde{p}_C$ are the maximum likelihood estimates of p_T and p_C, subject to the constraint that the odds ratio is θ, are computed as follows:

$$\tilde{p}_C = \left(-b + (b^2 - 4ac)\right)/2a \qquad \tilde{p}_T = \tilde{p}_C\theta/\left(1 + \tilde{p}_C(\theta - 1)\right)$$

where

$$a = n_C(\theta - 1)$$
$$b = n_T\theta + n_C - \hat{p}_T(\theta - 1)$$
$$c = -\hat{p}_T$$

By default, the score confidence interval includes a bias correction factor of $(n_T + n_C)/(n_T + n_C - 1)$ in the denominator of $Q(\theta)$.

The score confidence interval for the odds ratio consists of all values for θ in which the test statistic $Q(\theta)$ falls in the acceptance region $\left\{\theta : Q(\theta) < \chi^2_{1,\alpha}\right\}$ where $\chi^2_{1,\alpha}$ is the $100(1 - \alpha)$ percentile of the chi-square distribution with 1 degree of freedom. PROC FREQ finds the confidence limits by iterative computation.

Major advantage of the score confidence limits for the odds ratio is that it is computable despite any of the cell frequencies being equal to zero. For these intervals, when $\hat{\theta} = 0$, 0 is the lower limit and $\hat{\theta} = \infty$, ∞ is the upper limit (Agresti, 2013, p. 70).

Chapter 3 - Numerical Example

The following is a numerical example to get a better understanding of how noninferiority testing and test-based confidence intervals work. Suppose an investigational antibiotic is being compared to a standard antibiotic in uncomplicated bacterial sinusitis (Rothmann et al. 2012, p. 262). The outcome measured is the success of the drugs to cure the bacterial sinusitis. Cure rates for treatment and control groups were observed to be 89/100 (89%) and 92/100 (92%) respectively. Observed cure rates will be used to determine whether the investigational antibiotic is noninferior to the standard treatment in the market.

As described before, careful consideration must be made when determining the noninferiority margin. Rothmann et al. (2012) assume a noninferiority margin of the risk difference to be $-\delta = -0.10$. Defining an equivalent noninferiority margin for the relative risk or odds ratio is problematic without knowing the true proportion of success for the control. Therefore, we invoke standard conventions, instead of being driven by the data, when selecting the noninferiority margin for the relative risk and odds ratio. Chow and Liu (2009, p. 21) discuss the 80/125 rule which states "bioequivalence can be concluded if the average bioavailability of the test formulation is within (80%, 125%) that of the reference formulation." Using that notation, it is assumed that 20% would be the smallest unacceptable reduction for the relative risk which corresponds to a noninferiority margin where $\phi = 0.20$. For selecting the noninferiority margin of the odds ratio, we used Garrette's proposed lower margin of 0.5 (Ng, 2008, p. 5404) which corresponds to an acceptable reduction of 50% in the odds for the investigational antibiotic relative to standard ($\psi = 0.50$).

Calculations

The estimated risk difference is calculated to be $\hat{d} = 0.89 - 0.92 = -0.03$. The estimated relative risk is calculate to be $\widehat{RR} = 0.89/0.92 = 0.9674$. The estimated odds ratio is calculated to be $\widehat{OR} = 0.89(1 - 0.92)/0.92(1 - 0.89) = 0.7036$. Confidence intervals were calculated in SAS with a confidence coefficient of 90% to correspond to a noninferiority test with a 5% level of significance. Complete SAS code is provided in Appendix B (Numerical Example).

Risk Difference

The lower 90% confidence limits for the Wald method without and with a continuity correction are -0.0981 and -0.1081 respectively. The lower 90% confidence limit for the Farrington-Manning method is -0.1017. The lower 90% confidence limit for the Hauck-Anderson method is -0.1035. The lower 90% confidence limits for the Newcombe method without and with a continuity correction is -0.1009 and -0.1078 respectively. Using the noninferiority margin of $-\delta = -0.10$, we would conclude that the investigational antibiotic is noninferior only for the Wald confidence interval without a continuity correction.

Relative Risk

The lower 90% confidence limit for the asymptotic (Wald) method is 0.8971. Using the noninferiority margin of $1 - \phi = 0.80$, we would conclude that the investigational antibiotic is noninferior at the 5% significance level.

Odds Ratio

The lower 90% confidence limit for the asymptotic (Wald) method is 0.3153. The lower 90% confidence limit for the Score method is 0.3198. Using the noninferiority margin of $1 - \psi = 0.50$, we would be unable to conclude that the investigational antibiotic is noninferior at the 5% significance level.

Chapter 4 - Simulation Studies

A Monte Carlo simulation was programmed to determine the performance between asymptotic noninferiority methods based on the calculations of risk differences (with and without continuity correction), relative risks, and odd ratios from two independent samples. Investigation was done to compare type I error rates, power when the true proportions were exactly the same, and power when the true proportion for treatment group was less than the control, but not meaningfully inferior, which we refer to as the midpoint scenario. Generation of data, noninferiority analysis, and summary of results for the simulation studies were all performed within SAS/STAT® software version 9.4 (SAS Institute Inc., 2013).

Simulation of Data

Sample sizes for the control and treatment groups were assumed to be the same $(n = n_T = n_C)$, but could be generalized to unequal cases with slight modification of the code. Sizes ($n =$ 200, 600, and 1500) were picked to be representative of small, medium, and large clinical trials you would see in real life. The proportions of success for the control group span the middle spectrum of potential values ($p_c =$ 0.20, 0.35, 0.50, 0.65, and 0.80). The choice of sample size and proportion of success for the control group were strategic in assuring that asymptotic properties hold, i.e. having $n * p_i \geq 5$ and $n * (1 - p_i) \geq 5$ for $i = T, C$. This avoided situations where rare events were simulated and other (exact or Bayesian) methods may be more appropriate to deal with potential issues of sparseness. Various noninferiority margins were considered by letting $\phi =$ 0.05, 0.10, 0.15, 0.20, and 0.25. Using algebraic manipulation and the relationship between proportion of success for the control in accordance to the noninferiority margin for the relative risk, it can be shown that the noninferiority margins for risk difference and odds ratio are $-\delta = -\phi * p_C$ and $1 - \psi = 1 - \left[(1 - \phi)\,(1 - p_C)/\left(1 - p_C(1 - \phi)\right)\right]$ respectively.

Each simulation setting combination (75 total) was replicated 5,000 times. Separate randomly generated seeds for the treatment and control groups, seed1 and seed2 respectively, were generated as the greatest integer from a uniform distribution over the interval of 0 to 100,000 for each simulation setting. Using the described settings, random binomial count data for the treatment $(x_i \sim Binomial(n, p_T, seed1))$ and control $(y_i \sim Binomial(n, p_C, seed2))$

groups were independently generated by the RANBIN function for each replication (Simulation: $i = 1,2, \dots, 5000$).

Type I Error Simulation Study

Type I error is assessed by setting the true proportion for the treatment group to be equal to the respective noninferiority margin, i.e. on the boundary of the null hypothesis parameter space, where the type I error rate is maximized. This means the probability of success for the treatment group was calculated as follows: $p_T = (1 - \phi)p_C$. Table A.1 displays a complete list of settings used for the type I error simulation study.

Maximal Power Simulation Study

The goal of this study is to estimate power when the true proportion for the treatment group is equal to the control, i.e. $p_T = p_C$. Roebruck and Kühn (1995) refer to this scenario as maximal power. Table A.11 displays a complete list of settings used for this simulation study.

Midpoint Power Simulation Study

Here we study the power when the true proportion for the treatment group is less than the control, but not meaningfully inferior. The probability of success for the treatment group was calculated as follows: $p_T = (1 - 0.5\phi)p_C$, i.e. the midpoint between p_C and the boundary of the noninferiority margin on the probability scale (Figure 4.1). We will refer to this as the midpoint scenario. Table A.21 displays a complete list of settings used for this simulation study.

Figure 4.1 Parameter space of simulation study

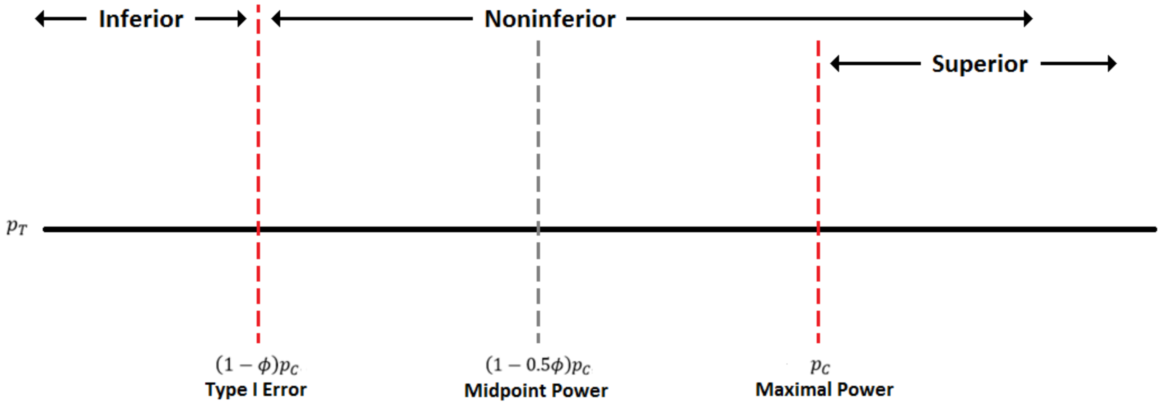


Table 4.1 Form of simulated data for each simulation setting

Simulation	Group	Outcome	Count
1	Treatment	Success	x_1
1	Treatment	Failure	$n - x_1$
1	Control	Success	y_1
1	Control	Failure	$n - y_1$
⋮	⋮	⋮	⋮
5000	Control	Failure	$n - y_{5000}$

Implementation

A control dataset was initially generated for the various simulation settings to be used with each simulation study (Type I Error, Maximal Power, & Midpoint Power). The control datasets contained the number of iterations (iter) to be simulated; sample size (n); proportion of success for treatment (p_T) and control (p_C) group; values used to calculate the noninferiority margin for risk difference (δ), relative risk (ϕ), and odds ratio (ψ); and seeds for treatment (seed1) and control group (seed2); and csv file names for simulated data and noninferiority results. The control dataset for each simulation study was exported as separate csv file for future reference. Macros were created to efficiently loop through the control datasets and “fetch” (Ruegsegger, 2009) the specified parameter for each simulation setting used with the generation of data, noninferiority analysis, and summary of results.

Count data was generated for each simulation setting, transposed into the cell count form to be analyzed with PROC FREQ (Table 4.1), and lastly exported into separate csv files with distinct names associated to values of settings used. Complete SAS code is provided in Appendix C (Generation of Simulated Data). Asymptotic assumptions of the simulated data were checked through the creation of bar charts and histograms using the success count data for the two groups.

A similar process was repeated to perform noninferiority testing on the three simulation studies. Complete SAS code is provided in Appendix C (Noninferiority Testing). Control datasets for each simulation study were imported back into SAS and used to store datasets already simulated for each simulation setting. Noninferiority analysis was performed in three parts using the PROC FREQ procedure. Initially, general descriptive statistics (i.e. estimated proportion of success for the treatment and control group, risk difference, relative risk, odds ratio, etc.) were calculated for each simulation, along with the 90% asymptotic (Wald) confidence intervals for the relative risk. This was achieved by specifying the RISKDIFF and RELRISK statistics options in the TABLES statement and setting $\alpha = 0.10$. Next, confidence intervals for the risk difference were calculated by performing noninferiority tests (NONINF) preprogrammed into the RISKDIFF statistics option. This was achieved by specifying multiple TABLES statements corresponding to the six different methods (METHOD=), specified noninferiority margin (MARGIN=), and default value of $\alpha = 0.05$ to get corresponding 90% confidence intervals. Lastly, confidence intervals for the odds ratio were calculated by using the ODDSRATIO (OR) option of the TABLES statement. This was achieved by setting $\alpha = 0.10$ and using two TABLES statements corresponding to the Wald and Score methods (CL=). Categorical variables for each method were created to indicate the result of each test based off the simulation's lower confidence limit. Data from noninferiority testing with each measure was transposed and merged together before being converted into an exportable csv file.

The results of the noninferiority testing for each simulation setting were summarized to determine the respective type I error rates or power percentages. Complete SAS code is provided in Appendix C (Noninferiority Results). This was achieved by converting the aforementioned categorical variable into an indicator variable and preforming PROC MEANS to get an overall count of conclusions drawn. Results were compiled into a consolidated data set that included corresponding simulation setting values. Ensuing graphics were created from the summarized results. Complete SAS code is provided in Appendix D (Noninferiority Graphics).

Results

Any conclusions drawn are under the assumptions for which each simulation study were specified. One must not generalized or extrapolated the results to other circumstances, e.g. when the proportions are very close to 0 or 1, as they may not hold.

Type I Error Simulation Study

Estimated type I error rates with each method for risk difference, relative risk, and odds ratio from the Type I Error simulation study are displayed in Tables A.2-A.10. The raw data is presented in a graphical form (Figures 4.2-4.6) to aid in comparison. Each point is representative of a simulation setting used. Figure 4.2 displays the estimated type I error rates as a percent for risk difference using (Wald, Farrington-Manning, and Newcombe) methods without a continuity correction. Looking at the results, we see that the various methods hold close to their nominal type I error rate of 5% (red reference line) and in most cases fall within the 95% confidence bounds for Monte Carlo simulation, which correspond to a margin of error of approximately $100 * 1.96 * \sqrt{0.05\,(1-0.05)/5000}\,\% = 0.6\%$ (blue reference lines). Figure 4.3 displays the estimated type I error rates for risk difference using (Wald, Hauck-Anderson, and Newcombe) methods with a continuity correction. It becomes clearly apparent that the addition of a continuity correction causes the type I error rates to be more conservative (smaller) than expected and therefore will be omitted from comparison of power. The Hauck-Anderson method appears to be the most immune out of the three. Figure 4.4 displays the estimated type I error rates for relative risk using asymptotic (Wald) method. The type I error rates seen for relative risk are comparable to what was seen with the risk difference without a continuity correction. Figure 4.5 displays the estimated type I error rates for odds ratio using Wald and Score methods. Both methods perform comparable to one another and hold their nominal type I error rates. Figure 4.6 displays the estimated type I error rates using asymptotic (Wald) method for risk difference, relative risk, and odds ratio. This graphic shows the across performance evaluation of three statistics used to measure the difference between the treatment group and control. All three statistics show similar performance, hold close to their true type I error rates of 5%, and in most cases fall within the Monte Carlo simulation confidence bounds.

Maximal Power Simulation Study

Estimated power as a percentage with each method for risk difference, relative risk, and odds ratio from the Maximal Power simulation study are displayed in Tables A.12-A.20. Figure 4.7 displays the estimated power using asymptotic (Wald) method for risk difference, relative risk, and odds ratio. As expected, the power increases with larger sample sizes, proportions of success for the control group, and values for ϕ. Under ideal circumstances it would be preferred to have a power above 80%, indicated by the red reference line. Furthermore, each simulation setting combination has a maximum Monte Carlo simulation margin of error no more than $100 * 1.96 * \sqrt{0.5\,(1-0.5)/5000}\,\% = 1.4\%$. *Ceteris paribus*, it is clear that the relative risk statistic performs the best, since it has the greatest power under all simulation settings.

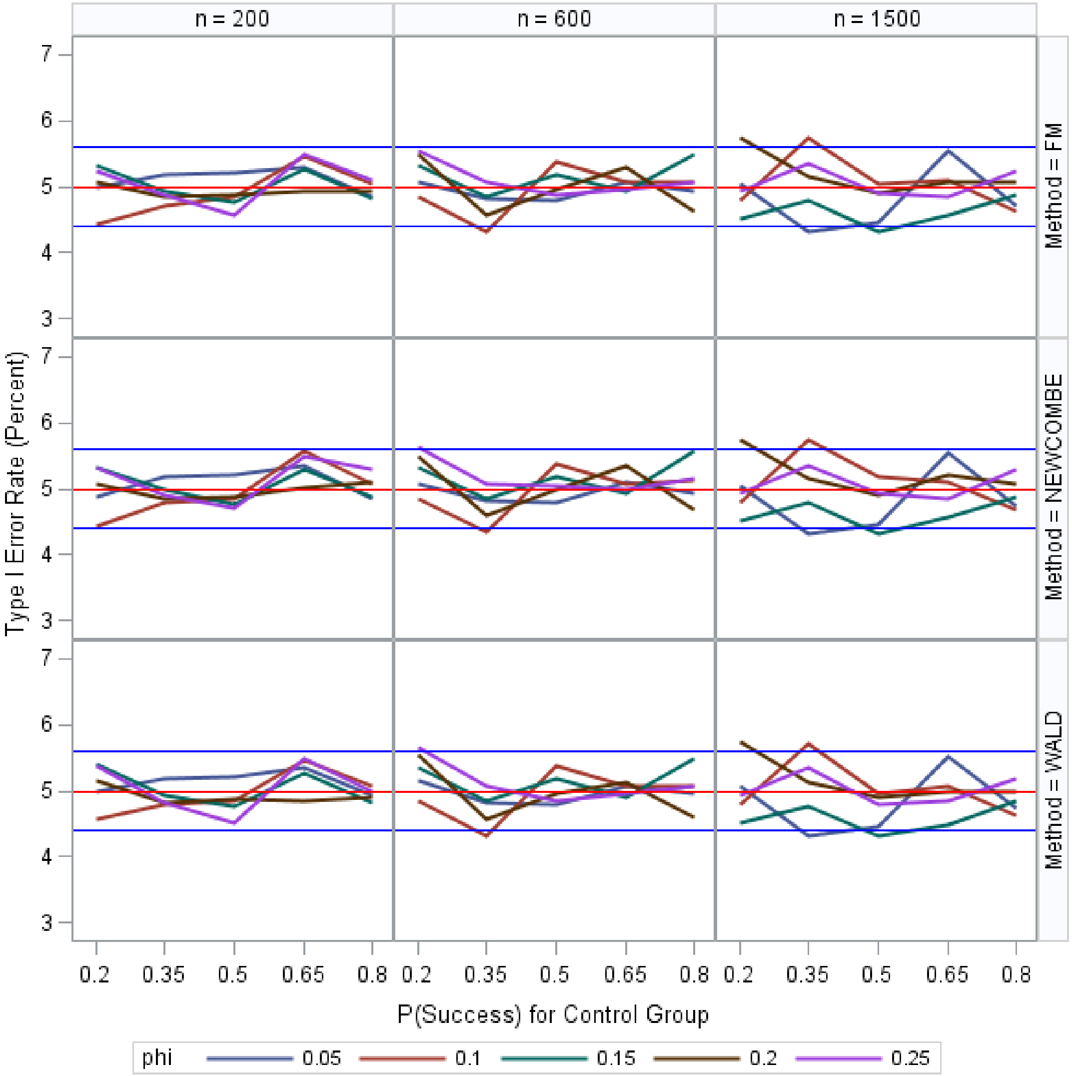
Midpoint Power Simulation Study

Estimated power as a percentage with each method for risk difference, relative risk, and odds ratio from the Midpoint Power simulation study are displayed in Tables A.22-A.30. Patterns seen are similar to the Maximal Power simulation study (Figure 4.8). *Ceteris paribus*, it is clear that the relative risk statistic seems to perform the best, since it has the greatest power under all simulation settings.

Summary

From the results seen over the three simulation studies, the relative risk statistic appears to be the best statistic when performing noninferiority testing on two binomial proportions between 0.2 and 0.8. This was concluded since the relative risk holds its nominal type I error rate while achieving higher power under most simulation settings. This concurs with what Rothmann et al. (2012, p. 212-213) state about how the relative risk is perceived to be more consistent across different patient populations with different event rates than the risk difference. Furthermore, the events were not rare enough, by how the simulation studies were conducted, for the relative risk to be outperformed by the odds ratio.

Figure 4.2 Estimates of type I error rate (percent) for the risk difference without a continuity correction

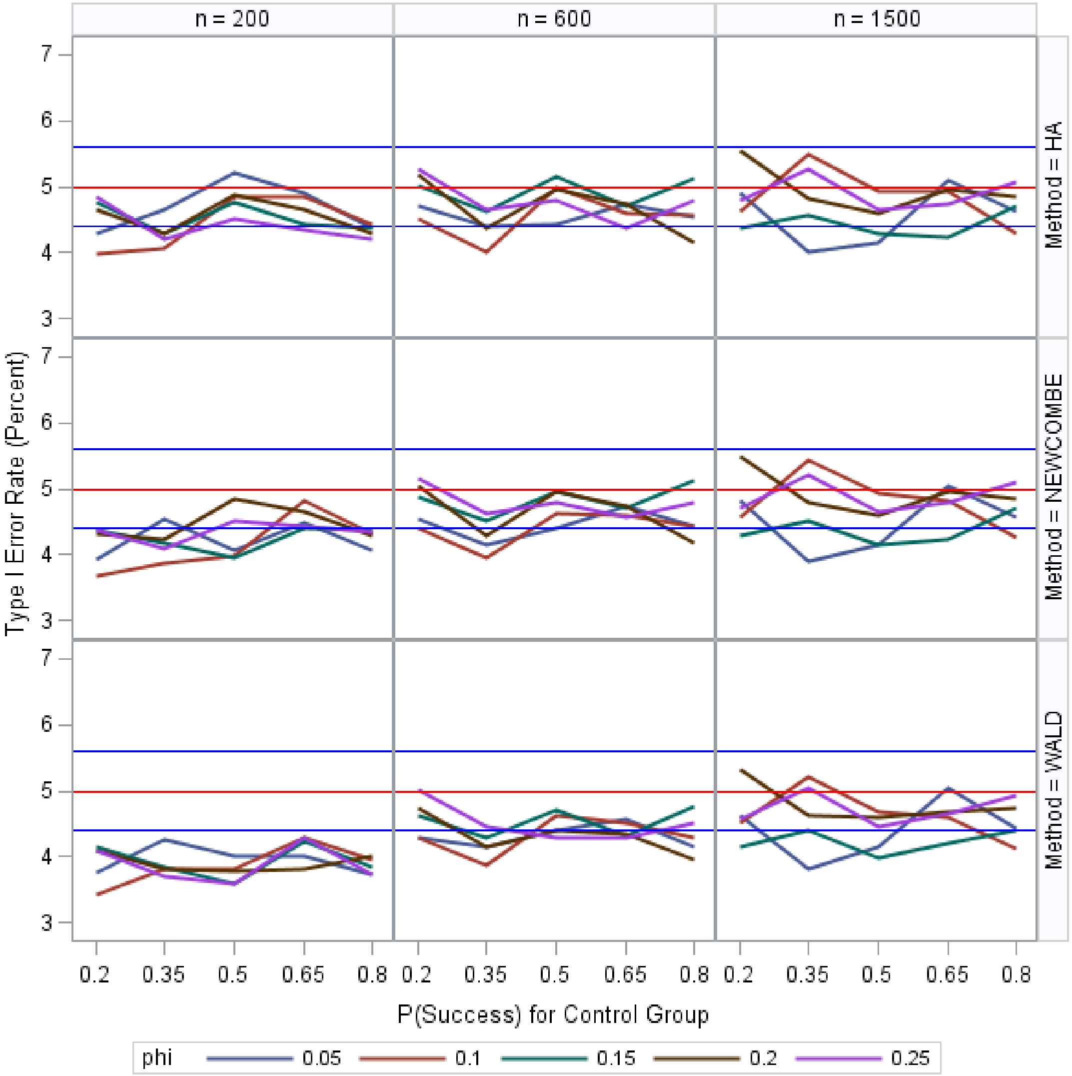


[1]

---

[1] Horizontal red reference line indicates nominal type I error rate of 5%. Horizontal blue reference lines indicate 95% Monte Carlo simulation confidence bounds for nominal type I error rate.

Figure 4.3 Estimates of type I error rate (percent) for the risk difference with a continuity correction

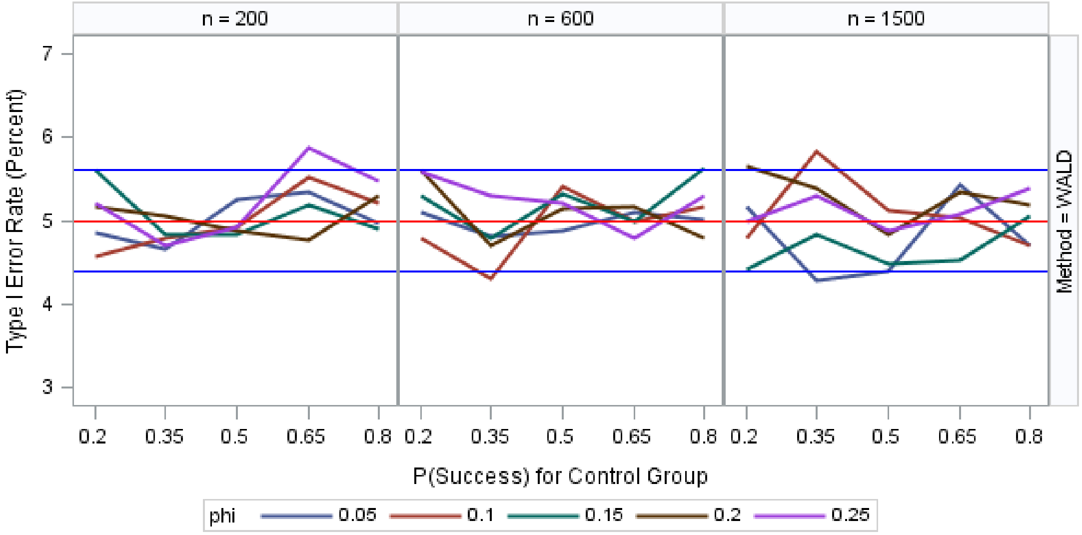


[2]

---

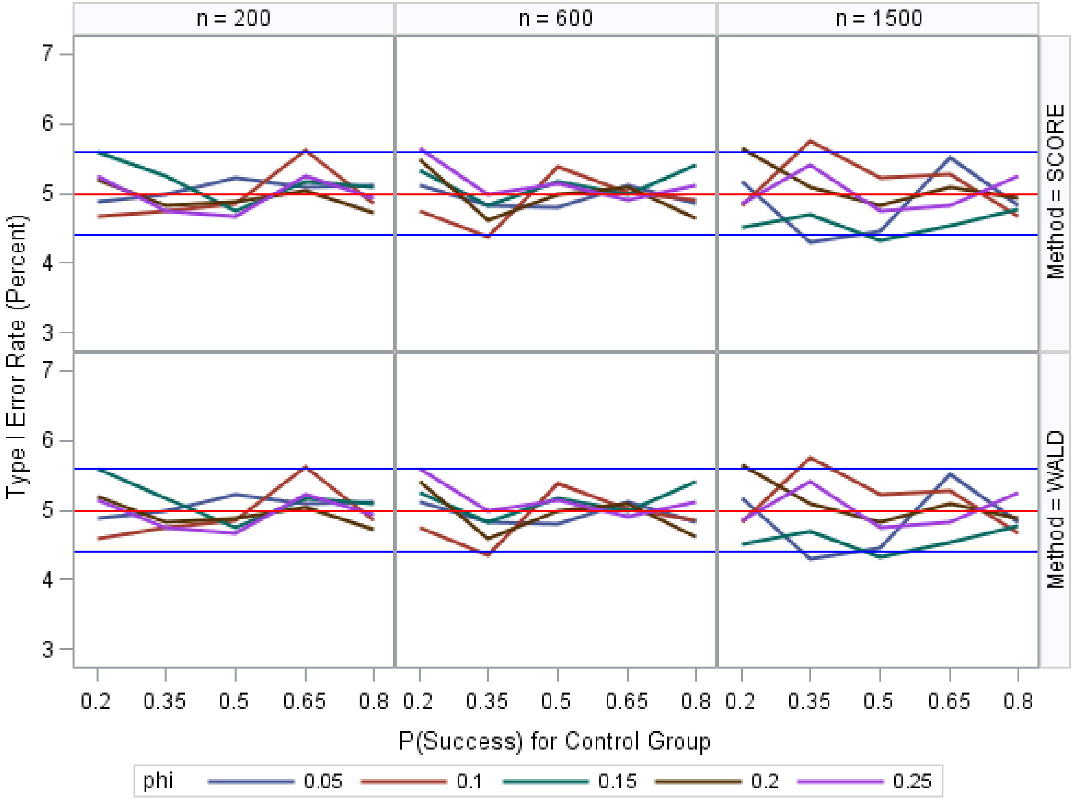
[2] Horizontal red reference line indicates nominal type I error rate of 5%. Horizontal blue reference lines indicate 95% Monte Carlo simulation confidence bounds for nominal type I error rate.

Figure 4.4 Estimates of type I error rate (percent) for the relative risk



[3]

Figure 4.5 Estimates of type I error rate (percent) for the odds ratio



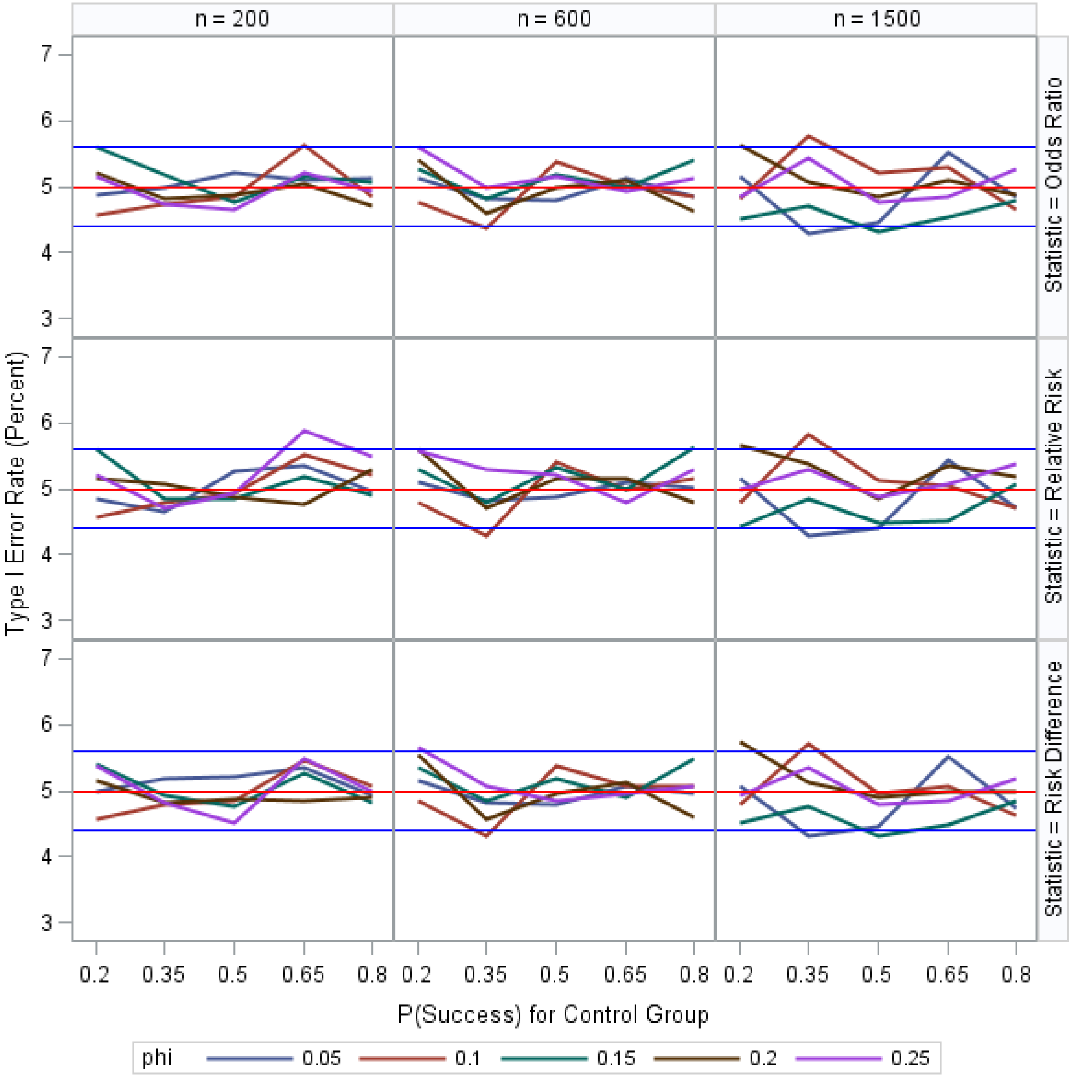
[4]

---

[3] Horizontal red reference line indicates nominal type I error rate of 5%. Horizontal blue reference lines indicate 95% Monte Carlo simulation confidence bounds for nominal type I error rate.

[4] Horizontal red reference line indicates nominal type I error rate of 5%. Horizontal blue reference lines indicate 95% Monte Carlo simulation confidence bounds for nominal type I error rate.

Figure 4.6 Estimates of type I error rate (percent) across statistics using asymptotic (Wald) method

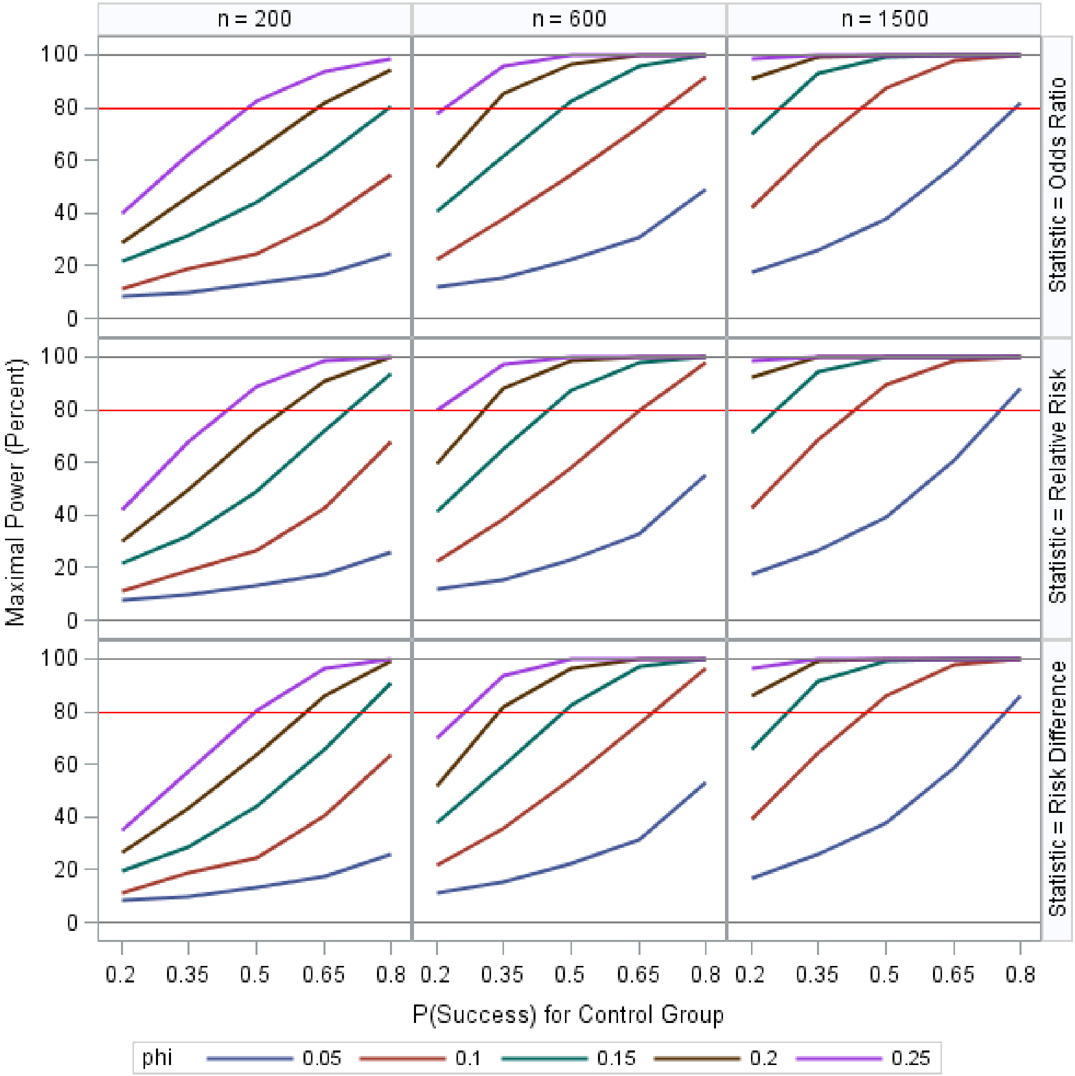


[5]

---

[5] Horizontal red reference line indicates nominal type I error rate of 5%. Horizontal blue reference lines indicate 95% Monte Carlo simulation confidence bounds for nominal type I error rate.

Figure 4.7 Estimates of maximal power (percent) across statistics using asymptotic (Wald) method

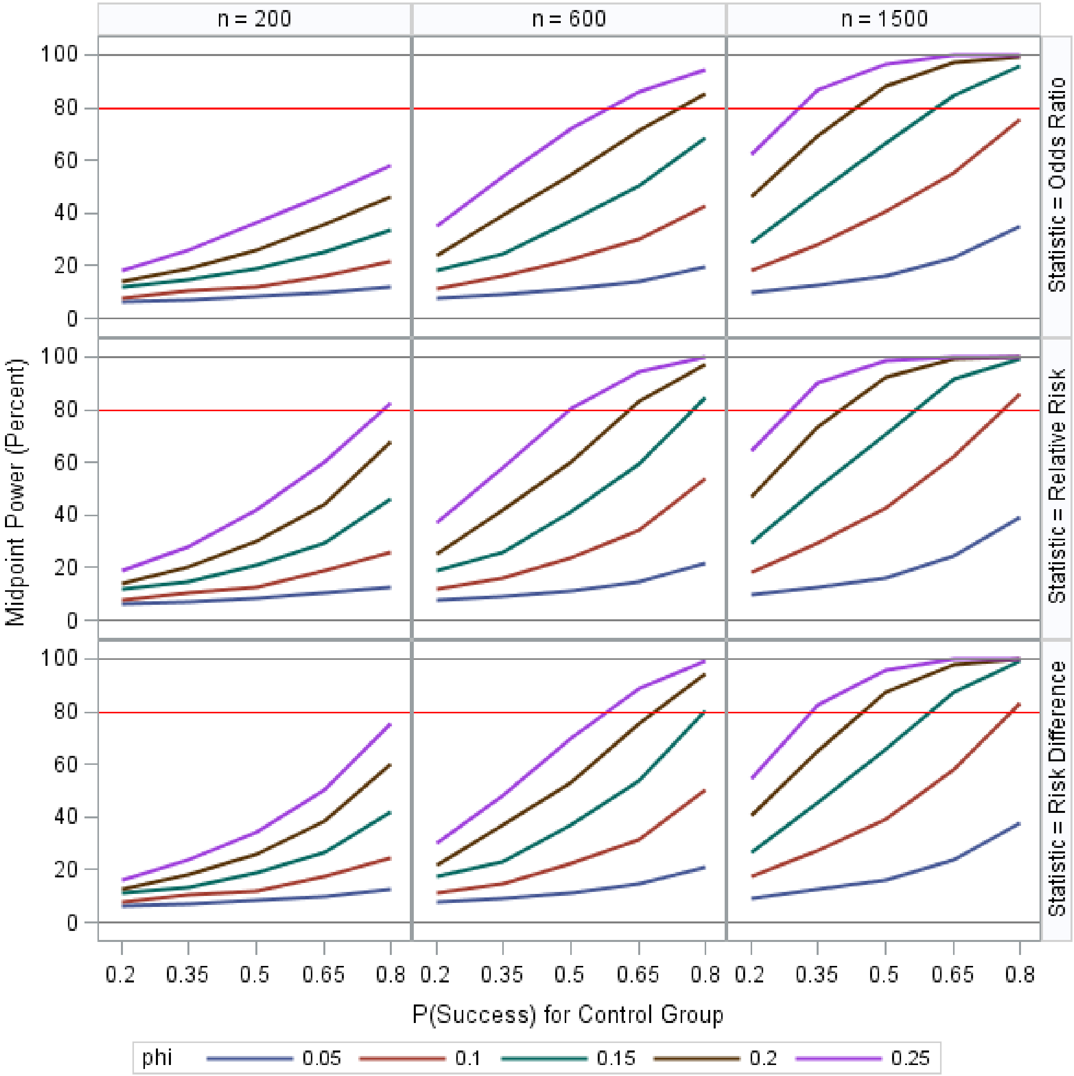


6

---

[6] Horizontal red reference line indicates power threshold of 80%. Each simulation setting combination has a maximum Monte Carlo simulation margin of error no more than 1.4%.

Figure 4.8 Estimates of midpoint power (percent) across statistics using asymptotic (Wald) method



[7]

---

[7] Horizontal red reference line indicates power threshold of 80%. Each simulation setting combination has a maximum Monte Carlo simulation margin of error no more than 1.4%.

Chapter 5 - Conclusion

Testing for noninferiority is an important tool for establishing the efficacy of an investigational treatment in comparison to a standard treatment, especially when the similarity of the two interventions prevents the demonstration of superiority. In the case of comparing two binomial proportions, there are three commonly used comparisons of differences: risk difference, relative risk, and odds ratio. Additionally, there are several statistical methods for assessing these comparisons.

In this report, we investigated these methods with respect to type I error rates and power under various sample sizes ($n =$ 200, 600, and 1500), true proportions of success for the control group ($p_c =$ 0.20, 0.35, 0.50, 0.65, and 0.80), and true differences in proportions as well as an assortment of choices for the noninferiority margin ($\phi =$ 0.05, 0.10, 0.15, 0.20, and 0.25). Simulation results indicate most analysis methods have comparable type I error rates. However, the implementation of the continuity correction inside of SAS for the risk difference tends to deflate the nominal type I error rate, especially for smaller sample sizes. The asymptotic (Wald) method based on relative risk had higher power under most circumstances. Due to its ease of interpretation, we recommend using the relative risk and its associated asymptotic confidence interval in the establishing noninferiority of a binomial proportion between 0.2 and 0.8.

Bibliography

Agresti, A. (2013). *Categorical data analysis* (3rd ed.). Hoboken, NJ: Wiley-Interscience.

Altman, D. G., & Bland, J. M. (1995). Absence of evidence is not evidence of absence. *BMJ: British Medical Journal, 311*(7003), p. 485. Retrieved from http://www.jstor.org/stable/29728438 (Last accessed 12/10/2014)

Chow, S., & Liu, J. (2009). *Design and analysis of bioavailability and bioequivalence studies* (Third ed.) Chapman & Hall/CRC Biostatistics Series.

Newcombe, R. G. (1998). Interval estimation for the difference between independent proportions: Comparison of eleven methods. *Statistics in Medicine, 17*(8), 873-890. doi:10.1002/(SICI)1097-0258(19980430)17:8<873::AID-SIM779>3.0.CO;2-I

Ng, T. (2008). Noninferiority hypotheses and choice of noninferiority margin. *Statistics in Medicine, 27*(26), 5392-5406. doi:10.1002/sim.3367

Roebruck, P., & Kühn, A. (1995). Comparison of tests and sample size formulae for proving therapeutic equivalence based on the difference of binomial probabilities. *Statistics in Medicine, 14*(14), 1583-1594. doi:10.1002/sim.4780141409

Rothmann, M., Wiens, B., & Chan, I. (2012). *Design and analysis of non-inferiority trials* Chapman & Hall/CRC Biostatistics Series.

Ruegsegger, S. (2009). Using datasets to define macro loops and local macro variables. *NorthEast SAS Users Group,* Burlington, VT. Retrieved from http://www.nesug.org/Proceedings/nesug09/bb/bb08.pdf (Last accessed 12/09/2014)

SAS/STAT® software, Version 9.4. Copyright © 2013. SAS Institute Inc., Cary, NC, USA.

SAS Institute Inc. 2013. SAS/STAT® 13.1 User's Guide. Cary, NC: SAS Institute Inc.

Schuirmann, D. J. (1987). A comparison of the two one- sided tests procedure and the power approach for assessing the equivalence of average bioavailability. *Journal of Pharmacokinetics and Biopharmaceutics, 15*(6), 657-680.

Walker, E., & Nowacki, A. (2011). Understanding equivalence and noninferiority testing. *Journal of General Internal Medicine, 26*(2), 192-196. doi:10.1007/s11606-010-1513-8

Appendix A - Supplemental Tables

Type I Error Simulation Study

Table A.1 Simulation settings used to generate data for type I error simulation study

iter	n	p_T	p_C	δ (delta)	ϕ (phi)	ψ (psi)	seed1	seed2
5000	200	0.19	0.2	0.01	0.05	0.0617	40451	61115
5000	200	0.18	0.2	0.02	0.1	0.122	19771	37218
5000	200	0.17	0.2	0.03	0.15	0.1807	98083	60060
5000	200	0.16	0.2	0.04	0.2	0.2381	5872	91199
5000	200	0.15	0.2	0.05	0.25	0.2941	64732	86536
5000	200	0.3325	0.35	0.0175	0.05	0.0749	65114	81031
5000	200	0.315	0.35	0.035	0.1	0.146	9412	87339
5000	200	0.2975	0.35	0.0525	0.15	0.2135	44131	86352
5000	200	0.28	0.35	0.07	0.2	0.2778	54395	96908
5000	200	0.2625	0.35	0.0875	0.25	0.339	69794	94459
5000	200	0.475	0.5	0.025	0.05	0.0952	94707	36593
5000	200	0.45	0.5	0.05	0.1	0.1818	26120	73393
5000	200	0.425	0.5	0.075	0.15	0.2609	18526	98483
5000	200	0.4	0.5	0.1	0.2	0.3333	24881	67474
5000	200	0.375	0.5	0.125	0.25	0.4	57355	90117
5000	200	0.6175	0.65	0.0325	0.05	0.1307	46870	58584
5000	200	0.585	0.65	0.065	0.1	0.241	10577	7856
5000	200	0.5525	0.65	0.0975	0.15	0.3352	66558	86389
5000	200	0.52	0.65	0.13	0.2	0.4167	8355	14157
5000	200	0.4875	0.65	0.1625	0.25	0.4878	83383	18514
5000	200	0.76	0.8	0.04	0.05	0.2083	16483	32436
5000	200	0.72	0.8	0.08	0.1	0.3571	88426	15034
5000	200	0.68	0.8	0.12	0.15	0.4688	70612	39418
5000	200	0.64	0.8	0.16	0.2	0.5556	64409	76264
5000	200	0.6	0.8	0.2	0.25	0.625	64713	63165

iter	n	p_T	p_C	δ **(delta)**	ϕ **(phi)**	ψ **(psi)**	**seed1**	**seed2**
5000	600	0.19	0.2	0.01	0.05	0.0617	28937	64546
5000	600	0.18	0.2	0.02	0.1	0.122	26264	61913
5000	600	0.17	0.2	0.03	0.15	0.1807	89532	59419
5000	600	0.16	0.2	0.04	0.2	0.2381	47288	60016
5000	600	0.15	0.2	0.05	0.25	0.2941	71755	64177
5000	600	0.3325	0.35	0.0175	0.05	0.0749	79096	46374
5000	600	0.315	0.35	0.035	0.1	0.146	29526	26754
5000	600	0.2975	0.35	0.0525	0.15	0.2135	18183	79790
5000	600	0.28	0.35	0.07	0.2	0.2778	31018	26768
5000	600	0.2625	0.35	0.0875	0.25	0.339	34254	27723
5000	600	0.475	0.5	0.025	0.05	0.0952	49577	99313
5000	600	0.45	0.5	0.05	0.1	0.1818	46919	93258
5000	600	0.425	0.5	0.075	0.15	0.2609	66041	24865
5000	600	0.4	0.5	0.1	0.2	0.3333	46804	92356
5000	600	0.375	0.5	0.125	0.25	0.4	30386	51141
5000	600	0.6175	0.65	0.0325	0.05	0.1307	22758	39931
5000	600	0.585	0.65	0.065	0.1	0.241	55372	70973
5000	600	0.5525	0.65	0.0975	0.15	0.3352	92665	83594
5000	600	0.52	0.65	0.13	0.2	0.4167	51360	76790
5000	600	0.4875	0.65	0.1625	0.25	0.4878	48804	44325
5000	600	0.76	0.8	0.04	0.05	0.2083	30723	8535
5000	600	0.72	0.8	0.08	0.1	0.3571	58032	34865
5000	600	0.68	0.8	0.12	0.15	0.4688	51831	49175
5000	600	0.64	0.8	0.16	0.2	0.5556	47784	49371
5000	600	0.6	0.8	0.2	0.25	0.625	52048	56054
5000	1500	0.19	0.2	0.01	0.05	0.0617	66134	96489
5000	1500	0.18	0.2	0.02	0.1	0.122	39983	47385
5000	1500	0.17	0.2	0.03	0.15	0.1807	97832	16162

iter	n	p_T	p_C	δ (delta)	ϕ (phi)	ψ (psi)	seed1	seed2
5000	1500	0.16	0.2	0.04	0.2	0.2381	53767	95715
5000	1500	0.15	0.2	0.05	0.25	0.2941	6421	96345
5000	1500	0.3325	0.35	0.0175	0.05	0.0749	18786	56680
5000	1500	0.315	0.35	0.035	0.1	0.146	83415	39340
5000	1500	0.2975	0.35	0.0525	0.15	0.2135	24330	37555
5000	1500	0.28	0.35	0.07	0.2	0.2778	42610	83802
5000	1500	0.2625	0.35	0.0875	0.25	0.339	92958	33068
5000	1500	0.475	0.5	0.025	0.05	0.0952	62381	55685
5000	1500	0.45	0.5	0.05	0.1	0.1818	38314	9410
5000	1500	0.425	0.5	0.075	0.15	0.2609	71709	90512
5000	1500	0.4	0.5	0.1	0.2	0.3333	63703	25523
5000	1500	0.375	0.5	0.125	0.25	0.4	87929	98397
5000	1500	0.6175	0.65	0.0325	0.05	0.1307	92078	18161
5000	1500	0.585	0.65	0.065	0.1	0.241	56363	27139
5000	1500	0.5525	0.65	0.0975	0.15	0.3352	59924	36938
5000	1500	0.52	0.65	0.13	0.2	0.4167	39082	21116
5000	1500	0.4875	0.65	0.1625	0.25	0.4878	6101	1988
5000	1500	0.76	0.8	0.04	0.05	0.2083	96035	39225
5000	1500	0.72	0.8	0.08	0.1	0.3571	6609	51506
5000	1500	0.68	0.8	0.12	0.15	0.4688	83919	20303
5000	1500	0.64	0.8	0.16	0.2	0.5556	260	92366
5000	1500	0.6	0.8	0.2	0.25	0.625	39554	78985

Table A.2 Estimates of type I error rate (percent) using Wald confidence limits for the risk difference without a continuity correction

		$n = 200$					$n = 600$					$n = 1500$				
p_C		**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**
ϕ	**0.05**	4.98	5.22	5.22	5.34	4.94	5.14	4.82	4.80	5.12	4.96	5.10	4.32	4.46	5.56	4.74
	0.10	4.58	4.78	4.86	5.46	5.06	4.88	4.36	5.38	5.08	5.12	4.88	5.82	5.34	5.12	4.80
	0.15	5.40	5.00	4.76	5.26	4.82	5.36	4.86	5.18	4.92	5.54	4.54	4.82	4.32	4.58	4.92
	0.20	5.16	4.84	4.88	4.86	4.90	5.54	4.60	4.96	5.22	4.62	5.88	5.18	4.90	5.06	5.06
	0.25	5.38	4.82	4.52	5.48	4.98	5.70	5.08	4.86	4.96	5.06	5.00	5.34	4.90	4.84	5.28

Table A.3 Estimates of type I error rate (percent) using Farrington-Manning confidence limits for the risk difference

		$n = 200$					$n = 600$					$n = 1500$				
p_C		**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**
ϕ	**0.05**	4.98	5.18	5.22	5.34	4.86	5.14	4.82	4.80	5.14	4.96	5.10	4.32	4.46	5.56	4.74
	0.10	4.58	4.72	4.86	5.56	5.04	4.86	4.36	5.38	5.08	5.12	4.88	5.82	5.34	5.12	4.80
	0.15	5.40	5.00	4.76	5.26	4.82	5.36	4.86	5.18	4.94	5.54	4.54	4.84	4.32	4.58	4.96
	0.20	5.08	4.84	4.88	4.92	4.94	5.54	4.62	4.98	5.34	4.66	5.86	5.22	4.90	5.32	5.08
	0.25	5.32	4.88	4.56	5.48	5.10	5.66	5.08	5.00	4.98	5.08	4.98	5.38	4.94	4.86	5.30

Table A.4 Estimates of type I error rate (percent) using Newcombe Score confidence limits for the risk difference without a continuity correction

		$n = 200$					$n = 600$					$n = 1500$				
p_C		**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**
ϕ	**0.05**	4.98	5.18	5.22	5.34	4.94	5.14	4.82	4.80	5.16	5.02	5.10	4.32	4.46	5.56	4.74
	0.10	4.58	4.78	4.86	5.72	5.06	4.86	4.36	5.38	5.08	5.18	4.88	5.84	5.34	5.12	4.82
	0.15	5.40	5.00	4.76	5.28	4.88	5.36	4.88	5.18	5.00	5.62	4.54	4.86	4.32	4.58	5.00
	0.20	5.16	4.86	4.90	5.06	5.10	5.54	4.64	4.98	5.36	4.70	5.88	5.22	4.90	5.34	5.12
	0.25	5.38	4.94	4.76	5.48	5.28	5.66	5.08	5.26	5.00	5.18	5.00	5.38	4.94	4.88	5.38

Table A.5 Estimates of type I error rate (percent) using Wald confidence limits for the risk difference with a continuity correction

		$n = 200$					$n = 600$					$n = 1500$				
p_C		**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**
ϕ	**0.05**	3.78	4.26	4.00	4.06	3.74	4.38	4.16	4.40	4.60	4.24	4.72	3.84	4.14	5.04	4.50
	0.10	3.58	3.80	3.82	4.42	4.10	4.32	3.88	4.62	4.56	4.30	4.56	5.26	4.92	4.64	4.16
	0.15	4.14	3.84	3.58	4.24	3.84	4.72	4.32	4.70	4.34	4.80	4.22	4.44	3.98	4.22	4.50
	0.20	4.08	3.86	3.78	3.84	4.00	4.80	4.16	4.40	4.46	4.00	5.46	4.74	4.60	4.76	4.76
	0.25	4.10	3.76	3.62	4.30	3.72	5.04	4.48	4.34	4.28	4.50	4.62	5.16	4.52	4.64	4.98

Table A.6 Estimates of type I error rate (percent) using Hauck-Anderson confidence limits for the risk difference

		$n = 200$					$n = 600$					$n = 1500$				
p_C		**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**
ϕ	**0.05**	4.30	4.66	5.22	4.90	4.36	4.72	4.46	4.80	4.76	4.60	4.94	4.08	4.46	5.24	4.62
	0.10	3.98	4.06	4.86	4.86	4.42	4.58	4.04	5.38	4.66	4.64	4.66	5.54	4.92	5.02	4.38
	0.15	4.76	4.36	4.76	4.42	4.48	5.10	4.62	5.18	4.74	5.12	4.40	4.66	4.32	4.24	4.80
	0.20	4.66	4.38	4.88	4.66	4.28	5.22	4.42	4.96	4.74	4.24	5.58	4.90	4.62	4.96	4.88
	0.25	4.84	4.20	4.52	4.34	4.20	5.34	4.68	4.78	4.52	4.80	4.82	5.28	4.66	4.82	5.14

Table A.7 Estimates of type I error rate (percent) using Newcombe Score confidence limits for the risk difference with a continuity correction

		$n = 200$					$n = 600$					$n = 1500$				
p_C		**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**
ϕ	**0.05**	3.92	4.62	4.06	4.48	4.24	4.56	4.28	4.40	4.74	4.42	4.86	3.96	4.14	5.10	4.60
	0.10	3.68	3.94	4.02	4.82	4.32	4.40	3.96	4.74	4.60	4.50	4.62	5.48	4.92	4.96	4.32
	0.15	4.38	4.18	4.14	4.40	4.36	4.88	4.50	5.14	4.74	5.12	4.30	4.58	4.32	4.24	4.80
	0.20	4.34	4.22	4.88	4.66	4.28	5.04	4.38	4.96	4.74	4.24	5.54	4.82	4.62	4.96	4.90
	0.25	4.42	4.08	4.52	4.42	4.34	5.16	4.66	4.78	4.78	4.86	4.78	5.28	4.66	4.84	5.14

Table A.8 Estimates of type I error rate (percent) using Wald confidence limits for the relative risk

		$n = 200$					$n = 600$					$n = 1500$				
p_C		**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**
ϕ	**0.05**	4.86	4.66	5.26	5.48	4.96	5.10	4.82	4.88	5.10	5.02	5.16	4.36	4.42	5.46	4.74
	0.10	4.58	4.78	4.92	5.54	5.20	4.80	4.36	5.40	5.00	5.24	4.80	5.82	5.14	5.12	4.74
	0.15	5.60	4.84	4.84	5.18	5.02	5.36	4.84	5.40	5.02	5.64	4.42	4.88	4.48	4.54	5.16
	0.20	5.16	5.06	4.96	4.92	5.30	5.60	4.70	5.16	5.22	4.86	5.66	5.38	4.88	5.40	5.20
	0.25	5.20	4.70	4.92	5.88	5.56	5.58	5.36	5.22	4.84	5.36	4.98	5.30	4.88	5.08	5.48

Table A.9 Estimates of type I error rate (percent) using Wald confidence limits for the odds ratio

		$n = 200$					$n = 600$					$n = 1500$				
p_C		**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**
ϕ	**0.05**	4.88	4.98	5.22	5.10	5.12	5.12	4.82	4.80	5.18	4.86	5.16	4.30	4.46	5.52	4.86
	0.10	4.58	4.74	4.86	5.62	4.86	4.76	4.38	5.38	5.02	4.90	4.84	5.76	5.28	5.28	4.66
	0.15	5.60	5.18	4.76	5.16	5.08	5.32	4.82	5.18	4.98	5.40	4.52	4.70	4.32	4.54	4.82
	0.20	5.20	4.82	4.88	5.04	4.72	5.44	4.62	5.00	5.10	4.64	5.64	5.08	4.86	5.16	5.02
	0.25	5.20	4.74	4.70	5.24	4.94	5.66	4.98	5.14	4.96	5.12	4.88	5.42	4.76	4.92	5.26

Table A.10 Estimates of type I error rate (percent) using Score confidence limits for the odds ratio

		$n = 200$					$n = 600$					$n = 1500$				
p_C		**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**
ϕ	**0.05**	4.88	4.98	5.22	5.10	5.12	5.12	4.82	4.80	5.18	4.86	5.16	4.30	4.46	5.52	4.86
	0.10	4.66	4.74	4.86	5.62	4.86	4.76	4.38	5.38	5.02	4.90	4.84	5.76	5.28	5.28	4.66
	0.15	5.60	5.24	4.76	5.16	5.08	5.32	4.82	5.18	4.98	5.42	4.52	4.70	4.32	4.54	4.82
	0.20	5.20	4.82	4.88	5.04	4.72	5.50	4.62	5.00	5.10	4.64	5.64	5.08	4.86	5.16	5.02
	0.25	5.26	4.74	4.66	5.24	4.94	5.66	5.00	5.14	4.96	5.12	4.90	5.42	4.76	4.92	5.26

Maximal Power Simulation Study

Table A.11 Simulation settings used to generate data for maximal power simulation study

iter	n	p_T	p_C	δ (delta)	ϕ (phi)	ψ (psi)	seed1	seed2
5000	200	0.2	0.2	0.01	0.05	0.0617	40451	61115
5000	200	0.2	0.2	0.02	0.1	0.122	19771	37218
5000	200	0.2	0.2	0.03	0.15	0.1807	98083	60060
5000	200	0.2	0.2	0.04	0.2	0.2381	5872	91199
5000	200	0.2	0.2	0.05	0.25	0.2941	64732	86536
5000	200	0.35	0.35	0.0175	0.05	0.0749	65114	81031
5000	200	0.35	0.35	0.035	0.1	0.146	9412	87339
5000	200	0.35	0.35	0.0525	0.15	0.2135	44131	86352
5000	200	0.35	0.35	0.07	0.2	0.2778	54395	96908
5000	200	0.35	0.35	0.0875	0.25	0.339	69794	94459
5000	200	0.5	0.5	0.025	0.05	0.0952	94707	36593
5000	200	0.5	0.5	0.05	0.1	0.1818	26120	73393
5000	200	0.5	0.5	0.075	0.15	0.2609	18526	98483
5000	200	0.5	0.5	0.1	0.2	0.3333	24881	67474
5000	200	0.5	0.5	0.125	0.25	0.4	57355	90117
5000	200	0.65	0.65	0.0325	0.05	0.1307	46870	58584
5000	200	0.65	0.65	0.065	0.1	0.241	10577	7856
5000	200	0.65	0.65	0.0975	0.15	0.3352	66558	86389
5000	200	0.65	0.65	0.13	0.2	0.4167	8355	14157
5000	200	0.65	0.65	0.1625	0.25	0.4878	83383	18514
5000	200	0.8	0.8	0.04	0.05	0.2083	16483	32436
5000	200	0.8	0.8	0.08	0.1	0.3571	88426	15034
5000	200	0.8	0.8	0.12	0.15	0.4688	70612	39418
5000	200	0.8	0.8	0.16	0.2	0.5556	64409	76264
5000	200	0.8	0.8	0.2	0.25	0.625	64713	63165
5000	600	0.2	0.2	0.01	0.05	0.0617	28937	64546

iter	n	p_T	p_C	δ (delta)	ϕ (phi)	ψ (psi)	seed1	seed2
5000	600	0.2	0.2	0.02	0.1	0.122	26264	61913
5000	600	0.2	0.2	0.03	0.15	0.1807	89532	59419
5000	600	0.2	0.2	0.04	0.2	0.2381	47288	60016
5000	600	0.2	0.2	0.05	0.25	0.2941	71755	64177
5000	600	0.35	0.35	0.0175	0.05	0.0749	79096	46374
5000	600	0.35	0.35	0.035	0.1	0.146	29526	26754
5000	600	0.35	0.35	0.0525	0.15	0.2135	18183	79790
5000	600	0.35	0.35	0.07	0.2	0.2778	31018	26768
5000	600	0.35	0.35	0.0875	0.25	0.339	34254	27723
5000	600	0.5	0.5	0.025	0.05	0.0952	49577	99313
5000	600	0.5	0.5	0.05	0.1	0.1818	46919	93258
5000	600	0.5	0.5	0.075	0.15	0.2609	66041	24865
5000	600	0.5	0.5	0.1	0.2	0.3333	46804	92356
5000	600	0.5	0.5	0.125	0.25	0.4	30386	51141
5000	600	0.65	0.65	0.0325	0.05	0.1307	22758	39931
5000	600	0.65	0.65	0.065	0.1	0.241	55372	70973
5000	600	0.65	0.65	0.0975	0.15	0.3352	92665	83594
5000	600	0.65	0.65	0.13	0.2	0.4167	51360	76790
5000	600	0.65	0.65	0.1625	0.25	0.4878	48804	44325
5000	600	0.8	0.8	0.04	0.05	0.2083	30723	8535
5000	600	0.8	0.8	0.08	0.1	0.3571	58032	34865
5000	600	0.8	0.8	0.12	0.15	0.4688	51831	49175
5000	600	0.8	0.8	0.16	0.2	0.5556	47784	49371
5000	600	0.8	0.8	0.2	0.25	0.625	52048	56054
5000	1500	0.2	0.2	0.01	0.05	0.0617	66134	96489
5000	1500	0.2	0.2	0.02	0.1	0.122	39983	47385
5000	1500	0.2	0.2	0.03	0.15	0.1807	97832	16162
5000	1500	0.2	0.2	0.04	0.2	0.2381	53767	95715

iter	**n**	p_T	p_C	δ **(delta)**	ϕ **(phi)**	ψ **(psi)**	**seed1**	**seed2**
5000	1500	0.2	0.2	0.05	0.25	0.2941	6421	96345
5000	1500	0.35	0.35	0.0175	0.05	0.0749	18786	56680
5000	1500	0.35	0.35	0.035	0.1	0.146	83415	39340
5000	1500	0.35	0.35	0.0525	0.15	0.2135	24330	37555
5000	1500	0.35	0.35	0.07	0.2	0.2778	42610	83802
5000	1500	0.35	0.35	0.0875	0.25	0.339	92958	33068
5000	1500	0.5	0.5	0.025	0.05	0.0952	62381	55685
5000	1500	0.5	0.5	0.05	0.1	0.1818	38314	9410
5000	1500	0.5	0.5	0.075	0.15	0.2609	71709	90512
5000	1500	0.5	0.5	0.1	0.2	0.3333	63703	25523
5000	1500	0.5	0.5	0.125	0.25	0.4	87929	98397
5000	1500	0.65	0.65	0.0325	0.05	0.1307	92078	18161
5000	1500	0.65	0.65	0.065	0.1	0.241	56363	27139
5000	1500	0.65	0.65	0.0975	0.15	0.3352	59924	36938
5000	1500	0.65	0.65	0.13	0.2	0.4167	39082	21116
5000	1500	0.65	0.65	0.1625	0.25	0.4878	6101	1988
5000	1500	0.8	0.8	0.04	0.05	0.2083	96035	39225
5000	1500	0.8	0.8	0.08	0.1	0.3571	6609	51506
5000	1500	0.8	0.8	0.12	0.15	0.4688	83919	20303
5000	1500	0.8	0.8	0.16	0.2	0.5556	260	92366
5000	1500	0.8	0.8	0.2	0.25	0.625	39554	78985

Table A.12 Estimates of maximal power (percent) using Wald confidence limits for the risk difference without a continuity correction

		$n = 200$					$n = 600$					$n = 1500$				
p_C		**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**
ϕ	**0.05**	8.34	10.14	13.14	17.18	25.58	11.58	15.20	22.34	31.78	53.40	16.78	25.72	37.86	59.42	86.26
	0.10	11.56	18.76	24.54	40.40	63.82	21.60	36.00	54.26	75.64	96.60	39.74	64.74	87.06	98.18	100.00
	0.15	19.36	28.64	43.82	65.54	91.16	37.58	59.32	82.62	96.86	100.00	65.86	91.76	99.34	100.00	100.00
	0.20	26.30	43.52	63.56	86.04	99.12	51.82	81.72	96.20	99.94	100.00	86.50	98.98	100.00	100.00	100.00
	0.25	35.04	57.14	80.18	96.24	99.96	70.08	93.40	99.66	100.00	100.00	96.36	99.96	100.00	100.00	100.00

Table A.13 Estimates of maximal power (percent) using Farrington-Manning confidence limits for the risk difference

		$n = 200$					$n = 600$					$n = 1500$				
p_C		**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**
ϕ	**0.05**	8.34	10.02	13.14	17.18	25.22	11.56	15.20	22.34	31.80	53.40	16.78	25.72	37.86	59.44	86.26
	0.10	11.56	18.72	24.54	40.48	63.36	21.60	36.00	54.26	75.68	96.56	39.74	64.74	87.06	98.18	100.00
	0.15	19.36	28.64	43.82	65.76	91.16	37.58	59.40	82.62	96.88	100.00	65.68	91.78	99.34	100.00	100.00
	0.20	25.90	43.52	63.56	86.22	99.10	51.82	81.78	96.20	99.94	100.00	86.50	98.98	100.00	100.00	100.00
	0.25	34.70	57.42	80.18	96.38	99.96	70.08	93.48	99.66	100.00	100.00	96.36	99.96	100.00	100.00	100.00

Table A.14 Estimates of maximal power (percent) using Newcombe Score confidence limits for the risk difference without a continuity correction

		$n = 200$					$n = 600$					$n = 1500$				
p_C		**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**
ϕ	**0.05**	8.34	10.02	13.14	17.18	25.58	11.56	15.20	22.34	31.80	53.40	16.78	25.72	37.86	59.44	86.26
	0.10	11.56	18.76	24.54	40.50	63.82	21.60	36.12	54.26	75.92	96.60	39.74	64.74	87.06	98.20	100.00
	0.15	19.36	28.64	43.82	66.02	91.20	37.58	59.40	82.62	96.92	100.00	65.86	91.80	99.34	100.00	100.00
	0.20	26.30	43.66	63.56	86.38	99.14	51.82	81.86	96.20	99.94	100.00	86.50	98.98	100.00	100.00	100.00
	0.25	35.04	57.54	80.18	96.52	99.96	70.16	93.50	99.66	100.00	100.00	96.40	99.96	100.00	100.00	100.00

Table A.15 Estimates of maximal power (percent) using Wald confidence limits for the risk difference with a continuity correction

		$n = 200$					$n = 600$					$n = 1500$				
p_C		**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**
ϕ	**0.05**	6.68	8.18	11.26	14.82	22.06	10.38	13.66	20.44	29.88	50.64	15.42	24.42	36.18	57.68	85.48
	0.10	9.68	16.08	21.36	36.70	59.16	19.26	33.72	52.20	73.90	95.94	38.20	63.58	86.32	97.94	100.00
	0.15	15.76	24.98	39.90	61.30	89.14	35.18	56.48	81.46	96.54	100.00	63.80	91.12	99.24	99.98	100.00
	0.20	22.44	39.20	60.12	83.84	98.86	49.42	80.18	95.74	99.94	100.00	85.30	98.88	100.00	100.00	100.00
	0.25	30.62	52.84	77.54	95.34	99.92	67.80	92.80	99.62	100.00	100.00	96.00	99.96	100.00	100.00	100.00

Table A.16 Estimates of maximal power (percent) using Hauck-Anderson confidence limits for the risk difference

		$n = 200$					$n = 600$					$n = 1500$				
p_C		**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**
ϕ	**0.05**	7.28	9.56	13.14	16.22	23.80	10.88	14.16	22.34	30.34	52.18	16.08	24.86	37.86	58.60	85.84
	0.10	10.48	16.76	24.54	37.80	61.32	20.62	35.24	54.26	75.12	96.18	39.00	64.24	86.32	98.10	100.00
	0.15	17.52	27.20	43.82	64.04	90.16	36.30	57.20	82.62	96.64	100.00	64.86	91.62	99.34	99.98	100.00
	0.20	23.80	40.38	63.56	84.52	98.98	50.56	81.20	96.20	99.94	100.00	85.94	98.94	100.00	100.00	100.00
	0.25	32.62	55.66	80.18	95.94	99.94	68.82	92.94	99.66	100.00	100.00	96.14	99.96	100.00	100.00	100.00

Table A.17 Estimates of maximal power (percent) using Newcombe Score confidence limits for the risk difference with a continuity correction

		$n = 200$					$n = 600$					$n = 1500$				
p_C		**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**
ϕ	**0.05**	6.94	9.16	11.28	15.80	23.28	10.70	13.78	20.44	30.12	51.54	15.80	24.58	36.18	58.40	85.66
	0.10	9.94	16.48	21.50	37.42	60.52	19.86	34.92	52.20	74.98	96.14	38.58	63.96	86.32	98.10	100.00
	0.15	16.60	26.80	40.32	63.86	89.82	35.74	56.94	81.54	96.60	100.00	64.32	91.58	99.28	99.98	100.00
	0.20	23.36	39.86	63.56	84.52	98.96	50.06	81.14	96.20	99.94	100.00	85.62	98.94	100.00	100.00	100.00
	0.25	32.10	55.52	80.18	96.00	99.94	68.44	92.86	99.66	100.00	100.00	96.08	99.96	100.00	100.00	100.00

Table A.18 Estimates of maximal power (percent) using Wald confidence limits for the relative risk

		$n = 200$					$n = 600$					$n = 1500$				
p_C		**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**
ϕ	**0.05**	7.92	9.68	13.40	17.76	26.14	11.94	15.28	22.94	32.78	54.98	17.30	26.36	39.52	60.98	87.94
	0.10	11.54	19.22	26.44	43.48	67.90	22.68	38.74	58.00	79.38	97.88	42.82	68.56	89.74	98.84	100.00
	0.15	21.90	31.94	49.10	71.80	93.92	41.30	64.90	87.52	98.22	100.00	71.66	94.74	99.62	100.00	100.00
	0.20	29.78	49.96	71.74	91.22	99.76	59.60	88.02	98.44	99.98	100.00	92.18	99.68	100.00	100.00	100.00
	0.25	41.72	67.52	88.58	98.76	100.00	79.38	97.40	99.98	100.00	100.00	98.86	100.00	100.00	100.00	100.00

Table A.19 Estimates of maximal power (percent) using Wald confidence limits for the odds ratio

		$n = 200$					$n = 600$					$n = 1500$				
p_C		**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**
ϕ	**0.05**	8.12	9.82	13.14	16.48	24.30	11.82	15.26	22.34	30.56	49.02	17.30	25.80	37.86	57.82	82.08
	0.10	11.54	18.90	24.54	37.28	54.48	22.06	37.60	54.26	72.78	91.88	41.98	66.72	87.06	97.64	99.88
	0.15	21.88	31.16	43.82	61.74	80.74	40.42	61.80	82.62	95.48	99.80	70.08	93.18	99.44	99.96	100.00
	0.20	28.84	46.68	63.56	81.80	94.34	57.72	84.98	96.62	99.88	100.00	91.12	99.50	100.00	100.00	100.00
	0.25	39.70	62.36	82.50	93.88	98.64	77.56	95.76	99.74	99.98	100.00	98.46	99.98	100.00	100.00	100.00

Table A.20 Estimates of maximal power (percent) using Score confidence limits for the odds ratio

		$n = 200$					$n = 600$					$n = 1500$				
p_C		**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**
ϕ	**0.05**	8.12	9.82	13.14	16.48	24.30	11.82	15.26	22.34	30.56	49.02	17.30	25.80	37.86	57.82	82.08
	0.10	11.88	18.90	24.54	37.28	54.58	22.06	37.60	54.26	72.78	91.90	41.98	66.72	87.06	97.64	99.90
	0.15	21.88	31.18	43.82	61.74	80.86	40.44	61.80	82.62	95.48	99.80	70.08	93.18	99.44	99.96	100.00
	0.20	28.84	46.68	63.56	81.80	94.50	57.86	84.98	96.62	99.88	100.00	91.12	99.50	100.00	100.00	100.00
	0.25	40.14	62.36	82.30	93.88	98.64	77.56	95.76	99.74	99.98	100.00	98.48	99.98	100.00	100.00	100.00

Midpoint Power Simulation Study

Table A.21 Simulation settings used to generate data for midpoint power simulation study

iter	n	p_T	p_C	δ (delta)	ϕ (phi)	ψ (psi)	seed1	seed2
5000	200	0.195	0.2	0.01	0.05	0.0617	40451	61115
5000	200	0.19	0.2	0.02	0.1	0.122	19771	37218
5000	200	0.185	0.2	0.03	0.15	0.1807	98083	60060
5000	200	0.18	0.2	0.04	0.2	0.2381	5872	91199
5000	200	0.175	0.2	0.05	0.25	0.2941	64732	86536
5000	200	0.34125	0.35	0.0175	0.05	0.0749	65114	81031
5000	200	0.3325	0.35	0.035	0.1	0.146	9412	87339
5000	200	0.32375	0.35	0.0525	0.15	0.2135	44131	86352
5000	200	0.315	0.35	0.07	0.2	0.2778	54395	96908
5000	200	0.30625	0.35	0.0875	0.25	0.339	69794	94459
5000	200	0.4875	0.5	0.025	0.05	0.0952	94707	36593
5000	200	0.475	0.5	0.05	0.1	0.1818	26120	73393
5000	200	0.4625	0.5	0.075	0.15	0.2609	18526	98483
5000	200	0.45	0.5	0.1	0.2	0.3333	24881	67474
5000	200	0.4375	0.5	0.125	0.25	0.4	57355	90117
5000	200	0.63375	0.65	0.0325	0.05	0.1307	46870	58584
5000	200	0.6175	0.65	0.065	0.1	0.241	10577	7856
5000	200	0.60125	0.65	0.0975	0.15	0.3352	66558	86389
5000	200	0.585	0.65	0.13	0.2	0.4167	8355	14157
5000	200	0.56875	0.65	0.1625	0.25	0.4878	83383	18514
5000	200	0.78	0.8	0.04	0.05	0.2083	16483	32436
5000	200	0.76	0.8	0.08	0.1	0.3571	88426	15034
5000	200	0.74	0.8	0.12	0.15	0.4688	70612	39418
5000	200	0.72	0.8	0.16	0.2	0.5556	64409	76264
5000	200	0.7	0.8	0.2	0.25	0.625	64713	63165
5000	600	0.195	0.2	0.01	0.05	0.0617	28937	64546

iter	n	p_T	p_C	δ (delta)	ϕ (phi)	ψ (psi)	seed1	seed2
5000	600	0.19	0.2	0.02	0.1	0.122	26264	61913
5000	600	0.185	0.2	0.03	0.15	0.1807	89532	59419
5000	600	0.18	0.2	0.04	0.2	0.2381	47288	60016
5000	600	0.175	0.2	0.05	0.25	0.2941	71755	64177
5000	600	0.34125	0.35	0.0175	0.05	0.0749	79096	46374
5000	600	0.3325	0.35	0.035	0.1	0.146	29526	26754
5000	600	0.32375	0.35	0.0525	0.15	0.2135	18183	79790
5000	600	0.315	0.35	0.07	0.2	0.2778	31018	26768
5000	600	0.30625	0.35	0.0875	0.25	0.339	34254	27723
5000	600	0.4875	0.5	0.025	0.05	0.0952	49577	99313
5000	600	0.475	0.5	0.05	0.1	0.1818	46919	93258
5000	600	0.4625	0.5	0.075	0.15	0.2609	66041	24865
5000	600	0.45	0.5	0.1	0.2	0.3333	46804	92356
5000	600	0.4375	0.5	0.125	0.25	0.4	30386	51141
5000	600	0.63375	0.65	0.0325	0.05	0.1307	22758	39931
5000	600	0.6175	0.65	0.065	0.1	0.241	55372	70973
5000	600	0.60125	0.65	0.0975	0.15	0.3352	92665	83594
5000	600	0.585	0.65	0.13	0.2	0.4167	51360	76790
5000	600	0.56875	0.65	0.1625	0.25	0.4878	48804	44325
5000	600	0.78	0.8	0.04	0.05	0.2083	30723	8535
5000	600	0.76	0.8	0.08	0.1	0.3571	58032	34865
5000	600	0.74	0.8	0.12	0.15	0.4688	51831	49175
5000	600	0.72	0.8	0.16	0.2	0.5556	47784	49371
5000	600	0.7	0.8	0.2	0.25	0.625	52048	56054
5000	1500	0.195	0.2	0.01	0.05	0.0617	66134	96489
5000	1500	0.19	0.2	0.02	0.1	0.122	39983	47385
5000	1500	0.185	0.2	0.03	0.15	0.1807	97832	16162
5000	1500	0.18	0.2	0.04	0.2	0.2381	53767	95715

iter	n	p_T	p_C	δ (delta)	ϕ (phi)	ψ (psi)	seed1	seed2
5000	1500	0.175	0.2	0.05	0.25	0.2941	6421	96345
5000	1500	0.34125	0.35	0.0175	0.05	0.0749	18786	56680
5000	1500	0.3325	0.35	0.035	0.1	0.146	83415	39340
5000	1500	0.32375	0.35	0.0525	0.15	0.2135	24330	37555
5000	1500	0.315	0.35	0.07	0.2	0.2778	42610	83802
5000	1500	0.30625	0.35	0.0875	0.25	0.339	92958	33068
5000	1500	0.4875	0.5	0.025	0.05	0.0952	62381	55685
5000	1500	0.475	0.5	0.05	0.1	0.1818	38314	9410
5000	1500	0.4625	0.5	0.075	0.15	0.2609	71709	90512
5000	1500	0.45	0.5	0.1	0.2	0.3333	63703	25523
5000	1500	0.4375	0.5	0.125	0.25	0.4	87929	98397
5000	1500	0.63375	0.65	0.0325	0.05	0.1307	92078	18161
5000	1500	0.6175	0.65	0.065	0.1	0.241	56363	27139
5000	1500	0.60125	0.65	0.0975	0.15	0.3352	59924	36938
5000	1500	0.585	0.65	0.13	0.2	0.4167	39082	21116
5000	1500	0.56875	0.65	0.1625	0.25	0.4878	6101	1988
5000	1500	0.78	0.8	0.04	0.05	0.2083	96035	39225
5000	1500	0.76	0.8	0.08	0.1	0.3571	6609	51506
5000	1500	0.74	0.8	0.12	0.15	0.4688	83919	20303
5000	1500	0.72	0.8	0.16	0.2	0.5556	260	92366
5000	1500	0.7	0.8	0.2	0.25	0.625	39554	78985

Table A.22 Estimates of midpoint power (percent) using Wald confidence limits for the risk difference without a continuity correction

		$n = 200$					$n = 600$					$n = 1500$				
p_C		**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**
ϕ	**0.05**	6.54	7.26	8.38	10.10	12.56	7.78	9.02	11.04	14.54	21.36	9.54	12.42	15.84	23.90	38.14
	0.10	7.46	10.42	11.78	17.78	24.70	11.26	15.16	22.06	31.24	50.74	17.66	27.60	40.74	58.44	83.62
	0.15	10.96	13.76	18.60	26.38	41.90	17.26	23.32	37.34	53.92	80.42	26.74	45.40	65.48	87.76	98.96
	0.20	12.96	18.40	25.74	38.74	60.32	21.94	37.02	52.80	75.80	94.68	40.76	65.16	87.42	97.94	99.96
	0.25	16.46	23.46	34.46	50.64	75.38	30.38	48.66	70.20	89.08	99.32	54.58	82.44	95.82	99.86	100.00

Table A.23 Estimates of midpoint power (percent) using Farrington-Manning confidence limits for the risk difference

		$n = 200$					$n = 600$					$n = 1500$				
p_C		**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**
ϕ	**0.05**	6.54	7.16	8.38	10.10	12.28	7.76	9.02	11.04	14.54	21.36	9.54	12.42	15.84	23.92	38.12
	0.10	7.46	10.38	11.78	17.98	24.56	11.26	15.16	22.06	31.28	50.68	17.66	27.60	40.74	58.54	83.62
	0.15	10.96	13.76	18.60	26.46	41.90	17.26	23.32	37.34	54.36	80.42	26.64	45.44	65.48	87.80	98.98
	0.20	12.68	18.40	25.74	39.42	60.32	21.94	37.06	52.80	75.82	94.76	40.68	65.22	87.42	97.94	99.96
	0.25	16.20	23.66	34.48	50.74	75.38	30.36	48.80	70.22	89.66	99.34	54.54	82.48	95.90	99.86	100.00

Table A.24 Estimates of midpoint power (percent) using Newcombe Score confidence limits for the risk difference without a continuity correction

		$n = 200$					$n = 600$					$n = 1500$				
p_C		**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**
ϕ	**0.05**	6.54	7.16	8.38	10.10	12.56	7.76	9.02	11.04	14.56	21.40	9.54	12.42	15.84	23.94	38.16
	0.10	7.46	10.42	11.78	18.10	24.70	11.26	15.30	22.06	31.38	50.78	17.66	27.60	40.74	58.56	83.72
	0.15	10.96	13.76	18.60	26.56	42.24	17.26	23.40	37.34	54.54	80.56	26.74	45.46	65.48	87.84	99.00
	0.20	12.96	18.46	25.76	39.78	60.74	21.94	37.18	52.80	75.92	94.86	40.76	65.34	87.42	97.96	99.96
	0.25	16.46	23.84	34.52	50.96	76.06	30.38	48.92	70.34	89.80	99.36	54.64	82.52	96.04	99.86	100.00

Table A.25 Estimates of midpoint power (percent) using Wald confidence limits for the risk difference with a continuity correction

		$n = 200$					$n = 600$					$n = 1500$				
p_C		**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**
ϕ	**0.05**	5.04	6.00	6.96	8.42	9.92	6.84	8.06	10.16	12.98	19.14	8.96	11.60	15.14	22.80	36.76
	0.10	5.98	8.58	10.02	15.08	21.42	10.18	13.64	20.62	29.32	47.60	16.70	26.34	39.50	56.86	82.26
	0.15	8.62	11.22	15.84	23.02	37.86	15.34	21.60	35.42	51.48	78.66	25.06	44.14	64.40	86.86	98.84
	0.20	10.52	15.34	22.66	34.90	56.22	19.74	35.02	50.70	74.02	93.82	39.00	63.80	86.72	97.68	99.96
	0.25	13.32	20.66	30.56	46.18	71.60	28.02	46.38	68.06	87.68	99.14	52.70	81.64	95.60	99.86	100.00

Table A.26 Estimates of midpoint power (percent) using Hauck-Anderson confidence limits for the risk difference

		$n = 200$					$n = 600$					$n = 1500$				
p_C		**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**
ϕ	**0.05**	5.54	6.68	8.38	9.50	11.06	7.22	8.38	11.04	13.50	20.24	9.28	12.00	15.84	23.40	37.48
	0.10	6.48	8.96	11.78	15.74	22.80	10.76	14.30	22.06	30.46	49.04	17.10	27.20	39.50	57.26	82.88
	0.15	9.48	12.52	18.60	25.12	39.28	16.26	22.60	37.34	52.90	79.52	25.94	44.54	65.48	87.54	98.92
	0.20	11.62	16.62	25.74	36.26	58.20	20.64	35.66	52.80	74.54	94.22	39.80	64.40	86.72	97.78	99.96
	0.25	14.74	21.92	34.46	48.24	73.36	29.08	47.80	70.20	88.72	99.28	53.56	81.94	95.82	99.86	100.00

Table A.27 Estimates of midpoint power (percent) using Newcombe Score confidence limits for the risk difference with a continuity correction

		$n = 200$					$n = 600$					$n = 1500$				
p_C		**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**
ϕ	**0.05**	5.26	6.48	7.00	9.24	10.70	7.06	8.16	10.16	13.38	19.78	9.14	11.76	15.14	23.20	37.22
	0.10	6.12	8.76	10.12	15.66	22.50	10.42	14.18	20.64	30.08	48.52	16.90	26.98	39.50	57.14	82.72
	0.15	9.02	12.18	16.56	24.78	38.94	15.92	22.14	36.16	52.88	79.34	25.64	44.42	65.38	87.50	98.90
	0.20	11.16	16.20	25.74	36.26	57.92	20.20	35.56	52.80	74.72	94.22	39.44	64.26	86.72	97.78	99.96
	0.25	14.14	21.88	34.46	48.88	73.54	28.68	47.66	70.20	88.72	99.28	53.20	81.90	95.82	99.86	100.00

Table A.28 Estimates of midpoint power (percent) using Wald confidence limits for the relative risk

		$n = 200$					$n = 600$					$n = 1500$				
p_C		**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**
ϕ	**0.05**	6.22	6.74	8.54	10.36	12.76	7.94	9.00	11.22	14.84	21.98	9.72	12.66	16.34	24.72	39.24
	0.10	7.46	10.48	12.28	18.92	26.22	11.60	16.06	23.48	34.04	54.12	18.50	29.38	42.88	62.38	86.42
	0.15	11.98	14.52	20.72	29.24	46.64	18.68	26.18	41.56	59.42	84.94	29.46	50.60	71.04	91.36	99.36
	0.20	14.14	20.52	30.40	44.22	67.62	24.92	42.04	60.32	82.94	97.10	47.08	73.12	92.18	98.96	100.00
	0.25	19.08	28.08	42.30	60.06	82.52	36.80	58.12	80.34	94.70	99.90	64.30	90.06	98.74	99.98	100.00

Table A.29 Estimates of midpoint power (percent) using Wald confidence limits for the odds ratio

		$n = 200$					$n = 600$					$n = 1500$				
p_C		**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**
ϕ	**0.05**	6.42	6.96	8.38	9.80	12.06	7.90	9.00	11.04	14.06	19.44	9.72	12.44	15.84	23.34	34.80
	0.10	7.46	10.38	11.78	16.46	21.58	11.38	15.90	22.06	30.00	42.56	18.26	28.22	40.72	55.60	75.40
	0.15	11.98	14.72	18.60	25.14	33.90	18.32	24.70	37.34	50.62	68.40	28.62	47.94	66.60	84.84	95.62
	0.20	13.88	19.42	25.74	35.50	46.34	23.86	39.52	54.28	71.64	85.06	46.02	69.56	88.04	96.88	99.30
	0.25	18.06	25.86	36.52	46.66	57.80	35.38	53.70	72.12	85.64	94.22	62.20	86.74	96.40	99.70	100.00

Table A.30 Estimates of midpoint power (percent) using Score confidence limits for the odds ratio

		$n = 200$					$n = 600$					$n = 1500$				
p_C		**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**	**0.20**	**0.35**	**0.50**	**0.65**	**0.80**
ϕ	**0.05**	6.42	6.96	8.38	9.80	12.06	7.90	9.00	11.04	14.06	19.44	9.72	12.44	15.84	23.34	34.80
	0.10	7.64	10.38	11.78	16.46	21.60	11.38	15.90	22.06	30.00	42.56	18.26	28.22	40.72	55.60	75.40
	0.15	11.98	14.82	18.60	25.14	33.90	18.32	24.70	37.34	50.62	68.42	28.62	47.94	66.60	84.84	95.62
	0.20	13.88	19.42	25.74	35.50	46.66	24.02	39.52	54.28	71.64	85.06	46.02	69.56	88.04	96.88	99.30
	0.25	18.34	25.86	36.14	46.66	58.64	35.38	53.76	72.12	85.64	94.22	62.22	86.74	96.40	99.70	100.00

Appendix B - Numerical Example Code

Below is SAS code used to perform noninferiority testing on the provided number example from Rothmann et. al (2012).

```
***********************************************************************************;
* Nick Bloedow's Master Project: Numerical Example (Rothmann et. al 2012, p. 262) *;
***********************************************************************************;

/*
Notes: Numerical example from p. 262 of Rothmann et. al (2012)
       Looks at an investigational antibiotic being compared to a standard antibiotic in uncomplicated
         bacterial sinusitis.
       Cure rates for treatment and control groups were observed to be 89/100 (89%) and 92/100 (92%)
         respectively.
*/


  *Creates the hypothetical data (Example) in cell count form to be easily read by PROC FREQ;
    data Example;
      title "Bacterial Sinusitis Antibiotic";
      input Group $ Outcome $ Count;
      datalines ;
      Treatment Success 89
      Treatment Failure 11
      Control   Success 92
      Control   Failure 08
      ;
    run;


  *Extracts Key Statistics from the 2x2 Table for Group*Outcome and Calculates Noninferiority Confidence
     Limits for Relative Risk;
    proc freq data=Example order=data;
      tables Group*Outcome / riskdiff relrisk norow nocol nopercent alpha=0.10 ;
      output out=risk_stat n rsk11 rsk21 rdif1 rrc1 or;
      weight Count / zeros;
      title2 "Simulation Statistics/Noninferiority Testing (Relative Risk)";
    run;
```

```
*Extracts the Risk Difference confidence limits for different methods preprogammed into SAS;
  proc freq data=Example order=data;
    tables Group*Outcome / riskdiff(noninf method=WALD     margin=0.10)         alpha=.05;
    tables Group*Outcome / riskdiff(noninf method=FM       margin=0.10)         alpha=.05;
    tables Group*Outcome / riskdiff(noninf method=NEWCOMBE margin=0.10)         alpha=.05;
    tables Group*Outcome / riskdiff(noninf method=WALD     margin=0.10 correct) alpha=.05;
    tables Group*Outcome / riskdiff(noninf method=HA       margin=0.10 )        alpha=.05;
    tables Group*Outcome / riskdiff(noninf method=NEWCOMBE margin=0.10 correct) alpha=.05;
    ods output Pdiffnoninf=noninf_rdiff ;
    weight Count / zeros;
      title2 "Noninferiority Testing (Risk Difference)";
  run;


*Extracts the Odds Ratio confidence limits for different methods preprogammed into SAS;
  proc freq data=Example order=data;
    tables Group*Outcome / OR(CL=WALD) alpha=.10;
    tables Group*Outcome / OR(CL=SCORE) alpha=.10;
    ods output OddsRatioCLs=noninf_or;
    weight Count / zeros;
    title2 "Noninferiority Testing (Odds Ratio)";
  run;
```

Appendix C - Simulation Code

Generation of Simulation Data

Below is SAS code used to generate simulated data for noninferiority testing.

```
**********************************************************************************************************;
* Nick Bloedow's Master Project: Generates Simulated Count Data from Two Independent Binomial            *;
*                                Distributions (Type I Error, Maximal Power, & Midpoint Power)           *;
**********************************************************************************************************;

/*
Notes: Generates simulated count data from two independent binomial distributions (Treatment & Control)
       Binomial distributions have the same number of individuals (n), but same true proportions of
         success (pT & pC)
       Can use data to test for equivalence and noninferiority between sample proportions for Type I Error,
         Maximal Power, & Midpoint Power
*/



/*
  Macro to Generates Simulated Count Data from Two Independent Binomial Distributions for Type I Error,
    Maximal Power, & Midpoint Power

    iter = number of iterations for each simulation
    n = size of individuals in each group
    pT = true proportion of success for Treatment group
    pC = true proportion of success for Control group
    phi = specified percentage decrease between Treatment & Control group
    seed1 = specified randomization seed for the Treatment group
    seed2 = specified randomization seed for the Control group
*/
```

```
%macro sim_iter_n_pC_phi(iter,n,pT,pC,phi,seed1,seed2);

*Titles pertaining to specified parameters used in the simulation of data;
  title2 "Number of Iterations = &iter";
  title3 "Group Size = &n";
  title4 "True Proportions: Treatment = &pT    Control = &pC";
  title5 "Delta = &delta    Phi = &phi    Psi = &psi";
  title6 "Seeds: Treatment = &seed1    Control = &seed2";
  title7 ;


*Template for generate a bar chart for checking the distribution of simulated counts for each group;
  proc template;
  define statgraph BarChart;
  dynamic _X _Title;
  begingraph;
     entrytitle halign=center _Title;
     layout overlay / yaxisopts=(griddisplay=on)
             xaxisopts=(type=discrete discreteopts=(tickvaluefitpolicy=thin)
                         display=(TICKS TICKVALUES LINE));
     barchart x=_X / name='bar' stat=pct ;
     discretelegend 'bar' / opaque=true border=true halign=right
          valign=top across=1 location=inside;
       ods graphics / ANTIALIASMAX=&iter;
       endlayout;
  endgraph;
  end;
  run;
```

```
*Template for generate a histogram (w/ normal density curve) to checking the distribution of simulated
   counts for each group;
  proc template;
  define statgraph Histogram;
  dynamic _X _Title;
  begingraph;
     entrytitle halign=center _Title;
     layout overlay / yaxisopts=(griddisplay=on)
             xaxisopts=(type=linear display=(TICKS TICKVALUES LINE ));
     histogram _X / name='hist' legendlabel='Sample';
     densityplot _X / name='density' normal() lineattrs=(color=blue) legendlabel='Density';
     discretelegend 'hist' 'density' / opaque=true border=true halign=right
          valign=top across=1 location=inside;
       ods graphics / ANTIALIASMAX=&iter;
     endlayout;
  endgraph;
  end;
  run;


*Generates a data set (Treatment) with the simulated counts of success and failures for the Treatment
   Group;
  data Treatment;
    do Simulation=1 to &iter;
      Group='Treatment';
      Success=ranbin(&seed1,&n,&pT);
      Failure=&n-Success;
    output;
    end;
  run;

*Transposes Treatment data set to be analyzed later with PROC FREQ;
  proc transpose data=Treatment out=Trans_treat(rename=(col1=Count)) name=Outcome;
    var Success Failure;
    by Simulation Group;
  run;

  data Trans_treat;
    set Trans_treat;
    label Outcome=' ';
  run;
```

```
*Creates a Bar Chart using the simulated number of successes for the Treatment Group;
  proc sgrender data=Treatment template=BarChart;
     dynamic _X="Success" _Title="Sampling Distribution for Treatment Group";
  run;

*Creates a Histogram (w/ Normal Density Plot)using the simulated number of successes for the Treatment
   Group;
  proc sgrender data=Treatment template=Histogram;
     dynamic _X="Success" _Title="Sampling Distribution for Treatment Group";
  run;


*Generates a data set (Control) with the simulated counts of success and failures for the Control Group;
  data Control;
    do Simulation=1 to &iter;
      Group='Control';
      Success=ranbin(&seed2,&n,&pC);
      Failure=&n-Success;
    output;
    end;
  run;

*Transposes Control data set to be analyzed later with PROC FREQ;
  proc transpose data=Control out=Trans_control(rename=(col1=Count)) name=Outcome;
    var Success Failure;
    by Simulation Group;
  run;

  data Trans_control;
    set Trans_control;
    label Outcome=' ';
  run;


*Creates a Bar Chart using the simulated number of successes for the Control Group;
  proc sgrender data=Control template=BarChart;
     dynamic _X="Success" _Title="Sampling Distribution for Control Group";
  run;
```

```
*Creates a Histogram (w/ Normal Density Plot) using the simulated number of successes for the Control
   Group;
  proc sgrender data=Control template=Histogram;
     dynamic _X="Success" _Title="Sampling Distribution for Control Group";
  run;


*Creates a combined simulated dataset (sim_&now) for the number of success and failures for the both
   groups;
  data sim_&now;
    set Trans_treat Trans_control;
  run;

*Sorts combined simulated dataset by Simulation and Group (descending order) for later analysis using
   PROC FREQ;
  proc sort data=sim_&now;
    by Simulation descending Group;
  run;


*Creates a Side-by-Side Bar Chart using the simulated number of successes for the Control & Treatment
   Group;
  proc sgpanel data=sim_&now;
    panelby Group / novarname;
      where Outcome="Success";
    colaxis type=discrete fitpolicy=thin;
    vbar Count / stat=pct nooutline;
      title7 "Sampling Distribution for Successes";
  run;

*Creates a Side-by-Side Histogram (w/ Normal Density Plot) using the simulated number of successes for
   the Control & Treatment Group;
  proc sgpanel data=sim_&now;
    panelby Group / novarname;
      where Outcome="Success";
    histogram Count / binwidth=1 nooutline;
    density   Count / lineattrs=(color=mediumblue);
      title7 "Sampling Distribution for Successes";
  run;

%mend sim_iter_n_pC_phi;
```

```
/*
  Macro to Import Specified Parameters to use when generating each Simulation
    settings = control dataset that has the specified simulation settings (w/ proper column names)
*/

  %macro sim_loop(settings);
    %local dsid rc now rows cols;
    %let dsid = %sysfunc(open(&settings));
      %let now=0;
    %let rows=%sysfunc(attrn(&dsid,nobs));  %* loops;
    %let cols=%sysfunc(attrn(&dsid,nvars));  %* vars;

    %* loops = rows;
    %do %while (%sysfunc(fetch(&dsid)) = 0);  %* outer loop across rows;

      %let now=%eval(&now + 1);

      %* get vars from cols;
      %do c = 1 %to &cols;  %* inner nested loop;
        %local v t;
        %let v=%sysfunc(varname(&dsid,&c));
        %local &v;
        %let t = %sysfunc(vartype(&dsid, &c));  %* N or C;
        %let &v = %sysfunc(getvar&t(&dsid, &c));  %* show var and value in log;
        %end;

      %put ** loop # &now of &rows **;

        *Generates simulated count data using specified parameter for both binomial distribution;
        %sim_iter_n_pC_phi(iter=&iter,n=&n,pT=&pT,pC=&pC,phi=&phi,seed1=&seed1,seed2=&seed2)

        *Exports the generated dataset for each simulation into a separate csv file
           (sim_iter_n_pT_pC_phi.csv);
        proc export data=Work.sim_&now outfile="E:\Simulation Study\Data\SAS Output\&sim_data"
          dbms=csv replace;
          run;
```

```
      *Deletes temporary data sets used for creating the final dataset for each simulation;
        proc delete data=Work.Treatment;
        run;

        proc delete data=Work.Trans_treat;
        run;

        proc delete data=Work.Control;
        run;

        proc delete data=Work.Trans_control;
        run;

    %out:
    %end; %* while fetch loop;

  %let rc = %sysfunc(close(&dsid));

%mend sim_loop;
```

```
ods rtf file = "E:\Simulation Study\Data\SAS Output\Type I Error\Generated Data (Type I Error).doc";

  *Creates a controls dataset (sim_setting) that has the necessary specifications for each simulation;
    data sim_setting;
      do iter=5000;
        do n=200,600,1500;
          do pC=0.2 to 0.8 by 0.15;
            do phi=0.05 to 0.25 by 0.05;
             pT=round(((1-phi)*pC),0.0001);
               delta=round((pC-pT),0.0001);;
               psi=round(1-((pT*(1-pC))/(pC*(1-pT))),0.0001);
               seed1=floor(100000*ranuni(13));
               seed2=floor(100000*ranuni(24));
               sim_data=cat('Type I Error\sim_',put(iter,z4.),'_',put(n,z4.),'_',put(pT,f6.4),'_',
                          put(pC,f4.2),'_',put(phi,f4.2),'_',put(seed1,z5.),'_',put(seed2,z5.),'.csv');
               noninf_results=cat('Type I Error\noninf_',put(iter,z4.),'_',put(n,z4.),'_',put(pT,f6.4),'_',
                              put(pC,f4.2),'_',put(phi,f4.2),'_',put(seed1,z5.),'_',put(seed2,z5.),'.csv');
               output;
            end;
          end;
      end;
      end;
    run;

  *Exports the control dataset into a separate csv file (Simulation Settings.csv) for future reference;
    proc export data=Work.sim_setting outfile="E:\Simulation Study\Data\SAS Output\Type I Error\Simulation
                                                    Settings (Type I Error).csv" dbms=csv replace;
    run;

  *Prints out of the specified simulation settings;
    proc print data=sim_setting;
      var iter n pT pC delta phi psi seed1 seed2;
      title1 "Simulation Study (Type I Error)";
      title2 'Specified Settings';
    run;

  *Executes macro (%sim_loop) to generate the simulated data for all specified settings;
    title  "Simulation Study (Type I Error)";
    %sim_loop(sim_setting);

ods rtf close;
```

```
ods rtf file = "E:\Simulation Study\Data\SAS Output\Power\Generated Data (Maximal Power).doc";

  *Creates a controls dataset (sim_setting) that has the necessary specifications for each simulation;
    data sim_setting;
      do iter=5000;
        do n=200,600,1500;
          do pC=0.2 to 0.8 by 0.15;
            do phi=0.05 to 0.25 by 0.05;
             pT=pC;
               delta=round((pC-((1-phi)*pC)),0.0001);
               psi=round(1-((((1-phi)*pC)*(1-pC))/(pC*(1-((1-phi)*pC)))),0.0001);
               seed1=floor(100000*ranuni(13));
               seed2=floor(100000*ranuni(24));
               sim_data=cat('Maximal Power\sim_',put(iter,z4.),'_',put(n,z4.),'_',put(pC,f4.2),
                          '_',put(phi,f4.2),'_',put(seed1,z5.),'_',put(seed2,z5.),'.csv');
               noninf_results=cat('Maximal Power\noninf_',put(iter,z4.),'_',put(n,z4.),'_',put(pC,f4.2),
                               '_',put(phi,f4.2),'_',put(seed1,z5.),'_',put(seed2,z5.),'.csv');
               output;
            end;
          end;
      end;
      end;
    run;

  *Exports the control dataset into a separate csv file (Simulation Settings.csv) for future reference;
    proc export data=Work.sim_setting outfile="E:\Simulation Study\Data\SAS Output\Maximal Power\Simulation
                                                    Settings (Maximal Power).csv" dbms=csv replace;
    run;

  *Prints out of the specified simulation settings;
    proc print data=sim_setting;
      var iter n pT pC delta phi psi seed1 seed2;
      title1 "Simulation Study (Maximal Power)";
      title2 'Specified Settings';
    run;

  *Executes macro (%sim_loop) to generate the simulated data for all specified settings;
    title  "Simulation Study (Maximal Power)";
    %sim_loop(sim_setting);

ods rtf close;
```

```
ods rtf file = "E:\Simulation Study\Data\SAS Output\Midpoint\Generated Data (Midpoint Power).doc";

  *Creates a controls dataset (sim_setting) that has the necessary specifications for each simulation;
    data sim_setting;
      do iter=5000;
        do n=200,600,1500;
          do pC=0.2 to 0.8 by 0.15;
            do phi=0.05 to 0.25 by 0.05;
             pT=((1-(.5*phi))*pC);
               delta=round((pC-((1-phi)*pC)),0.0001);
               psi=round(1-((((1-phi)*pC)*(1-pC))/(pC*(1-((1-phi)*pC)))),0.0001);
               seed1=floor(100000*ranuni(13));
               seed2=floor(100000*ranuni(24));
               sim_data=cat('Midpoint Power\sim_',put(iter,z4.),'_',put(n,z4.),'_',put(pT,f7.5),'_',
                          put(pC,f4.2),'_',put(phi,f4.2),'_',put(seed1,z5.),'_',put(seed2,z5.),'.csv');
               noninf_results=cat('Midpoint Power\noninf_',put(iter,z4.),'_',put(n,z4.),'_',put(pT,f7.5),
                          '_',put(pC,f4.2),'_',put(phi,f4.2),'_',put(seed1,z5.),'_',put(seed2,z5.),'.csv');
               output;
            end;
          end;
      end;
      end;
    run;

  *Exports the control dataset into a separate csv file (Simulation Settings.csv) for future reference;
    proc export data=Work.sim_setting outfile="E:\Simulation Study\Data\SAS Output\Midpoint Power\
                                                    Simulation Settings (Midpoint).csv" dbms=csv replace;
    run;

  *Prints out of the specified simulation settings;
    proc print data=sim_setting;
      var iter n pT pC delta phi psi seed1 seed2;
      title1 "Simulation Study (Midpoint Power)";
      title2 'Specified Settings';
    run;

  *Executes macro (%sim_loop) to generate the simulated data for all specified settings;
    title  "Simulation Study (Midpoint Power)";
    %sim_loop(sim_setting);

ods rtf close;
```

Noninferiority Testing

Below is SAS code used to perform the noninferiority testing on the previously simulated data.

```
**********************************************************************************************************;
* Nick Bloedow's Master Project: Noninferiority Testing on Two Binomial Proportions                      *;
*                                  (Type I Error, Maximal Power, & Midpoint Power)                       *;
**********************************************************************************************************;

/*
Notes: Performs Noninferiority (One-Sided Equivalence) Testing on Two Proportions
       Uses 90% Confidence Limits based on Risk Difference, Relative Risk, & Odds Ratio
       Implements different test methods preprogrammed inside of SAS
*/



/*
  Macro to Perform Noninferiority Testing based on 90% Confidence Limits for Risk Difference, Relative
    Risk, & Odds Ratio

    Tests used for Risk Difference = Wald (w/ & w/o CC), Hauck-Anderson (HA), Farrington-Manning (FM),
      and Newcombe (w/ & w/o CC)
    Tests used for Relative Risk = Wald (Asymptotic)
    Tests used for Odds Ratio = Wald and SCORE

    iter = number of iterations for simulation
    n = size of individuals in each group
    pT = true proportion of success for Treatment group
    pC = true proportion of success for Control group
    phi = specified percentage decrease between Treatment & Control group
    seed1 = specified seed for randomization with the Treatment group
    seed2 = specified seed for randomization with the Control group
*/
```

```
%macro noninf_iter_n_pC_phi(iter,n,pT,pC,phi,seed1,seed2);

*Titles pertaining to specified parameters used in the simulation of data;
  title2 "Number of Iterations = &iter";
  title3 "Group Size = &n";
  title4 "True Proportions: Treatment = &pT    Control = &pC";
  title5 "Delta = &delta    Phi = &phi    Psi = &psi";
  title6 "Seeds: Treatment = &seed1    Control = &seed2";
  title7 " ";


*Restricts SAS to exclude all generated ODS output;
  ods exclude all;


*Extracts Key Statistics from the 2x2 Table for Group*Outcome and Calculates Noninferiority Confidence
   Limits for Relative Risk;
  proc freq data=sim_&now order=data;
  tables Group*Outcome / riskdiff relrisk norow nocol nopercent alpha=0.10 ;
    output out=risk_stat n rsk11 rsk21 rdif1 rrc1 or;
    weight Count / zeros;
    by Simulation;
    title7 "Simulation Statistics";
  run;

*Creates master data set (noninf_&now) with renamed variables pertaining to simulation settings,
   estimates of statistics (Risk Difference, Relative Risk, & Odds Ratio), and Relative Risk CLs;
  data noninf_&now;
    set risk_stat;

      ITERATION=&iter;
      GROUP=&n;

      PT=Round(&pT,0.0001);
      PC=Round(&pC,0.0001);

      DELTA=&delta;
      PHI=&phi;
      PSI=&psi;
```

```
  _RISK1_=Round(_RSK11_,0.0001);
  _RISK2_=Round(_RSK21_,0.0001);

  _RDIFF_=Round(_RDIF1_,0.0001);

  _RR_=Round(_RRC1_,0.0001);

  _OR_=Round(_RROR_,0.0001);


  LNIM_RR=Round((1-&phi),0.0001);

  LCL_RR_WALD=Round(L_RRC1,0.0001);

  if ( LNIM_RR<LCL_RR_WALD ) then Noninf_RR_WALD="NON"; else Noninf_RR_WALD="INF";


  label ITERATION='Number of Iterations Used in Simulation'
        GROUP='Number of Subjects in Each Group'
        PT='True Population Proportion of Success (Treatment)'
        PC='True Population Proportion of Success (Control)'
        DELTA='Risk Difference Margin'
        PHI='Relative Risk Percentage Decrease'
        PSI='Odds Ratio Percentage Decrease'
        _RISK1_='Proportion of Success (Treatment)'
        _RISK2_='Proportion of Success (Control)'
        _RDIFF_='Risk Difference of Success (T-C)'
        _RR_='Relative Risk of Success'
        _OR_='Odds Ratio of Success'
        LNIM_RR='Lower Noninferiority Margin for Relative Risk'
        LCL_RR_WALD='Lower 90% WALD Confidence Limit for Relative Risk'
        Noninf_RR_WALD='Noninferiority Test Result for Relative Risk [WALD Method]';

  keep Simulation ITERATION GROUP PT PC DELTA PHI PSI _RISK1_ _RISK2_ _RDIFF_ _RR_ _OR_
         LNIM_RR LCL_RR_WALD Noninf_RR_WALD;

run;
```

```
*Extracts the Risk Difference confidence limits for different methods preprogammed into SAS;
  proc freq data=sim_&now order=data;
    tables Group*Outcome / riskdiff(noninf method=WALD     margin=&delta)         alpha=.05;
    tables Group*Outcome / riskdiff(noninf method=FM       margin=&delta)         alpha=.05;
    tables Group*Outcome / riskdiff(noninf method=NEWCOMBE margin=&delta)         alpha=.05;
    tables Group*Outcome / riskdiff(noninf method=WALD     margin=&delta correct) alpha=.05;
    tables Group*Outcome / riskdiff(noninf method=HA       margin=&delta )        alpha=.05;
    tables Group*Outcome / riskdiff(noninf method=NEWCOMBE margin=&delta correct) alpha=.05;
    ods output Pdiffnoninf=noninf_rdiff ;
    weight Count / zeros;
    by Simulation;
    title7 "Noninferiority Testing (Risk Difference)";
  run;

*Creates a list (rdiff_tests) of methods corresponding to the Confidence Limits used for
   testing Noninferiority with Risk Difference statistic;
  data rdiff_tests;
    do Simulation=1 to &iter by 1;
      length method $12;
        method='WALD';
          output;
          method='FM';
          output;
          method='NEWCOMBE';
          output;
          method='WALD_CC';
          output;
          method='HA';
          output;
          method='NEWCOMBE_CC';
          output;
    end;
  run;

*Merges methods names with calculated confidence limits for Risk Difference;
  data noninf_rdiff;
    merge noninf_rdiff rdiff_tests;
    by Simulation;
  run;
```

```
*Transposes Noninferiority Statistics (Lower Confidence Limit) with Risk Difference into separate columns
   for each Method;
  proc transpose data=noninf_rdiff out=LCL_RDIFF(drop=_name_ _label_) prefix=LCL_RDIFF_;
      by Simulation;
      id method;
      var LowerCL;
    run;

*Merges Noninferiority statistics for Risk Difference with rest of relevant information (noninf_&now);
  data noninf_&now;
    merge noninf_&now LCL_RDIFF;
    by Simulation;

    LNIM_RDIFF=Round(-&delta,0.0001);

    LCL_RDIFF_WALD=Round(LCL_RDIFF_WALD,0.0001);
    LCL_RDIFF_FM=Round(LCL_RDIFF_FM,0.0001);
    LCL_RDIFF_NEWCOMBE=Round(LCL_RDIFF_NEWCOMBE,0.0001);
    LCL_RDIFF_WALD_CC=Round(LCL_RDIFF_WALD_CC,0.0001);
    LCL_RDIFF_HA=Round(LCL_RDIFF_HA,0.0001);
    LCL_RDIFF_NEWCOMBE_CC=Round(LCL_RDIFF_NEWCOMBE_CC,0.0001);

      if ( LNIM_RDIFF<LCL_RDIFF_WALD )        then Noninf_DIFF_WALD="NON";
        else Noninf_DIFF_WALD="INF";
      if ( LNIM_RDIFF<LCL_RDIFF_FM )          then Noninf_DIFF_FM="NON";
        else Noninf_DIFF_FM="INF";
      if ( LNIM_RDIFF<LCL_RDIFF_NEWCOMBE )    then Noninf_DIFF_NEWCOMBE="NON";
        else Noninf_DIFF_NEWCOMBE="INF";
      if ( LNIM_RDIFF<LCL_RDIFF_WALD_CC )     then Noninf_DIFF_WALD_CC="NON";
        else Noninf_DIFF_WALD_CC="INF";
      if ( LNIM_RDIFF<LCL_RDIFF_HA )          then Noninf_DIFF_HA="NON";
        else Noninf_DIFF_HA="INF";
      if ( LNIM_RDIFF<LCL_RDIFF_NEWCOMBE_CC ) then Noninf_DIFF_NEWCOMBE_CC="NON";
        else Noninf_DIFF_NEWCOMBE_CC="INF";
```

```
    label LNIM_RDIFF='Lower Noninferiority Margin for Risk Difference'
          LCL_RDIFF_WALD='Lower 90% WALD (Asymptotic) Confidence Limit for Risk Difference'
          LCL_RDIFF_FM='Lower 90% FM Confidence Limit for Risk Difference'
          LCL_RDIFF_NEWCOMBE='Lower 90% NEWCOMBE Confidence Limit for Risk Difference'
          LCL_RDIFF_WALD_CC='Lower 90% WALD Confidence Limit (w/CC) for Risk Difference'
          LCL_RDIFF_HA='Lower 90% HA Confidence Limit for Risk Difference'
          LCL_RDIFF_NEWCOMBE_CC='Lower 90% NEWCOMBE Confidence Limit (w/CC) for Risk Difference'
          Noninf_DIFF_WALD='Noninferiority Test Result for Risk Difference [WALD Method]'
          Noninf_DIFF_FM='Noninferiority Test Result for Risk Difference [FM Method]'
          Noninf_DIFF_NEWCOMBE='Noninferiority Test Result for Risk Difference [NEWCOMBE Method]'
          Noninf_DIFF_WALD_CC='Noninferiority Test Result for Risk Difference [WALD Method (w/CC)]'
          Noninf_DIFF_HA='Noninferiority Test Result for Risk Difference [HA Method]'
          Noninf_DIFF_NEWCOMBE_CC='Noninferiority Test Result for Risk Difference
                                   [NEWCOMBE Method (w/CC)]';

  run;


*Extracts the Odds Ratio confidence limits for different methods preprogammed into SAS;
  proc freq data=sim_&now order=data;
      tables Group*Outcome / OR(CL=WALD) alpha=.10;
      tables Group*Outcome / OR(CL=SCORE) alpha=.10;
      ods output OddsRatioCLs=noninf_or;
    weight Count / zeros;
    by Simulation;
    title7 "Noninferiority Testing (Odds Ratio)";
  run;

*Creates a list (or_tests) of methods corresponding to the Confidence Limits used for testing
   Noninferiority with Odds Ratio statistic;
    data or_tests;
    do Simulation=1 to &iter by 1;
      length method $7;
        method='WALD';
        output;
        method='SCORE';
        output;
    end;
    run;
```

```
*Merges methods names with calculated confidence limits for Odds Ratio;
  data noninf_or;
    merge noninf_or or_tests;
    by Simulation;
  run;

*Transposes Noninferiority Statistics (Lower Confidence Limit) with Odds Ratio into separate columns
   for each Method;
  proc transpose data=noninf_or out=LCL_OR(drop=_name_ _label_) prefix=LCL_OR_;
      by Simulation;
      id method;
      var LowerCL;
    run;

*Merges Noninferiority statistics for Odds Ratio with rest of relevant information (noninf_&now);
  data noninf_&now;
    merge noninf_&now LCL_OR;
    by Simulation;

    LNIM_OR=Round((1-&psi),0.0001);

    LCL_OR_WALD=Round(LCL_OR_WALD,0.0001);
    LCL_OR_SCORE=Round(LCL_OR_SCORE,0.0001);

    if ( LNIM_OR<LCL_OR_WALD )  then Noninf_OR_WALD="NON";  else Noninf_OR_WALD="INF";
    if ( LNIM_OR<LCL_OR_SCORE ) then Noninf_OR_SCORE="NON"; else Noninf_OR_SCORE="INF";

    label LNIM_OR='Lower Noninferiority Margin for Odds Ratio'
          LCL_OR_WALD='Lower 90% WALD Confidence Limit for Odds Ratio'
          LCL_OR_SCORE='Lower 90% SCORE Confidence Limit for Odds Ratio'
          Noninf_OR_WALD='Noninferiority Test Result for Odds Ratio [WALD Method]'
          Noninf_OR_SCORE='Noninferiority Test Result for Odds Ratio [SCORE Method]';

  run;


%mend noninf_iter_n_pC_phi;
```

```
/*
  Macro to Peform Noninferiority Testing on all specified Simulation Settings
    settings = control dataset that has the specified simulation settings (w/ proper column names)
*/

  %macro noninf_loop(settings);
    %local dsid rc now rows cols;
    %let dsid = %sysfunc(open(&settings));
      %let now=0;
    %let rows=%sysfunc(attrn(&dsid,nobs));  %* loops;
    %let cols=%sysfunc(attrn(&dsid,nvars));  %* vars;

    %* loops = rows;
    %do %while (%sysfunc(fetch(&dsid)) = 0);  %* outer loop across rows;

      %let now=%eval(&now + 1);

      %* get vars from cols;
      %do c = 1 %to &cols;  %* inner nested loop;
        %local v t;
        %let v=%sysfunc(varname(&dsid,&c));
        %local &v;
        %let t = %sysfunc(vartype(&dsid, &c));  %* N or C;
        %let &v = %sysfunc(getvar&t(&dsid, &c));  %* show var and value in log;
        %end;

      %put ** loop # &now of &rows **;

        *Imports the generated dataset for each simulation;
          proc import out=Work.sim_&now datafile="E:\Simulation Study\Data\SAS Output\&sim_data"
            dbms=csv replace;
            getnames=yes;
          run;

        *Performs Noninferiority Testing using Risk Difference, Relative Risk, and Odds Ratio on
           each simulated dataset;
        %noninf_iter_n_pC_phi(iter=&iter,n=&n,pT=&pT,pC=&pC,phi=&phi,seed1=&seed1,seed2=&seed2)
```

```
      *Exports the final noninferiority results dataset for each simulation into a separate
         csv file (noninf_iter_n_pT_pC_phi.csv);
      proc export data=Work.noninf_&now
        outfile="E:\Simulation Study\Analysis\Noninferiority Testing\SAS Output\&noninf_results"
        dbms=csv replace;
        run;

      *Deletes temporary data sets used for creating the final noninferiority results dataset
         for each simulation;
        proc delete data=Work.Risk_stat;
        run;

        proc delete data=Work.noninf_rdiff;
        run;

        proc delete data=Work.rdiff_tests;
        run;

        proc delete data=Work.LCL_RDIFF;
        run;

        proc delete data=Work.noninf_or;
        run;

        proc delete data=Work.or_tests;
        run;

        proc delete data=Work.LCL_OR;
        run;

    %out:
    %end; %* while fetch loop;

  %let rc = %sysfunc(close(&dsid));

%mend noninf_loop;
```

```
  ods rtf file = "E:\Simulation Study\Analysis\Noninferiority Testing\SAS Output\Type I Error\
                      Noninferiority Testing (Type I Error).doc";

  *Imports the control dataset [Simulation Settings (Type I Error).csv] into SAS for reference;
    proc import out=Work.sim_typeIerror
      datafile="E:\Simulation Study\Data\SAS Output\Type I Error\Simulation Settings (Type I Error).csv"
      dbms=csv replace;
      getnames=yes;
    run;

  *Executes macro (%noninf_loop) to perform noninferiority testing on the simulated data from prespecified
     settings for Type I Error;
    title 'Noninferiority Testing (Type I Error)';
    %noninf_loop(Work.sim_typeIerror);

ods rtf close;



ods rtf file = "E:\Simulation Study\Analysis\Noninferiority Testing\SAS Output\Maximal Power\
                      Noninferiority Testing (Maximal Power).doc";

  *Imports the control dataset [Simulation Settings (Maximal Power).csv] into SAS for reference;
    proc import out=Work.sim_power
      datafile="E:\Simulation Study\Data\SAS Output\Power\Simulation Settings (Maximal Power).csv"
      dbms=csv replace;
      getnames=yes;
    run;

  *Executes macro (%noninf_loop) to perform noninferiority testing on the simulated data from prespecified
     settings for Maximal Power;
    title 'Noninferiority Testing (Maximal Power)';
    %noninf_loop(Work.sim_power);

ods rtf close;
```

```
ods rtf file = "E:\Simulation Study\Analysis\Noninferiority Testing\SAS Output\Midpoint Power\
                     Noninferiority Testing (Midpoint Power).doc";

  *Imports the control dataset [Simulation Settings (Midpoint Power).csv] into SAS for reference;
    proc import out=Work.sim_midpoint
      datafile="E:\Simulation Study\Data\SAS Output\Midpoint Power\
                     Simulation Settings (Midpoint Power).csv"
      dbms=csv replace;
      getnames=yes;
    run;

  *Executes macro (%noninf_loop) to perform noninferiority testing on the simulated data from prespecified
     settings for Midpoint Power;
    title 'Noninferiority Testing (Midpoint Power)';
    %noninf_loop(Work.sim_midpoint);

ods rtf close;
```

Noninferiority Results

Below is SAS Code used to summarize the results obtained from noninferiority testing with the simulated data.

```
*****************************************************************************************************;
* Nick Bloedow's Master Project: Results of Noninferiority Testing on Two Binomial Proportions         *;
*                                  (Type I Error, Power, & Midpoint)                                    *;
*****************************************************************************************************;

/*
Notes: Summarizes the Results of Noninferiority (One-Sided Equivalence) Testing on Two Binomial Proportions
       Creates exportable dataset (results_ind) summarizing the individual performance of methods
*/



/*
  Macro to Summarize the results of Noninferiority Tests based on 90% Confidence Limits with
    Risk Difference, Relative Risk, & Odds Ratio

    Tests used for Risk Difference = Wald (w/ & w/o CC), Hauck-Anderson (HA), Farrington-Manning (FM),
      and Newcombe (w/ & w/o CC)
    Tests used for Relative Risk = Wald (Asymptotic)
    Tests used for Odds Ratio = Wald and SCORE

    iter = number of iterations for simulation
    n = size of individuals in each group
    pT = true proportion of success for Treatment group
    pC = true proportion of success for Control group
    phi = specified percentage decrease between Treatment & Control group
    seed1 = specified seed for randomization with the Treatment group
    seed2 = specified seed for randomization with the Control group

*/
```

```
%macro results_iter_n_pC_phi(iter,n,pT,pC,phi,seed1,seed2);

*Titles pertaining to specified parameters used in the simulation of data;
  title2 "Number of Iterations = &iter";
  title3 "Group Size = &n";
  title4 "True Proportions: Treatment = &pT    Control = &pC";
  title5 "Delta = &delta    Phi = &phi    Psi = &psi";
  title6 "Seeds: Treatment = &seed1    Control = &seed2";
  title7 " ";


*Creates a control dataset (ind_methods) of individual methods and variable names used with
   Risk Difference, Relative Risk, & Odds Ratio;
  data ind_methods;

      length Method $20;
      length VarName $24;

        Method='1_DIFF_WALD';
        VarName='Noninf_DIFF_WALD';
            output;

        Method='2_DIFF_FM';
        VarName='Noninf_DIFF_FM';
            output;

        Method='3_DIFF_NEWCOMBE';
        VarName='Noninf_DIFF_NEWCOMBE';
            output;

        Method='4_DIFF_WALD_CC';
        VarName='Noninf_DIFF_WALD_CC';
            output;

        Method='5_DIFF_HA';
        VarName='Noninf_DIFF_HA';
            output;

        Method='6_DIFF_NEWCOMBE_CC';
        VarName='Noninf_DIFF_NEWCOMBE_CC';
```

```
            output;

        Method='7_RR_WALD';
        VarName='Noninf_RR_WALD';
            output;

        Method='8_OR_WALD';
        VarName='Noninf_OR_WALD';
            output;

        Method='9_OR_SCORE';
        VarName='Noninf_OR_SCORE';
            output;

  run;

*Executes macro (%ind_loop) to summarize individual performance of noninferiority methods on
   simulated data from prespecified settings;
  %ind_loop(ind_methods)

*Sorts the outputted dataset (OneWay_Freq) by Result;
  proc sort data=OneWay_Freq;
      by Result;
  run;

*Transposes the outputted dataset (OneWay_Freq) according to Method by Result for Frequency &
   Percent statistics;
  proc transpose data=OneWay_Freq out=perform_ind(rename=(_NAME_=Stat));
    by Result;
    id Method;
    var Frequency Percent;
  run;

*Deletes temporary data set (OneWay_Freq) summarizing the individual performance for methods within
   each simulation;
  proc delete data=OneWay_Freq;
  run;
```

```
*Creates a final dataset (perform_ind) summarizing the performance of individual methods for each
   simulation;
  data perform_ind;
    set perform_ind;
    Simulation=&now;
    label Stat=" ";
  run;


%mend results_iter_n_pC_phi;
```

```
/*
  Macro to summarize results of individual methods with Noninferiority Testing on each simulation
    settings = control dataset that has the specified methods used with noninferiority testing
*/

  %macro ind_loop(settings);
    %local ind_dsid ind_rc ind_loop ind_rows ind_cols;
    %let ind_dsid = %sysfunc(open(&settings));
      %let ind_loop=0;
    %let ind_rows=%sysfunc(attrn(&ind_dsid,nobs));  %* loops;
    %let ind_cols=%sysfunc(attrn(&ind_dsid,nvars));  %* vars;

    %* loops = rows;
    %do %while (%sysfunc(fetch(&ind_dsid)) = 0);  %* outer loop across rows;

      %let ind_loop=%eval(&ind_loop + 1);

      %* get vars from cols;
      %do c = 1 %to &ind_cols;  %* inner nested loop;
        %local v t;
        %let v=%sysfunc(varname(&ind_dsid,&c));
        %local &v;
        %let t = %sysfunc(vartype(&ind_dsid, &c));  %* N or C;
        %let &v = %sysfunc(getvar&t(&ind_dsid, &c));  %* show var and value in log;
        %end;

      %put ** subloop # &ind_loop of &ind_rows **;

      *Creates indicator variables (INF & NON) for evaluating the individual performance of a
         noninferiority method;
        data OneWay_Ind;
          set noninf_&now;
          length Method $24;
          Method="&Method";

          if &VarName="INF"  then INF=1; else INF=0;
          if &VarName="NON"  then NON=1; else NON=0;

          keep Method INF NON;
```

```
  run;
*Sums the one-way counts for results of an individual method with noninferiority testing;
    proc means data=OneWay_Ind sum;
      class Method;
      var INF NON;
      ods output Summary=OneWay_Sum;
    run;


*Creates final dataset (OneWay_Final) for the performance of an individual method with noninferiority
   testing;
    data OneWay_Final;
    set OneWay_Sum;

    length Method $24;
    length Result $13;
    Method="&Method";

      Result="Inferior";
        Frequency=INF_Sum;
        Percent=100*round((INF_Sum/(INF_SUM+NON_SUM)),0.0001);
        output;

      Result="Noninferior";
        Frequency=NON_Sum;
        Percent=100*round((NON_Sum/(INF_SUM+NON_SUM)),0.0001);
        output;

      keep Method Result Frequency Percent;

    run;

  *Appends the current summary of an individual method into a master dataset (OneWay_Final);
  proc append base=OneWay_Freq data=OneWay_Final;
    run;
```

```
      *Deletes temporary data sets used to summarize the individual performance for methods within
         each simulation;
        proc delete data=Work.OneWay_Ind;
        run;

        proc delete data=Work.OneWay_Sum;
        run;

        proc delete data=Work.OneWay_Final;
        run;

    %out:
    %end; %* while fetch loop;

  %let ind_rc = %sysfunc(close(&ind_dsid));

%mend ind_loop;
```

```
/*
  Macro to Summarize the Noninferiority Results on all specified Simulation Settings
    settings = control dataset that has the specified simulation settings (w/ proper column names)
*/

  %macro results_loop(settings);
    %local dsid rc now rows cols;
    %let dsid = %sysfunc(open(&settings));
      %let now=0;
    %let rows=%sysfunc(attrn(&dsid,nobs));  %* loops;
    %let cols=%sysfunc(attrn(&dsid,nvars));  %* vars;

    %* loops = rows;
    %do %while (%sysfunc(fetch(&dsid)) = 0);  %* outer loop across rows;

      %let now=%eval(&now + 1);

      %* get vars from cols;
      %do c = 1 %to &cols;  %* inner nested loop;
        %local v t;
        %let v=%sysfunc(varname(&dsid,&c));
        %local &v;
        %let t = %sysfunc(vartype(&dsid, &c));  %* N or C;
        %let &v = %sysfunc(getvar&t(&dsid, &c));  %* show var and value in log;
        %end;

      %put ** loop # &now of &rows **;

        *Imports the generated dataset for each simulation;
          proc import out=Work.noninf_&now
            datafile="E:\Simulation Study\Analysis\Noninferiority Testing\SAS Output\&noninf_results"
            dbms=csv replace;
            getnames=yes;
          run;

        *Summarizes the noninferiority (individual, across methods, & overall) results for each simulation;
        %results_iter_n_pC_phi(iter=&iter,n=&n,pT=&pT,pC=&pC,phi=&phi,seed1=&seed1,seed2=&seed2)
```

```
      *Appends the individual noninferiority results for each simulation into a master dataset
        (results_ind);
        proc append base=results_ind      data=perform_ind;
        run;

      *Deletes temporary data sets used for creating the final noninferiority results dataset for
         each simulation;
        proc delete data=Work.ind_methods;
        run;

        proc delete data=Work.perform_ind;
        run;

    %out:
    %end; %* while fetch loop;

  %let rc = %sysfunc(close(&dsid));

%mend results_loop;
```

```
ods rtf file = "E:\Simulation Study\Results\Noninferiority Testing\SAS Output\Type I Error\
                     Noninferiority Results (Type I Error).doc";

  *Imports the control dataset [Simulation Settings (Type I Error).csv] into SAS for reference with
     the Type I Error simulations;
    proc import out=Work.sim_typeIerror
      datafile="E:\Simulation Study\Data\SAS Output\Type I Error\Simulation Settings (Type I Error).csv"
      dbms=csv replace;
      getnames=yes;
    run;

  *Creates a variable (Simulation) corresponding to simulation number associated with the Type I Error
     simulation control dataset;
    data sim_typeIerror;
      set sim_typeIerror;
      Simulation=_n_;
    run;

  *Executes macro (%results_loop) to summarize results of noninferiority testing on the simulated data
     from prespecified settings;
    title 'Noninferiority Results (Type I Error)';
    %results_loop(Work.sim_typeIerror);

  *Merges the individual results (results_ind) with the simulation settings for the type I error analysis
     (sim_typeIerror) by simulation number;
    data results_ind;
      merge results_ind sim_typeIerror;
      by Simulation;
      drop Simulation sim_data noninf_results;
    run;

  *Exports the final noninferiority individual results dataset (results_ind) from type I error simulation
     into a separate csv file;
    proc export data=Work.results_ind
      outfile="E:\Simulation Study\Results\Noninferiority Testing\SAS Output\Type I Error\
                    Noninferiority Results_Type I Error (Individual).csv" dbms=csv replace;
    run;

ods rtf close;
```

```
ods rtf file = "E:\Simulation Study\Results\Noninferiority Testing\SAS Output\Maximal Power\
                     Noninferiority Results (Maximal Power).doc";

  *Imports the control dataset [Simulation Settings (Power).csv] into SAS for reference;
    proc import out=Work.sim_power
      datafile="E:\Simulation Study\Data\SAS Output\Maximal Power\Simulation Settings (Maximal Power).csv"
      dbms=csv replace;
      getnames=yes;
    run;

  *Creates a variable (Simulation) corresponding to simulation number associated with the Power simulation
     control dataset;
    data sim_power;
      set sim_power;
      Simulation=_n_;
    run;

  *Executes macro (%results_loop) to summarize results of noninferiority testing on the simulated data
     from prespecified settings;
    title 'Noninferiority Results (Maximal Power)';
    %results_loop(Work.sim_power);

  *Merges the individual results (results_ind) with the simulation settings for the power analysis
     (sim_power) by simulation number;
    data results_ind;
      merge results_ind sim_power;
      by Simulation;
      drop Simulation sim_data noninf_results;
    run;

  *Exports the final noninferiority individual results dataset (results_ind) from power simulation into
     a separate csv file;
    proc export data=Work.results_ind
      outfile="E:\Simulation Study\Results\Noninferiority Testing\SAS Output\Maximal Power\
                    Noninferiority Results_Maximal Power (Individual).csv" dbms=csv replace;
    run;

ods rtf close;
```

```
ods rtf file = "E:\Simulation Study\Results\Noninferiority Testing\SAS Output\Midpoint Power\
                     Noninferiority Results (Midpoint).doc";

  *Imports the control dataset [Simulation Settings (Midpoint Power).csv] into SAS for reference;
    proc import out=Work.sim_midpoint
      datafile="E:\Simulation Study\Data\SAS Output\Midpoint Power\
                     Simulation Settings (Midpoint Power).csv"
      dbms=csv replace;
      getnames=yes;
    run;

  *Creates a variable (Simulation) corresponding to simulation number associated with the Midpoint
     simulation control dataset;
    data sim_midpoint;
      set sim_midpoint;
      Simulation=_n_;
    run;

  *Executes macro (%results_loop) to summarize results of noninferiority testing on the simulated data from
     prespecified settings;
    title 'Noninferiority Results (Midpoint Power)';
    %results_loop(Work.sim_midpoint);

  *Merges the individual results (results_ind) with the simulation settings for the midpoint power analysis
     (sim_midpoint) by simulation number;
    data results_ind;
      merge results_ind sim_midpoint;
      by Simulation;
      drop Simulation sim_data noninf_results;
    run;

  *Exports the final noninferiority individual results dataset (results_ind) from midpoint power simulation
     into a separate csv file;
    proc export data=Work.results_ind
      outfile="E:\Simulation Study\Results\Noninferiority Testing\SAS Output\Midpoint Power\
                    Noninferiority Results_Midpoint Power (Individual).csv" dbms=csv replace;
    run;

ods rtf close;
```

Appendix D - Supplemental Graphics Code

Contour & 3D Plots of Noninferiority Parameters

Below is R Code used to create contour and 3D plots to visualize the possible range of values for noninferiority parameters (Risk Difference, Relative Risk, and Odds Ratio).

```
#############################
#  Nicholas Bloedow         #
#  Kansas State University  #
#  Master's Report          #
#  Advisor: Dr. Vahl        #
#############################

                               #####################################################
                               #  Contour & 3D Plots of Noninferiority Parameters  #
                               #####################################################

##################
#  General Case  #
##################

# Generates ranges of possible values for Risk Difference, Relative Risk, and Odds Ratio corresponding
#   to subspace of pT & pC

   pT<-rep(seq(0.05,0.95,0.01),each=91)
   pC<-rep(seq(0.05,0.95,0.01),91)


   diff<-(pT-pC)
   RD<-array(diff,c(91,91))


   relrisk<-(pT/pC)
   RR<-array(relrisk,c(91,91))


   oddsratio<-((pT*(1-pC))/(pC*(1-pT)))
   OR<-array(oddsratio,c(91,91))
```

```
# Creates 2D Contour Plots to show range for Noninferiority parameters corresponding to subspace of pT & pC

   library(plot3D)

   par(mfrow = c(1,3))

   image2D(RD, x=seq(0.05,0.95,0.01), y=seq(0.05,0.95,0.01),
        xlab = "P(Success) for Control", ylab = "P(Success) for Treatment",
        contour=list(levels=c(-0.75,-0.5,-0.25,0,0.25,0.5,0.75),col="grey", lwd=2),
        shade=0.1, main="Risk Difference", clab=" ")

   image2D(RR, x=seq(0.05,0.95,0.01), y=seq(0.05,0.95,0.01),
        xlab = "P(Success) for Control", ylab = "P(Success) for Treatment",
        contour=list(levels=c(8,4,2,(4/3),1,0.75,0.5,0.25,0.125),col="grey", lwd=2),
        shade=0.1, main="Relative Risk", clab=" ")

   image2D(OR, x=seq(0.05,0.95,0.01), y=seq(0.05,0.95,0.01),
        xlab = "P(Success) for Control", ylab = "P(Success) for Treatment",
        contour=list(levels=c(50,20,10,4,2,1,0.5,0.25,0.10,0.05,0.02),col="grey", lwd=2),
        shade=0.1, main="Odds Ratio", clab=" ")



# Creates 3D Plots to show range for Noninferiority parameters corresponding to subspace of pT & pC

   par(mfrow = c(1,3))

   persp3D(z = RD, xlab = "P(Success) for Control", bty = "bl2",
        ylab = "P(Success) for Treatment", zlab = "Risk Difference", main = "Risk Difference",
        expand = 0.5, d = 2, phi = 20, theta = 30, resfac = 2,
        contour = list(col = "grey",levels=c(-0.75,-0.5,-0.25,0,0.25,0.5,0.75),side = c("zmin", "z")),
        zlim = c(-1,1), colkey = list(side = 1, length = 0.5))

   persp3D(z = RR, xlab = "P(Success) for Control", bty = "bl2",
        ylab = "P(Success) for Treatment", zlab = "Relative Risk", main = "Relative Risk",
        expand = 0.5, d = 2, phi = 20, theta = 30, resfac = 2,
        contour = list(col="grey",levels=c(8,4,2,(4/3),1,0.75,0.5,0.25,0.125),side = c("zmin", "z")),
        zlim = c(0,20), colkey = list(side = 1, length = 0.5))

   persp3D(z = OR, xlab = "P(Success) for Control", bty = "bl2",
        ylab = "P(Success) for Treatment", zlab = "Odds Ratio", main = "Odds Ratio",
        expand = 0.5, d = 2, phi = 20, theta = 30, resfac = 2,
        contour = list(col="grey",levels=c(50,20,10,5,2,1,0.5,0.25,0.10,0.05,0.02),side = c("zmin", "z")),
        zlim = c(0,360), colkey = list(side = 1, length = 0.5))
```

Noninferiority Graphics

Below is SAS Code used in creating graphics to summarize the results obtained from noninferiority testing on the simulated data.

```
*******************************************************************************************************;
* Nick Bloedow's Master Project: Graphics from Noninferiority Testing on Two Binomial Proportions        *;
*                                   (Type I Error, Maximal Power, & Midpoint Power)                      *;
*******************************************************************************************************;

/*
Notes: Generates Graphics Displaying Results from Noninferiority Testing on Two Binomial Proportions
         associated with previously discussed Simulation Study
*/


ods rtf file = " E:\Simulation Study\Graphics\Noninferiority Testing\SAS Output\
                      Noninferiority Graphics.doc";


  *Imports the Type I Error results (results_typeIerror) from the Noniferirority Testing with simulated
     data into SAS;
    proc import out=Work.results_typeIerror
      datafile="E:\Simulation Study\Results\Noninferiority Testing\SAS Output\Type I Error\
                     Noninferiority Results_Type I Error (Individual).csv" dbms=csv replace;
      getnames=yes;
    run;

  *Isolates results pertaining to the noninferiority percentages and creates variable corresponding to
     simulation number;
      data results_typeIerror;
        set results_typeIerror;
        where Result="Noninferior" & Stat="Percent";
        Simulation=_n_;
      run;
```

```
*Transposes the Type I Error noninferiority results into the proper form for generating lattice plots;
  proc transpose data=results_typeIerror out=trans_typeIerror(rename=(Col1=TypeIError)) name=Variable;
    by Simulation;
    var _1_DIFF_WALD _2_DIFF_FM _3_DIFF_NEWCOMBE _4_DIFF_WALD_CC _5_DIFF_HA _6_DIFF_NEWCOMBE_CC
        _7_RR_WALD _8_OR_WALD _9_OR_SCORE ;
  run;

*Creates a data set (panel_typeIerror) with descriptive variable names and simulation settings used for
   Noninferiority Testing;
    data panel_typeIerror;
      merge trans_typeIerror results_typeIerror(keep=simulation iter n pT pC delta phi psi);
      by Simulation;
      label Variable=' ';
      Stat_Method=substr(Variable,4);
      Statistic=substr(Stat_Method,1,index(Stat_Method,'_') -1 );
      Method=trim(left(substr(Stat_Method, index(Stat_Method,'_') + 1)));
      if Statistic="DIFF" then Statistic="Risk Difference";
      if Statistic="RR"   then Statistic="Relative Risk";
      if Statistic="OR"   then Statistic="Odds Ratio";
      if Method="WALD_CC" then Method="WALD";
      if Method="NEWCOMBE_CC" then Method="NEWCOMBE";
      drop Simulation;
    run;

*Sorts the Type I Error panel data by Group Size, Variable Name, Phi, and then pC;
    proc sort data=panel_typeIerror;
      by n Variable phi pC ;
    run;
```

```
ods html file="Noninferiority Graphics (Type I Error).html"
         gpath="E:\Simulation Study\Graphics\Noninferiority Testing\SAS Output\Type I Error";

 *Creates a Lattice Plot of Vertical Bar Charts of Type I Error Rate across pC and phi for all Group Size
    and Methods;
   ods graphics on / height=700px width=700px imagefmt=png imagename="TypeIError_Bar_All" border=off;

   proc sgpanel data=panel_typeIerror;
     panelby pC phi / layout=lattice columns=5 rows=5 ;
     colaxis display=none;
     rowaxis label="Type I Error Rate (Percent)" Min=3 Max=7 Values=(3 to 7 by 1);
     vbar Variable / response=TypeIError group=Stat_Method dataskin=pressed ;
     refline 4.4 5.6 / axis=y lineattrs=(color=blue thickness=1);
     refline 0 / axis=y lineattrs=(color=gray thickness=1);
     refline 5 / axis=y lineattrs=(color=red thickness=1);
     by n;
   run;

   ods graphics off;


 *Creates a Lattice Plot of Line Graphs for Type I Error Rate across Group Size and Method for Risk
    Difference;
   ods graphics on / height=600px width=600px imagefmt=png imagename="TypeIError_RiskDiff_woCC"
     border=off;

   proc sgpanel data=panel_typeIerror;
     where Stat_Method="DIFF_WALD" | Stat_Method="DIFF_FM" | Stat_Method="DIFF_NEWCOMBE";
     panelby n Method / layout=lattice columns=3 rows=3;
     colaxis label="P(Success) for Control Group" ;
     rowaxis label="Type I Error Rate (Percent)" Min=3 Max=7 Values=(3 to 7 by 1);
     vline pC / response=TypeIError group=phi lineattrs=(thickness=2);
     refline 4.4 5.6 / axis=y lineattrs=(color=blue thickness=1);
     refline 0 / axis=y lineattrs=(color=gray thickness=1);
     refline 5 / axis=y lineattrs=(color=red thickness=1);
   run;

   ods graphics off;
```

```
*Creates a Lattice Plot of Line Graphs for Type I Error Rate across Group Size and Method for Risk
   Difference (w/ Continuity Correction);
  ods graphics on / height=600px width=600px imagefmt=png imagename="TypeIError_RiskDiff_wCC" border=off;

  proc sgpanel data=panel_typeIerror;
    where Stat_Method="DIFF_WALD_CC" | Stat_Method="DIFF_HA" | Stat_Method="DIFF_NEWCOMBE_CC";
    panelby n Method / layout=lattice columns=3 rows=3;
    colaxis label="P(Success) for Control Group" ;
    rowaxis label="Type I Error Rate (Percent)" Min=3 Max=7 Values=(3 to 7 by 1);
    vline pC / response=TypeIError group=phi lineattrs=(thickness=2);
    refline 4.4 5.6 / axis=y lineattrs=(color=blue thickness=1);
    refline 0 / axis=y lineattrs=(color=gray thickness=1);
    refline 5 / axis=y lineattrs=(color=red thickness=1);
  run;

  ods graphics off;


*Creates a Lattice Plot of Line Graphs for Type I Error Rate across Group Size and Method for Relative
   Risk;
  ods graphics on / height=300px width=600px imagefmt=png imagename="TypeIError_RelRisk" border=off;

  proc sgpanel data=panel_typeIerror;
    where Stat_Method="RR_WALD";
    panelby n Method / layout=lattice columns=3 rows=1;
    colaxis label="P(Success) for Control Group" ;
    rowaxis label="Type I Error Rate (Percent)" Min=3 Max=7 Values=(3 to 7 by 1);
    vline pC / response=TypeIError group=phi lineattrs=(thickness=2);
    refline 4.4 5.6 / axis=y lineattrs=(color=blue thickness=1);
    refline 0 / axis=y lineattrs=(color=gray thickness=1);
    refline 5 / axis=y lineattrs=(color=red thickness=1);
  run;

  ods graphics off;
```

```
*Creates a Lattice Plot of Line Graphs for Type I Error Rate across Group Size and Method for Odds Ratio;
  ods graphics on / height=450px width=600px imagefmt=png imagename="TypeIError_OddsRatio" border=off;

  proc sgpanel data=panel_typeIerror;
    where Stat_Method="OR_WALD" | Stat_Method="OR_SCORE";
    panelby n Method / layout=lattice columns=3 rows=2;
    colaxis label="P(Success) for Control Group" ;
    rowaxis label="Type I Error Rate (Percent)" Min=3 Max=7 Values=(3 to 7 by 1);
    vline pC / response=TypeIError group=phi lineattrs=(thickness=2);
    refline 4.4 5.6 / axis=y lineattrs=(color=blue thickness=1);
    refline 0 / axis=y lineattrs=(color=gray thickness=1);
    refline 5 / axis=y lineattrs=(color=red thickness=1);
  run;

  ods graphics off;


*Creates a Lattice Plot of Vertical Bar Charts of Type I Error Rate across pC and phi for all Group Sizes
   and Statistics with Asymptotic (WALD) Method;
  ods graphics on / height=600px width=600px imagefmt=png imagename="TypeIError_Bar_Across" border=off;

  proc sgpanel data=panel_typeIerror;
    where Stat_Method="DIFF_WALD" | Stat_Method="RR_WALD" | Stat_Method="OR_WALD";
    panelby pC phi / layout=lattice columns=5 rows=5 ;
    colaxis display=none;
    rowaxis label="Type I Error Rate (Percent)" Min=3 Max=7 Values=(3 to 7 by 1);
    vbar Statistic / response=TypeIError group=Statistic dataskin=pressed ;
    refline 4.4 5.6 / axis=y lineattrs=(color=blue thickness=1);
    refline 0 / axis=y lineattrs=(color=gray thickness=1);
    refline 5 / axis=y lineattrs=(color=red thickness=1);
    by n;
  run;

  ods graphics off;
```

```
 *Creates a Lattice Plot of Line Graphs for Type I Error Rate across Group Size and Statistics with
    Asymptotic (WALD) Method;
   ods graphics on / height=600px width=600px imagefmt=png imagename="TypeIError_Line_Across" border=off;

   proc sgpanel data=panel_typeIerror;
     where Stat_Method="DIFF_WALD" | Stat_Method="RR_WALD" | Stat_Method="OR_WALD";
     panelby n Statistic / layout=lattice columns=3 rows=3 ;
     colaxis label="P(Success) for Control Group" ;
     rowaxis label="Type I Error Rate (Percent)" Min=3 Max=7 Values=(3 to 7 by 1);
     vline pC / response=TypeIError group=phi lineattrs=(thickness=2) ;
     refline 4.4 5.6 / axis=y lineattrs=(color=blue thickness=1);
     refline 0 / axis=y lineattrs=(color=gray thickness=1);
     refline 5 / axis=y lineattrs=(color=red thickness=1);
   run;

   ods graphics off;


ods html close;
```

```
*Imports the Maximal Power results (results_power) from the Noninferiority Testing with simulated data
   into SAS;
  proc import out=Work.results_power
    datafile="E:\Simulation Study\Results\Noninferiority Testing\SAS Output\Maximal Power\
                   Noninferiority Results_Maximal Power (Individual).csv" dbms=csv replace;
    getnames=yes;
  run;

*Isolates results pertaining to the noninferiority percentages and creates variable corresponding to
   simulation number;
    data results_power;
      set results_power;
      where Result="Noninferior" & Stat="Percent";
      Simulation=_n_;
    run;

*Transposes the Maximal Power noninferiority results into the proper form for generating lattice plots;
  proc transpose data=results_power out=trans_power(rename=(Col1=Power)) name=Variable;
    by Simulation;
    var _1_DIFF_WALD _2_DIFF_FM _3_DIFF_NEWCOMBE _4_DIFF_WALD_CC _5_DIFF_HA _6_DIFF_NEWCOMBE_CC
        _7_RR_WALD _8_OR_WALD _9_OR_SCORE ;
  run;

*Creates a data set (panel_power) with descriptive variable names and simulation settings used for
   Noninferiority Testing;
    data panel_power;
      merge trans_power results_power(keep=simulation iter n pT pC delta phi psi);
      by Simulation;
      label Variable=' ';
      Stat_Method=substr(Variable,4);
      Statistic=substr(Stat_Method,1,index(Stat_Method,'_') -1 );
      Method=trim(left(substr(Stat_Method, index(Stat_Method,'_') + 1)));
      if Statistic="DIFF" then Statistic="Risk Difference";
      if Statistic="RR"   then Statistic="Relative Risk";
      if Statistic="OR"   then Statistic="Odds Ratio";
      if Method="WALD_CC" then Method="WALD";
      if Method="NEWCOMBE_CC" then Method="NEWCOMBE";
      drop Simulation;
    run;
```

```
*Sorts the Maximal Power panel data by Group Size, Variable Name, Phi, and then pC;
    proc sort data=panel_power;
      by n Variable phi pC ;
    run;


ods html file="Noninferiority Graphics (Maximal Power).html"
         gpath="E:\Simulation Study\Graphics\Noninferiority Testing\SAS Output\Maximal Power";

*Creates a Lattice Plot of Vertical Bar Charts of Maximal Power across pC and phi for all Group Size
   and Methods;
  ods graphics on / height=700px width=700px imagefmt=png imagename="Power_Bar_All" border=off;

  proc sgpanel data=panel_power;
    panelby pC phi / layout=lattice columns=5 rows=5 ;
    colaxis display=none;
    rowaxis label="Maximal Power (Percent)" Min=0 Max=100 Values=(0 to 100 by 20);
    vbar Variable / response=Power group=Stat_Method dataskin=pressed ;
    refline 0 100 / axis=y lineattrs=(color=gray thickness=1);
    refline 80 / axis=y lineattrs=(color=red thickness=1);
    by n;
  run;

 ods graphics off;

*Creates a Lattice Plot of Line Graphs for Maximal Power across Group Size and Method for Risk
   Difference;
  ods graphics on / height=600px width=600px imagefmt=png imagename="Power_RiskDiff_woCC" border=off;

  proc sgpanel data=panel_power;
    where Stat_Method="DIFF_WALD" | Stat_Method="DIFF_FM" | Stat_Method="DIFF_NEWCOMBE";
    panelby n Method / layout=lattice columns=3 rows=3;
    colaxis label="P(Success) for Control Group" ;
    rowaxis label=" Maximal Power (Percent)" Min=0 Max=100 Values=(0 to 100 by 20);
    vline pC / response=Power group=phi lineattrs=(thickness=2);
    refline 0 100 / axis=y lineattrs=(color=gray thickness=1);
    refline 80 / axis=y lineattrs=(color=red thickness=1);
  run;

  ods graphics off;
```

```
*Creates a Lattice Plot of Line Graphs for Maximal Power across Group Size and Method for Risk Difference
   (w/ Continuity Correction);
  ods graphics on / height=600px width=600px imagefmt=png imagename="Power_RiskDiff_wCC" border=off;

  proc sgpanel data=panel_power;
    where Stat_Method="DIFF_WALD_CC" | Stat_Method="DIFF_HA" | Stat_Method="DIFF_NEWCOMBE_CC";
    panelby n Method / layout=lattice columns=3 rows=3;
    colaxis label="P(Success) for Control Group" ;
    rowaxis label="Maximal Power (Percent)" Min=0 Max=100 Values=(0 to 100 by 20);
    vline pC / response=Power group=phi lineattrs=(thickness=2);
    refline 0 100 / axis=y lineattrs=(color=gray thickness=1);
    refline 80 / axis=y lineattrs=(color=red thickness=1);
  run;

  ods graphics off;


*Creates a Lattice Plot of Line Graphs for Maximal Power across Group Size and Method for Relative Risk;
  ods graphics on / height=300px width=600px imagefmt=png imagename="Power_RelRisk" border=off;

  proc sgpanel data=panel_power;
    where Stat_Method="RR_WALD";
    panelby n Method / layout=lattice columns=3 rows=1;
    colaxis label="P(Success) for Control Group" ;
    rowaxis label="Maximal Power (Percent)" Min=0 Max=100 Values=(0 to 100 by 20);
    vline pC / response=Power group=phi lineattrs=(thickness=2);
    refline 0 100 / axis=y lineattrs=(color=gray thickness=1);
    refline 80 / axis=y lineattrs=(color=red thickness=1);
  run;

  ods graphics off;
```

```
*Creates a Lattice Plot of Line Graphs for Maximal Power across Group Size and Method for Odds Ratio;
  ods graphics on / height=450px width=600px imagefmt=png imagename="Power_OddsRatio" border=off;

  proc sgpanel data=panel_power;
    where Stat_Method="OR_WALD" | Stat_Method="OR_SCORE";
    panelby n Method / layout=lattice columns=3 rows=2;
    colaxis label="P(Success) for Control Group" ;
    rowaxis label="Maximal Power (Percent)" Min=0 Max=100 Values=(0 to 100 by 20);
    vline pC / response=Power group=phi lineattrs=(thickness=2);
    refline 0 100 / axis=y lineattrs=(color=gray thickness=1);
    refline 80 / axis=y lineattrs=(color=red thickness=1);
  run;

 ods graphics off;


*Creates a Lattice Plot of Vertical Bar Charts of Maximal Power across pC and phi for all Group Sizes and
   Statistics with Asymptotic (WALD) Method;
  ods graphics on / height=600px width=600px imagefmt=png imagename="Power_Bar_Across" border=off;

  proc sgpanel data=panel_power;
      where Stat_Method="DIFF_WALD" | Stat_Method="RR_WALD" | Stat_Method="OR_WALD";
    panelby pC phi / layout=lattice columns=5 rows=5 ;
      colaxis display=none;
      rowaxis label="Maximal Power (Percent)" Min=0 Max=100 Values=(0 to 100 by 20);
    vbar Statistic / response=Power group=Statistic dataskin=pressed ;
      refline 0 100 / axis=y lineattrs=(color=gray thickness=1);
      refline 80 / axis=y lineattrs=(color=red thickness=1);
      by n;
  run;

  ods graphics off;
```

```
*Creates a Lattice Plot of Line Graphs for Maximal Power across Group Size and Statistics for
   Asymptotic (WALD) Method;
  ods graphics on / height=600px width=600px imagefmt=png imagename="Power_Line_Across" border=off;

  proc sgpanel data=panel_power;
    where Stat_Method="DIFF_WALD" | Stat_Method="RR_WALD" | Stat_Method="OR_WALD";
    panelby n Statistic / layout=lattice columns=3 rows=3 ;
    colaxis label="P(Success) for Control Group" ;
    rowaxis label="Maximal Power (Percent)" Min=0 Max=100 Values=(0 to 100 by 20);
    vline pC / response=Power group=phi lineattrs=(thickness=2);
    refline 0 100 / axis=y lineattrs=(color=gray thickness=1);
    refline 80 / axis=y lineattrs=(color=red thickness=1);
  run;

  ods graphics off;

ods html close;
```

```
*Imports the Midpoint Power results (results_midpoint) from the Noninferiority Testing with simulated
   data into SAS;
  proc import out=Work.results_midpoint
    datafile="E:\Simulation Study\Results\Noninferiority Testing\SAS Output\Midpoint Power\
                   Noninferiority Results_Midpoint Power (Individual).csv" dbms=csv replace;
    getnames=yes;
  run;

*Isolates results pertaining to the noninferiority percentages and creates variable corresponding to
   simulation number;
    data results_midpoint;
      set results_midpoint;
      where Result="Noninferior" & Stat="Percent";
      Simulation=_n_;
    run;

*Transposes the Midpoint Power noninferiority results into the proper form for generating lattice plots;
  proc transpose data=results_midpoint out=trans_midpoint(rename=(Col1=Power)) name=Variable;
    by Simulation;
    var _1_DIFF_WALD _2_DIFF_FM _3_DIFF_NEWCOMBE _4_DIFF_WALD_CC _5_DIFF_HA _6_DIFF_NEWCOMBE_CC
        _7_RR_WALD _8_OR_WALD _9_OR_SCORE ;
  run;

*Creates a data set (panel_midpoint) with descriptive variable names and simulation settings used for
   Noninferiority Testing;
    data panel_midpoint;
      merge trans_midpoint results_midpoint(keep=simulation iter n pT pC delta phi psi);
      by Simulation;
      label Variable=' ';
      Stat_Method=substr(Variable,4);
      Statistic=substr(Stat_Method,1,index(Stat_Method,'_') -1 );
      Method=trim(left(substr(Stat_Method, index(Stat_Method,'_') + 1)));
      if Statistic="DIFF" then Statistic="Risk Difference";
      if Statistic="RR"   then Statistic="Relative Risk";
      if Statistic="OR"   then Statistic="Odds Ratio";
      if Method="WALD_CC" then Method="WALD";
      if Method="NEWCOMBE_CC" then Method="NEWCOMBE";
      drop Simulation;
    run;
```

```
*Sorts the Midpoint Power panel data by Group Size, Variable Name, Phi, and then pC;
    proc sort data=panel_midpoint;
      by n Variable phi pC ;
    run;

ods html file="Noninferiority Graphics (Midpoint Power).html"
         gpath="E:\Simulation Study\Graphics\Noninferiority Testing\SAS Output\Midpoint Power";

*Creates a Lattice Plot of Vertical Bar Charts of Midpoint Power across pC and phi for all Group Size
   and Methods;
  ods graphics on / height=700px width=700px imagefmt=png imagename="Midpoint_Bar_All" border=off;

  proc sgpanel data=panel_midpoint;
    panelby pC phi / layout=lattice columns=5 rows=5 ;
    colaxis display=none;
    rowaxis label="Midpoint Power (Percent)" Min=0 Max=100 Values=(0 to 100 by 20);
    vbar Variable / response=Power group=Stat_Method dataskin=pressed ;
    refline 0 100 / axis=y lineattrs=(color=gray thickness=1);
    refline 80 / axis=y lineattrs=(color=red thickness=1);
    by n;
  run;

 ods graphics off;


*Creates a Lattice Plot of Line Graphs for Midpoint Power across Group Size and Method for Risk
   Difference;
  ods graphics on / height=600px width=600px imagefmt=png imagename="Midpoint_RiskDiff_woCC" border=off;

  proc sgpanel data=panel_midpoint;
    where Stat_Method="DIFF_WALD" | Stat_Method="DIFF_FM" | Stat_Method="DIFF_NEWCOMBE";
    panelby n Method / layout=lattice columns=3 rows=3;
    colaxis label="P(Success) for Control Group" ;
    rowaxis label="Midpoint Power (Percent)" Min=0 Max=100 Values=(0 to 100 by 20);
    vline pC / response=Power group=phi lineattrs=(thickness=2);
    refline 0 100 / axis=y lineattrs=(color=gray thickness=1);
    refline 80 / axis=y lineattrs=(color=red thickness=1);
  run;

  ods graphics off;
```

```
*Creates a Lattice Plot of Line Graphs for Midpoint Power across Group Size and Method for Risk
   Difference (w/ Continuity Correction);
  ods graphics on / height=600px width=600px imagefmt=png imagename="Midpoint_RiskDiff_wCC" border=off;

  proc sgpanel data=panel_midpoint;
    where Stat_Method="DIFF_WALD_CC" | Stat_Method="DIFF_HA" | Stat_Method="DIFF_NEWCOMBE_CC";
    panelby n Method / layout=lattice columns=3 rows=3;
    colaxis label="P(Success) for Control Group" ;
    rowaxis label="Midpoint Power (Percent)" Min=0 Max=100 Values=(0 to 100 by 20);
    vline pC / response=Power group=phi lineattrs=(thickness=2);
    refline 0 100 / axis=y lineattrs=(color=gray thickness=1);
    refline 80 / axis=y lineattrs=(color=red thickness=1);
  run;

  ods graphics off;


*Creates a Lattice Plot of Line Graphs for Midpoint Power across Group Size and Method for Relative Risk;
  ods graphics on / height=300px width=600px imagefmt=png imagename="Midpoint_RelRisk" border=off;

  proc sgpanel data=panel_midpoint;
    where Stat_Method="RR_WALD";
    panelby n Method / layout=lattice columns=3 rows=1;
    colaxis label="P(Success) for Control Group" ;
    rowaxis label="Midpoint Power (Percent)" Min=0 Max=100 Values=(0 to 100 by 20);
    vline pC / response=Power group=phi lineattrs=(thickness=2);
    refline 0 100 / axis=y lineattrs=(color=gray thickness=1);
    refline 80 / axis=y lineattrs=(color=red thickness=1);
  run;

  ods graphics off;
```

```
*Creates a Lattice Plot of Line Graphs for Midpoint Power across Group Size and Method for Odds Ratio;
  ods graphics on / height=450px width=600px imagefmt=png imagename="Midpoint_OddsRatio" border=off;

  proc sgpanel data=panel_midpoint;
    where Stat_Method="OR_WALD" | Stat_Method="OR_SCORE";
    panelby n Method / layout=lattice columns=3 rows=2;
    colaxis label="P(Success) for Control Group" ;
    rowaxis label="Midpoint Power (Percent)" Min=0 Max=100 Values=(0 to 100 by 20);
    vline pC / response=Power group=phi lineattrs=(thickness=2);
    refline 0 100 / axis=y lineattrs=(color=gray thickness=1);
    refline 80 / axis=y lineattrs=(color=red thickness=1);
  run;

  ods graphics off;


*Creates a Lattice Plot of Vertical Bar Charts of Midpoint Power across pC and phi for all Group Sizes
   and Statistics with Asymptotic (WALD) Method;
  ods graphics on / height=600px width=600px imagefmt=png imagename="Midpoint_Bar_Across" border=off;

  proc sgpanel data=panel_midpoint;
    where Stat_Method="DIFF_WALD" | Stat_Method="RR_WALD" | Stat_Method="OR_WALD";
    panelby pC phi / layout=lattice columns=5 rows=5 ;
    colaxis display=none;
    rowaxis label="Midpoint Power (Percent)" Min=0 Max=100 Values=(0 to 100 by 20);
    vbar Statistic / response=Power group=Statistic dataskin=pressed ;
    refline 0 100 / axis=y lineattrs=(color=gray thickness=1);
    refline 80 / axis=y lineattrs=(color=red thickness=1);
    by n;
  run;

  ods graphics off;
```

```
  *Creates a Lattice Plot of Line Graphs for Midpoint Power across Group Size and Statistics for
     Asymptotic (WALD) Method;
    ods graphics on / height=600px width=600px imagefmt=png imagename="Midpoint_Line_Across" border=off;

    proc sgpanel data=panel_midpoint;
      where Stat_Method="DIFF_WALD" | Stat_Method="RR_WALD" | Stat_Method="OR_WALD";
      panelby n Statistic / layout=lattice columns=3 rows=3 ;
      colaxis label="P(Success) for Control Group" ;
      rowaxis label="Midpoint Power (Percent)" Min=0 Max=100 Values=(0 to 100 by 20);
      vline pC / response=Power group=phi lineattrs=(thickness=2);
      refline 0 100 / axis=y lineattrs=(color=gray thickness=1);
      refline 80 / axis=y lineattrs=(color=red thickness=1);
    run;

    ods graphics off;


  ods html close;


ods rtf close;
```